Classical communication with indefinite causal order for N completely depolarizing channels

Sk Sazim,[1,*] Michal Sedlak,[1,†] Kratveer Singh,[2] and Arun Kumar Pati[3]

[1] *RCQI, Institute of Physics, Slovak Academy of Sciences, 845 11 Bratislava, Slovakia*
[2] *Indian Institute of Science Education and Research, Bypass Road, Bhauri, Bhopal 462066, India*
[3] *QIC Group, Harish-Chandra Research Institute, Homi Bhabha National Institute, Allahabad 211019, India*

(Received 28 April 2021; accepted 2 June 2021; published 16 June 2021)

If two identical copies of a completely depolarizing channel are put into a superposition of their possible causal orders, they can transmit nonzero classical information. Here we study how well we can transmit classical information with N depolarizing channels put in superposition of M causal orders via a quantum SWITCH. We calculate the Holevo quantity if the superposition uses only cyclic permutations of channels and find that it increases with M and it is independent of N. For a qubit it never reaches 1 if we are increasing M. On the other hand, the classical capacity decreases with the dimension d of the message system. Further, for $N = 3$ and 4 we studied the superposition of all causal orders and uniformly superposed causal orders belonging to different cosets created by a cyclic permutation subgroup.

DOI: 10.1103/PhysRevA.103.062610

I. INTRODUCTION

In classical information theory, it is assumed that the information carriers are deterministic, and the transmission lines are used in a definite configuration in space as well as in fixed time [1,2]. However, physical systems obey principles of quantum theory and they offer resources which are not available in its classical counterpart. These unique resources can be harnessed to achieve communication protocols which are impossible in classical information theory [3–5]. These findings led to a complete revolution in quantum information theory [6,7]. Still, quantum information theory assumes that the channels maintain a specific order in space and time. However, quantum theory allows the configurations where channels themselves are in superposition [8,9]. Moreover, recently, it was realized that the superposition can exist also in the order of channels in time, in a scenario known as indefinite causal order or a quantum SWITCH [10–14]. In a quantum SWITCH, the relative order of the two channels is indefinite and gives rise to quantum advantages in reducing communication complexity [15,16], improving channel discrimination [17,18], and quantum computing [19]. Moreover, several proposals for an experimental realization of a quantum SWITCH have actually been built and tested [20,21], suggesting that the notion is not just a theoretical possibility.

Recently, in Ref. [22] Ebler *et al.* showed that one may achieve nonzero classical communication rates using two completely depolarizing channels (CDCs) inserted into the quantum SWITCH, which was also experimentally demonstrated later in [23]. Likewise, it was reported that, using two completely entanglement breaking channels in a SWITCH, one may achieve perfect quantum communication [24–26]. After these findings, several applications of quantum SWITCH have been discovered in quantum metrology, quantum thermometry, and quantum information [27–34].

The extension of such settings beyond the superposition of two channels is an immediate and interesting generalization to make to see whether it provides a bigger communication advantage. However, such generalization comes with a serious concern, whether it is not beyond the experimental scope. In Ref. [35] Procopio *et al.* showed that there is an almost twofold increase in communication rate if a causal superposition of three channels is used instead of a two-channel causal superposition. On the other hand, the number of relevant configurations jumps from 2 to $3! = 6$. This makes experimental implementation very cumbersome but, nevertheless, possible. Furthermore, their numerical results suggests that usage of three channels in three cyclic causal orders gives gain similar to that in all 3! causal orders. This bolsters the idea that considering N causal orders for N channels should be efficient. An extension to N channels with $N!$ causal orders was proffered in Ref. [36]; however, Procopio *et al.* used a numerical approach to find the communication rates which might suffer from numerical errors. An analytical approach is in demand to delve deeper into these matters and to answer the following open questions: Can N channels in a quantum SWITCH allow perfect transmission of classical information? Can we achieve substantial gain in classical communication rates with an optimal number of causal orders in a quantum SWITCH? We answer these questions in detail in this paper.

Cyclic permutations of N elements form a subgroup of all N element permutations. In this paper we find that all cosets of a permutation group factorized with respect to cyclic permutations behave equivalently when they determine the casual order superposition of CDCs used in a quantum SWITCH. Therefore, we consider single-coset or multiple-coset causal superposition. We refer to causal orders of channels from a single coset as cyclic causal orders and we term the

*sk.sazimsq49@gmail.com
†michal.sedlak@savba.sk

superposition of causal orders of channels from more than one coset as noncyclic. We analytically find the classical communication capacity for N CDCs inserted into a quantum SWITCH, which superposes $M \in [2, N]$ cyclic causal orders. Similarly, we derive some results for $M \in [2, N!]$ noncyclic causal orders, when $N = 3$ and 4. We find that for the cyclic case the classical communication rate depends on $M \leqslant N$, the number of superposed cyclic permutations, but does not depend on N, the number of CDCs in the quantum SWITCH. If we keep on increasing M (and necessarily increasing also N), we observed that the communication rate increases rapidly with the increase in the number of causal orders, but it never reaches noiseless transmission. For example, it saturates at 0.311 bits for qubit systems. For the noncyclic case, the increase of the classical communication rate is not directly linked to M. On the other hand, we find that the classical communication rate decreases almost exponentially with the dimension of the message state, which seems a bit counterintuitive at first.

The rest of the paper is organized as follows. In the next section we briefly introduce the quantum SWITCH formalism. In Sec. III we present results for N completely depolarizing channels in a quantum SWITCH while superposing only cyclic causal orders. Section IV contains a detailed analysis of N CDCs with arbitrary noncyclic causal orders in a quantum SWITCH with special emphasis on $N = 3$ and 4. We summarize in Sec. V.

II. QUANTUM SWITCH AND QUANTUM CHANNELS IN THE SUPERPOSITION OF DIFFERENT CAUSAL ORDERS

Quantum communication devices can be modeled as quantum channels, i.e., a completely positive and trace preserving linear map $\Lambda : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H})$. Any such map admits Kraus decomposition, i.e., $\Lambda(\rho) = \sum_i K_i \rho K_i^\dagger$, where $\{K_i\}$ is a set of Kraus operators with $\sum_i K_i^\dagger K_i = \mathbb{I}$ and $\rho \in \mathcal{L}(\mathcal{H})$.

In this work we are considering a scenario where N channels are put into a coherent superposition of their differently ordered concatenations. Originally, a quantum SWITCH was used to construct a superposition of $N = 2$ causal orders [19]. In this case, the quantum SWITCH is a higher-order map which takes two channels as input and then outputs the superposition of their orders based on the state of the control qubit (see Fig. 1). Mathematically, the quantum SWITCH transforms two input channels Λ_1 and Λ_2, with Kraus decompositions $\{K_i^{(1)}\}$ and $\{K_i^{(2)}\}$, respectively, into the overall channel

$$S(\Lambda_1, \Lambda_2)(\cdot) = \sum_{ij} W_{ij}(\cdot)W_{ij}^\dagger,$$

whose Kraus operators W_{ij} are defined as

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

Note that though W_{ij} depends on the specific Kraus decomposition of channels Λ_1 and Λ_2, the effective quantum channel $S(\Lambda_1, \Lambda_2)$ depends only on the input channels, allowing the SWITCH to be a valid higher-order map [19]. We can extend the SWITCH formalism for more than two inputs, i.e., for $N > 2$ [36,37]. In this case, the extended SWITCH is a higher-order map which takes N channels as input and then outputs the superposition of orders based on the state of the control system that must have sufficiently high dimensionality. Then for N channels $\{\Lambda_p\}$ with Kraus representations $\{K_j^{(p)}\}$, the extended SWITCH will output an effective channel of the form

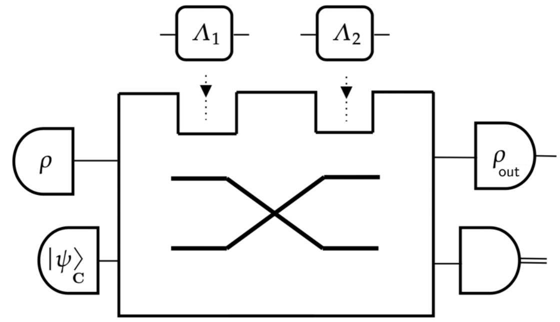

FIG. 1. Illustration of a quantum SWITCH. Based on the state of the control qubit $|\psi\rangle_c$, a SWITCH takes two channels Λ_1 and Λ_2 as input and outputs (i) either $\Lambda_1 \circ \Lambda_2$ or $\Lambda_2 \circ \Lambda_1$ if the control qubit is in $|0\rangle$ or $|1\rangle$, respectively, and (ii) a superposition of causal orders if $|\psi\rangle_c = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. Here ρ is the message quantum state and ρ_{out} is the final output after postselecting the control qubit.

$$S(\Lambda_1, \Lambda_2, \dots, \Lambda_N)(\cdot) = \sum_{ij\cdots\eta} W_{ij\cdots\eta}(\cdot)W_{ij\cdots\eta}^\dagger, \qquad (1)$$

whose Kraus operators $W_{ij\cdots\eta}$ are defined as

$$W_{ij\cdots\eta} = \sum_{\ell=0}^{M-1} |\ell\rangle\langle\ell| \otimes \mathcal{P}_\ell(K_i^{(1)}, K_j^{(2)}, \dots, K_\eta^{(N)}), \qquad (2)$$

where $M \in [2, N!]$ and $\mathcal{P}_\ell \in \mathbf{S}_N$ represents concatenation of N operators reordered according to the permutation j, e.g., $\mathcal{P}_0(K_i^{(1)}, K_j^{(2)}, \dots, K_\eta^{(N)}) = K_i^{(1)}K_j^{(2)} \cdots K_\eta^{(N)}$. For brevity, we will drop the upper index (p) in the rest of the paper.

To have a simple but sufficient picture suitable for further consideration, let us consider two unitary channels U_1 and U_2 and the control qubit in the state $|\psi\rangle_c = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. For the pure state $|\Psi\rangle$ of the message system (we often call it a target state as well), the output of the SWITCH will be a pure state $S(U_1, U_2)(|\psi\rangle_c \langle\psi| \otimes |\Psi\rangle \langle\Psi|) = |\xi\rangle \langle\xi|$, where $|\xi\rangle = \frac{1}{\sqrt{2}}(|0\rangle \otimes U_1U_2 |\Psi\rangle + |1\rangle \otimes U_2U_1 |\Psi\rangle)$. To see the interference phenomenon, one needs to measure the control qubit in the Fourier basis, i.e., $\{|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)\}$; then the resulting target state will take the form $|\Psi_\pm^f\rangle = \frac{1}{\sqrt{2}}(U_1U_2 \pm U_2U_1) |\Psi\rangle$.

Extending this idea to N channels is possible [36], and the number of possible causal orders increases to $M \in [2, N!]$. For the brevity of the explanation, let us consider N unitary channels $\{U_j\}$ and the control state $\frac{1}{\sqrt{M}} \sum_{j=0}^{M-1} |j\rangle$. Then the final target state after the measurement of the control system will be $|\Psi_k^f\rangle = \frac{1}{\sqrt{M}}(\sum_{j=0}^{M-1} \langle e_k |j\rangle \mathcal{P}_j(U_1, U_2, \dots, U_N) |\Psi\rangle$, where $\{|e_k\rangle \mid |e_k\rangle = \frac{1}{\sqrt{M}} \sum_{j=0}^{M-1} e^{2i\pi jk} |j\rangle\}$ is the Fourier basis of $\{|j\rangle\}$.

