


## Increasing communication capacity via superposition of order

K. Goswami,<sup>1,\*</sup> Y. Cao,<sup>2,3</sup> G. A. Paz-Silva,<sup>2</sup> J. Romero,<sup>1,†</sup> and A. G. White<sup>1</sup>

<sup>1</sup>Centre for Engineered Quantum Systems, School of Mathematics and Physics, University of Queensland, QLD 4072, Australia

<sup>2</sup>Centre for Quantum Dynamics, Griffith University, Queensland 4111, Australia

<sup>3</sup>State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China

 (Received 22 August 2019; revised 10 February 2020; accepted 25 June 2020; published 24 August 2020)

Classically, no information can be transmitted through a depolarising, which is a completely noisy, channel. We show that by combining a depolarising channel with another channel in an indefinite causal order—that is, when there is a superposition of the order that these two channels were applied—it becomes possible to transmit significant information. We consider two limiting cases. When both channels are fully depolarising, the ideal limit is communication of 0.049 bits; experimentally we achieve  $(3.4 \pm 0.2) \times 10^{-2}$  bits. When one channel is fully depolarising, and the other is a known unitary, the ideal limit is communication of 1 bit. We experimentally achieve  $0.64 \pm 0.02$  bits. Our results offer intriguing possibilities for future communication strategies beyond conventional quantum Shannon theory.

DOI: [10.1103/PhysRevResearch.2.033292](https://doi.org/10.1103/PhysRevResearch.2.033292)

### I. INTRODUCTION

Noise is ubiquitous: communication protocols aim to optimize the amount of information that can be sent through a channel with a given amount of noise. In the limit of a completely noisy channel, no information can be transmitted [1]. This is true even with a single quantum channel [2]. Surprisingly, quantum physics offers strategies to transmit information in the scenario of two noisy channels, e.g., via superposition of path [3–5] or via superposition of order [6–8]. In these strategies, the superposition in a control qubit determines the superposition in path or causal order. Placing two completely noisy channels in a superposition of paths—that is in different arms of an interferometer—allows some classical information to be communicated, at least 0.16 bits when the paths are equally weighted. Placing them instead in a superposition of causal order—that is the order in which the channels are applied is indefinite—allows up to 0.049 bits to be communicated when the orders are equally weighted [6]. Here we show that using superposition of causal order, a greater communication advantage can be achieved than superposition of paths [3], ideally up to 100% of information.

We use the quantum switch, a physically realizable process that simulates the superposition of causal order, which has been implemented in several photonic experiments [9–13]. We follow the experimental setup of [11]. We analyze, and experimentally demonstrate, communication through various

combinations of noisy and unitary channels in an indefinite causal order. We provide an example where perfect transmission is possible, given the freedom to choose the measurement basis of the control that determines the order. We outline the mathematical description of the channels and provide an experimental method to estimate the information-theoretic advantage quantified by the Holevo capacity,  $\chi$  [14].

### II. THEORETICAL BACKGROUND

We depict the communication protocol in Fig. 1. The sender maps a classical message into a quantum state  $\rho_t = |\psi_t\rangle\langle\psi_t|$ , which we will refer to as the target qubit. This state passes through two generalized Pauli channels  $\mathcal{N}_p$  and  $\mathcal{N}_q$ . We

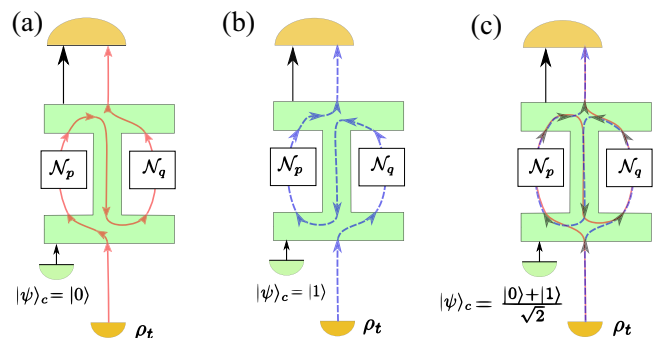


FIG. 1. Classical communication through indefinite causal order. The sender maps the classical message bit to a quantum state  $\rho_t$ —the target qubit.  $\rho_t$  passes through two noisy channels  $\mathcal{N}_p$  and  $\mathcal{N}_q$ . The order of the channels is controlled by a control qubit  $|\psi_c\rangle$ . (a) When the control qubit  $\rho_c = |\psi_c\rangle\langle\psi_c|$  is off, i.e.,  $|\psi_c\rangle = |0\rangle$ , the channels have a definite order  $\mathcal{N}_q \circ \mathcal{N}_p$ . (b) When the control qubit is on,  $|\psi_c\rangle = |1\rangle$ , the order is  $\mathcal{N}_p \circ \mathcal{N}_q$ . (c) When the control is in a superposition,  $|\psi_c\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ , the channels have an indefinite order.

\*k.goswami@uq.edu.au

†m.romero@uq.edu.au

describe the noisy channels acting on the target qubit  $\rho_t$  as

$$\mathcal{N}_p(\rho_t) = \sum_{i=0}^3 p_i \sigma_i \rho_t \sigma_i^\dagger, \quad (1)$$

$$\mathcal{N}_q(\rho_t) = \sum_{i=0}^3 q_i \sigma_i \rho_t \sigma_i^\dagger, \quad (2)$$

where  $\sum_i p_i = \sum_i q_i = 1$ . As the equation suggests, each Pauli channel is a probabilistic mixture of all three Pauli errors via the identifications:  $\sigma_1 \equiv \sigma_x$  (bit flip);  $\sigma_3 \equiv \sigma_z$  (phase flip);  $\sigma_2 \equiv \sigma_y$  (combination of bit flip and phase flip); and  $\sigma_0 \equiv I$ . The order that these two channels are applied to  $\rho_t$  is selected by a control qubit  $\rho_c$ . If the control is off,  $|0\rangle_c$ , then the order is  $\mathcal{N}_q \circ \mathcal{N}_p$ , i.e.,  $\mathcal{N}_p$  is before  $\mathcal{N}_q$ . If the control is on,  $|1\rangle_c$ , then  $\mathcal{N}_p \circ \mathcal{N}_q$ . However, if the control qubit is in a superposition state,  $|\psi\rangle_c = (|0\rangle + |1\rangle)_c / \sqrt{2}$ , the channels  $\mathcal{N}_p$  and  $\mathcal{N}_q$  have indefinite causal order. Specifically, given a target qubit  $\rho_t$  and a control qubit initially in the state  $\rho_c = |\psi_c\rangle\langle\psi_c|$ , where  $|\psi_c\rangle = \sqrt{\gamma}|0\rangle + \sqrt{1-\gamma}|1\rangle$ , the total output state of the switch becomes

$$\mathfrak{s}[\mathcal{N}_p, \mathcal{N}_q](\rho_c \otimes \rho_t) = \sum_{i,j} K_{ij}(\rho_c \otimes \rho_t) K_{ij}^\dagger \quad (3)$$

with

$$K_{ij} = p_i q_j (|0\rangle\langle 0| \otimes \sigma_i \sigma_j + |1\rangle\langle 1| \otimes \sigma_j \sigma_i). \quad (4)$$

In matrix representation, the overall output state becomes

$$\mathfrak{s}[\mathcal{N}_p, \mathcal{N}_q](\rho_c \otimes \rho_t) = \begin{pmatrix} A & B \\ B & \tilde{A} \end{pmatrix}, \quad (5)$$

with

$$\begin{aligned} A &= \gamma \mathcal{N}_q \circ \mathcal{N}_p(\rho_t), \\ \tilde{A} &= (1-\gamma) \mathcal{N}_p \circ \mathcal{N}_q(\rho_t), \\ B &= \sqrt{\gamma(1-\gamma)}(\epsilon_+(\rho_t) - \epsilon_-(\rho_t)), \end{aligned} \quad (6)$$

where  $(\epsilon_-)$   $\epsilon_+$  represents an auxiliary trace nonpreserving map composed of all the operators from  $\mathcal{N}_p$  and  $\mathcal{N}_q$  that (anticommute) commute. That is,

$$\epsilon_+(\rho_t) = \sum_{i=0}^3 p_i q_i \rho_t + \sum_{i=0}^3 r_{0i} \sigma_i \rho_t \sigma_i^\dagger, \quad (7)$$

$$\epsilon_-(\rho_t) = r_{12} \sigma_3 \rho_t \sigma_3^\dagger + r_{23} \sigma_1 \rho_t \sigma_1^\dagger + r_{31} \sigma_2 \rho_t \sigma_2^\dagger, \quad (8)$$

with  $r_{ij} = p_i q_j + p_j q_i$ . Note that  $\mathcal{N}_p \circ \mathcal{N}_q(\rho_t) = \epsilon_+(\rho_t) + \epsilon_-(\rho_t) = \mathcal{N}_q \circ \mathcal{N}_p(\rho_t)$ , any definite order of  $\mathcal{N}_p$  and  $\mathcal{N}_q$  will have the same effect on the target qubit. To take an extreme example, if  $\mathcal{N}_p$  and  $\mathcal{N}_q$  were both completely depolarising, either of the definite orders  $\mathcal{N}_p \circ \mathcal{N}_q(\rho_t)$  or  $\mathcal{N}_q \circ \mathcal{N}_p(\rho_t)$  will completely scramble the target qubit.

