Witnessing latent time correlations with a single quantum particle

Hlér Kristjánsson,1,2 Wenxu Mao (貌文徐),3,4 and Giulio Chiribella3,1,2,5,∗

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, NY 10044, United States of America
5 Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario, N2L 2Y5, Canada

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 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, 7]. For example, photons transmitted through an optical fibre are subject to random changes in their polarisation [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, 12–25].

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 travelling at different moments of time, as illustrated in Figure 1. Taking advantage of the time correlations in the noise, we show that it is possible to enhance the amount of information that a 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
to reproduce the advantages of the quantum SWITCH we show that

advantages in quantum communication \[42, 45–50\]. Here setups inspired by the quantum SWITCH correlated channels underlie all the existing experimental perposition of two alternative orders. In practice, time-
time correlations are essential in or-
tum

that combines two variable quantum channels in a su-
terial, they have been used to reproduce the action of the quantum SWITCH of two completely depolarising 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 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 favourable for the transmission of classical information: while the quantum SWITCH of two completely depolarising channels can at most yield 0.049 bits of classical communication \[45, 50\], a more sophisticated pattern of time correlations yields the communication of 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 Section \[II\] 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 Section \[III\] we consider the transmission of a single particle at a superposition of times, as in Figure \[1\] 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 Section \[IV\] we consider the network scenario of Figure \[2\] 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 Section \[V\] and conclude in Section \[VI\].

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 that transforms density matrices of the composite system $S_1 \otimes \cdots \otimes S_k$, where $S_j$ denotes the system sent at the $j$-th 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

$$\mathcal{R}(\rho_{12}) = \sum_{m,n} p(m,n) (U_m \otimes U_n) \rho_{12} (U_m \otimes U_n)^\dagger,$$  \hspace{1cm} (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 $\rho_{12}$ represents the joint state of the two systems sent at the two times $t_1$ and $t_2$, that is, $\rho_{12}$ 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 fibre, where the random unitary changes of the photon polarisation 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

$$\mathcal{R}_2(\rho) := \sum_m p_2(n) U_n \rho U_n^\dagger,$$

where $p_2(n)$ is a joint probability distribution of the unitaries at time $t_2$. 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 Figure 3). The superposition of transmission times could 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 localisation 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 polarisation modes in the first (second) time bin. Here we take the polarisation 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 $|\alpha|H_1|0\rangle_{V_1} + |\beta|H_1|1\rangle_{V_1} \otimes |0\rangle_{H_2}|0\rangle_{V_2}$ and states of the form $|0\rangle_{H_1}|0\rangle_{V_1} \otimes (|\alpha|H_2|0\rangle_{V_2} + |\beta|H_2|1\rangle_{V_2})$. 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 $|\psi_i\rangle \otimes |\text{vac}\rangle_2$ and states of the form $|\text{vac}\rangle_1 \otimes |\psi_i\rangle_2$, where, for $i \in \{1, 2\}$, $|\text{vac}\rangle_i := |0\rangle_i |0\rangle_{V_i}$ is the vacuum state of the modes in system $S_i$, and $|\psi_i\rangle := |\alpha|H_1|0\rangle_{V_1} + |\beta|H_1|1\rangle_{V_1} \otimes |0\rangle_{H_2}|0\rangle_{V_2}$ is a single-photon polarisation state. The change in the particle’s state upon the transmission is then computed by applying the chan-
nel (1) to the appropriate state. Generalising 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 polarisation. 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 polarisation. 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 times will then be described by states of the form $\alpha \langle \psi\rangle_1 \otimes |\text{vac}\rangle_2 + \beta |\text{vac}\rangle_1 \otimes \langle \psi\rangle_2$, 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 can generally depend on the correlations between the evolution of the particle and the vacuum state.

In the quantum optical example, each unitary $U_m$ can be realised by a Hamiltonian acting on the two polarisation modes associated to system $S_i$, $i \in \{1, 2\}$. For example, the unitary $Z \otimes e^{i\phi_m} |\text{vac}\rangle \langle \text{vac}|$ 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) polarisation, 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

\[
|\psi\rangle_1 \otimes |\text{vac}\rangle_2 \mapsto |\psi\rangle_M \otimes |0\rangle_C,
\]

\[
|\text{vac}\rangle_1 \otimes |\psi\rangle_2 \mapsto |\psi\rangle_M \otimes |1\rangle_C,
\]

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 $\omega$, the overall evolution is described by an effective channel $C_\omega$, which transforms a generic state $\rho$ of the message into the state

\[
C_\omega(\rho) := \sum_{m,n} p(m, n) W_{mn}(\rho \otimes \omega) W_{mn}^\dagger,
\]

where $W_{mn}$ is the unitary $W_{mn} := V_m e^{i\phi_n} |0\rangle \langle 0| + e^{i\phi_m} V_n |1\rangle \langle 1|$. 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

\[
C_\omega(\rho) = \frac{C(\rho) + G(\rho)}{2} \otimes \omega + \frac{C(\rho) - G(\rho)}{2} \otimes Z\omega Z,
\]

with

\[
C(\rho) := \sum_{m,n} p(m, n) V_m \rho V_m^\dagger
\]

and

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

(See Appendix A for the derivation.) Here, the map $C$ 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 $C$ 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 $G$ can generally depend on the correlations between the evolution of the particle at two mutually exclusive moments of time. We call $G$ 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 depolarising on the message-carrying sector $M$, that is,

\[
C_{|j\rangle\langle j|}(\rho) = \frac{I}{d} \otimes |j\rangle \langle j| \quad \forall \rho, \forall j \in \{0, 1\},
\]

where $C_{|j\rangle\langle j|}$ is the quantum channel obtained by plugging $\omega = |j\rangle \langle j|$ into Eq. (4). Eq. (8) implies that, whenever the particle is sent at a definite moment of time, the message is replaced by white noise. Accordingly, the channel $C$ in Eq. (6) is depolarising. When the probability distribution $p(m, n)$ is symmetric, Eq. (5) becomes

\[
C_\omega(\rho) = \frac{I/d + G(\rho)}{2} \otimes \omega + \frac{I/d - G(\rho)}{2} \otimes Z\omega Z.
\]
In the realisation 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, labelled by integers from 0 to \( d^2 - 1 \). For qubits, we will take \( \{ V_m \} \) to be the four Pauli matrices \( \{ I, X, Y, Z \} \), labelled 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 \( \phi \) amounts to requiring that the marginal probability distributions \( p_1(m) \) and \( p_2(n) \) be uniform, that is

\[
p_1(m) = p_2(n) = \frac{1}{d^2} \quad \forall m, n \in \{0, \ldots, d^2 - 1\}. \tag{10}
\]