III. USING CYCLIC ORDERS OF N COMPLETELY DEPOLARIZING CHANNELS IN QUANTUM SWITCH

A completely depolarizing channel in d dimensions can be described by

$$\Lambda(X) = \frac{1}{d^2}\sum_{i=1}^{d^2} U_i X U_i^{\dagger} = \frac{1}{d}\mathrm{Tr}[X]\mathbb{I}_d, \tag{3}$$

where $\{U_i \mid i = 1, 2, \dots, d^2\}$ are $d \times d$ unitary operators satisfying $\mathrm{Tr}[U_i^{\dagger}U_j] = d\delta_{ij}$, $\mathbb{I}_d$ is the identity operator of order d, and X is an arbitrary linear operator in d-dimensional Hilbert space. Direct transmission of information through single or several concatenated CDCs necessarily results in a zero classical communication rate. In contrast, it was shown that, given two identical CDCs labeled as Λ_1 and Λ_2 and a control qubit state $|\psi\rangle_c = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, there is a possibility of nonzero classical communication using a quantum SWITCH [22].

Here we generalize the scheme represented in [22] to N CDCs $\{\Lambda_i\}$. We consider first only $M \in [2, N]$ possible cyclic orders. Accordingly, the state of the control qubit is $|\tilde{\psi}\rangle_c = \frac{1}{\sqrt{M}}\sum_{j=0}^{M-1}|j\rangle$. Therefore, the Kraus operators of a channel resulting from N CDCs in an extended quantum SWITCH can be written as [see Eq. (2)]

$$K_{ij\cdots\eta} = \frac{1}{d^N}\sum_{\ell=0}^{M-1} |\ell\rangle\langle\ell| \otimes \mathcal{P}_{\ell}^{(c)}(U_i, U_j, \dots, U_\eta), \tag{4}$$

where $\mathcal{P}_{\ell}^{(c)}(U_i, U_j, \dots, U_\eta)$ defines the cyclic permutations of unitaries. For example, for $N = 3$ the cyclic permutations are $\mathcal{P}_0^{(c)}(U_i, U_j, U_k) = U_iU_jU_k$, $\mathcal{P}_1^{(c)}(U_i, U_j, U_k) = U_jU_kU_i$, and $\mathcal{P}_2^{(c)}(U_i, U_j, U_k) = U_kU_iU_j$. If the sender prepared the target system in the state ρ, then the receiver will receive the output from the quantum SWITCH as

$$\begin{aligned}\rho_M :&= S(\Lambda_1, \Lambda_2, \dots, \Lambda_N)(\rho_c \otimes \rho)\\ &= \frac{1}{Md^{2N}}\sum_{i,j,\dots,\eta}\sum_{l=0}^{M-1}\sum_{l'=0}^{M-1}|l\rangle\langle l'|\\ &\quad \otimes P_l^{(c)}(U_i, U_j, \dots, U_\eta)\rho(P_{l'}^{(c)}(U_i, U_j, \dots, U_\eta))^{\dagger},\end{aligned} \tag{5}$$

where $\rho_c = |\tilde{\psi}\rangle\langle\tilde{\psi}|_c$. For cyclic permutations of orders we will see below that only two types of contributions are present in the final output state: (i) diagonals ($l = l'$), which are proportional to $\mathbb{I}$, and (ii) off-diagonals ($l \neq l'$), which are proportional to ρ. All the diagonal terms are equivalent to the prototypal form

$$\begin{aligned}&\frac{1}{d^{2N}}\sum_{ij\cdots\eta=1}^{d^2}\overbrace{U_\eta\cdots U_j}^{N-1}U_i(\rho)U_i^{\dagger}\overbrace{U_j^{\dagger}\cdots U_\eta^{\dagger}}^{N-1}\\ &= \frac{1}{d^{2N-2}}\mathrm{Tr}(\rho)\sum_{j\cdots\eta=1}^{d^2}\overbrace{U_\eta\cdots U_j}^{N-1}\frac{\mathbb{I}}{d}\overbrace{U_j^{\dagger}\cdots U_\eta^{\dagger}}^{N-1}\\ &= \frac{1}{d^{2N-2}}d^{2(N-1)}\frac{\mathbb{I}}{d} = \frac{\mathbb{I}}{d}\end{aligned} \tag{6}$$

and all off-diagonal terms are equivalent to the form

$$\begin{aligned}&\frac{1}{d^{2N}}\sum_{ij\cdots\eta=1}^{d^2}\overbrace{U_\eta\cdots U_\mu}^{k}U_\ell\cdots U_jU_i\left(\rho\overbrace{U_\mu^{\dagger}\cdots U_\eta^{\dagger}}^{k}\right)U_i^{\dagger}U_j^{\dagger}\cdots U_\ell^{\dagger}\\ &= \frac{1}{d^{2N-2}}\sum_{j\cdots\eta=1}^{d^2}\mathrm{Tr}\left(\rho\overbrace{U_\mu^{\dagger}\cdots U_\eta^{\dagger}}^{k}\right)\overbrace{U_\eta\cdots U_\mu}^{k}U_\ell\\ &\times\cdots U_j\frac{\mathbb{I}}{d}U_j^{\dagger}\cdots U_\ell^{\dagger}\\ &= \frac{d^{2(N-k-1)}}{d^{2N-2}}\sum_{\mu\cdots\eta=1}^{d^2}\mathrm{Tr}\left[\left(\rho\overbrace{U_\mu^{\dagger}\cdots}^{k-1}\right)\frac{1}{\sqrt{d}}U_\eta^{\dagger}\right]\frac{1}{\sqrt{d}}U_\eta\left(\overbrace{\cdots U_\mu}^{k-1}\right)\\ &= \frac{1}{d^{2k}}d^{2(k-1)}\rho = \frac{\rho}{d^2},\end{aligned} \tag{7}$$

where $k \in \{1, \dots, N-1\}$ and we used $\sum_{i=1}^{d^2}\frac{1}{d}\mathrm{Tr}(XU_i^{\dagger})U_i = X$ and Eq. (3) multiple times. Note that the index k represents all possible cyclic k shifts. Therefore, for N CDCs in a SWITCH the final output state for $M \in [2, N]$ causal orders is

$$\rho_M = \frac{1}{M}\left(\mathbb{I} \otimes \frac{\mathbb{I}}{d} + \sum_{i\neq j}|i\rangle\langle j| \otimes \frac{\rho}{d^2}\right). \tag{8}$$

We note that the output density matrix does not depend on N and only its off-diagonal entries depend on ρ. Further, if the control system is measured in the Fourier basis and the outcome is known, the target system will regain a dependence on ρ and one can extract information about it [22]. Therefore, there is a possibility of nonzero classical communication according to the Holevo-Schumacher-Westmoreland theorem [38,39]. Let us note here that from Eq. (8) we see that distinguishability of any signal states will be strongly suppressed with increasing dimension d of the target system as the off-diagonal terms decay quadratically with d, making the signal just a small admixture (perturbation) to a white noise. As we will see later, this results in a steep decrease of classical communication capacity and for the relevant range of dimensions it qualitatively behaves as an exponential function. Thus, we refer to this decrease as an almost-exponential decrease. In what follows, we will quantitatively investigate the above scheme.

A. Method

The classical capacity of the quantum communication channel Λ is characterized by the Holevo quantity, which is defined as

$$\chi(\Lambda) = \max_{\{p_i\rho_i\}}\left[H(\Lambda(\rho)) - \sum_i p_iH(\Lambda(\rho_i))\right]. \tag{9}$$

In Ref. [22] Ebler *et al.* determined how to evaluate the Holevo quantity for two quantum channels when the information is sent through a pair of quantum channels processed by a quantum SWITCH. Since this method works with the channel induced by the SWITCH on the control plus the target system after the insertion of the depolarizing channels, it automatically works also for N channels in the generalized SWITCH.

Therefore, the Holevo quantity of N CDCs in the SWITCH is given by (we refer readers to the Supplemental Material of Ref. [22])

$$\chi^{(M)} = \log d + H(\tilde{\rho}_c(M)) - H_{\min}(\rho_M)\,\forall\, M \in [2, N], \quad (10)$$

where $\tilde{\rho}_c(M)$ is the reduced state of the control qubit after evolution [see Eq. (8)], i.e., $\tilde{\rho}_c(M) = \frac{1}{M}(\sum_i |i\rangle\langle i| + \frac{1}{d^2}\sum_{i\neq j}|i\rangle\langle j|)$, and $H_{\min}$ is the minimum output entropy of the effective channel, i.e., $H_{\min}(\rho_M) = \min_\rho H(\rho_M)$, with $H(\cdot)$ the von Neumann entropy. Note that we use logarithm base 2 in the whole paper. However, the main difficulty will be to calculate the eigenvalues of the $Md \times Md$ matrix, ρ_M, to evaluate $H_{\min}(\rho_M)$. First, we use the method given in Refs. [40,41] to evaluate the determinant of the matrix ρ_M,

$$\begin{aligned}\mathrm{Det}(\rho_M) =& \mathrm{Det}\left(\frac{1}{M}\left[\frac{\mathbb{I}}{d} - \frac{\rho}{d^2}\right]\right)^{\times(M-1)} \\ &\times \mathrm{Det}\left(\frac{1}{M}\left[\frac{\mathbb{I}}{d} + (M-1)\frac{\rho}{d^2}\right]\right), \end{aligned} \quad (11)$$

where $\mathrm{Det}(\cdot)^{\times(M-1)}$ indicates that there are $M-1$ products of the same determinant. Full details of the calculations for Eq. (11) and the following claim are given in Appendix A. The beautiful simplified form of the determinant of an actual $Md \times Md$ matrix in Eq. (11) tells us that finding the eigenvalues of an actual matrix is reduced to finding the eigenvalues of these small matrices, i.e., $\frac{1}{M}[\frac{\mathbb{I}}{d} - \frac{\rho}{d^2}]$ with degeneracy $M-1$ and $\frac{1}{M}[\frac{\mathbb{I}}{d} + (M-1)\frac{\rho}{d^2}]$ with degeneracy 1. As $[\mathbb{I}, \rho] = 0$, the eigenvalues of the matrix ρ_M will be the union of the eigenvalues of these two smaller matrices with their appropriate degeneracy. If we let $\{\lambda_i^+\}_{i=1}^d$ and $\{\lambda_i^-\}_{i=1}^d$ be the eigenvalues of $\frac{1}{M}[\frac{\mathbb{I}}{d} + (M-1)\frac{\rho}{d^2}]$ and $\frac{1}{M}[\frac{\mathbb{I}}{d} - \frac{\rho}{d^2}]$, respectively, then

$$\lambda_i^+ = \frac{1}{Md} + \frac{M-1}{Md^2}\lambda_i^\rho, \quad \lambda_i^- = \frac{1}{Md} - \frac{1}{Md^2}\lambda_i^\rho,$$

where $\{\lambda_i^\rho\}_{i=1}^d$ are eigenvalues of ρ. As $H_{\min}(\rho_M) = \min_\rho H(\rho_M)$, certainly the minima will be ascertained if $\lambda_i^\rho = 1$ and $\lambda_j^\rho = 0$ with $i \neq j$. Therefore, with the constraint that $\sum_i \lambda_i^\rho = 1$, we can find that

$$\begin{aligned}H_{\min}(\rho_M) = -\Bigg\{&\frac{d+(M-1)}{Md^2}\log\frac{d+(M-1)}{Md^2} \\ &+ \frac{(M-1)(d-1)}{Md^2}\log\frac{(d-1)}{Md^2} + \frac{d-1}{d}\log\frac{1}{Md}\Bigg\}.\end{aligned} \quad (12)$$

Now the remaining task is to find the expression for $H(\tilde{\rho}_c(M))$, which is given by

$$\begin{aligned}H(\tilde{\rho}_c(M)) = -\Bigg(&\frac{M-1+d^2}{Md^2}\log\frac{M-1+d}{Md^2} \\ &+ (M-1)\frac{d^2-1}{Md^2}\log\frac{d^2-1}{Md^2}\Bigg).\end{aligned} \quad (13)$$

With these expressions, we can evaluate the classical communication rate $\chi^{(M)}$ for N CDCs with the SWITCH from Eq. (10) for cyclic causal orders $M \in [2, N]$. Note that for $M = 2$ it reduces to the result for the $N = 2$ scenario as discussed in Ref. [22]. This observation tells us that the gain in classical communication depends only on the number of superposed causal orders M.