Interestingly, the output of a quantum switch [Eqs. (5) and (6)] implies that some information is contained in the control qubit. Depending on the outcome of a measurement in the  $\sigma_1^c$  basis of the control qubit, we obtain either of the conditional states  $\epsilon_+(\rho_t)$  or  $\epsilon_-(\rho_t)$ . This means that a  $\sigma_1^c$  measurement allows us to estimate information encoded in the target qubit, and confirm whether there is a communication advantage when  $\mathcal{N}_p$  and  $\mathcal{N}_q$  are in an indefinite order.

In our quantum switch experiment, the control qubit is the polarization of light. A  $\sigma_1^c$  measurement in this case is equivalent to a Stokes measurement  $S_2(\mathfrak{s}[\mathcal{N}_p, \mathcal{N}_q])$  at the output of the quantum switch  $\mathfrak{s}[\mathcal{N}_p, \mathcal{N}_q]$ . As Eq. (2) suggests, the channels  $\mathcal{N}_p$  and  $\mathcal{N}_q$  can be constructed from combinations of Pauli operations  $\{\sigma_i\}$  acting on the target qubit. In Appendix, we show how  $S_2(\mathfrak{s}[\mathcal{N}_p, \mathcal{N}_q])$  can be calculated from Stokes measurements  $S_2(\mathfrak{s}[\sigma_i, \sigma_j])$  at the output of a quantum switch,  $\mathfrak{s}[\sigma_i, \sigma_j]$ . This means that rather than physically implementing  $\mathfrak{s}[\mathcal{N}_p, \mathcal{N}_q]$ , we can simply use Stokes measurements from  $\mathfrak{s}[\sigma_i, \sigma_j]$ , so long as we keep track of the Pauli operations we perform, in effect using an additional memory. The detailed description of our quantum switch and how we implemented the operations  $\{\sigma_i\}$  on the transverse spatial mode—our target qubit—can be found in Appendix.

The Stokes parameters  $S_2(\mathfrak{s}[\sigma_i, \sigma_j])$  are important because we use these to obtain the output control qubit  $\tilde{\rho}_c$  after tracing out the target qubit. With knowledge of  $\tilde{\rho}_c$ , we can then calculate the Holevo capacity  $\chi(\mathfrak{s}[\mathcal{N}_p, \mathcal{N}_q])$ —a measure of the maximum amount of classical information that can be transferred through our arrangement of indefinitely ordered channels. This is given by [6]

$$\chi(\mathfrak{s}[\mathcal{N}_p, \mathcal{N}_q]) = 1 + H(\tilde{\rho}_c) - H^{\min}(\mathfrak{s}[\mathcal{N}_p, \mathcal{N}_q]), \quad (9)$$

where  $H^{\min}(\mathfrak{s}[\mathcal{N}_p, \mathcal{N}_q])$  is the minimum entropy of the two-qubit output of the quantum switch, and  $H(\tilde{\rho}_c)$  is the von-Neumann entropy of the output control qubit  $\tilde{\rho}_c$  (see Appendix for explicit expressions in terms of  $S_2(\mathfrak{s}[\sigma_i, \sigma_j])$ ).

### III. EXPERIMENT

In our experiment, we implemented the quantum switch with all 16 pairs of  $\sigma_i$  and  $\sigma_j$ , as shown in Table I. We measured  $S_2(\mathfrak{s}[\sigma_i, \sigma_j])$ , i.e., the diagonal and antidiagonal components of the polarization of the output light, for each pair  $\sigma_i, \sigma_j$ , with the control qubit initialized to  $|\psi\rangle_c = (|0\rangle + |1\rangle)_c / \sqrt{2}$ —diagonally polarized light. We measure over 10 s; the measured rate at the control output is around 100 000 counts per second. Table I summarizes the results. The first two columns are the ideal settings for the target operations on the transverse spatial mode,  $\sigma_i, \sigma_j$ , the third and fourth columns are the ideal output transformations,  $\tilde{\rho}_t, \tilde{\rho}_c$  of the target and control qubits, respectively. Figure 2 shows the action of the transformations from column 3 on our choice of input target qubit. The last two columns are the ideal and measured values of the Stokes parameter of the control output. We minimize the uncertainty in each measurement by accumulating a large number of counts. Computer-controlled waveplates, with angular orientation uncertainty of  $\pm(2.5 \times 10^{-4})^\circ$ , are used to measure in the diagonal/antidiagonal basis. Our measurements are limited by the nonideal interferometric visibility in our switch. We calculate the average visibility from the values mentioned in Table I, which is  $85 \pm 2\%$ . This is due to several factors. We use inverting prisms that have relatively coarse rotation accuracy  $\pm 1^\circ$  to perform the unitary operations on the spatial mode. Moreover, each optical element is not perfectly flat, introducing wavefront distortions that limit the final visibility.

TABLE I. Data used to calculate the Holevo capacity using Eq. (9) for the different combinations of unitary operations  $\sigma_i$  and  $\sigma_j$ .  $\tilde{\rho}_t$  and  $\tilde{\rho}_c$  are the output target and control state respectively.  $S_2(\mathfrak{s}[\sigma_i, \sigma_j])$  is the Stokes parameter obtained by measuring polarization of the control qubit in the diagonal/antidiagonal basis after the unitaries  $\sigma_i$  and  $\sigma_j$ . We list both the theoretically expected values,  $S_2(\mathfrak{s}[\sigma_i, \sigma_j])^{\text{theor.}}$ , and the experimentally measured values,  $S_2(\mathfrak{s}[\sigma_i, \sigma_j])^{\text{exp.}}$ , of this Stokes parameter. Error bars ( $1\sigma$ ) were calculated by propagation of error on the individual Stokes parameter with Poissonian counting statistics.

$\sigma_i$	$\sigma_j$	$\tilde{\rho}_t$	$\tilde{\rho}_c$	$S_2(\mathfrak{s}[\sigma_i, \sigma_j])^{\text{theor.}}$	$S_2(\mathfrak{s}[\sigma_i, \sigma_j])^{\text{exp.}}$
$\sigma_0$	$\sigma_0$	$ 1\rangle$	D	1	$0.8547 \pm 0.0006$
	$\sigma_1$	$ 0\rangle$	D	1	$0.8718 \pm 0.0005$
	$\sigma_2$	$-i 0\rangle$	D	1	$0.8792 \pm 0.0005$
	$\sigma_3$	$- 1\rangle$	D	1	$0.8823 \pm 0.0005$
$\sigma_1$	$\sigma_0$	$ 0\rangle$	D	1	$0.8459 \pm 0.0006$
	$\sigma_1$	$ 1\rangle$	D	1	$0.8439 \pm 0.0007$
	$\sigma_2$	$-i 1\rangle$	A	-1	$-0.8434 \pm 0.0006$
$\sigma_2$	$\sigma_0$	$-i 0\rangle$	D	1	$0.8473 \pm 0.0007$
	$\sigma_1$	$i 1\rangle$	A	-1	$-0.8600 \pm 0.0005$
	$\sigma_2$	$ 1\rangle$	D	1	$0.8447 \pm 0.0006$
$\sigma_3$	$\sigma_0$	$i 0\rangle$	A	-1	$-0.8278 \pm 0.0008$
	$\sigma_1$	$- 1\rangle$	D	1	$0.8316 \pm 0.0006$
	$\sigma_2$	$ 0\rangle$	A	-1	$-0.8228 \pm 0.0006$
$\sigma_3$	$\sigma_0$	$- 1\rangle$	D	1	$0.8316 \pm 0.0006$
	$\sigma_1$	$ 0\rangle$	A	-1	$-0.8228 \pm 0.0006$
	$\sigma_2$	$-i 0\rangle$	A	-1	$-0.8575 \pm 0.0006$
$\sigma_3$	$\sigma_0$	$ 1\rangle$	D	1	$0.8780 \pm 0.0005$