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 \( \sigma \) 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 \( G(\rho) \) is the completely depolarising channel \( G(\rho) = I/d \forall \rho \). Hence, the channel \( C_\omega \) in Eq. (9) is completely depolarising, and no information can be transmitted through it, no matter what state \( \omega \) 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 depolarising channels

Here we focus on the case where the message is a qubit \( (d = 2) \). Let \( \sigma \) 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

\[
G(\rho) = \frac{\rho X e^{i(\phi_1 - \phi_0)} + Y \rho Z e^{i(\phi_1 - \phi_2)} + \text{h.c.}}{4}, \tag{11}
\]

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

Note that \( G(\rho) \) 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 \( \omega \), 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

\[
G(\rho) = \frac{\{\rho, X\} - \{Z \rho Z, X\}}{4}, \tag{12}
\]

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

\[
G(|\pm\rangle\langle\pm|) = \pm \frac{I}{2}. \tag{13}
\]

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

\[
C_\omega(|\pm\rangle\langle\pm|) = \frac{I}{2} \otimes \omega_\pm, \tag{14}
\]

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

Putting the control in the state \( \omega = |\pm\rangle\langle+| \), one obtains the orthogonal output states \( \omega_+ = |\pm\rangle\langle\pm| \). 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 depolarising 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 factorises 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 \( C_\omega \) 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 \( C_\omega \) over arbitrary states \( \omega \) of the control system and over arbitrary (not necessarily random-unitary) realisations of the completely depolarising 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 optimisation.

#### 1. Analytical bound on the classical capacity

The derivation of the bound consists of three steps, whose details are provided in Appendix [3].

The first step is to prove that, in the lack of correlations and for message dimension \( d = 2 \), the channel \( C_\omega \) 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 \[ \text{Holevo capacity} \] \( \text{of} \). For a generic quantum channel \( \mathcal{E} \), the Holevo capacity is \( \chi(\mathcal{E}) = \max_{\{p_x, \rho_x\}} H \left[ \sum_x p_x \mathcal{E}(\rho_x) \right] - \sum_x p_x H(\rho_x) \), 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 \( \rho \), log denoting the logarithm in base 2.

The second step is to observe that state of the control that maximises the Holevo capacity of the channel \( \mathcal{C}_\omega \) is \( \omega = |+\rangle \langle +| \). This result holds for arbitrary message dimension \( d \geq 2 \), 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 \( d \geq 2 \), 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 \( 1/2 \) for every possible state \( \omega \). Hence, the perfect transmission of 1 bit achieved in Subsection [1113] is impossible in the lack of correlations.

### 2. Numerical evaluation of the capacity

The evaluation of the Holevo capacity involves an optimisation over all possible input ensembles. For quantum channels with \( d \)-dimensional input, the optimisation can be restricted to ensembles with up to \( d^2 \) linearly independent pure states \( \{\rho_{mn}\} \). In practice, however, the optimisation is often hard to carry out even in dimension \( d = 2 \). To make the optimisation feasible, we first show that in our case the optimisation can be reduced to an optimisation 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 [3].

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 \( |\phi_m\rangle \) with \( p(m,n) = 1/16 \). We set the state of the control to \( \omega = |+\rangle \langle +| \), which we know to guarantee the maximum Holevo information (cf. Lemma [3] in Appendix [3]).

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 \( \phi_0 \), can be set to 0 without loss of generality, as it represents a global phase. In Figure [11] 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 \( \phi_1, \phi_2 \), and \( \phi_3 \). The maximum over all possible choices of phases is 0.16 bits.

In Appendix [3] we also show that 0.16 bits is the maximum capacity achievable with arbitrary (not necessarily random unitary) channels that reduce to the depolarising channel in the one-particle subspace sector. The value 0.16 was previously found to be a lower bound to the classical capacity \( [51] \), 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 depolarising 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 optimisation 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 optimisation 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 \( [64][65] \), 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 \( p(m,n) = \delta_{m,n}\sigma(m)/4 \), where \( \sigma \) is the permutation that exchanges 0 with 1, and 2 with 3. This particular choice is interesting because as we have seen in Subsection [1113], 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...
Here, 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 realisations of the completely depolarising channel.

In the following, we provide two new 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 realisations of the completely depolarising channel.
and 0.042 bits for arbitrary realisations. This result shows that, when the quantum SWITCH is reproduced by the network in Figure 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 Figure 5 when the channels are completely depolarising and no correlation is present, that is, when $p_A(m,n) = p_B(m,n) = \delta_{n,\sigma(m)}/4$, where $\sigma$ 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.

and 0.024 bits for arbitrary realisations. This result shows that, when the quantum SWITCH is reproduced by the network in Figure 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 Figure 5 when the channels are completely depolarising and no correlation is present, that is, when $p_A(m,n) = p_B(m,n) = \delta_{n,\sigma(m)}/4$, where $\sigma$ 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 derivation of these results is provided in Appendix B.

Building on the above observations, we evaluate the capacity of the channel $E_\omega$ in Eq. (15) by scanning all possible values of the phases $\{\phi_m\}_{m=0}^3$. The result is the plot shown in Figure 6a. The largest classical capacity over all random unitary realisations 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 optimisation from random unitary realisations to arbitrary realisations of the completely depolarising channel. For this broader class of realisations, we numerically obtain that the maximum capacity is 0.024 bits.