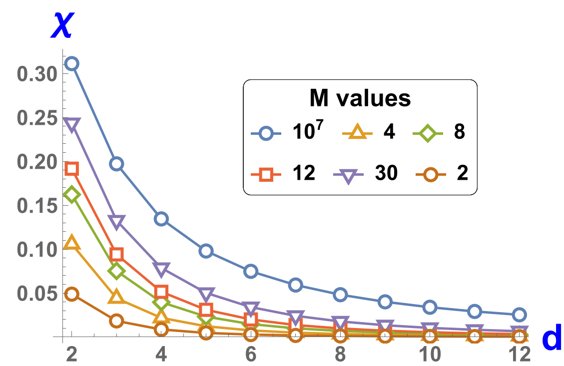

FIG. 2. Plot illustrating that the Holevo quantity for different number of causal orders M is almost exponentially decreasing with the dimension d of the target state ρ.

B. Results

Figure 2 illustrates the behavior of the Holevo quantity $\chi^{(M)}$ with respect to the dimension d of the input state ρ and the number of causal orders. We find that the classical communication capacity increases as we increase M; however, it decreases almost exponentially as d increases. Figure 3 shows the communication rates for different choices of (M, d). It is clear from the contour plot that the higher values of communication rates are achieved with smaller d values as well as higher M values. This means that using a quantum SWITCH with M causal orders in $d = 2$ will offer the maximum classical communication rate. However, we find that the communication rate saturates with the increase of M, indicating that it is not possible to reach perfect communication in the asymptotic limit, i.e., $\chi^{(\infty)} \neq 1$ (see Fig. 4). To prove this claim, we write down the expression for the Holevo

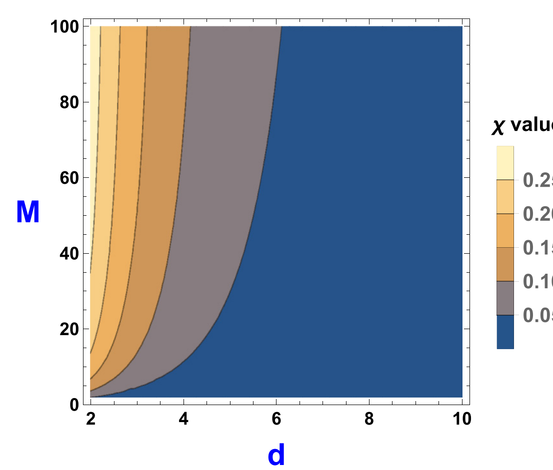

FIG. 3. Contour plot depicting the Holevo quantity for different number of causal orders M and the dimension d of the target state ρ. It can be seen that higher classical communication rates are achieved for lower d and higher M.

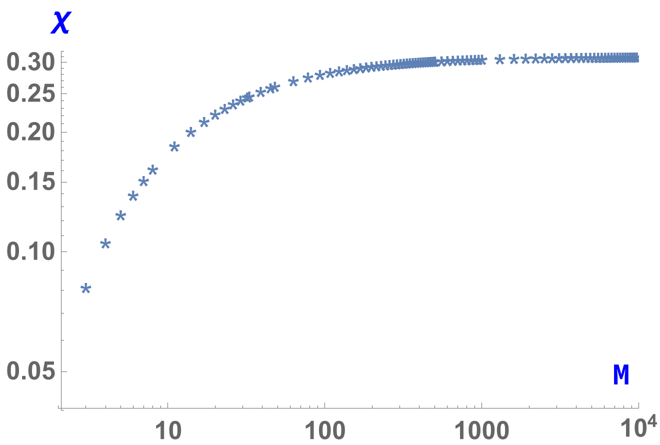

FIG. 4. Logarithmic plot depicting that the Holevo quantity is increasing for different number of causal orders M for $d = 2$. The communication rate is converging to the value of 0.311 bits per single system transmission.

quantity for $d = 2$, i.e.,

$$\chi^{(M)}|_{d=2} = 1 + \frac{1}{4}\left[\log_2 \frac{4}{27} + \left(1 + \frac{1}{M}\right)\log_2\left(1 + \frac{1}{M}\right) + \frac{1}{M}\log_2 \frac{27}{M^2} - \left(1 + \frac{3}{M}\right)\log_2\left(1 + \frac{3}{M}\right)\right].$$

Therefore, at $M \to \infty$, the Holevo quantity $\chi^{(\infty)}|_{d=2} \cong 1 + \frac{1}{4}\log_2 \frac{4}{27} \approx 0.311$ bits per transmission.

IV. GENERALIZATION TO VARIOUS COMBINATIONS OF CYCLIC AND NONCYCLIC CAUSAL ORDERS

There are $N!$ possible permutations of N elements; thus there exist $N!$ causal orders of N channels. However, above we considered only $M \leqslant N$ cyclic causal orders of channels, because the increase in communication rate is sufficiently substantial to notice and it becomes problematic to implement a superposition of $N!$ causal orders in an experiment.

In this section we discuss whether the extension from N cyclic causal orders to all $N!$ orders for N CDCs provides a significant improvement in the communication rates. In a recent work [36] Procopio *et al.* showed that communication rates do not increase evenly with the increase of the number of causal orders. We think that this behavior can be attributed to inevitable mixing of cyclic and noncyclic causal orders, which hinders the potential benefit due to the appearance of input-state-independent terms (see Appendix B). This is why we considered the cyclic orders separately.

By considering any order of action of N channels as the zeroth element and then performing cyclic permutations of this element, we can form a coset of cyclic causal orders (coset of a permutation group with respect to the subgroup of cyclic permutations of N elements) which contains N elements. In this way, we can identify $\frac{N!}{N} = (N-1)!$ such cosets in the set of all $N!$ permutations. We refer to these individual cosets as cyclic causal orders, and each of them yields the same classical communication rate, which we already evaluated in the preceding section.

TABLE I. Off-diagonal blocks in the output density matrix for N channels in the SWITCH when we are considering all possible causal orders, i.e., $M \in [2, N!]$. Note that off-diagonal terms within a single coset are always $\frac{\rho}{d^2}$ (see Appendix B).

N	Terms in off-diagonal block
2	$\frac{\rho}{d^2}$
3	$\frac{\rho}{d^2}, \frac{\mathbb{I}}{d^3}$
4	$\frac{\rho}{d^2}, \frac{\mathbb{I}}{d^3}, \frac{\rho}{d^4}$
$\vdots$	$\vdots$
$2k$	$\frac{\rho}{d^2}, \frac{\mathbb{I}}{d^3}, \frac{\rho}{d^4}, \dots, \frac{\mathbb{I}}{d^{N-1}}, \frac{\rho}{d^N}$
$2k+1$	$\frac{\rho}{d^2}, \frac{\mathbb{I}}{d^3}, \frac{\rho}{d^4}, \dots, \frac{\rho}{d^{N-1}}, \frac{\mathbb{I}}{d^N}$

However, if we consider the superposition of orders of elements from different cosets the situation becomes a bit demanding. Off-diagonal contributions to ρ_M that map between control system states belonging to a single coset act on the message system as $\frac{\rho}{d^2}$ [see Eq. (7)]; however, this is no longer the case if the control system states are not from the same coset.

Let us note that as the number of channels increases the number of types of off-diagonal contributions increases and will include the terms presented in Table I. For a derivation of the terms appearing in the table, we refer the reader to Appendix B. In accordance with Table I, for $N \geqslant 3$ we found that many cross-coset off-diagonal terms are proportional to $\mathbb{I}$ and the terms which are proportional to ρ have a decreasing weight factor as N increases. To illustrate the impact of the above findings, we present a case study for $N = 3$ and for $N = 4$.

A. Case study for $N = 3$

For three CDCs, there will be two cosets of cyclic causal orders. The Kraus operator for three CDCs with a SWITCH can be expressed as $d^3 K_{ijk} = \sum_{\ell=0}^{M-1} |\ell\rangle\langle\ell| \otimes \mathcal{P}_\ell(U_i, U_j, U_k)$, where $2 \leqslant M \leqslant 6$. For the message qubit prepared in ρ, the evolved state at the output of the quantum SWITCH is $\rho_{M_1,M_2} = S(\Lambda_1, \Lambda_2, \Lambda_3)(\rho_c \otimes \rho)$, where M_1 and M_2 denote the number of causal orders from two cosets, respectively, and $M_1 + M_2 = M$. In Appendix B 1 we find that

$$\rho_{M_1,M_2} = \frac{1}{M}\left\{\mathbb{I}_c \otimes \frac{\mathbb{I}}{d} + L_{M_1,M_2} \otimes \frac{\rho}{d^2} + B_{M_1,M_2} \otimes \frac{\mathbb{I}}{d^3}\right\}, \tag{14}$$

where $2 \leqslant M \leqslant 6$. The matrices L and B are $M \times M$ matrices of the form

$$\mathbf{L} = \begin{pmatrix} S_{M_1\times M_1} & \mathbf{0}_{M_1\times M_2} \\ \mathbf{0}_{M_2\times M_1} & S_{M_2\times M_2} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} \mathbf{0}_{M_1\times M_1} & \mathbf{1}_{M_1\times M_2} \\ \mathbf{1}_{M_2\times M_1} & \mathbf{0}_{M_2\times M_2} \end{pmatrix},$$

where the matrix $S = \sum_{i\neq j=0}^{m-1} |i\rangle\langle j|$, $\mathbf{0}$ is a null matrix, and $\mathbf{1}$ is a matrix with all entries equal to 1. For such a scenario, the general expression for the classical communication for three CDCs with $M \in [2, 6]$ causal orders can efficiently be written as before

$$\chi^{(M_1,M_2)} = \log d + H(\tilde{\rho}_c(M)) - H_{\min}(\rho_{M_1,M_2}), \tag{15}$$

TABLE II. Communication rates $\chi^{(M_1,M_2)}$ for three CDCs in a superposition of $M = M_1 + M_2$ causal orders from two cosets. Here we consider the message state to be a qubit ($d = 2$). The Holevo quantity is higher for the cases when either M_1 or M_2 is zero.

M_1 \ M_2	0	1	2	3
0		0	0.0488	0.0817
1	0	0	0.0334	0.0640
2	0.0488	0.0334	0.0524	0.0767
3	0.0817	0.0640	0.0767	0.0981

where $\tilde{\rho}_c(M)$ is the reduced state of the control system after evolution, i.e., $\tilde{\rho}_c(M) = \frac{1}{M}(\mathbb{I} + \frac{1}{d^2}\sum_{i\neq j}|i\rangle\langle j|)$.

To evaluate $H(\tilde{\rho}_c(M))$ and $H_{\min}(\rho_{M_1,M_2})$, we need to diagonalize the matrices $\tilde{\rho}_c(M)$ and ρ_{M_1,M_2}, respectively. We know that the expression for $H(\tilde{\rho}_c(M))$ is given in Eq. (13). However, diagonalizing ρ_{M_1,M_2} is much more complicated. Analytical diagonalization is done in Appendix B 1 using the method in Appendix C. It follows that for $M_1 = M_2$, $[L, B] = 0$ and therefore the matrices L and B are simultaneously diagonalizable. Consequently, $H_{\min}(\rho_{M_1,M_1})$ is given by

$$\begin{aligned}-H_{\min}(\rho_{M_1,M_1}) &= \lambda_+\log\lambda_+ + \lambda_-\log\lambda_- \\ &\quad + (d-1)\bigg[\lambda_+^0\log\lambda_+^0 + \lambda_-^0\log\lambda_-^0 \\ &\quad + (M-2)\bigg\{\frac{1}{Md^2}\log\frac{(d-1)}{Md^2} + \frac{1}{Md}\log\frac{1}{Md}\bigg\}\bigg],\end{aligned}$$

where $\lambda_\pm = \frac{d^2+(M_1-1)d\pm M_1}{Md^3}$ and $\lambda_\pm^0 = \frac{d^2\pm M_1}{Md^3}$ with $M = 2M_1$. We show in Appendix B 1 that the output states for $M_1 \neq M_2$ can also be diagonalized using the method presented in Appendix C.