**A. Two depolarizing channels**

Now that we can calculate the output control qubit from Table I, we can use Eq. (9) to estimate the Holevo capacity  $\chi$ . We compare the experimental and predicted Holevo capacities for several combinations of channels. We first consider a

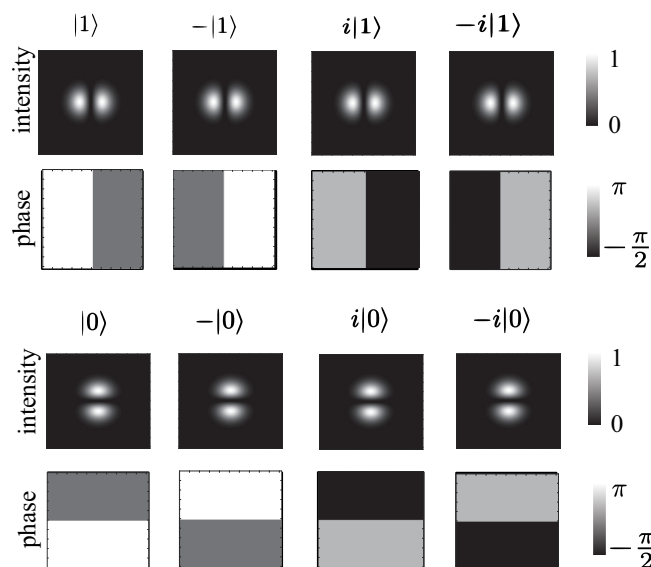


FIG. 2. Predicted spatial mode of target qubit outputs,  $\tilde{\rho}_t$ , after the Pauli operations  $\{\sigma_i, \sigma_j\}$  (see columns 1–3 of Table I). The input target qubit  $\rho_t$  is  $|1\rangle = \text{HG}_{10}$ , a first-order Hermite-Gaussian mode.

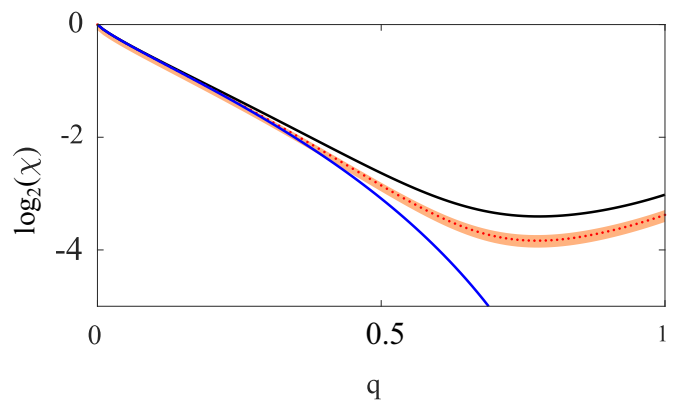


FIG. 3. Logarithm of Holevo capacity  $\chi$ , of two identical depolarising channels, versus depolarising channel strength,  $q$ . The blue line is the predicted Holevo capacity for the case of definite order. The black line is the predicted Holevo capacity for the case of indefinite order. The red dots are the measured values for Holevo capacity in our quantum switch; in excellent agreement with the predicted Holevo capacity for our experimental visibilities  $85 \pm 2\%$ , shown by the orange shaded area. Note that the minimum measured Holevo capacity of  $\chi = (2.1 \pm 0.2) \times 10^{-2}$  bits occurs at a depolarization  $q = 0.78$ , and that higher capacity of  $\chi = (3.4 \pm 0.2) \times 10^{-2}$  bits occurs at full depolarization,  $q = 1$ .

quantum switch  $\mathfrak{s}[\mathcal{N}, \mathcal{N}]$  of two depolarizing channels  $\mathcal{N}$  of identical strengths  $q$ , i.e., in Eq. (2), we set all coefficients to  $q/4$ :

$$\mathcal{N} = (1 - 3q/4)\rho_t + q/4(\sigma_1\rho_t\sigma_1 + \sigma_2\rho_t\sigma_2 + \sigma_3\rho_t\sigma_3), \quad (10)$$

This scenario was theoretically studied in Ref. [6]. Experimental and predicted Holevo capacities are shown in Fig. 3, which plots the logarithm of  $\chi$  against increasing  $q$ . The red circles are our measured values; the orange shaded region is the predicted Holevo capacity for visibilities of  $85 \pm 2\%$ ; the black curve is the Holevo capacity for an ideal quantum switch. The blue curve is the ideal Holevo capacity for two depolarising channels in some definite order which—as expected—decreases monotonically with increasing depolarising strength. In the limit of two fully depolarising channels,  $q = 1$ ,  $\chi = 0$  bits are transmitted using definite order. In this limit there is a clear advantage in using quantum channels: ideally  $4.9 \times 10^{-2}$  bits can be transmitted, we measure  $\chi = (3.4 \pm 0.2) \times 10^{-2}$  bits. This is a rather counterintuitive result as individually none of the channels can transmit any information.

This nonzero Holevo capacity can be understood intuitively from the fact that the output of a quantum switch with two depolarising channels is a statistical mixture of the output of a quantum switch with different Pauli operations  $\{\sigma_i\}$ . Some of these Pauli operations anticommute, hence superpositions of the order of anticommuting Pauli operations can preserve a finite amount of information in the target qubit.

This intuition helps us understand another striking prediction, which is that above some nonzero depolarization strength the Holevo capacity will increase. In the ideal case, we see that Holevo capacity attains a minimum value of  $3.3 \times 10^{-2}$  bits at  $q = 0.78$  and then the capacity increases to the limit of  $4.9 \times 10^{-2}$  bits at  $q = 1$ , in stark contrast to the classical

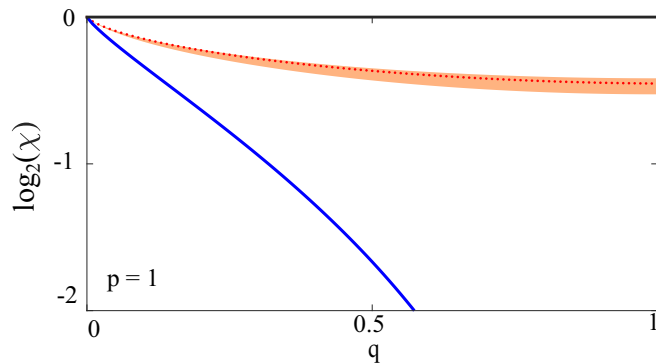


FIG. 4. Logarithm of Holevo capacity  $\chi$ , where one channel is a depolarising channel of varying strength  $q$ , and the other is a  $\sigma_3$  channel, the channel in Eq. (11) with  $p = 1$ . The solid black line [ $\log_2(\chi) = 0$ ] is the theoretical predictions for Holevo capacity for indefinite causal order. The orange shade show the uncertainty due to nonideal visibility in our experiment. The red dots are the experimentally measured data points and the blue line is the theoretical predictions for Holevo capacity for definite order. We show the special case for  $p = 1$ , i.e., the dephasing channel becomes a  $\sigma_3$  channel. Note that in this situation, indefinite causal order allows us to get full 1 bit information. Experimentally we measure a Holevo capacity of  $0.64 \pm 0.02$  bits at  $q = 1$ .

case of definite causal order which decreases monotonically to zero. Experimentally, we see this increase in information capacity from  $\chi = (2.1 \pm 0.2) \times 10^{-2}$  bits at  $q = 0.78$ , i.e., at its worst absolute performance it is  $13.5 \sigma$  above the classical performance at that value of  $q$ .