Summarising, the best classical capacity one can obtain by sending a single particle through the network in Figure 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 realisations of the completely depolarising 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 Figure 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 non-trivial 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 $A$ and $B$ is sometimes described informally as a “superposition of channels $AB$ and $BA$.” While this expression may be formally correct (at least according to a broad notion of superposition), it can be misleading if taken at face value, because it does not mention explicitly the requirement of correlations between the channels $A$ and $B$ 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 $\sigma$ is the permutation that
exchanges 0 with 1, and 2 with 3. This choice is motivated by the fact that the permutation $\sigma$ guarantees the maximum communication capacity in the case where a single time-correlated channel is used (cf. Subsection III B).

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

$$E_\omega(\rho) = \frac{t + K(\rho)}{2} \otimes \omega + \frac{t - K(\rho)}{2} \otimes Z \omega Z, \quad (20)$$

with

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

The derivation of this formula is provided in Appendix D. Note that the channel $E_\omega$ 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 $E_\omega$. 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 maximisation to orthogonal input ensembles. For the state of the control qubit, we pick $\omega = |+\rangle\langle+|$, which is the choice that maximises the Holevo capacity (cf. Lemma 3 in Appendix E).

The lower bound to the classical capacity is shown in Figure 6b 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 effect of noise on the control system, focussing in particular on dephasing noise, of the form

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

where $s \in [0,1/2]$ is a probability and $\omega$ 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 Figure 1. In this setting, the evolution experienced by a single particle is described by the channel

$$C_\omega := (I_M \otimes P)C_\omega, \quad (23)$$

obtained by dephasing the control system at the output of the channel $C_\omega$ in Eq. (5). By inserting the expression (6) into the above equation, it is immediate to see that the effect of dephasing is to dampen the interference term $G$ in the effective channel (5): specifically, the interference term changes from $G$ to $(1-2s)G$.

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

Figure 7 shows the behaviour 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 realisations of the completely depolalising channel (blue curve). Notably, the capacity for every fixed value of $s$ is achieved by the same realisation of the completely depolarising 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 Subsection III B. To this purpose, we numerically evaluate a lower bound to the Holevo capacity (and therefore...
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 realisation 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 analyse 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 analysed 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.

ACKNOWLEDGEMENTS

We acknowledge discussions with Robert Spekkens, Sandu Popescu, Paul Skrzypczyk, Debbie Leung, Aephraim Steinberg, Philippe Grangier, Philip Walther, Giulia Rubino, Caslav Brukner, David Schmidt, Chiara Macchiavello, Massimiliano F. Sacchi, Santiago Semper Llagostera, Robert Gardiner, Kwok Ho Wan, Raj Patel, Ian Walmsley, Giulio Amato, Kavan Modi, Guillaume Boisseau, Alastair Abbott, Marco Túlio Quintino, Fabio Costa, Daniel Ebler, Sina Salek, and Carlo Sparaciari. The numerical simulations presented in this paper were written using the Python software package QuTiP and the circuit diagrams were drawn using TikZiT. This work is supported by the National Natural Science Foundation of China through grant 11675136, the Hong Research Grant Council through grant 17307719, the Croucher Foundation, the HKU Seed Funding for Basic Research, the UK Engineering and Physical Sciences Research Council (EPSRC) through grant EP/R513295/1, and the Perimeter Institute for Theoretical Physics. Research at the Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Research, Innovation and Science. This publication was made possible through the support of the grants 60609 ‘Quantum Causal Structures’ and 61466 ‘The Quantum Information Structure of Spacetime (QISS)’ (qiss.fr) from the John Templeton Foundation. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.

[1] C. H. Bennett and G. Brassard, in Int. Conf. on Computers, Systems and Signal Processing (Bangalore, India, Dec. 1984) (1984) pp. 175–9.
[2] A. K. Ekert, Physical Review Letters 67, 661 (1991).
[3] P. W. Shor, Physical Review A 52, R2493 (1995).
[4] D. Gottesman, in Quantum information science and its contributions to mathematics, Proceedings of Symposia in Applied Mathematics, Vol. 68 (2010) pp. 13–58.
[5] D. A. Lidar and T. A. Brun, Quantum Error Correction (Cambridge University Press, 2013).
[6] C. Macchiavello and G. M. Palma, Physical Review A 65, 050301(R) (2002).
[7] D. Kretschmann and R. F. Werner, Physical Review A 72, 062323 (2005).
[8] F. Caruso, V. Giovannetti, C. Lupo, and S. Mancini, Reviews of Modern Physics 86, 1203 (2014).
[9] F. A. Pollock, C. Rodríguez-Rosario, T. Frauenheim, M. Paternostro, and K. Modi, Physical Review A 97, 012127 (2018).
[10] J. L. Ball and K. Banaszek, Journal of Physics A: Mathematical and General 39, L1 (2005).
[11] C. Bonato, M. Aspelmeyer, T. Jennewein, C. Pernechele, P. Villoresi, and A. Zeilinger, Optics Express 14, 10050 (2006).
[12] G. Chiribella, M. Dall’Arno, G. M. D’Ariano, C. Macchiavello, and P. Perinotti, Physical Review A 83, 052305 (2011).
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 \( \text{Chan}(S \rightarrow S') \) the set of quantum channels with input system \( S \) and (possibly different) output system \( S' \). When \( S = S' \) we will use the shorthand \( \text{Chan}(S) \).

The action of a quantum channel \( \mathcal{A} \) on a density matrix \( \rho \) can be conveniently written in the Kraus representation \( \mathcal{A}(\rho) = \sum_i A_i \rho A_i^\dagger \), where \( \{A_i\} \) is a (non-unique) set of operators, satisfying \( \sum_i A_i^\dagger A_i = I \).

A transmission line with \( k \) input/output ports is described by a \( k \)-partite quantum channel \( \mathcal{B} \in \text{Chan}(S^{(1)} \otimes \ldots \otimes S^{(k)}) \) with \( k \) input/output pairs \( (S^{(i)}, S^{(j)}) \) with \( j < i \).