To further elucidate our findings here, we compute the Holevo quantity for all (M_1, M_2) values in Table II for $d = 2$. We find that the Holevo quantity is higher for the case when either $M_1 = 0$ or $M_2 = 0$ compared to $M_1 \neq M_2$. However, we find that the Holevo quantity reaches its maximum for $M_1 = M_2 = 3$.

We also plot the Holevo quantity for different (M_1, M_2) values with respect to the dimension of the target state d in Fig. 5 and find that the Holevo quantity decreases with d in an exponential-like fashion. In Fig. 6 we plot the Holevo quantity of $(N = 3, M = 6)$ (all cosets) as well as $(N = 3, M = 3)$ (cyclic orders) against the dimension d. The figure shows that both scenarios yield the same communication rates except for $d = 2$, where the former dominates. This plot indicates that it might be efficient to consider only cyclic permutations (one coset).

B. Case study for $N = 4$

For four channels there are six $[(4-1)! = 6]$ cosets of cyclic causal orders. Let us denote by M_η the number of causal orders in a coset where $M_\eta \in [0, 3]$ for each $\eta \in [1, 6]$. Let us consider that the target state is ρ. Now if we consider $M = (\sum_\eta M_\eta)$ causal orders in a quantum SWITCH, the output

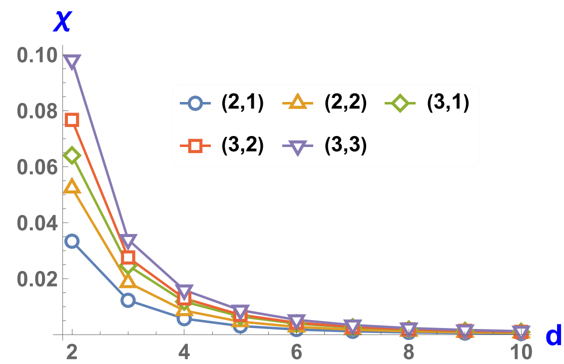

FIG. 5. Holevo quantity for $N = 3$ and different (M_1, M_2) values depending on the dimension d of the target state.

state will have the following: (i) Within each cyclic coset the diagonal terms are proportional to $\frac{\mathbb{I}}{d}$ and off-diagonal terms are proportional to $\frac{\rho}{d^2}$ and (ii) cross-coset off-diagonal terms are proportional to $\frac{\mathbb{I}}{d^3}$ as well as $\frac{\rho}{d^4}$ (for detailed calculation see Appendix B). We will use a particular way of listing causal orders for each coset. Namely, the first entry of each coset is linked with the order $\Lambda_1\Lambda_2\Lambda_3\Lambda_4$ by some permutation of the last three labels. In this way, the first element in each coset is starting with Λ_1 and the rest are constructed using cyclic permutations from the first element of the corresponding coset. For $\eta = 1, 2, \ldots, 6$, the zeroth elements are $\Lambda_1\Lambda_2\Lambda_3\Lambda_4$, $\Lambda_1\Lambda_2\Lambda_4\Lambda_3$, $\Lambda_1\Lambda_3\Lambda_2\Lambda_4$, $\Lambda_1\Lambda_3\Lambda_4\Lambda_2$, $\Lambda_1\Lambda_4\Lambda_2\Lambda_3$, and $\Lambda_1\Lambda_4\Lambda_3\Lambda_2$, respectively. Using this setup, we find that the output state after evolution driven by the SWITCH is given by

$$\rho_{M_\eta} = \frac{1}{M}\bigg\{\mathbb{I}_c\otimes\frac{\mathbb{I}}{d} + \bigg(\frac{L_{M_\eta}}{d^2} + \frac{Q_{M_\eta}}{d^4}\bigg)\otimes\rho + B_{M_\eta}\otimes\frac{\mathbb{I}}{d^3}\bigg\}, \tag{16}$$

where the matrices L, B, and Q are specified in Appendix B. Also, in this case, the general expression for the classical communication for four CDCs with $2 \leqslant M \leqslant 24$ causal orders can efficiently be written as

$$\chi^{(M_\eta)} = \log d + H(\tilde{\rho}_c(M_\eta)) - H_{\min}(\rho_{M_\eta}), \tag{17}$$

where $\tilde{\rho}_c(M_\eta)$ is the reduced state of the control qubit after evolution, i.e., $\tilde{\rho}_c(M_\eta) = \frac{1}{M}(\mathbb{I} + \frac{1}{d^2}\{L_{M_\eta} + B_{M_\eta}\} + \frac{1}{d^4}Q_{M_\eta})$.

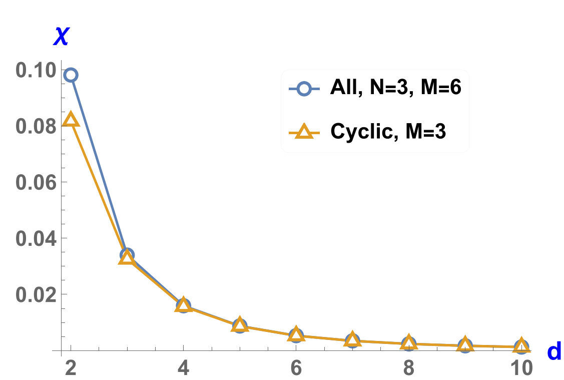

FIG. 6. Holevo quantity as a function of target state dimension d for $(N = 3, M = 6)$ and $(N = 3, M = 3)$.

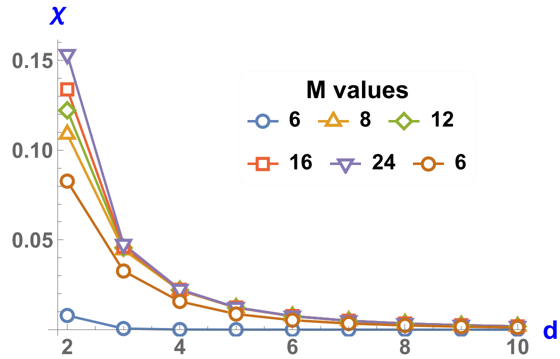

FIG. 7. Holevo quantity for $N = 4$ and different M_η values depending on the dimension d of the target state. See the text for a detailed description of the curves.

It is very hard to evaluate the Holevo quantity for arbitrary $\{M_\eta\}$, as diagonalizing ρ_{M_η} is usually hard analytically. Therefore, we mostly resort to a numerical approach. However, there are specific cases where it is possible to analytically diagonalize ρ_{M_η}, e.g., the scenario with $(M_3 = M_5 = 4; M = 8)$ (we refer readers to Appendix B 2 for a complete analysis).

To graphically illustrate our findings for $N = 4$ we present two plots. In Fig. 7 we plot dependence of the Holevo quantity on d for different M values, i.e., $M = \{6, 8, 12, 16, 24\}$. For $M = 6$, we consider two scenarios $(M_1 = 4, M_2 = 2)$ and $M_\eta = 1\,\forall\,\eta$ (red and blue lines with circles, respectively, in Fig. 7). In the former case, we consider a situation such that only two cross-coset off-diagonal terms are dependent on ρ. However, in the latter case, we consider the maximum number of ρ-dependent cross-coset off-diagonal terms (see Appendix B 2). Figure 7 shows that the Holevo quantity is decreasing in an exponential-like fashion with d. Unlike Fig. 5, here the Holevo quantities for all the plotted M_η are very close to each other except for $d = 2$. This is due to the presence of a big number of cross-coset off-diagonal terms $\frac{\rho}{d^4}$ in these scenarios. In Fig. 8 we plot the Holevo quantity for $(N = 4, M = 24)$ and its cyclic counterpart $(N = 4, M = 4)$ with respect to the message system dimension d. Analogously to Fig. 6, also here the plot shows that both scenarios provide the same communication rates except for $d = 2$. Comparing Figs. 6 and 8, we find that the gap between the Holevo quantities for $(N, M = N!)$ (all cosets) and $(N, M = N)$ (one coset) increases for $d = 2$ as we increase N from 3 to 4.

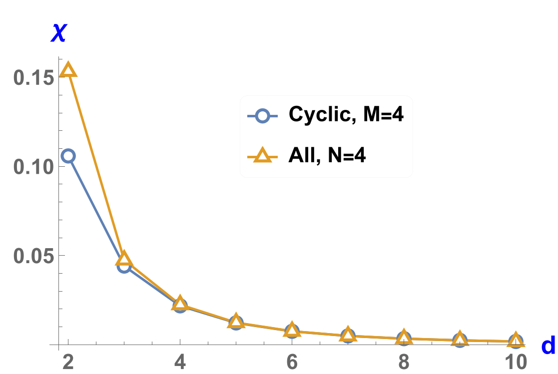

FIG. 8. Holevo quantity as a function of target state dimension d for $(N = 4, M = 24)$ and $(N = 4, M = 4)$.

V. CONCLUSION

A completely depolarizing channel erases all information about its input state and always prepares a completely mixed state. Thus, sequential application of two such channels on the same system in any fixed order must have a zero classical (or quantum) communication capacity. It was a rather surprising finding of Ebler *et al.* [22] that processing of two depolarizing channels by a quantum SWITCH followed by suitable control system measurement enables a nonzero classical communication rate.

In this paper we studied a generalization of this scenario to N completely depolarizing channels inserted into a (generalized) quantum SWITCH. A quickly growing number of possible causal orders ($N!$) might become a roadblock for experimental realization of the SWITCH; thus one might wonder if less demanding superpositions of causal orders could provide similar advantages. As previous numerical results for $N = 3$ show [35], $M = N$ superposed causal orders can provide almost the fully achievable communication rate. We provided analytical results for the transmission of classical information via superpositions of M cyclically permuted completely depolarizing channels. We found that the Holevo quantity is increasing with M and is independent of N. Surprisingly, the classical capacity decreases with the dimension d of the message system. We found that the classical communication rate for a qubit never reaches 1 if we are increasing M (and inevitably also N); it saturates at around 0.311 bits per transmission. Out of $N!$ possible causal orders for N channels there are N cyclic permutations forming a subgroup. Factoring the permutation group with respect to it, we obtained $(N-1)!$ cosets, each of which is shown to be equally usable for the investigated task. However, for general N we did not consider all possible causal orders together as the cross-coset off-diagonal terms are mostly independent of the message state. Instead, for $N = 3, 4$ we studied separately cyclic, all causal orders and various superpositions of causal orders consisting of a different number of terms from different cosets. For the causal orders from an arbitrary number of cosets, i.e., the noncyclic case, we found that with growing N the cross-coset off-diagonal terms have smaller scaling factors and are either proportional to identity or to the message state ρ. Therefore, we found that the Holevo quantity does not always increase as we increase the number of superposed causal orders. Our findings support the belief that considering superpositions of cyclic causal orders might yield almost optimal classical communication rates for N CDCs in the SWITCH.

Note added. Recently, we noticed a similar work by Chiribella *et al.* [42] which independently derives one of our results. They also claim that the classical capacity of N channels with a superposition of cyclic orders is exactly equal to the Holevo quantity, which will also strengthen our results in Sec. III.

ACKNOWLEDGMENTS

S.S. acknowledges financial support through the Štefan Schwarz stipend from Slovak Academy of Sciences, Bratislava. M.S. and S.S. acknowledge financial support through the projects OPTIQUTE (Grant No. APVV-18-0518) and HOQIP (Grant No. VEGA 2/0161/19). M.S. was also supported by Grant No. 61466 from the John Templeton Foundation, as part of the "The Quantum Information Structure of Spacetime" project. The opinions expressed in this publication are those of the author(s) and do not necessarily reflect the views of the John Templeton Foundation. We thank Mario Ziman for valuable discussions. S.S. thanks Nidhin for help.