We look at Eq. (9) to understand this behavior: for low depolarization strengths, the minimum entropy  $H^{\min}(\mathfrak{s}[\mathcal{N}_p, \mathcal{N}_q])$  increases more rapidly than the von Neumann entropy  $H(\tilde{\rho}_c)$ . This means that for low  $q$ , the rate of depolarization of the composite system—target and control—is faster than the rate of depolarization of the control qubit. However, at  $q \approx 0.78$ ,  $H(\tilde{\rho}_c)$  begins to increase more rapidly than  $H^{\min}(\mathfrak{s}[\mathcal{N}_p, \mathcal{N}_q])$  so the depolarization rate of the control overtakes the depolarization rate of the composite system, and the information revival occurs.

### B. Depolarizing and dephasing channel

Consequently, as a second example we consider the quantum switch  $\mathfrak{s}[\mathcal{N}, \mathcal{M}]$  composed of a depolarizing channel  $\mathcal{N}$  of strength  $q$  Eq. (10) and a dephasing channel  $\mathcal{M}$  of strength  $p$  given by

$$\mathcal{M} = [(1 - p)\rho_t + p\sigma_3\rho_t\sigma_3]. \quad (11)$$

When  $q = 1$ , any definite order of these two channels results in a fully depolarizing channel, regardless of the value of  $p$ , thus  $\chi = 0$ . However, with the additional freedom to prepare and measure the control qubit, the Holevo capacity of these two channels in indefinite order,  $\chi(\mathfrak{s}[\mathcal{N}, \mathcal{M}]) \equiv \chi_i$  can be nonzero and is a function of  $p$ . In fact, this combination achieves the maximum value of 1 bit when  $p = 1$  regardless of the depolarization strength  $q$ , as shown in Fig. 4 (black line). We emphasize the contrast with the zero capacity of the definitely ordered channels at  $q = 1$ , Fig. 4 (blue line).

This unit Holevo capacity can be understood by noting that when  $p = 1$ ,  $\mathcal{M}$  is a unitary channel and the conditional states become

$$\begin{aligned} \epsilon_+(\rho_t) &= (q/4)\rho_t + (1 - 3q/4)\sigma_3\rho_t\sigma_3, \\ \epsilon_-(\rho_t) &= q/4(\sigma_1\rho_t\sigma_1 + \sigma_2\rho_t\sigma_2), \end{aligned} \quad (12)$$

and thus from Eq. (9)  $\chi_i = 1 + H(\tilde{\rho}_c) - H_{\rho_t}^{\min}[\epsilon_+(\rho_t) \oplus \epsilon_-(\rho_t)]$ , where  $H(\cdot)$  is again the von Neumann entropy and the minimization is done over all possible target states. Notice that the states that minimize the entropy of  $\epsilon_+(\rho_t) \oplus \epsilon_-(\rho_t)$  are the eigenvectors of  $\sigma_3$ . Take  $\rho = |0\rangle\langle 0|$  as an example,  $\epsilon_+(\rho_t) = (1 - q/2)|0\rangle\langle 0|$ , while  $\epsilon_-(\rho_t)$  will be  $(q/2)|1\rangle\langle 1|$  in which case,  $H[\epsilon_+(\rho) \oplus \epsilon_-(\rho)] = H(q/2)$ , and thus the Holevo-capacity of the quantum switch becomes 1 bit.<sup>1</sup> We compare our experimental and predicted  $\chi_i$  in Fig. 4, red dots are experimental values and the orange shaded regions accounts for nonideal visibility. For full depolarization at  $q = 1$ , we show  $0.64 \pm 0.02$  bits compared to the ideal case of  $\chi_i = 1$ . Interestingly, this ideal capacity is strictly larger than the one achieved in Ref. [15] given the same channels in a path superposition. More strongly, in Appendix, we prove that these channels in a superposition of path cannot lead to unit capacity. We also compare predicted and experimental  $\chi_i$  for other dephasing strengths  $p$  in Appendix.

We note that when [6] proposed communication advantage from indefinite causal order they did not claim that the advantage is unique to coherent superposition of order. Subsequently, there has been an active discussion on the origin of the communication advantage [15–17] with consensus yet to be reached. The experimental developments in the present work and [13] are further motivation to develop a well-defined resource communication theory featuring coherence.

## IV. CONCLUSION

Quantum mechanics allows indefinite causal order, which allows us to communicate up to 1 (0.049) bits of information through one (two) fully depolarising channels. This is possibly useful for communication through turbid media: there have been proposals for quantum communication protocols using a Sagnac interferometer (see Ref. [18] for a review) reminiscent of the quantum switch implementation of Wei *et al.* [12]. We can imagine a situation where one arm of the Sagnac loop consists of a unitary channel in the laboratory and the other arm goes through the atmosphere, e.g., ground-satellite-ground path. Placing these channels in indefinite causal order will enable to transmission of qudits in the transverse spatial mode even though the atmosphere is naturally noisy for these qudits. We can also use this idea for secret sharing, where a specific combination of channels, when connected indefinitely, can transfer information between two parties, but in a scenario of eavesdropping, any intervention will break the

<sup>1</sup>This maximum advantage over definite order is also attained by having an indefinite order of one fully depolarising channel and a unitary  $\sigma_2$  or  $\sigma_3$  channel, when the input states are the eigenstates of  $\sigma_2$  and  $\sigma_3$ , respectively.



“indefiniteness,” and the message will remain scrambled. We can also think of imaging through scattering media where the information being communicated is in the modulation of the spatial mode, e.g., imaging through skin. In all these cases, placing the noisy channel in indefinite causal order with a clean channel will provide information that would normally be entirely lost.

**ACKNOWLEDGMENTS**

We thank Howard Wiseman for suggesting this collaboration. We also thank Giulio Chiribella for a helpful discussion. K.G. thanks the organizers of the Quantum Information Structure Of Spacetime, 2020 workshop where valuable insights related to this work were discussed. This work has been supported by the Australian Research Council (ARC) by Centre of Excellence for Engineered Quantum Systems (EQUS, CE170100009), a DECRA grant (DE160100409), Advance Queensland Funding, and L’Oreal-UNESCO FWIS grant for J.R.; a DECRA grant (DE170100088) for GPS; Y.C. is supported by a CSC scholarship; the University of Queensland by a Vice-Chancellor’s Senior Research and Teaching Fellowship for A.G.W.; and 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).

**APPENDIX**

**1. Estimating the Stokes parameter of the quantum switch**

With the Pauli decomposition discussed above, we now show how measurements of the control qubit,  $\tilde{\rho}_c$ , can be used to estimate the Holevo capacity. Since our control qubit is polarization, we express the output control in terms of the Stokes vector  $(S_1, S_2, S_3)$  [19]:

$$\tilde{\rho}_c = \frac{1}{2} \begin{pmatrix} 1 + S_1 & S_2 + i S_3 \\ S_2 - i S_3 & 1 - S_1 \end{pmatrix}. \tag{A1}$$

From Eq. (5), if we measure the output control state,  $\tilde{\rho}_c$  becomes

$$\begin{aligned} \tilde{\rho}_c &= \begin{pmatrix} \text{Tr}(A) & \text{Tr}(B) \\ \text{Tr}(B) & \text{Tr}(\tilde{A}) \end{pmatrix} \\ &= \begin{pmatrix} \gamma & \text{Tr}(B) \\ \text{Tr}(B) & 1 - \gamma \end{pmatrix}, \end{aligned} \tag{A2}$$

where the second equality comes from the fact that, Pauli channels are trace preserving. It is easy to see that  $\text{Tr}(B)$  is a real number, so comparing Eq. (A1) and (A2) at  $\gamma = 1/2$ , we can see that  $S_2 = 2 \text{Tr}(B)$  and  $S_1 = S_3 = 0$ . Thus we consider the effect of the quantum switch on the  $S_2$  component of the control, note that to distinguish the operations on the control and target, we use  $\sigma_0^c, \sigma_1^c, \sigma_2^c$ , and  $\sigma_3^c$  for the control qubit and  $\sigma_0, \sigma_1, \sigma_2$ , and  $\sigma_3$  for the target qubit,

$$\begin{aligned} S_2(\mathfrak{s}[\mathcal{N}_p, \mathcal{N}_q]) &= \text{Tr} \{ (\sigma_1^c \otimes \sigma_0) \cdot \mathfrak{s}[\mathcal{N}_p, \mathcal{N}_q] \} \\ &= \sum_{i,j} p_i q_j \text{Tr} \{ (\sigma_1^c \otimes \sigma_0) \cdot \mathfrak{s}[\sigma_i, \sigma_j] \} \\ &= \sum_{i,j} p_i q_j S_2(\mathfrak{s}[\sigma_i, \sigma_j]). \end{aligned} \tag{A3}$$

The above equation shows that with knowledge of the control qubit for individual combinations of  $\sigma_i, \sigma_j$  we can get the  $S_2$  of the switch with channels  $\mathcal{N}_p$  and  $\mathcal{N}_q$ .