A transmission line that can be used \( k \) times in succession is described by \( k \)-step quantum channel \( \mathcal{B} \) (also known as a quantum \( k \)-comb \[31, 32\]). 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 \( j < i \).

We will denote the set of \( k \)-step quantum channels as \( \text{Chan}(S^{(1)} \rightarrow S^{(1)} \otimes \ldots \otimes S^{(k)}) \), or simply \( \text{Chan}(S^{(1)} \rightarrow S^{(k)}) \) when the input and output of each pair coincide. For \( k = 2 \), an example of 2-step quantum channel is illustrated in Figure 8.

The possibility that no particle is sent through a port of a device can be described using the notion of vacuum extension \[32\]. Consider first a single-port device, described by an ordinary quantum channel \( \mathcal{A} \in \text{Chan}(S) \). 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 \( \mathcal{H}_\text{Vac} \), which is orthogonal to the one-particle sector \( S \). Overall, the device acts on an extended system \( S = S \otimes \mathcal{H}_\text{Vac} \), 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 \( \tilde{\mathcal{A}} \) of \( \mathcal{A} \) is any channel which acts as \( \mathcal{A} \) (respectively, \( I\mathcal{H}_\text{Vac} \)) when the input is a state in sector \( S \) (respectively, \( \mathcal{H}_\text{Vac} \)). The Kraus operators of \( \tilde{\mathcal{A}} \) are \( \tilde{A}_i = A_i \otimes |\text{vac}\rangle \langle \text{vac}| \), 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.

Consider a transmission line described by a bipartite channel \( \mathcal{B} \in \text{Chan}(S^{(1)} \otimes S^{(2)}) \). A vacuum extension of the channel \( \mathcal{B} \) is another bipartite channel \( \tilde{\mathcal{B}} \in \text{Chan}(S^{(1)} \otimes \mathcal{H}_\text{Vac} \otimes S^{(2)}) \), acting on the extended systems \( S^{(1)} \otimes \mathcal{H}_\text{Vac} \) and \( S^{(2)} \otimes \mathcal{H}_\text{Vac} \). In general, the systems \( S^{(1)} \), \( S^{(2)} \) can represent the systems accessible at the same location at two consecutive moment 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/output ports of our multiport device.

2. A single particle travelling 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 \( S^{(1)} \otimes S^{(2)} \) contains a
no-particle sector $\text{Vac}^{(1)} \otimes \text{Vac}^{(2)}$, 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 $\mathcal{S}(B)$, and define it as the restriction of $\mathcal{B}$ to the one-particle sector, regarded as isomorphic to the composite system “message + control.” Explicitly, the action of the superposition channel is defined as

$$\mathcal{S}(\mathcal{B}) := \mathcal{U}^\dagger \circ \mathcal{B} \circ \mathcal{U},$$  \hspace{1cm} (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

$$U(|\psi\rangle_M \otimes |0\rangle_C) := |\psi\rangle_S^{(1)} \otimes |\text{Vac}\rangle_S^{(2)},$$

$$U(|\psi\rangle_M \otimes |1\rangle_C) := |\text{Vac}\rangle_S^{(1)} \otimes |\psi\rangle_S^{(2)}.$$  \hspace{1cm} (A2)

Mathematically, the transformation $\mathcal{S} : \mathcal{Chan}(\mathcal{S}^{(1)} \otimes \mathcal{S}^{(2)}) \rightarrow \mathcal{Chan}(M \otimes C)$ is a quantum supermap, that is, a transformation from quantum channels to quantum channels satisfying appropriate consistency requirements \cite{23,22,64}. An illustration of the supermap $\mathcal{S}$ is provided in subfigure (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 subfigure (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 $\mathcal{B} \in \mathcal{Chan}(\mathcal{S}^{(1)} \otimes \cdots \otimes \mathcal{S}^{(k)})$, with $S^{(1)} \cong S^{(2)} \cong \cdots \cong S^{(k)}$, and with vacuum extension $\mathcal{B} \in \mathcal{Chan}(\mathcal{S}^{(1)} \otimes \cdots \otimes \mathcal{S}^{(k)})$. The superposition channel is then defined as the restriction of $\mathcal{B}$ to the one-particle sector

$$k \sum_{j=1}^k \text{Vac}^{(1)} \otimes \cdots \otimes \text{Vac}^{(j-1)} \otimes S^{(j)} \otimes \text{Vac}^{(j+1)} \otimes \cdots \otimes \text{Vac}^{(k)} \cong M \otimes C,$$  \hspace{1cm} (A3)

where $C$ is now a $k$-dimensional control system.

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

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

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

$$\tilde{V}_m := U_m = V_m \oplus e^{i\phi_m} |\text{Vac}\rangle \langle \text{Vac}|,$$

(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

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

(A6)

with $\tilde{V}_m(\cdot) = \tilde{V}_m(\cdot)\tilde{V}_m^\dagger$, which is equivalent to Equation (4) in the main text, with $U_m = \tilde{V}_m$.

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

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

Explicitly, we have the expression

$$\mathcal{S}(\tilde{\mathcal{R}})(\rho \otimes \omega) = \sum_{m,n} C_{mn} (\rho \otimes \omega) C_{mn}^\dagger,$$  \hspace{1cm} (A8)

where $\rho$ (respectively, $\omega$) is an arbitrary state of the message (respectively, control), and

$$C_{mn} := \sqrt{p(m,n)} e^{i\phi_n} V_m \otimes |0\rangle\langle 0| + \sqrt{p(m,n)} V_n e^{i\phi_m} \otimes |1\rangle\langle 1|,$$  \hspace{1cm} (A9)
\( e^{i\phi_m} \) being the vacuum amplitude in Eq. (A5). Eq. (A8) coincides with Equation (14) in the main text, with \( \mathcal{Z} := \mathcal{S}(\mathcal{R}) \) and \( W_{mn} := C_{mn}/\sqrt{p(m,n)} \).