APPENDIX A: CALCULATIONS FOR N CHANNELS WITH CYCLIC ORDERS IN A QUANTUM SWITCH

Initially, we calculated the determinant of the matrix from Eq. (8) for small values of M using Eq. (C1). This allowed us to anticipate its form for general M. However, the following lemma will be proved in a simpler way using the properties of the block circulant matrix [43].

Lemma 1. For a Hermitian matrix ρ, the determinant of the $Md \times Md$ matrix ρ_M defined in Eq. (8) is

$$\mathrm{Det}(\rho_M) = \mathrm{Det}\left(\frac{1}{M}\left[\frac{\mathbb{I}}{d} + (M-1)\frac{\rho}{d^2}\right]\right) \times \mathrm{Det}\left(\frac{1}{M}\left[\frac{\mathbb{I}}{d} - \frac{\rho}{d^2}\right]\right)^{\times(M-1)}, \tag{A1}$$

where $\mathrm{Det}(\cdot)^{\times(M-1)}$ indicates that there are $M-1$ products of the same determinant.

Proof. We will begin here by mentioning some properties of the block circulant matrix [43]. A block circulant matrix $\mathbf{C}$ of the form

$$\mathbf{C} = \begin{pmatrix} A_0 & A_1 & \cdots & A_{M-2} & A_{M-1} \\ A_{M-1} & A_0 & \cdots & A_{M-3} & A_{M-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ A_2 & A_3 & \cdots & A_0 & A_1 \\ A_1 & A_2 & \cdots & A_{M-1} & A_0 \end{pmatrix}$$

can be written in the form

$$\mathbf{C} = \sum_{i=0}^{M-1} \mathbf{P_i} \otimes A_i,$$

where $\mathbf{P_i}$ are permutation matrices. Now we can block diagonalize $\mathbf{C}$ using the properties of $\mathbf{P_i}$ matrices as mentioned in Ref. [43]. We can decompose the matrix ρ_M in Eq. (8) into the form

$$\begin{aligned} \rho_M &= \mathbf{P_0} \otimes \frac{\mathbb{I}}{Md} + \left(\sum_{i=1}^{M-1} \mathbf{P_i}\right) \otimes \frac{\rho}{Md^2} \\ &= \mathbf{P_0} \otimes \frac{\mathbb{I}}{Md} + \mathbf{S} \otimes \frac{\rho}{Md^2}, \end{aligned} \tag{A2}$$

where $\mathbf{S} = \sum_{i=1}^{M-1} \mathbf{P_i}$ is a symmetric matrix with the entries, $\mathbf{S}_{ij} = 1 - \delta_{ij}$,

$$\mathbf{S} = \begin{pmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \cdots & 0 & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{pmatrix}_{M\times M}. \tag{A3}$$

Therefore, the block diagonalization of ρ_M will depend on the properties of the matrix $\mathbf{S}$ alone as $\mathbf{P_0} = \mathbb{I}$. Then we can define a matrix function $G : \mathcal{X} \to \mathcal{M}^2$ such that

$$G(x) = x^0 \otimes \frac{\mathbb{I}}{Md} + x^1 \otimes \frac{\rho}{Md^2}, \tag{A4}$$

where each element of an object in $\mathcal{X}$ is mapped to a $d \times d$ matrix. If x is a (complex) number then the symbol $\otimes$ will just be a product and the zeroth power of x is 1. Hence, we can show that $\rho_M = G(\mathbf{S})$, since $\mathbf{S}^0 = \mathbb{I}$.

Next we diagonalize the matrix $\mathbf{S}$. Its characteristic equation is $\mathrm{Det}(\mathbf{S} - \lambda\mathbb{I}) = 0 \Rightarrow \{\lambda - (M-1)\}(\lambda+1)^{M-1} = 0$, i.e., eigenvalues of $\mathbf{S}$ are $M-1$ with degeneracy 1 and -1 with degeneracy $M-1$. So we can find a unitary T which diagonalizes $\mathbf{S}$ such that $T^\dagger \mathbf{S} T = \mathrm{diag}(M-1, -1, -1, \ldots, -1)$. Consequently, the matrix of the form $\tilde{T} = T \otimes \mathbb{I}$ block diagonalizes ρ_M as

$$\tilde{T}^\dagger \rho_M \tilde{T} = \mathbf{P_0} \otimes \frac{\mathbb{I}}{Md} + T^\dagger \mathbf{S} T \otimes \frac{\rho}{Md^2},$$

i.e., we can write $\tilde{T}^\dagger \rho_M \tilde{T} = \mathrm{diag}[G(M-1), G(-1), \ldots G(-1)]$. Hence, we can conclude the validity of Eq. (A1). ■

Equipped with Lemma 1 and its proof, it is an easy task to show that the characteristic equation for ρ_M is of the form

$$\mathrm{Det}(\rho_M - \lambda\mathbb{I}) = 0 \Rightarrow \mathrm{Det}\left(\frac{1}{M}\left[\frac{\mathbb{I}}{d} + (M-1)\frac{\rho}{d^2}\right] - \lambda\mathbb{I}\right) \times \mathrm{Det}\left(\frac{1}{M}\left[\frac{\mathbb{I}}{d} - \frac{\rho}{d^2}\right] - \lambda\mathbb{I}\right)^{\times(M-1)} = 0. \tag{A5}$$

Therefore, the eigenvalues of ρ_M can be obtained as the union of eigenvalues of matrices $\frac{1}{M}[\frac{\mathbb{I}}{d} + (M-1)\frac{\rho}{d^2}]$ with degeneracy 1 and $\frac{1}{M}[\frac{\mathbb{I}}{d} - \frac{\rho}{d^2}]$ with degeneracy $M-1$.

APPENDIX B: ALL POSSIBLE CAUSAL ORDERS

Here we sketch a proof for the entries in Table I. Below, in Eq. (B1), we show that for more than two channels, if we go beyond cyclic causal orders, the off-diagonal terms are also proportional to $\frac{\mathbb{I}}{d^3}$ along with the terms proportional to $\frac{\rho}{d^2}$. Also, in Eq. (B2) we show that some of the off-diagonal terms for more than three channels are proportional to $\frac{\rho}{d^4}$. Using these observations, we sketch how off-diagonal terms appear for arbitrary N. Let us begin with the following term:

$$\frac{1}{d^{2N}}\sum_{ij\cdots\eta}\overbrace{U_\eta\cdots U_\ell}^{N-3}U_kU_jU_i(\rho U_j^\dagger U_k^\dagger)U_i^\dagger\overbrace{U_\ell^\dagger\cdots U_\eta^\dagger}^{N-3} = \frac{1}{d^{2N}}\cdot d^2\cdot d^2\cdot d^{2(N-3)}\mathrm{Tr}[\rho]\frac{\mathbb{I}}{d} = \frac{\mathbb{I}}{d^3}. \tag{B1}$$

Next consider the off-diagonal term

$$\frac{1}{d^{2N}}\sum_{ij\cdots\eta}U_j\overbrace{U_\eta\cdots U_m}^{N-4}U_kU_\ell U_i(\rho U_j^\dagger U_k^\dagger U_\ell^\dagger)U_i^\dagger\overbrace{U_m^\dagger\cdots U_\eta^\dagger}^{N-4} = \frac{1}{d^{2N}}\cdot d^2\cdot d^2\cdot d^{2(N-4)}\rho = \frac{\rho}{d^4}. \tag{B2}$$

It is evident from the above calculation that the term $\frac{\rho}{d^4}$ will not occur if $N < 4$. Now notice the similarity between Eqs. (B1) and (B2). They are quite similar except for the positioning of U_j. Using this connection, we consider the term

$$\frac{1}{d^{2N}}\sum_{ij\cdots\eta}\overbrace{U_\eta\cdots U_n}^{N-5}U_jU_mU_kU_\ell U_i(\rho U_j^\dagger U_m^\dagger U_k^\dagger U_\ell^\dagger)U_i^\dagger\overbrace{U_n^\dagger\cdots U_\eta^\dagger}^{N-5} = \frac{1}{d^{2N}}\cdot d^2\cdot d^2\cdot d^2\cdot d^{2(N-5)}\mathrm{Tr}[\rho]\frac{\mathbb{I}}{d} = \frac{\mathbb{I}}{d^5}.$$

The above term will not occur when $N < 5$. Therefore, we can infer the terms proportional to $\frac{\mathbb{I}}{d^{2k}}$ and $\frac{\rho}{d^{2k}}$, where $k \in \mathbb{Z}^+$:

$$\frac{1}{d^{2N}}\sum_{ij\cdots\eta}\overbrace{U_\eta\cdots U_\mu}^{N-(2k+1)}U_j\overbrace{U_p\cdots U_k}^{2(k-1)}U_\ell U_i\left(\rho U_j^\dagger\overbrace{U_p^\dagger\cdots U_k^\dagger}^{2(k-1)}U_\ell^\dagger\right)U_i^\dagger\overbrace{U_\mu^\dagger\cdots U_\eta^\dagger}^{N-(2k+1)} = \frac{1}{d^{2N}}\cdot d^4\cdot d^{2(k-1)}d^{2[N-(2k+1)]}\mathrm{Tr}[\rho]\frac{\mathbb{I}}{d} = \frac{\mathbb{I}}{d^{2k+1}}.$$

Again, we can note that the above term will not occur if $N < (2k+1)$. Analogously, we have the term

$$\frac{1}{d^{2N}}\sum_{ij\cdots\eta}U_j\overbrace{U_\eta\cdots U_\mu}^{N-2k}\overbrace{U_p\cdots U_k}^{2(k-1)}U_\ell U_i\left(\rho U_j^\dagger\overbrace{U_p^\dagger\cdots U_k^\dagger}^{2(k-1)}U_\ell^\dagger\right)U_i^\dagger\overbrace{U_\mu^\dagger\cdots U_\eta^\dagger}^{N-2k} = \frac{1}{d^{2N}}\cdot d^2\cdot d^{2(k-1)}\cdot d^{2(N-2k)}\rho = \frac{\rho}{d^{2k}}.$$

Thus, we showed the existence of the terms presented in Table I.

1. Three channels in a SWITCH

For three CDCs, there are two cosets ($\frac{6}{3} = 2$) of cyclic orders. Let us consider the following permutation of channels $\Lambda_2\Lambda_1\Lambda_3$. Now applying cyclic permutations to it, we get a coset $\{\Lambda_2\Lambda_1\Lambda_3, \Lambda_1\Lambda_3\Lambda_2, \Lambda_3\Lambda_2\Lambda_1\}$. One can also see that the rest of the channel orders form the other coset. The remaining task is to see how a message state behaves when it is sent through the superposition of $M_1 + M_2 = M$ causal orders. It is easy to see that off-diagonal elements [in the sense of Eq. (5)] within a single coset will contribute in the same way as in Eq. (8). Next we need to investigate the off-diagonal terms between two cosets. By a calculation analogous to Eq. (7), which we summarize below, we found that these are proportional to identity, implying that they will not contribute to the classical communication. In particular, there are two types of terms that can be evaluated as

$$\frac{1}{d^6}\sum_{ijk}U_iU_jU_k\rho U_j^\dagger U_k^\dagger U_i^\dagger = \frac{1}{d^4}\sum_{ik}U_i\left(\frac{1}{d^2}\sum_j U_j(U_k\rho)U_j^\dagger\right)U_k^\dagger U_i^\dagger = \frac{1}{d^4}\sum_{ik}U_i\left(\mathrm{Tr}[\rho U_k]\frac{\mathbb{I}}{d}\right)U_k^\dagger U_i^\dagger = \frac{1}{d^4}\sum_i U_i\rho U_i^\dagger = \frac{\mathbb{I}}{d^3},$$

$$\frac{1}{d^6}\sum_{ijk}U_iU_jU_k\rho U_i^\dagger U_j^\dagger U_k^\dagger = \frac{1}{d^4}\sum_{ik}U_i\left(\frac{1}{d^2}\sum_j U_j(U_k\rho U_i^\dagger)U_j^\dagger\right)U_k^\dagger = \frac{1}{d^4}\sum_{ik}U_i\left(\mathrm{Tr}[U_i^\dagger(U_k\rho)]\frac{\mathbb{I}}{d}\right)U_k^\dagger = \frac{1}{d^4}\sum_k U_k\rho U_k^\dagger = \frac{\mathbb{I}}{d^3},$$

where we used $\frac{1}{d}\sum_k(\mathrm{Tr}[\rho U_k])U_k^\dagger = \rho$ and $\frac{1}{d}\sum_i(\mathrm{Tr}[U_i^\dagger(U_k\rho)])U_i = U_k\rho$.