From Eq. (A3), we get  $\tilde{\rho}_c$ , and we can calculate the von Neumann entropy  $H(\tilde{\rho}_c)$ , which is necessary to obtain the Holevo capacity as shown in Eq. (9). The other quantity needed to evaluate (9) is the minimum entropy of the total output  $H^{\min}(\mathfrak{s}[\mathcal{N}_p, \mathcal{N}_q])$ . In order to calculate this, let us first write the action of the quantum switch, in terms of individual combinations of Pauli operations:

$$\begin{aligned} \mathfrak{s}[\mathcal{N}_p, \mathcal{N}_q] (\rho_c \otimes \rho_t) &= \sum_{i,j} p_i q_j (\gamma |0\rangle\langle 0|_c \otimes \sigma_i \sigma_j \rho_t \sigma_i^\dagger \sigma_j^\dagger \\ &\quad + (1 - \gamma) |1\rangle\langle 1|_c \otimes \sigma_j \sigma_i \rho_t \sigma_i^\dagger \sigma_j^\dagger \\ &\quad + \sqrt{\gamma(1 - \gamma)} |0\rangle\langle 1|_c \otimes \sigma_i \sigma_j \rho_t \sigma_i^\dagger \sigma_j^\dagger \\ &\quad + \sqrt{\gamma(1 - \gamma)} |1\rangle\langle 0|_c \otimes \sigma_j \sigma_i \rho_t \sigma_i^\dagger \sigma_j^\dagger) \\ &= \sum_{i,j} p_i q_j \mathfrak{s}[\sigma_i, \sigma_j] (\rho_c \otimes \rho_t). \end{aligned} \tag{A4}$$

From Eq. (A4), we notice that from pairwise combinations of  $\sigma_i$  and  $\sigma_j$  operators we can find out the output  $\mathfrak{s}[\mathcal{N}_p, \mathcal{N}_q] (\rho_c \otimes \rho_t)$  and its minimum entropy  $H^{\min}(\mathfrak{s}[\mathcal{N}_p, \mathcal{N}_q])$ , given we use an optimized target state  $\rho_t$ . Note that, the operations in  $\{\sigma_i\}$  either commute or anti-commute. We can construct  $\mathfrak{s}[\sigma_i, \sigma_j] (\rho_c \otimes \rho_t)$  by projecting the control qubit into the diagonal/antidiagonal basis which results to a product state. Denoting the anticommutator by  $\{\dots\}$  and the commutator by  $[\dots]$ , we have, with  $\gamma = 1/2$ ,

$$\mathfrak{s}[\sigma_i, \sigma_j] (\rho_c \otimes \rho_t) = \begin{cases} |+\rangle\langle +|_c \otimes \{\sigma_i, \sigma_j\} \rho_t \{\sigma_i, \sigma_j\}^\dagger \\ \text{for } [\sigma_i, \sigma_j] = 0 \\ |-\rangle\langle -|_c \otimes [\sigma_i, \sigma_j] \rho_t [\sigma_i, \sigma_j]^\dagger \\ \text{for } \{\sigma_i, \sigma_j\} = 0 \end{cases}.$$

The value of the Stokes parameter  $S_2(\mathfrak{s}[\sigma_i, \sigma_j])$ , which is 1 (−1) for commuting (anticommuting) operations in the ideal case, can be experimentally obtained by noting that  $S_2(\mathfrak{s}[\sigma_i, \sigma_j]) = \text{Tr}[(\sigma_1^c \otimes \sigma_0) \mathfrak{s}[\sigma_i, \sigma_j] (\rho_c \otimes \rho_t)]$ , i.e., expectation value of the operator  $\sigma_1^c \otimes \sigma_0$ .

**2. Experimental details**

Our experimental setup is adapted from the quantum switch of Ref. [11], removing cylindrical lenses and using only inverting prisms, since only Pauli operations were necessary, Fig. 5. Our input light is diagonally polarized ( $|\rho_c\rangle = (|0\rangle + |1\rangle)/\sqrt{2} = D$ ), and is in the first-order spatial mode ( $|\rho_t\rangle = |1\rangle = |\text{HG}_{10}\rangle$ ).

We realize the unitary operations  $\{\sigma_i\}$  using a pair of rotating prisms [20] as shown in Fig. 5. A mechanical rotation of the inverting prism results in a rotation of the incoming spatial mode of the photon, the outputs of the combined operation  $\{\sigma_i \circ \sigma_j\}$  are shown in Fig. 2.

To implement  $\sigma_i, \sigma_j$ , we use up to two rotating prisms [11] in each arm. Each prism is oriented at a physical angle  $\theta$ , which reflects and rotates an incoming beam by  $2\theta$ . The action of the rotating prism on our target qubit subspace is

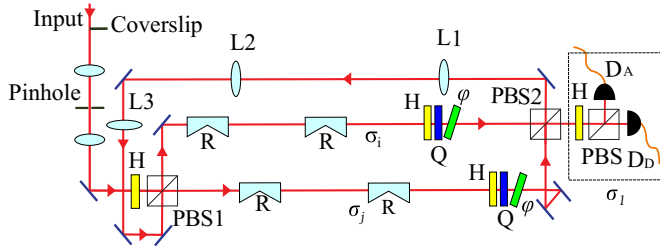


FIG. 5. Schematic of quantum switch. R is the rotating prism (A5), H and Q are the half and quarter waveplates, respectively, and  $\varphi$  are the phase plates. The control qubit is the polarization,  $|\psi_c\rangle = D$ . The polarization of the light controls order of Pauli operations  $\{\sigma_i\}$  acting on the photonic spatial mode  $|\psi_r\rangle = \text{HG}_{10}$ , for horizontal polarization  $H$ , the order is  $\sigma_i \circ \sigma_j$ , for the vertical polarization  $V$  the order is  $\sigma_j \circ \sigma_i$ . The polarization  $D$  ensures superposition of the orders.  $X$  is a polarization measurement, determining the Stokes parameter of the measured photon in the diagonal/antidiagonal basis. Lenses L1 to L3 form a telescope for mode-matching.

represented by the following unitary operation:

$$R(\theta) = \begin{pmatrix} -\cos 2\theta & \sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix}. \quad (\text{A5})$$

We impart the global phase  $\varphi$  by a tilted phase plate. We write the transformation performed by the pair of prisms and the phase plate as

$$U(\varphi, \theta_1, \theta_2) = e^{i\varphi} R(\theta_2) R(\theta_1). \quad (\text{A6})$$

We place the phase plate only when we are doing  $\sigma_2$  operation. For operations,  $\sigma_1$  and  $\sigma_3$  we need only one rotating prism, and for  $\sigma_0$  and  $\sigma_2$  we need both prisms. We achieve this by moving the second rotating prism via a translation stage. Table II are the angles and phases used to implement the Pauli operations. Since this rotation also changes the polarization of the field—the control qubit—which is not desired, this rotation is canceled after the two rotating prisms by the actions of the half- (H), and quarter- (Q) waveplates and the phase plate ( $\phi$ ), see Fig. 5.