4. Derivation of Eq. [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

\[
\mathcal{S}(\mathcal{R}) = \frac{\mathcal{R}_1 + \mathcal{G}}{2} \otimes I + \frac{\mathcal{R}_1 - \mathcal{G}}{2} \otimes \mathcal{Z},
\]

(A10)

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

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

(A11)

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

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

(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 depolarising in the one-particle subspace. This section makes use of the notation introduced in Appendix A.

1. **Proof that the superposition of uncorrelated completely depolarising channels is entanglement-breaking**

Let \( \mathcal{A}(\cdot) = \sum_{m=0}^{r-1} A_m(\cdot) A_m^\dagger \in \text{Chan}(S) \) be a generic quantum channel, and let \( \tilde{\mathcal{A}} \in \text{Chan}(\tilde{S}) \) be a vacuum extension of \( \mathcal{A} \). Using Eq. (A11), we obtain

\[
\mathcal{S}(\tilde{\mathcal{A}} \otimes \tilde{\mathcal{A}}) = \frac{\mathcal{A}(\rho) + F \rho F^\dagger}{2} \otimes I + \frac{\mathcal{A}(\rho) - F \rho F^\dagger}{2} \otimes \mathcal{Z}.
\]

(B1)

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

\[
F := \sum_m \sigma_m A_m
\]

(B2)

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

Now, let \( \mathcal{A} \) be the completely depolarising channel \( \mathcal{D} : \rho \mapsto I/d \), with vacuum extension \( \tilde{\mathcal{D}} \). For a fixed state \( \omega \) of the control system, consider the effective channel defined by

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

(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 2.** Let \( \mathcal{D} \) be a completely depolarising channel with vacuum extension \( \tilde{\mathcal{D}} \) and vacuum interference operator \( F \). Then, the operator norm of \( F \) satisfies the inequality \( ||F||_\infty \leq 1/\sqrt{d} \).

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

\[
||F||_\infty = \max_{\{|v|:||v||=1\}} \max_{\{|w|:||w||=1\}} \langle v|F|w \rangle
\]

(B4)

and

\[
||\langle v|F|w \rangle|| \leq \sqrt{\sum_i |\alpha_i|^2} \sqrt{\sum_j \langle v|A_j|w \rangle \langle w|A_j^\dagger|v \rangle} = \sqrt{\langle v|D(\langle w|w \rangle)|v \rangle}.
\]

(B5)

If \( \mathcal{D} \) is the completely depolarising channel, then \( D(\langle w|w \rangle) = I/d \) and therefore the bound becomes

\[
||\langle v|F|w \rangle|| \leq \sqrt{1/d} \text{ which implies } ||F||_\infty \leq 1/\sqrt{d}.
\]

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 to 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

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

(B6)

with \( G_F := (F \otimes I)(|\Phi^+\rangle)(|\Phi^+\rangle)(F \otimes I)^\dagger \).
We now show that the operators $\frac{IG}{d} \pm GF$ are proportional to states with positive partial transpose. To this purpose, note that the partial transpose of $GF$ on the second space is

$$G_F^T = (F \otimes I)^{\text{SWAP}} d (F \otimes I)^\dagger.$$  \hfill (B7)

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

$$\langle \Psi | (G_F^T)^2 |\Psi\rangle \leq \frac{\langle \Psi | (FF^\dagger \otimes I) |\Psi\rangle}{d} \leq \frac{\|FF^\dagger\|_\infty}{d} = \frac{\|F\|_\infty^2}{d} \leq \frac{1}{d^2}.$$  \hfill (B8)

where the first inequality follows from Schwarz’ inequality, and the last inequality follows from Lemma 2.

Using Eq. (B8), we obtain the relation

$$\langle \Psi | \left( \frac{I \otimes I}{d^2} \pm GF \right)^2 |\Psi\rangle \geq \frac{1}{d^2} - \langle \Psi | G_F^T |\Psi\rangle \geq 0.$$  \hfill (B9)

Since $|\Psi\rangle$ is an arbitrary vector, we conclude that the operator $\left( \frac{IG}{d} \pm GF \right)^2$ has positive partial transpose. For $d = 2$, the Peres-Horodecki criterion \cite{Peres_Horodecki}, guarantees that $\frac{IG}{d} \pm GF$ is proportional to a separable state. Hence, the whole output state is separable.

2. Optimal control state for maximizing the Holevo capacity

Proposition 1 implies that the classical capacity of the channel $C_{\omega,F}$ is equal to its Holevo capacity (see \cite{Holevo_capacity}). Here we show that the Holevo capacity is maximised by the state $\omega = |+\rangle \langle +|$. In fact, we prove a more general result:

**Lemma 3.** Let $C_\omega$ be an arbitrary channel of the form

$$C_\omega := L_+ (\rho) \otimes \omega + L_- (\rho) \otimes Z \omega Z,$$  \hfill (B10)

where $L_\pm$ are arbitrary linear maps. Then, for every density matrix $\omega$, the Holevo capacity satisfies the bound

$$\chi(C_\omega) \leq \chi(C_{|+\rangle \langle +|}).$$

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

$$C_\omega = (I_M \otimes \mathcal{P}_\omega) \circ C_{|+\rangle \langle +|},$$  \hfill (B11)

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

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

for an arbitrary state $\gamma$. Hence, we have $\chi(C_\omega) = \chi \left( (I_M \otimes \mathcal{P}_\omega) \circ C_{|+\rangle \langle +|} \right) \leq \chi(C_{|+\rangle \langle +|}).$ \hfill (B13)

Lemma 3 holds in particular for

1. the channel $C_{\omega,F}$ defined in Eq. (B13)
2. the channel $C_\omega$ defined in Eq. (9) of the main text
3. the channel $E_{\omega,F}$ defined in Eq. (20) of the main text.