Hence, the final output state is

$$\rho_{M_1,M_2} = \frac{1}{M}\left\{\mathbb{I}_c \otimes \frac{\mathbb{I}}{d} + \left(\sum_{\ell\neq\ell'=0}^{M_1-1} |\ell\rangle\langle\ell'|_c + \sum_{\mu\neq\mu'=M_1}^{M-1} |\mu\rangle\langle\mu'|_c\right) \otimes \frac{\rho}{d^2} + \sum_{\ell,\mu}(|\ell\rangle\langle\mu|_c + |\mu\rangle\langle\ell|_c) \otimes \frac{\mathbb{I}}{d^3}\right\}. \tag{B3}$$

To find the eigenvalues of the matrix in Eq. (B3) for general M_1 and M_2, we rewrite it in the block form, i.e.,

$$\rho_{M_1,M_2} = \frac{1}{M}\left\{\mathbb{I}_c \otimes \frac{\mathbb{I}}{d} + L_{M_1,M_2} \otimes \frac{\rho}{d^2} + B_{M_1,M_2} \otimes \frac{\mathbb{I}}{d^3}\right\}, \tag{B4}$$

where the $M \times M$ matrices L and B have the form

$$\mathbf{L} = \begin{pmatrix} S_{M_1\times M_1} & \mathbf{0}_{M_1\times M_2} \\ \mathbf{0}_{M_2\times M_1} & S_{M_2\times M_2}\end{pmatrix}_{M\times M}, \quad \mathbf{B} = \begin{pmatrix} \mathbf{0}_{M_1\times M_1} & \mathbf{1}_{M_1\times M_2} \\ \mathbf{1}_{M_2\times M_1} & \mathbf{0}_{M_2\times M_2}\end{pmatrix}_{M\times M},$$

where the matrix S is defined in Eq. (A3), $\mathbf{0}$ is a null matrix, and $\mathbf{1}$ is a matrix with all entries 1. For $M_1 = M_2$, $[L, B] = 0$, so these two matrices are simultaneously diagonalizable using an $M \otimes M$ unitary matrix U, i.e., $U^\dagger LU = \mathrm{diag}(M_1 - 1, M_1 - 1, -1, \dots, -1)$ and $U^\dagger BU = \mathrm{diag}(-M_1, M_1, 0, \dots, 0)$. Consequently, the unitary matrix $\tilde{U} = U \otimes \mathbb{I}$ will block diagonalize the matrix ρ^{M_1,M_1}, i.e.,

$$\tilde{U}^\dagger \rho_{M_1,M_1}\tilde{U} = \frac{1}{M}\left\{\mathbb{I}_c \otimes \frac{\mathbb{I}}{d} + U^\dagger LU \otimes \frac{\rho}{d^2} + U^\dagger BU \otimes \frac{\mathbb{I}}{d^3}\right\}. \tag{B5}$$

This enables us to find the characteristic equations for the above Block diagonal matrix as

$$\begin{aligned}\mathrm{Det}(\rho_{M_1,M_1} - \lambda\mathbb{I}) = 0 \Rightarrow &\mathrm{Det}\left(\frac{1}{M}\left[\frac{\mathbb{I}}{d} + (M_1 - 1)\frac{\rho}{d^2} + M_1\frac{\mathbb{I}}{d^3}\right] - \lambda\mathbb{I}\right) \\ &\times \mathrm{Det}\left(\frac{1}{M}\left[\frac{\mathbb{I}}{d} + (M_1 - 1)\frac{\rho}{d^2} - M_1\frac{\mathbb{I}}{d^3}\right] - \lambda\mathbb{I}\right) \times \mathrm{Det}\left(\frac{1}{M}\left[\frac{\mathbb{I}}{d} - \frac{\rho}{d^2}\right] - \lambda\mathbb{I}\right)^{\times(M-2)} = 0.\end{aligned} \tag{B6}$$

Therefore, the eigenvalues of ρ_{M_1,M_1} are the union of eigenvalues of $\frac{1}{M}[\frac{\mathbb{I}}{d} + (M_1 - 1)\frac{\rho}{d^2} \pm M_1\frac{\mathbb{I}}{d^3}]$ with degeneracy 1 and $\frac{1}{M}[\frac{\mathbb{I}}{d} - \frac{\rho}{d^2}]$ with degeneracy $M - 2$.

However, for $M_1 \neq M_2$, we will resort to the technique described in Appendix C. In this scenario, we have three distinct cases, i.e., $(M_1 = 2, M_2 = 1)$, $(M_1 = 3, M_2 = 1)$, and $(M_1 = 3, M_2 = 2)$. To diagonalize the output matrix ρ_{M_1,M_2} for these cases, we consider the characteristic equation $\mathrm{Det}(\rho_{M_1,M_2} - \lambda\mathbb{I}) = 0$ and solve it to find eigenvalues.

Case (2,1). The characteristic equation for this case can be written using Appendix C,

$$\mathrm{Det}\left(\frac{1}{3}\left[\frac{\mathbb{I}}{d} - \frac{\rho}{d^2}\right] - \lambda\mathbb{I}\right) \times \mathrm{Det}\left(\frac{1}{9}\left[\frac{a^2\mathbb{I}}{d^2} + \frac{a\rho}{d^3} - \frac{2\mathbb{I}}{d^6}\right]\right) = 0, \tag{B7}$$

where $a = 1 - 3d\lambda$.

Proof. Using Eq. (C1), the characteristic determinant reduces to $\mathrm{Det}(\rho_{2,1} - \lambda\mathbb{I}) = \mathrm{Det}[\mathbf{X}_{11}^{(2)}]\mathrm{Det}[\mathbf{X}_{22}^{(1)}]\mathrm{Det}[\mathbf{X}_{33}^{(0)}]$. Again using Eq. (C2), we find $\mathbf{X}_{11}^{(2)} = \mathbf{X}_{11}^{(1)} - \mathbf{X}_{12}^{(1)}(\mathbf{X}_{22}^{(1)})^{-1}\mathbf{X}_{21}^{(1)}$, and as in this case $[\mathbf{X}_{22}^{(1)}, \mathbf{X}_{21}^{(1)}] = 0$, then $\mathrm{Det}[\mathbf{X}_{11}^{(2)}]\mathrm{Det}[\mathbf{X}_{22}^{(1)}] = \mathrm{Det}[\mathbf{X}_{11}^{(1)}\mathbf{X}_{22}^{(1)} - \mathbf{X}_{12}^{(1)}\mathbf{X}_{21}^{(1)}]$. Now we notice that $\mathbf{X}_{11}^{(1)} = \mathbf{X}_{22}^{(1)}$ and $\mathbf{X}_{12}^{(1)} = \mathbf{X}_{21}^{(1)}$, which means that

$$\mathrm{Det}\big[\mathbf{X}_{11}^{(2)}\big]\mathrm{Det}\big[\mathbf{X}_{22}^{(1)}\big] = \mathrm{Det}\big[\mathbf{X}_{11}^{(1)} - \mathbf{X}_{12}^{(1)}\big]\mathrm{Det}\big[\mathbf{X}_{11}^{(1)} + \mathbf{X}_{12}^{(1)}\big],$$

where we have used the fact that $[\mathbf{X}_{11}^{(1)}, \mathbf{X}_{12}^{(1)}] = 0$. At this stage we will consider the product of determinants $\mathrm{Det}[\mathbf{X}_{11}^{(1)} + \mathbf{X}_{12}^{(1)}]\mathrm{Det}[\mathbf{X}_{33}^{(0)}]$. Applying Eq. (C1), we can simplify it as

$$\mathrm{Det}\big[\mathbf{X}_{11}^{(1)} + \mathbf{X}_{12}^{(1)}\big]\mathrm{Det}\big[\mathbf{X}_{33}^{(0)}\big] = \mathrm{Det}\big[\big\{\mathbf{X}_{11}^{(0)}\big\}^2 + \mathbf{X}_{12}^{(0)}\mathbf{X}_{11}^{(0)} - 2\big\{\mathbf{X}_{13}^{(0)}\big\}^2\big],$$

where we have used the fact that $\mathbf{X}_{11}^{(0)} = \mathbf{X}_{33}^{(0)} = \frac{a\mathbb{I}}{3d}$ and $\mathbf{X}_{13}^{(0)} = \mathbf{X}_{31}^{(0)} = \mathbf{X}_{32}^{(0)} = \frac{\mathbb{I}}{3d^3}$. Noticing that $\mathbf{X}_{12}^{(0)} = \frac{\rho}{3d^2}$ and $\mathrm{Det}(\mathbf{X}_{11}^{(1)} - \mathbf{X}_{12}^{(1)}) = \mathrm{Det}(\frac{a\mathbb{I}}{3d} - \frac{\rho}{3d^2})$, we complete the proof. ■

Therefore, the eigenvalues of the matrix $\rho_{2,1}$ are the union of eigenvalues of matrices $\frac{1}{3}[\frac{\mathbb{I}}{d} - \frac{\rho}{d^2}]$ and $\frac{1}{3}[\frac{\mathbb{I}}{d} + \frac{\rho}{2d^2} \pm \frac{1}{2}\sqrt{\frac{8\mathbb{I}}{d^6} + \frac{\rho^2}{d^4}}]$. Note that the above method and results directly apply also for the case $\rho_{1,2}$ due to symmetry with respect to the exchange of M_1 and M_2.

Case (3,1). Using the method described in Appendix C, we find that the characteristic equation for this case takes the form

$$\mathrm{Det}\left(\frac{1}{4}\left[\frac{\mathbb{I}}{d} - \frac{\rho}{d^2}\right] - \lambda\mathbb{I}\right)^{\times 2} \times \mathrm{Det}\left(\frac{1}{16}\left[\frac{b^2\mathbb{I}}{d^2} + \frac{2b\rho}{d^3} - \frac{3\mathbb{I}}{d^6}\right]\right) = 0, \tag{B8}$$

where $b = 1 - 4d\lambda$.