We emphasize that in our architecture, the sender cannot access the individual channels—channels  $\mathcal{N}_p$  or  $\mathcal{N}_q$ —without using the other one. That is, when we consider combinations of channels where one of the channels is unitary, this unitary channel cannot be accessed without also going through the noisy channel. This is certainly not a unique realization of indefinite causal order. For example, one can use a Sagnac interferometer to achieve the same indefinite causal order as in Ref. [12].

TABLE II. Phases and angles for the unitary operations realized in our experiment, given by Eq. (A6).

Unitary	$\varphi$	$\theta_1$	$\theta_2$
$\sigma_0$	0	$\frac{\pi}{2}$	$\frac{\pi}{2}$
$\sigma_1$	0	$\frac{\pi}{4}$	—
$\sigma_2$	$\frac{\pi}{2}$	$\frac{\pi}{2}$	$\frac{\pi}{4}$
$\sigma_3$	0	$\frac{\pi}{2}$	—

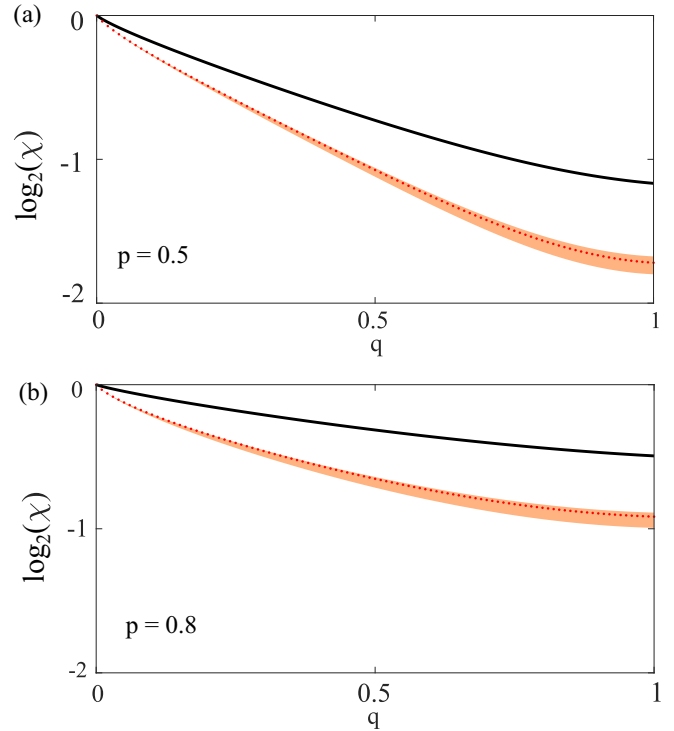


FIG. 6. Logarithm of Holevo capacity  $\chi$ , where one channel is a depolarising channel of varying strength  $q$ , and the other one is a dephasing channel of strength  $p$ . The solid black lines are the theoretical predictions for Holevo capacity for indefinite causal order. The orange shades show the uncertainty due to nonideal visibility in our experiment. The red dots are the experimentally measured data points. (a) shows the case when  $p = 0.5$ . In this case, the ideal value is  $\chi_i = 0.31$  bits, whereas we experimentally measure  $0.179 \pm 0.006$  bits. (b) is the case when  $p = 0.8$ . Here, the  $\chi_i$  is 0.61 bits and we measure  $0.42 \pm 0.01$  bits.

### 3. Effects of experimental imperfections

We note that the Pauli operations we are implementing by the rotating prisms, can be nonideal due to uncertainty of the angles. This can affect the experiment in two ways. First, it leads to nonzero  $S_1$  and  $S_3$  Stokes parameters in the output control qubit and second, these nonzero terms contribute to  $H^{\min}(\mathfrak{s}[\mathcal{N}_p, \mathcal{N}_q])$ . To account for this issue, we note that the uncertainty associated with our rotation mounts is  $\pm 1^\circ$ . We numerically introduce random uniformly distributed error of  $\pm 1^\circ$  to the angles of the rotating prisms. We repeat the simulation for 500 iterations and measure the capacity in each run. We confirm that the contribution of the random errors are well within the range of our experimental visibilities, which is reflected in the orange shades of the graphs in Figs. 3, 4, 6, and 7.

### 4. More general combinations of depolarising and dephasing channels

In the main section, we have introduced a combination of depolarising channel of strength  $q$  and a dephasing channel of strength  $p$  as shown in Eq. (11). We have shown the special case for  $p = 1$  where it is possible to achieve unit Holevo capacity. In this section we show more general situations where

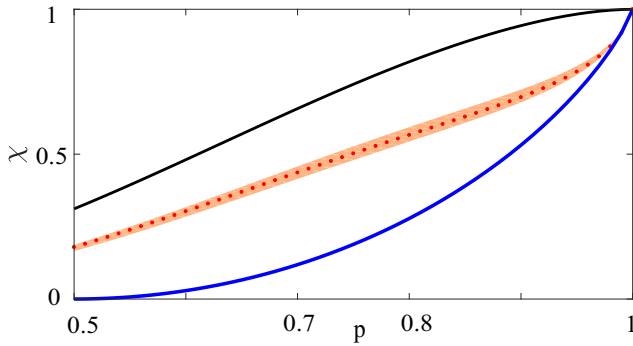


FIG. 7. Experimental Holevo capacity  $\chi$  for  $p \geq 0.5$  region, where one channel is a bit-flip channel and the other one is a dephasing channel with varying strength  $p$  as in the Eq. (A7). The solid black line is the theoretical predictions for Holevo capacity for indefinite causal order. The blue line is the theoretical prediction for Holevo capacity for definite causal order. The red dots are the experimental data points and the orange shade is the expected range due to nonideal visibility. At  $p = 0.5$ , the experimentally measured Holevo capacity is  $0.179 \pm 0.006$  bits, whereas  $\chi_d = 0$ .

$p \leq 1$ . In Fig. 8, the black lines show the cases for either  $p = 0.5, 0.8$ , or  $1$  combined with a depolarising channel of strength  $q$  in an indefinite order. The blue line shows the capacity when these channels are in definite order, which regardless of the value of  $p$  monotonically decreases to zero when  $q = 1$ . In contrast, with indefinite order arrangements the predicted Holevo capacities are 0.31 bits, 0.61 bits and 1 bit for  $p = 0.5, 0.8$ , and  $1$ , respectively. In Fig. 6, we plot the logarithm of Holevo capacities as a function of both depolarising channel strength  $q$  for 2 different strengths of the dephasing channel,  $p = 0.5$  and  $0.8$ . The black curves on the graphs show the predicted Holevo capacity for the indefinite order, the orange shade is the uncertainty due to the fringe visibility and the red dots are experimental data points. In the case of  $p = 0$ , the

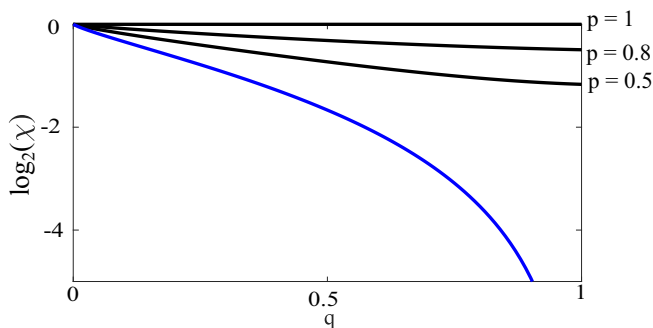


FIG. 8. Predicted logarithm of Holevo capacity  $\chi$  for the channels in Eq. (11) vs depolarization noise strength,  $q$ . The solid black lines are for indefinite order in the cases of  $p = 0.5, 0.8$ , and  $1$ . The solid blue line are for definite order, regardless of the value of  $p$ , the combination of the dephasing and depolarising channels scramble the information equally. At the full depolarization strength,  $q = 1$ ,  $\chi_d = 0$ , whereas  $\chi_i$  are 0.31 bits, 0.61 bits, and 1 bit in case of  $p = 0.5, 0.8$ , and  $1$ , respectively.