3. Bound on the Holevo capacity

**Proposition 4.** The Holevo capacity of the channel $C_{\omega,F}$ defined in Eq. (B3) is upper bounded as

$$\chi(C_{\omega,F}) \leq \frac{\log(2d)}{d} + \frac{1}{2} + \frac{1}{2} \log \frac{1 + \|F\|_\infty^2}{2} \log \frac{\|F\|_\infty^2}{2} + \frac{1}{2} - \frac{\|F\|_\infty^2}{2} \log \frac{\|F\|_\infty^2}{2},$$  \hfill (B14)

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 $C_\omega$ is upper bounded by the Holevo capacity of the channel $C_{|+\rangle \langle +|,F}$ (Lemma 3). Hence, it is enough to prove the bound for the channel $C_{|+\rangle \langle +|}$. Note that the output of channel $C_{|+\rangle \langle +|,F}$ has dimension $2d$. For a generic channel $\mathcal{E}$ with $(2d)$-dimensional output, the Holevo capacity is upper bounded as

$$\chi(\mathcal{E}) \leq \log(2d) - \min_{\rho} H \left[ \mathcal{E}(\rho) \right],$$  \hfill (B15)

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

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

$$C_{|+\rangle \langle +|,F}(\rho) = \frac{1}{2} + F \rho F^\dagger \otimes |+\rangle \langle +| + \frac{1}{2} - F \rho F^\dagger \otimes |-\rangle \langle -|,$$  \hfill (B16)

as one can deduce from Eqs. (B3) and (B11).

In the case of a pure state $\rho = |\psi\rangle \langle \psi|$, we write $F |\varphi\rangle = k |\varphi\rangle$, where $|\varphi\rangle$ is a unit vector and $k$ is a normalisation constant. With this notation, we obtain

$$C_{|+\rangle \langle +|,F}(|\psi\rangle \langle \psi|) = \frac{1}{2} + k^2 |\varphi\rangle \langle \varphi| + \frac{1}{2} P_\perp \otimes |+\rangle \langle +| + \frac{1}{2} - k^2 |\varphi\rangle \langle \varphi| + \frac{1}{2} P_\perp \otimes |-\rangle \langle -|,$$  \hfill (B17)

for every where $|\varphi\rangle$ is a unit vector and $k$ is a normalisation constant.
with \( P_\perp := I - |\varphi\rangle \langle \varphi | \). The von Neumann entropy of this state is
\[
H \left[ C_{(+)}(+|F)(|\psi\rangle \langle \psi|) \right] = \frac{1}{2} \log \frac{1}{\sin^2 \theta} - \frac{1}{2} \log \frac{1}{\sin^2 \theta} - \frac{1}{2} \log \frac{1}{\sin^2 \theta}.
\]

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

**Lemma 6.** Without loss of generality, the operator \( F \) that maximises the Holevo information of the channel \( C_{\omega,F} \) can be taken to be of the form \( F = a |0\rangle \langle 0| + b |1\rangle \langle 1| \), with \( a^2 + b^2 \leq 1/d \), \( a, b \geq 0 \).

**Proof.** Using the singular value decomposition, \( F \) can be written as \( F = UFV^\dagger \), where \( U \) and \( V \) are suitable unitary matrices, and \( F^\dagger \) is diagonal in the basis \( \{|0\rangle, |1\rangle\} \). Now the capacity of the channel \( C_{\omega,F} \) is equal to the capacity of the channel \( C_{\omega,F^\dagger} = (U \otimes I_C)^\dagger \circ C_{\omega,F} \circ V^\dagger \), where \( U^\dagger \) and \( V^\dagger \) are the inverses of the unitary channels associated to the unitary matrices \( U \) and \( V \), respectively, and \( I_C \) is the identity channel on the control system. Notice that \( F^\dagger \) is also a vacuum interference operator associated to the completely depolarising channel. Hence, the maximisation 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 depolarising channel, the vacuum interference operator \( F \) must satisfy the condition \( \text{Tr} F^\dagger F \leq 1/d \) [11]. For an operator of the form \( F = a |0\rangle \langle 0| + b |1\rangle \langle 1| \), this implies the inequality \( |a| \leq 1/d \), \( |b| \leq 1/d \). Finally, we show that \( a, b \) can be restricted to positive numbers. Let \( W = a |0\rangle \langle 0| + b |1\rangle \langle 1| \), where \( a' = a/b \). Then \( F'' := WF = FW = |a'\rangle \langle 0| + |0\rangle \langle 1| \). The capacity of the channel \( C_{\omega,F''} \) is equal to the capacity of the channel \( C_{\omega,F} \). Therefore, a maximisation of the Holevo capacity can be restricted to vacuum interference operators with positive coefficients in the computational basis.

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

**Lemma 7.** When the operator \( F \) is diagonal in the computational basis, the input ensemble that maximises the Holevo information after application of the channel \( C_{\omega,F} \) can be chosen without loss of generality to be of the form
\[
\left\{ \frac{p_x}{d}, M^\dagger |\psi_x\rangle \langle \psi_x| M^\dagger \right\}_{x \in \{0, \ldots, d-1\}, j \in \{0, \ldots, d-1\}} ,
\]
where \( (p_x)_{x \in \{0, \ldots, d-1\}} \) 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.** Since \( F \) is diagonal, the channel \( C_{\omega,F} \) has the covariance property
\[
C_{\omega,F} \circ \mathbf{U}_\theta = (U_\theta \otimes I_C) \circ C_{\omega,F} \quad \forall \theta ,
\]
where \( \theta = (\theta_0, \theta_1, \ldots, \theta_{d-1}) \) is a vector of \( d \) phases, and \( U_\theta \) is the unitary channel associated to the unitary matrix
\[ U_\theta = \sum_{m=0}^{d-1} e^{i \theta_m} |m\rangle\langle m|. \]

Note that, in particular, we have

\[ C_{\omega,F} \circ \mathcal{M}^j = (\mathcal{M}^j \otimes \mathcal{I}_C) \circ C_{\omega,F} \quad j \in \{0, \ldots, d-1\}, \]  

(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

\[ E := \left\{ \frac{p_x}{d}, \mathcal{M}^j(\rho_x) \right\}_{x \in X, j \in \{0, \ldots, d-1\}}, \]  