Proof. Using Eq. (C1), the characteristic determinant reduces to $\mathrm{Det}(\rho_{3,1}-\lambda\mathbb{I}) = \mathrm{Det}[\mathbf{X}_{11}^{(3)}]\mathrm{Det}[\mathbf{X}_{22}^{(2)}]\mathrm{Det}[\mathbf{X}_{33}^{(1)}]\mathrm{Det}[\mathbf{X}_{44}^{(0)}]$. Again using Eq. (C2), we find $\mathbf{X}_{11}^{(3)} = \mathbf{X}_{11}^{(2)} - \mathbf{X}_{12}^{(2)}(\mathbf{X}_{22}^{(2)})^{-1}\mathbf{X}_{21}^{(2)}$, and as in this case $[\mathbf{X}_{22}^{(2)},\mathbf{X}_{21}^{(2)}]=0$, then $\mathrm{Det}[\mathbf{X}_{11}^{(3)}]\mathrm{Det}[\mathbf{X}_{22}^{(2)}] = \mathrm{Det}[\mathbf{X}_{11}^{(2)}\mathbf{X}_{22}^{(2)} - \mathbf{X}_{12}^{(2)}\mathbf{X}_{21}^{(2)}]$. Now we notice that $\mathbf{X}_{11}^{(2)} = \mathbf{X}_{22}^{(2)}$ and $\mathbf{X}_{12}^{(2)} = \mathbf{X}_{21}^{(2)}$, which means that

$$\mathrm{Det}\big[\mathbf{X}_{11}^{(3)}\big]\mathrm{Det}\big[\mathbf{X}_{22}^{(2)}\big] = \mathrm{Det}\big[\mathbf{X}_{11}^{(2)} - \mathbf{X}_{12}^{(2)}\big]\mathrm{Det}\big[\mathbf{X}_{11}^{(2)} + \mathbf{X}_{12}^{(2)}\big],$$

where we have used the fact that $[\mathbf{X}_{11}^{(2)},\mathbf{X}_{12}^{(2)}]=0$. At this stage we will consider the product of determinants $\mathrm{Det}[\mathbf{X}_{11}^{(2)} + \mathbf{X}_{12}^{(2)}]\mathrm{Det}[\mathbf{X}_{33}^{(1)}]$. Applying Eq. (C1), we can simplify it as

$$\begin{aligned}\mathrm{Det}\big[\mathbf{X}_{11}^{(2)} + \mathbf{X}_{12}^{(2)}\big]\mathrm{Det}\big[\mathbf{X}_{33}^{(1)}\big] &= \mathrm{Det}\big[\{\mathbf{X}_{11}^{(1)}\}^2 + \mathbf{X}_{12}^{(1)}\mathbf{X}_{11}^{(1)} - 2\{\mathbf{X}_{12}^{(1)}\}^2\big]\\ &= \mathrm{Det}\big[\mathbf{X}_{11}^{(1)} - \mathbf{X}_{12}^{(1)}\big]\mathrm{Det}\big[\mathbf{X}_{11}^{(1)} + 2\mathbf{X}_{12}^{(1)}\big],\end{aligned}$$

where we have used the fact that $\mathbf{X}_{11}^{(1)} = \mathbf{X}_{33}^{(1)}$ and $\mathbf{X}_{12}^{(1)} = \mathbf{X}_{13}^{(1)} = \mathbf{X}_{31}^{(1)} = \mathbf{X}_{32}^{(1)}$. Now the product $\mathrm{Det}[\mathbf{X}_{11}^{(1)} + 2\mathbf{X}_{12}^{(1)}]\mathrm{Det}[\mathbf{X}_{44}^{(0)}]$ simplifies to

$$\mathrm{Det}\big[\mathbf{X}_{11}^{(1)} + 2\mathbf{X}_{12}^{(1)}\big]\mathrm{Det}\big[\mathbf{X}_{44}^{(0)}\big] = \mathrm{Det}\big[\{\mathbf{X}_{11}^{(0)}\}^2 + 2\mathbf{X}_{12}^{(0)}\mathbf{X}_{11}^{(0)} - 3\{\mathbf{X}_{14}^{(0)}\}^2\big],$$

where we have used the fact that $\mathbf{X}_{11}^{(0)} = \mathbf{X}_{44}^{(0)} = \frac{b\mathbb{I}}{4d}$ and $\mathbf{X}_{14}^{(0)} = \mathbf{X}_{41}^{(0)} = \mathbf{X}_{42}^{(0)} = \frac{\mathbb{I}}{4d^3}$. Noticing that $\mathbf{X}_{12}^{(0)} = \frac{\rho}{4d^2}$ and $\mathrm{Det}[\mathbf{X}_{11}^{(2)} - \mathbf{X}_{12}^{(2)}] = \mathrm{Det}[\mathbf{X}_{11}^{(1)} - \mathbf{X}_{12}^{(1)}] = \mathrm{Det}(\frac{b\mathbb{I}}{4d} - \frac{\rho}{4d^2})$, we complete the proof. ■

Therefore, the eigenvalues of the matrix $\rho_{3,1}$ are the union of eigenvalues of matrices $\frac{1}{4}[\frac{\mathbb{I}}{d} - \frac{\rho}{d^2}]$ with degeneracy 2 and $\frac{1}{4}[\frac{3\mathbb{I}}{d} + \frac{4\rho}{d^2} \pm \frac{1}{2}\sqrt{\frac{3\mathbb{I}}{d^6} + \frac{\rho^2}{d^4} - \frac{7\mathbb{I}}{16d^2} - \frac{\rho}{2d^3}}]$. Note that, due to symmetry, the above applies also to the case of $\rho_{1,3}$.

Case (3,2). Using the method described in Appendix C, we find that the characteristic equation for this case takes the form

$$\mathrm{Det}\left(\frac{1}{5}\left[\frac{\mathbb{I}}{d} - \frac{\rho}{d^2}\right] - \lambda\mathbb{I}\right)^{\times 3} \times \mathrm{Det}\left(\frac{1}{25}\left[\frac{c^2\mathbb{I}}{d^2} + \frac{3c\rho}{d^3} + \frac{2\rho^2}{d^4} - \frac{6\mathbb{I}}{d^6}\right]\right) = 0, \tag{B9}$$

where $c = 1 - 5d\lambda$.

Proof. Using Eq. (C1), the characteristic determinant reduces to $\mathrm{Det}(\rho_{3,2}-\lambda\mathbb{I}) = \prod_{k=1}^{5}\mathrm{Det}[\mathbf{X}_{kk}^{(5-k)}]$.

Again using Eq. (C2), we find $\mathbf{X}_{11}^{(4)} = \mathbf{X}_{11}^{(3)} - \mathbf{X}_{12}^{(3)}(\mathbf{X}_{22}^{(3)})^{-1}\mathbf{X}_{21}^{(3)}$, and as in this case $[\mathbf{X}_{22}^{(3)},\mathbf{X}_{21}^{(3)}]=0$, then $\mathrm{Det}[\mathbf{X}_{11}^{(4)}]\mathrm{Det}[\mathbf{X}_{22}^{(3)}] = \mathrm{Det}[\mathbf{X}_{11}^{(3)}\mathbf{X}_{22}^{(3)} - \mathbf{X}_{12}^{(3)}\mathbf{X}_{21}^{(3)}]$. Now we notice that $\mathbf{X}_{11}^{(3)} = \mathbf{X}_{22}^{(3)}$ and $\mathbf{X}_{12}^{(3)} = \mathbf{X}_{21}^{(3)}$, which means that

$$\mathrm{Det}\big[\mathbf{X}_{11}^{(4)}\big]\mathrm{Det}\big[\mathbf{X}_{22}^{(3)}\big] = \mathrm{Det}\big[\mathbf{X}_{11}^{(3)} - \mathbf{X}_{12}^{(3)}\big]\mathrm{Det}\big[\mathbf{X}_{11}^{(3)} + \mathbf{X}_{12}^{(3)}\big],$$

where we have used the fact that $[\mathbf{X}_{11}^{(3)},\mathbf{X}_{12}^{(3)}]=0$. At this stage we will consider the product of determinants $\mathrm{Det}[\mathbf{X}_{11}^{(3)} + \mathbf{X}_{12}^{(3)}]\mathrm{Det}[\mathbf{X}_{33}^{(2)}]$. Applying Eq. (C1), we can simplify it as

$$\begin{aligned}\mathrm{Det}\big[\mathbf{X}_{11}^{(3)} + \mathbf{X}_{12}^{(3)}\big]\mathrm{Det}\big[\mathbf{X}_{33}^{(2)}\big] &= \mathrm{Det}\big[\{\mathbf{X}_{11}^{(2)}\}^2 + \mathbf{X}_{12}^{(2)}\mathbf{X}_{11}^{(2)} - 2\{\mathbf{X}_{13}^{(2)}\}^2\big]\\ &= \mathrm{Det}\big[\mathbf{X}_{11}^{(2)} - \mathbf{X}_{12}^{(1)}\big]\mathrm{Det}\big[\mathbf{X}_{11}^{(1)} + 2\mathbf{X}_{12}^{(1)}\big],\end{aligned}$$

where we have used the fact that $\mathbf{X}_{11}^{(2)} = \mathbf{X}_{33}^{(2)}$ and $\mathbf{X}_{12}^{(2)} = \mathbf{X}_{13}^{(2)} = \mathbf{X}_{31}^{(2)} = \mathbf{X}_{32}^{(2)}$. Next the product $\mathrm{Det}[\mathbf{X}_{11}^{(2)} + 2\mathbf{X}_{12}^{(2)}]\mathrm{Det}[\mathbf{X}_{44}^{(1)}]\mathrm{Det}[\mathbf{X}_{55}^{(0)}]$ simplifies to

$$\begin{aligned}\mathrm{Det}\big[\mathbf{X}_{11}^{(2)} + 2\mathbf{X}_{12}^{(2)}\big]\mathrm{Det}\big[\mathbf{X}_{44}^{(1)}\big]\mathrm{Det}\big[\mathbf{X}_{55}^{(0)}\big] &= \mathrm{Det}\big[\{\mathbf{X}_{11}^{(1)} + 2\mathbf{X}_{12}^{(1)}\}\mathbf{X}_{44}^{(1)} - 3\{\mathbf{X}_{14}^{(1)}\}^2\big]\mathrm{Det}\big[\mathbf{X}_{55}^{(0)}\big]\\ &= \mathrm{Det}\big[\{\mathbf{X}_{11}^{(1)} + 2\mathbf{X}_{12}^{(1)}\}\{\mathbf{X}_{44}^{(0)}\mathbf{X}_{55}^{(0)} - \mathbf{X}_{45}^{(0)}\mathbf{X}_{54}^{(0)}\} - 3\mathbf{X}_{14}^{(1)}\{\mathbf{X}_{14}^{(0)}\mathbf{X}_{55}^{(0)} - \mathbf{X}_{15}^{(0)}\mathbf{X}_{54}^{(0)}\}\big],\end{aligned}$$

where in the first line we have used the fact that $\mathbf{X}_{14}^{(1)} = \mathbf{X}_{41}^{(1)} = \mathbf{X}_{42}^{(1)}$. We know that $\mathbf{X}_{44}^{(0)} = \frac{c\mathbb{I}}{5d}$, $\mathbf{X}_{45}^{(0)} = \mathbf{X}_{54}^{(0)} = \frac{\rho}{5d^2}$, and $\mathbf{X}_{14}^{(0)} = \mathbf{X}_{15}^{(0)} = \frac{\mathbb{I}}{5d^3}$. After few tedious algebraic steps, we find that $\mathbf{X}_{11}^{(1)} = \frac{c\mathbb{I}}{5d} - \frac{\mathbb{I}}{5cd^3}$, $\mathbf{X}_{12}^{(1)} = \frac{\rho}{5d^2} - \frac{\mathbb{I}}{5cd^3}$, and $\mathbf{X}_{14}^{(1)} = \frac{\mathbb{I}}{5d^3} - \frac{\rho}{5cd^4}$. Putting all these in the above equation, we find the simplification

$$\mathrm{Det}\big[\mathbf{X}_{11}^{(2)} + 2\mathbf{X}_{12}^{(2)}\big]\mathrm{Det}\big[\mathbf{X}_{44}^{(1)}\big]\mathrm{Det}\big[\mathbf{X}_{55}^{(0)}\big] = \mathrm{Det}\left(\frac{1}{5}\left[\frac{\mathbb{I}}{d} - \frac{\rho}{d^2}\right] - \lambda\mathbb{I}\right)\mathrm{Det}\left(\frac{1}{25}\left[\frac{c^2\mathbb{I}}{d^2} + \frac{3c\rho}{d^3} + \frac{2\rho^2}{d^4} - \frac{6\mathbb{I}}{d^6}\right]\right).$$

Noticing that $\mathrm{Det}(\mathbf{X}_{11}^{(2)} - \mathbf{X}_{12}^{(2)}) = \mathrm{Det}(\mathbf{X}_{11}^{(3)} - \mathbf{X}_{12}^{(3)}) = \mathrm{Det}(\frac{c\mathbb{I}}{5d} - \frac{\rho}{5d^2})$, we complete our proof. ■

Therefore, the eigenvalues of the matrix $\rho_{3,2}$ are the union of eigenvalues of matrices $\frac{1}{5}[\frac{\mathbb{I}}{d} - \frac{\rho}{d^2}]$ with degeneracy 3 and $\frac{1}{5}[\frac{\mathbb{I}}{d} + \frac{3\rho}{2d^2} \pm \frac{1}{2}\sqrt{\frac{24\mathbb{I}}{d^6} + \frac{(9-6d-8\rho)\rho}{d^4}}]$. Due to symmetry, the above results apply also for the case of $\rho_{2,3}$.