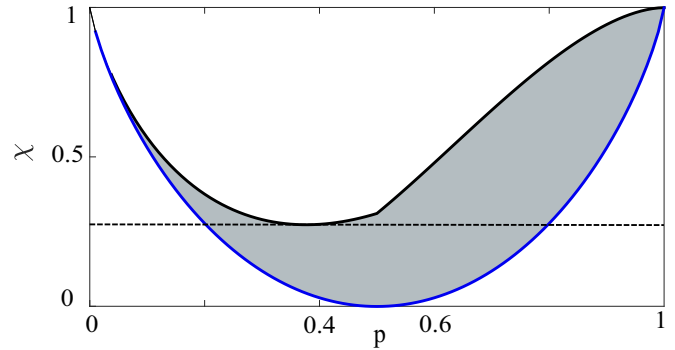


FIG. 9. Predicted Holevo capacity  $\chi$  for the channels in Eq. (A7) vs noise parameter,  $p$ . The solid black line is the Holevo capacity of depolarising channels in indefinite order and the solid blue line is for the case of definite order. We observe a knee at  $p = 0.5$ , this is because the optimum state for  $p \leq 0.5$  are the eigenvectors of the operator  $\sigma_1$  and  $\sigma_2$ . On the other hand, for  $p \geq 0.5$  the optimum state is the eigenvector of  $\sigma_3$ . At  $p = 0.5$ , the definite-ordered channel becomes completely depolarising, making  $\chi_d = 0$ , whereas at this point  $\chi_i = 0.31$  bits. In the region of  $p \geq 0.5$ , the maximum difference between the two cases is 0.55 bits, occurring when  $p = 0.75$ . The black horizontal dashed line denotes the minimum capacity of the indefinite-ordered case, 0.27 bits when  $p = 0.37$ .

dephasing channel behaves as an identity channel and the input state  $\rho_i$  is equally scrambled in both definite and indefinite order cases, thus resulting no overall advantage. However, in other cases there is nonzero information-theoretic advantage. At a full depolarising strength of  $q = 1$ : in the case of  $p = 0.5$ , the predicted  $\chi_i = 0.31$  bits, whereas we experimentally show  $0.179 \pm 0.006$  bits; at  $p = 0.8$ , the predicted  $\chi_i = 0.61$  bits, we experimentally show  $0.42 \pm 0.01$  bits.

### 5. Bit-flip and bit-phase-flip channels

Consider two noisy channels, a bit-flip and a bit-phase-flip channel with strength  $p$  given, correspondingly, by

$$N_p^1(\rho_i) = (1 - p)\rho_i + p\sigma_1\rho_i\sigma_1, \\ N_p^2(\rho_i) = (1 - p)\rho_i + p\sigma_2\rho_i\sigma_2. \tag{A7}$$

Then,

$$\epsilon_+(\rho_i) = (1 - p)^2\rho_i + p(1 - p)(\sigma_1\rho_i\sigma_1 + \sigma_2\rho_i\sigma_2), \\ \epsilon_-(\rho_i) = p^2\sigma_3\rho_i\sigma_3. \tag{A8}$$

It should be highlighted that the outcome corresponding to the  $|- \rangle \langle -|$ , namely, the  $\epsilon_-(\rho_i)/p^2$  channel, is actually a purely unitary channel, capable of achieving perfect transmission despite both channels being noisy (as was also pointed out recently in Ref. [7]). However, here we are interested in the Holevo capacity resulting from considering both possible outcomes weighted by their corresponding probabilities, rather than the post-selected  $\epsilon_-(\rho_i)$ . We solve the corresponding minimization problem to compute the Holevo capacity. Figure 9 plots the Holevo capacity  $\chi$  against different noise parameters  $p$  for both the cases of noisy channels in a definite order, and in an indefinite order. The black curve is the

Holevo capacity for indefinite order, while the blue curve is for definite order. As shown by the difference of these two curves, indefinite order provides an advantage over the definite-order case. Indefinite order allows transmission of 0.27 bits or more over the whole range of  $p$  values. The minimum capacity of 0.27 bits which occurs at 0.37, is higher than the capacity of the definite-order case from  $p = 0.2$  to 0.8, 60% of the range of  $p$ . The maximum advantage  $\max[(\chi_i - \chi_d)] = 0.55$  bits occurs at  $p = 0.75$ , where the Holevo capacity for the quantum switch and definite order cases are 0.75 and 0.19 bits, respectively. We can see there is a knee in the  $\chi_i$  at noise parameter  $p = 0.5$ . This is because when  $p \leq 0.5$ , the eigenvectors of  $\sigma_1$  or  $\sigma_2$  achieve the minimum entropy  $H^{\min}(\mathfrak{s}[\mathcal{N}_p^1, \mathcal{N}_p^2])$ , while for  $p \geq 0.5$ , the corresponding optimum state turns out to be the eigenvectors of  $\sigma_3$ . When  $p = 0.5$ , where the classical order results in full depolarization, the capacity resulting from indefinite order is  $\chi_i \approx 0.31$  bits. In this case, any of the eigenstates of the Pauli operator achieves the maximum capacity thus increasing the domain over which we can optimally encode the input.

### 6. Holevo capacity from path superposition

Although the advantage in communication is not unique to superposition of order, there is strong reason to believe that the perfect transmission discussed above is not possible via a superposition of paths, where we place two noisy channels in two arms of an interferometer [15]. The optimized Holevo capacity that we obtain for a fully depolarising channel and a unitary channel in path superposition is 0.75 bits. Moreover, we prove that reaching a value of 1 bit is impossible.

In our communication task, we consider the target system, the control system and two Pauli channels to transfer the target qubit. The control enables the target to go through a superposition of paths. In our experiment,  $|\psi_c\rangle$  is set to be  $|+\rangle$ , the target state is optimized to achieve the maximum communication capacity, and a generalized measurement on the control system is applied after noisy channels.

We consider our Pauli channels to be the depolarising channel  $\mathcal{N}$ , as shown in Eq. (10), and  $\sigma_3$ . Following Ref. [15], after tracing out the environment, the output control and target state takes the following form:

$$\rho_{\text{tot}}^{(ct)} = \frac{1}{2}(|0\rangle\langle 0|^c \otimes \mathcal{N}(\rho_t) + |1\rangle\langle 1|^c \otimes \sigma_3 \rho_t \sigma_3 + |0\rangle\langle 1|^c \otimes T_0 \rho_t \sigma_3 + |1\rangle\langle 0|^c \otimes \sigma_3 \rho_t T_0^\dagger), \quad (\text{A9})$$

where  $T_0 = \sum_i e_i K_i$ ,  $K_i$  are arbitrary choices for Kraus operators for the depolarising channel with the constraint  $\sum_i |e_i|^2 = 1$ . Exploiting the unitary freedom in the operator-sum representation (see, for example, Ref. [21]), one has that  $e_i K_i = e_i \sum_j U_{i,j} \frac{\sigma_j}{2}$ , with  $U_{i,j}$  the entries of a unitary matrix, and thus  $T_0 = \sum_j (\sum_i e_i U_{i,j}) \frac{\sigma_j}{2} \equiv \sum_j f_j \sigma_j / 2$ , with  $\sum_j |f_j|^2 = 1$ . Now if we measure the control qubit in the  $\{M_+, M_-\}$  basis,

$$\begin{aligned} M_+ &= \cos \alpha |0\rangle + \sin \alpha |1\rangle, \\ M_- &= -\sin \alpha |0\rangle + \cos \alpha |1\rangle, \end{aligned} \quad (\text{A10})$$

the two possible output states, with corresponding probabilities  $p_+$  and  $p_-$ , are given by

$$\begin{aligned} p_+ \rho_+^t &= \frac{1}{4} \left[ \cos^2 \alpha \frac{\sigma_0}{2} + \sin^2 \alpha \sigma_3 \rho_t \sigma_3 \right. \\ &\quad \left. + \frac{\cos \alpha \sin \alpha}{2} \sum_i (f_i \sigma_i \rho_t \sigma_3 + f_i^* \sigma_3 \rho_t \sigma_i) \right], \\ p_- \rho_-^t &= \frac{1}{4} \left[ \sin^2 \alpha \frac{\sigma_0}{2} + \cos^2 \alpha \sigma_3 \rho_t \sigma_3 \right. \\ &\quad \left. - \frac{\cos \alpha \sin \alpha}{2} \sum_i (f_i \sigma_i \rho_t \sigma_3 + f_i^* \sigma_3 \rho_t \sigma_i) \right], \end{aligned} \quad (\text{A11})$$

where  $f_i^*$  is the complex conjugate of  $f_i$ .