(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 \( |X| \leq d \) without loss of generality. Let \( E \) be an optimal covariant ensemble, and let

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

(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 maximisation 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) \( (P_{x,j})_{x \in X, j \in \{0, \ldots, d-1\}} \), via the correspondence

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

(C7)

Since the correspondence is linear, the extreme ensembles are in one-to-one correspondence with the extreme POVMs. The latter have been characterised by one of us in Ref. [64], where it was shown that a necessary condition for extremality is that the ranks of the operators \( \xi_x \), denoted by \( r_x \), satisfy the condition

\[ \sum_{x \in X} r_x^2 \leq \sum_{\mu} m_\mu^2, \]  

(C8)

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

\[ \sum_{x \in X} r_x^2 \leq d. \]  

(C9)

In particular, this means that the number of non-zero operators \( \xi_x \) 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 maximisation of the Holevo information can be restricted without loss of generality to covariant ensembles with \( |X| \leq d \).

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 \( E = \{ p_x/d, \mathcal{M}^j(\psi_x) \langle \psi_x | \mathcal{M}^j | \psi_x \rangle \} \), let us expand each state as \( |\psi_x \rangle = \sum_{m} c_m |e^{i \theta_x m} \rangle \), where \( \{ \theta_x, m \} \) are suitable phases. Then, we can define the new states \( |\psi'_x \rangle := U_\theta |\psi_x \rangle \), with \( \theta := (\theta_{x,0}, \ldots, \theta_{x,d-1}) \).

By construction, these states have positive coefficients in the computational basis, and the corresponding ensemble \( E' = \{ p_x/d, \mathcal{M}^j(\psi'_x) \langle \psi'_x | \mathcal{M}^j | \psi'_x \rangle \} \) gives rise to the same Holevo information as \( E \), when fed into the channel \( C_{\omega,F} \).

\[ \square \]

**Corollary 8.** When the operator \( F \) is diagonal in the computational basis, the Holevo capacity of the channel \( C_{\omega,F} \) is given by

\[ \chi(C_{\omega,F}) = \max_{\{ p_x, \psi_x \}} \left\{ H \left[ C_{\omega,F} \left( \sum_{x,m} p_x |\psi_x \rangle \langle \psi_x | \mathcal{M}^j | \psi_x \rangle \langle \psi_x | \mathcal{M}^j | \psi_x \rangle \right) \right] \right\}, \]  

(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 \( C_{\omega,F} \) to the pure state ensemble in Lemma 7 using the relations,

\[ H[C_{\omega,F}(\mathcal{M}^j|\psi\rangle\langle\psi|\mathcal{M}^j)] = H[C_{\omega,F}(|\psi\rangle\langle\psi|)], \]  

(C11)

\[ \frac{1}{d} \sum_{j=0}^{d-1} \mathcal{M}^j|\psi\rangle\langle\psi|\mathcal{M}^j = \sum_{m=0}^{d-1} |\psi(m)|^2 |m\rangle\langle m|, \]  

(C12)

valid for every vector \( |\psi\rangle \).

\[ \square \]

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

**Theorem 9.** For every vacuum extension of the completely depolarising channel and for every state of the control qubit, the classical capacity of the channel resulting from the superposition of two independent depolarising qubit channels is upper bounded as
\[
C(C_\omega,F) \leq \max_{a \geq 0,b \geq 0 \atop a^2 + b^2 \leq 1/2} \max_{0 \leq q,p,\rho_1 \leq 1} H[C_\omega,F(\rho_1)] \\
- qH[C_\omega,F(|\psi_0\rangle\langle\psi_0|)] - (1-q)H[C_\omega,F(|\psi_1\rangle\langle\psi_1|)] ,
\]
(13)

Proof. For \(d = 2\), Proposition \ref{prop:max} guarantees that the channel \(C_\omega,F\) is entanglement breaking, and therefore its classical capacity is equal to the Holevo capacity. Lemma \ref{lem:holevo} guarantees that the maximum of the Holevo capacity \(\omega\) is attained by the state \(|+\rangle\langle+|\). Then, Lemma \ref{lem:bound} guarantees the maximum of the Holevo capacity of the channel \(C_{|+\rangle\langle+|,F}\) can be obtained with a diagonal operator \(F = a|0\rangle\langle0| + b|1\rangle\langle1|\), \(a, b \geq 0\). The Holevo capacity of \(C_{|+\rangle\langle+|,F}\) can be computed explicitly using Corollary \ref{cor:explicit} with

\[
|\psi_0\rangle := \sqrt{p_0}|0\rangle + \sqrt{1-p_0}|1\rangle , \\
|\psi_1\rangle := \sqrt{p_1}|1\rangle + \sqrt{1-p_1}|0\rangle .
\]
(15)

Finally, an upper bound is obtained by relaxing the constraint on \(a\) and \(b\) to \(a^2 + b^2 \leq 1/2\) (Lemma \ref{lem:upper}). \(\square\)

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

We now connect the 2-step channels \(\tilde{A}\) and \(\tilde{B}\) 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 Figure 10. This particular composition of two 2-step channels is described by a supermap \(Z\) that maps pairs of channels in \(\text{Chan}(\tilde{A}^{(1)}, \tilde{A}^{(2)}) \times \text{Chan}(\tilde{B}^{(1)}, \tilde{B}^{(2)})\) into bipartite channels in \(\text{Chan}(\tilde{A}^{(1)} \otimes \tilde{B}^{(1)} \rightarrow \tilde{B}^{(2)} \otimes \tilde{A}^{(2)})\).