Now, achieving a unit Holevo capacity implies that both the input and output state must be pure. In the Bloch representation, this implies that

$$\sum_i |r_i^{(+)}|^2 = 1, \quad (\text{A12})$$

$$\sum_i |r_i^{(-)}|^2 = 1, \quad (\text{A13})$$

$$\sum_i |r_i^{(\text{in})}|^2 = 1, \quad (\text{A14})$$

where  $\{r_i^{(+)}\}$ ,  $\{r_i^{(-)}\}$ , and  $\{r_i^{(\text{in})}\}$  are the Bloch vector components of the projected states  $\rho_+^t$ ,  $\rho_-^t$ , and the input target state,  $\rho_t$  respectively. We proceed by writing  $f_i$  into its real and imaginary parts, i.e.,  $f_j = f_{j,R} + i f_{j,I}$ . Then, from  $\sum |r_i^{(+)}|^2 = \sum |r_i^{(-)}|^2$ , we find that

$$\sin 2\alpha (r_3^{(\text{in})} f_{0,R} + f_{3,R} + r_1^{(\text{in})} f_{2,I} - r_2^{(\text{in})} f_{1,I}) + \cos 2\alpha = 0, \quad (\text{A15})$$

which can be replaced into  $\sum |r_i^{(+)}|^2 + \sum |r_i^{(-)}|^2 = 2$  to obtain

$$\begin{aligned} \sin^2 2\alpha &[(r_2^{(\text{in})} f_{0,I} - r_3^{(\text{in})} f_{1,R} + r_1^{(\text{in})} f_{3,R} + f_{2,I})^2 \\ &+ (r_1^{(\text{in})} f_{0,I} + r_3^{(\text{in})} f_{2,R} + f_{1,I} - r_2^{(\text{in})} f_{3,R})^2 \\ &+ (f_{0,R} + r_1^{(\text{in})} f_{1,R} + r_2^{(\text{in})} f_{2,R} + r_3^{(\text{in})} f_{3,R})^2] + \cos 4\alpha = 2. \end{aligned} \quad (\text{A16})$$

Next we proceed to show that Eq. (A16) cannot hold under conditions (A14) and  $\sum_i |f_i|^2 = \sum_i \sum_{k \in \{R,I\}} f_{i,k}^2 = 1$ . Note that for this purpose, it is sufficient to prove that

$$\begin{aligned} F(r, f) &= (r_2^{(\text{in})} f_{0,I} - r_3^{(\text{in})} f_{1,R} + r_1^{(\text{in})} f_{3,R} + f_{2,I})^2 \\ &+ (r_1^{(\text{in})} f_{0,I} + r_3^{(\text{in})} f_{2,R} + f_{1,I} - r_2^{(\text{in})} f_{3,R})^2 \\ &+ (f_{0,R} + r_1^{(\text{in})} f_{1,R} + r_2^{(\text{in})} f_{2,R} + r_3^{(\text{in})} f_{3,R})^2 < 3. \end{aligned} \quad (\text{A17})$$

Exploiting the relation that  $2xy \leq x^2 + y^2$  for any real  $x$  and  $y$ , Eq. (A14), and that for that for  $|r_i| \leq 1$  one has that  $r_i \leq |r_i|$  and  $|r_i|^2 \leq |r_1|$ , we have



$$\begin{aligned}
F(r, f) &\leq (1 + r_1^{(\text{in})} + r_2^{(\text{in})} + |r_3^{(\text{in})}|^2) f_{0,I}^2 + (1 + r_1^{(\text{in})} + r_2^{(\text{in})} + r_3^{(\text{in})})(f_{0,R}^2 + f_{1,I}^2 + f_{2,I}^2 + f_{3,R}^2) \\
&\quad + (1 + r_1^{(\text{in})} + |r_2^{(\text{in})}|^2 + r_3^{(\text{in})}) f_{1,R}^2 + (1 + |r_1^{(\text{in})}|^2 + r_2^{(\text{in})} + r_3^{(\text{in})}) f_{2,R}^2 \\
&\leq (1 + |r_1^{(\text{in})}| + |r_2^{(\text{in})}| + |r_3^{(\text{in})}|) \sum_i |f_i|^2 \\
&\leq 1 + \sqrt{3 \sum_i |r_i^{(\text{in})}|^2} \leq 1 + \sqrt{3} < 3,
\end{aligned} \tag{A18}$$

as desired. This supports the idea that  $\mathcal{N}$  and a unitary channel in a superposition of paths cannot lead

to a pure output and thus cannot achieve unit Holevo capacity.

- 
- [1] C. E. Shannon, *Bell Syst. Tech. J.* **27**, 379 (1948).  
[2] J. Preskill, [arXiv:1604.07450](https://arxiv.org/abs/1604.07450).  
[3] A. A. Abbott, C. Giarmatzi, F. Costa, and C. Branciard, *Phys. Rev. A* **94**, 032131 (2016).  
[4] N. Gisin, N. Linden, S. Massar, and S. Popescu, *Phys. Rev. A* **72**, 012338 (2005).  
[5] G. Chiribella and H. Kristjánsson, *Proc. R. Soc. A* **475**, 20180903 (2019).  
[6] D. Ebler, S. Salek, and G. Chiribella, *Phys. Rev. Lett.* **120**, 120502 (2018).  
[7] S. Salek, D. Ebler, and G. Chiribella, [arXiv:1809.06655](https://arxiv.org/abs/1809.06655).  
[8] G. Chiribella, M. Banik, S. S. Bhattacharya, T. Guha, M. Alimuddin, A. Roy, S. Saha, S. Agarwal, and G. Kar, [arXiv:1810.10457](https://arxiv.org/abs/1810.10457).  
[9] L. M. Procopio, A. Moqanaki, M. Araújo, F. Costa, I. Alonso Calafell, E. G. Dowd, D. R. Hamel, L. A. Rozema, Č. Brukner, and P. Walther, *Nat. Commun.* **6**, 7913 (2015).  
[10] G. Rubino, L. A. Rozema, A. Feix, M. Araújo, J. M. Zeuner, L. M. Procopio, Č. Brukner, and P. Walther, *Sci. Adv.* **3**, e1602589 (2017).  
[11] K. Goswami, C. Giarmatzi, M. Kewming, F. Costa, C. Branciard, J. Romero, and A. G. White, *Phys. Rev. Lett.* **121**, 090503 (2018).  
[12] K. Wei, N. Tischler, S.-R. Zhao, Y.-H. Li, J. M. Arrazola, Y. Liu, W. Zhang, H. Li, L. You, Z. Wang, Y.-A. Chen, B. C. Sanders, Q. Zhang, G. J. Pryde, F. Xu, and J.-W. Pan, *Phys. Rev. Lett.* **122**, 120504 (2019).  
[13] Y. Guo, X.-M. Hu, Z.-B. Hou, H. Cao, J.-M. Cui, B.-H. Liu, Y.-F. Huang, C.-F. Li, G.-C. Guo, and G. Chiribella, *Phys. Rev. Lett.* **124**, 030502 (2020).  
[14] A. S. Holevo, *J. Problems Inform. Transmission* **9**, 177 (1973).  
[15] A. A. Abbott, J. Wechs, D. Horsman, M. Mhalla, and C. Branciard, [arXiv:1810.09826](https://arxiv.org/abs/1810.09826).  
[16] P. A. Guérin, G. Rubino, and C. Brukner, *Phys. Rev. A* **99**, 062317 (2019).  
[17] H. Kristjánsson, G. Chiribella, S. Salek, D. Ebler, and M. Wilson, [arXiv:1910.08197](https://arxiv.org/abs/1910.08197).  
[18] W. A. de Brito and R. V. Ramos, *J. Mod. Opt.* **55**, 1231 (2008).  
[19] G. G. Stokes, *Trans. Cambridge Philos. Soc.* **9**, 399 (1852).  
[20] J. Leach, J. Courtial, K. Skeldon, S. M. Barnett, S. Franke-Arnold, and M. J. Padgett, *Phys. Rev. Lett.* **92**, 013601 (2004).  
[21] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2010).