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 \(Z(\tilde{R}_A, \tilde{R}_B)\). Following Eq. (A1), the evolution of the particle is described by the superposition channel

\[
S[Z(\tilde{A}, \tilde{B})] := U^* \circ Z(\tilde{A}, \tilde{B}) \circ U ,
\]
(1)

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

Let us apply the above construction to the special case where the channels \(\tilde{A}\) and \(\tilde{B}\) are of the random unitary form

\[
\tilde{A} = \tilde{R}_A := \sum_{m,n} \rho_{A}(m,n) \tilde{V}_m^{(A)} \otimes \tilde{V}_n^{(A)} \\
\tilde{B} = \tilde{R}_B := \sum_{k,l} \rho_{B}(k,l) \tilde{V}_k^{(B)} \otimes \tilde{V}_l^{(B)} ,
\]
(2)

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

\[
\tilde{V}_m^{(A)} := V_m^{(A)} \otimes e^{i\phi_m^{(A)}} |\text{vac}\rangle \langle \text{vac}| \\
\tilde{V}_k^{(B)} := V_k^{(B)} \otimes e^{i\phi_k^{(B)}} |\text{vac}\rangle \langle \text{vac}| ,
\]
(3)

respectively. With this choice, we have

\[
Z(\tilde{R}_A, \tilde{R}_B) = \sum_{m,n,k,l} \rho_{A}(m,n)\rho_{B}(k,l) (\tilde{V}_l^{(B)} \circ \tilde{V}_k^{(A)}) \otimes (\tilde{V}_n^{(A)} \circ \tilde{V}_m^{(B)}) .
\]
(4)
We now restrict our attention to the case where \( K \) is the linear map defined by

\[
\tilde{V}_{m,n,k,l} := \sum_{\pi} \rho_{B}(m,n) p_{B}(k,l) V_{m,n,k,l}^{\dagger} \rho V_{m,n,k,l} =: \rho_{V,m,n,k,l}.
\]

with

\[
W_{mnkl} := V_{m}^{(B)} V_{n}^{(A)} e^{i(\phi_{m}^{(B)} + \phi_{n}^{(A)})} \otimes |0\rangle\langle 0|
+ V_{n}^{(A)} V_{k}^{(B)} e^{i(\phi_{n}^{(A)} + \phi_{k}^{(B)})} \otimes |1\rangle\langle 1|. \tag{D6}
\]

This proves Equation (15) in the main text.

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

For the control (in this case the path of the particle) initialised in the state \( \omega \), the superposition channel specified by the vacuum extension \( Z(\tilde{R}, \tilde{R}) \) is given by

\[
\mathcal{S}\left[Z(\tilde{R}, \tilde{R})\right] = \sum_{m,n,k,l} p_{A}(m,n) p_{B}(k,l) W_{mnkl} \rho_{\tilde{V},m,n,k,l} \rho_{V,m,n,k,l}.
\]

We now restrict our attention to the case where

1. the two channels \( \tilde{R} \) and \( \tilde{R} \) 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 \( \rho \), and the effective channel can be rewritten as

\[
\mathcal{S}\left[Z(\tilde{R}, \tilde{R})\right] = \mathcal{R}(\rho) + \mathcal{K}(\rho) = \mathcal{R}(\rho) + \mathcal{K}(\rho).
\]

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

\[
\mathcal{S}\left[Z(\tilde{R}, \tilde{R})\right] = \frac{I/d + \mathcal{K}(\rho)}{2} \otimes \omega + \frac{I/d - \mathcal{K}(\rho)}{2} \otimes Z\omega Z, \tag{D10}
\]

with

\[
\mathcal{K}(\rho) = \frac{1}{d^{2}} \sum_{m,k} V_{m}^{(B)} V_{m}^{(A)} \rho_{V_{m,k}} V_{m}^{(B)} V_{m}^{(A)} e^{i\left[\phi_{m}^{(A)} + \phi_{m}^{(B)} - \phi_{m}^{(A)} - \phi_{m}^{(B)}\right]}.
\]

Setting \( d = 2 \) and choosing \( \sigma \) to be the permutation that exchanges 0 with 1, and 2 with 3, we obtain Eqs. (20)–(21) of the main text.
Appendix E: Proofs of the statements in Subsection IV B

Here we consider the scenario of Figure 11 in the special case where the 2-step channels \( \hat{A} \) and \( \hat{B} \) are of the product form \( \hat{A} = \hat{A}_1 \otimes \hat{A}_2 \) and \( \hat{B} = \hat{B}_1 \otimes \hat{B}_2 \), respectively. In this case, the combination of the channels in the network of Figure 11 gives the bipartite channel

\[
Z(\hat{A} \otimes \hat{B}) = \tilde{B}_2 \hat{A}_1 \otimes \tilde{A}_2 \tilde{B}_1 .
\]

(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

\[
S \left[ Z(\hat{A} \otimes \hat{B}) \right] = S(\tilde{B}_2 \hat{A}_1 \otimes \tilde{A}_2 \tilde{B}_1) ,
\]

(E2)

where \( S \) is the supermap defined in Eq. (A1).

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

\[
S(\hat{D}^2 \otimes \hat{D}^2)(\rho \otimes \omega) = \frac{I/d + F \rho F^\dagger}{2} \otimes \omega + \frac{I/d - F^2 \rho F^2 \dagger}{2} \otimes Z \omega Z ,
\]

(E3)

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

Note that one has the equality

\[
S(\hat{D}^2 \otimes \hat{D}^2)(\rho \otimes \omega) \equiv C_{\omega,F^2}(\rho) ,
\]

(E5)

using the notation of Eq. (B3). That is, in the lack of correlations the configuration of channels depicted in Figure 11 gives rise to the effective channel in Equation (B1), with \( F \) replaced by \( F^2 \). This means that all of the results in Appendices B–C apply to this scenario as well, with \( F \) replaced by \( F^2 \). In particular, the classical capacity can be determined numerically using Theorem 9 with the maximisation constraint now being that for the vacuum interference operator \( F^2 = g|0\rangle\langle 0| + h|1\rangle\langle 1| \), \( g+h \leq 1/d \), where \( g, h \geq 0 \).

The classical capacity of the channels \( C_{\omega,F} \) and \( C_{\omega,F^2} \) can be evaluated numerically. For the cases where each completely depolarising channel is implemented by a random unitary channel (cf. Eqs. (4) and (15), respectively, in the main text), Figure 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 \( \phi_1, \phi_2, \phi_3 \) as shown in Figs. 4 and 6.