The reduced density matrix for the control system after the evolution can be calculated using the prescription given in Ref. [22] and is given by

$$\tilde{\rho}_c(M) = \frac{1}{M}\left\{\mathbb{I}_c + \frac{1}{d^2}\sum_{i\neq j}^{M-1}|i\rangle\langle j|\right\}, \tag{B10}$$

whose eigenvalues are $\frac{1}{M}(1-\frac{1}{d^2})$ with degeneracy $M-1$ and $\frac{1}{M}(1+\frac{M-1}{d^2})$ with degeneracy one. Note that we can retrieve the special cases of the cyclic orders by replacing either $M_1 = 0$ or $M_2 = 0$.

2. The $N = 4$ case

To show that $\frac{\rho}{d^4}$ terms occur in an off-diagonal block, we will pick two instances

$$\begin{aligned}\frac{1}{d^8}\sum_{ijk\ell} U_jU_kU_\ell U_i(\rho U_j^\dagger U_\ell^\dagger)U_i^\dagger U_k^\dagger &= \frac{1}{d^8}\cdot d^2\sum_{jk\ell}U_jU_kU_\ell\left(\mathrm{Tr}[\rho U_j^\dagger U_\ell^\dagger]\frac{\mathbb{I}}{d}\right)U_k^\dagger \\ &= \frac{1}{d^6}\sum_{jk}U_jU_k(\rho U_j^\dagger)U_k^\dagger = \frac{1}{d^6}\cdot d^2\sum_j U_j\mathrm{Tr}[\rho U_j^\dagger]\frac{\mathbb{I}}{d} = \frac{\rho}{d^4},\end{aligned}$$

where we have used the fact that $\frac{1}{d}\sum_k(\mathrm{Tr}[\rho U_k^\dagger])U_k = \rho$, and another off-diagonal term

$$\begin{aligned}&\frac{1}{d^8}\sum_{ijk\ell}U_jU_\ell U_kU_i(\rho U_j^\dagger U_k^\dagger U_\ell)U_i^\dagger = \frac{1}{d^8}\cdot d^2\sum_{jk\ell}U_jU_\ell U_k\left(\mathrm{Tr}[\rho U_j^\dagger U_\ell^\dagger U_k^\dagger]\frac{\mathbb{I}}{d}\right) \\ &= \frac{1}{d^6}\sum_{j\ell}U_jU_\ell(\rho U_j^\dagger)U_\ell^\dagger = \frac{1}{d^6}\cdot d^2\sum_j U_j\mathrm{Tr}[\rho U_j^\dagger]\frac{\mathbb{I}}{d} = \frac{\rho}{d^4},\end{aligned}$$

where we have used the fact that $\frac{1}{d}\sum_k(\mathrm{Tr}[\rho U_j^\dagger U_\ell^\dagger U_k^\dagger])U_k = \rho U_j^\dagger U_\ell^\dagger$.

Using the ordering within cosets described in the main text, we find that the output state after evolution is given by

$$\rho_{M_\eta} = \frac{1}{M}\left\{\mathbb{I}_c\otimes\frac{\mathbb{I}}{d} + L_{M_\eta}\otimes\frac{\rho}{d^2} + B_{M_\eta}\otimes\frac{\mathbb{I}}{d^3} + Q_{M_\eta}\otimes\frac{\rho}{d^4}\right\}, \tag{B11}$$

where the matrices L, B, and Q are given by

$$\mathbf{L} = \begin{pmatrix} S_{M_1\times M_1} & \mathbf{0}_{M_1\times M_2} & \mathbf{0}_{M_1\times M_3} & \cdots & \mathbf{0}_{M_1\times M_6} \\ \mathbf{0}_{M_2\times M_1} & S_{M_2\times M_2} & \mathbf{0}_{M_2\times M_3} & \cdots & \mathbf{0}_{M_2\times M_6} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0}_{M_5\times M_1} & \mathbf{0}_{M_5\times M_2} & \cdots & S_{M_5\times M_5} & \mathbf{0}_{M_5\times M_6} \\ \mathbf{0}_{M_6\times M_1} & \mathbf{0}_{M_6\times M_2} & \cdots & \mathbf{0}_{M_6\times M_5} & S_{M_6\times M_6}\end{pmatrix},$$

and for the sake of simplicity we write the form of B and Q when $M = 4!$, i.e.,

$$\mathbf{B} = \begin{pmatrix} \mathbf{0} & B_1 & B_2 & B_3 & B_4 & B_7 \\ B_1 & \mathbf{0} & B_3 & B_6 & B_2 & B_4 \\ B_2 & B_3 & \mathbf{0} & B_4 & B_5 & B_1 \\ B_3 & B_6 & B_4 & \mathbf{0} & B_1 & B_2 \\ B_4 & B_2 & B_5 & B_1 & \mathbf{0} & B_3 \\ B_7 & B_4 & B_1 & B_2 & B_3 & \mathbf{0}\end{pmatrix}, \quad \mathbf{Q} = \begin{pmatrix} \mathbf{0} & Q_1 & Q_2 & Q_3 & Q_4 & Q_7 \\ Q_1 & \mathbf{0} & Q_3 & Q_6 & Q_2 & Q_4 \\ Q_2 & Q_3 & \mathbf{0} & Q_4 & Q_5 & Q_1 \\ Q_3 & Q_6 & Q_4 & \mathbf{0} & Q_1 & Q_2 \\ Q_4 & Q_2 & Q_5 & Q_1 & \mathbf{0} & Q_3 \\ Q_7 & Q_4 & Q_1 & Q_2 & Q_3 & \mathbf{0}\end{pmatrix},$$

where B_i and Q_i are defined as

$$B_1 = \begin{pmatrix}1&0&1&1\\0&1&0&0\\1&0&1&1\\1&0&1&1\end{pmatrix}, \quad B_5 = \begin{pmatrix}1&1&0&0\\1&1&0&0\\0&0&1&1\\0&0&1&1\end{pmatrix}, \quad Q_1 = \begin{pmatrix}0&1&0&0\\1&0&1&1\\0&1&0&0\\0&1&0&0\end{pmatrix}, \quad Q_5 = \begin{pmatrix}0&0&1&1\\0&0&1&1\\1&1&0&0\\1&1&0&0\end{pmatrix}.$$

In addition, $B_p = \pi_p[B_1]$ for $p = 2, 3, 4$ and $B_p = \pi_p[B_5]$ for $p = 6, 7$; similarly, $Q_p = \pi_p[Q_1]$ for $p = 2, 3, 4$ and $Q_p = \pi_p[Q_5]$ for $p = 6, 7$, where π_p are permutation operators which permute among the columns as well as rows of the target matrix. For our scenario, $\pi_2 = \pi^{\text{row}}_{1\leftrightarrow 2}\cdot\pi^{\text{col}}_{1\leftrightarrow 2}$, $\pi_3 = \pi_7 = \pi^{\text{row}}_{2\leftrightarrow 4}\cdot\pi^{\text{col}}_{2\leftrightarrow 4}$, and $\pi_4 = \pi_6 = \pi^{\text{row}}_{2\leftrightarrow 3}\cdot\pi^{\text{col}}_{2\leftrightarrow 3}$, where $i\leftrightarrow j$ denotes exchange between the specific labels. Therefore, we find that in each row (and column) in ρ_{M_η}, there are 12 entries proportional to $\frac{\mathbb{I}}{d^3}$ and 8 entries proportional to $\frac{\rho}{d^4}$. Note, however, that our analysis of the Holevo quantity is invariant under arbitrary arrangement of elements in a particular coset. The reduced density matrix for the control qubit after the evolution can be calculated using the prescription given in Ref. [22] and is given by

$$\tilde{\rho}_c(M_\eta) = \frac{1}{M}\left\{\mathbb{I}_c + \frac{1}{d^2}(L_{M_\eta} + B_{M_\eta}) + \frac{1}{d^4}Q_{M_\eta}\right\}. \tag{B12}$$

It would be too tedious to diagonalize analytically the output matrix in Eq. (B11) for $M = 4!$ for arbitrary d. Instead, we will use numerical methods to diagonalize the output state in order to calculate the Holevo quantity of the effective channel. However, there are specific cases where we can easily take the analytical routes. Specifically, we consider some nontrivial situations such as the following.

Case $M = 8$. We consider a scenario where most of the off-diagonal elements are proportional to ρ. Such a situation occurs when $M_3 = M_5 = 4$ and other $M_\eta = 0$. Therefore, the output density matrix is given by

$$\rho_{4,4,\vec{0}} = \frac{1}{8}\left(\mathbb{I}_c \otimes \frac{\mathbb{I}}{d} + L_{4,4} \otimes \frac{\rho}{d^2} + B_{4,4} \otimes \frac{\mathbb{I}}{d^3} + Q_{4,4} \otimes \frac{\rho}{d^4}\right), \tag{B13}$$

where the coefficient matrices L, B, and Q are

$$L_{4,4} = \begin{pmatrix} S_{4\times4} & \mathbf{0}_{4\times4} \\ \mathbf{0}_{4\times4} & S_{4\times4}\end{pmatrix}, \quad B_{4,4} = \begin{pmatrix} \mathbf{0}_{4\times4} & B_5 \\ B_5 & \mathbf{0}_{4\times4}\end{pmatrix}, \quad Q_{4,4} = \begin{pmatrix} \mathbf{0}_{4\times4} & Q_5 \\ Q_5 & \mathbf{0}_{4\times4}\end{pmatrix}.$$

As $[L_{4,4}, B_{4,4}] = [L_{4,4}, Q_{4,4}] = [B_{4,4}, Q_{4,4}] = 0$, the matrices L, B, and Q are simultaneously diagonalizable, i.e., there exists a unitary matrix $U_{4,4}$ such that $U^\dagger LU = \mathrm{diag}(3, 3, -1, \dots, -1)$ and $U^\dagger BU = U^\dagger QU = \mathrm{diag}(-2, 2, 0, \dots, 0)$.

Case $M = 6$. We can choose many scenarios here with different M_η. The most trivial case can occur when $M_\eta = 1\,\forall\,\eta$. However, there exist two interesting scenarios $(M_1 = 3, M_4 = 3)$ and $(M_3 = 4, M_5 = 2)$.

Example 1. We consider those causal orders for $(M_1 = 3, M_4 = 3)$ such that the output state is exactly that of $(N = 3, M = 6)$ given in Eq. (B4). Further, we know that the output state is exactly diagonalizable using simultaneously the diagonalization method.

Example 2. Here we choose those causal orders for $(M_3 = 4, M_5 = 2)$ such that we get the output state

$$\rho_{4,2,\vec{0}} = \frac{1}{6}\left(\mathbb{I}_c \otimes \frac{\mathbb{I}}{d} + L_{4,2} + \otimes\frac{\rho}{d^2} + B_{4,2} \otimes \frac{\mathbb{I}}{d^3} + Q_{4,2} \otimes \frac{\rho}{d^4}\right),$$

where the coefficient matrices (L, B, and Q) are

$$L_{4,4} = \begin{pmatrix} S_{4\times4} & \mathbf{0}_{4\times2} \\ \mathbf{0}_{2\times4} & S_{2\times2}\end{pmatrix}, \quad B_{4,4} = \begin{pmatrix} \mathbf{0}_{4\times4} & \tilde{B}_5^T \\ \tilde{B}_5 & \mathbf{0}_{2\times2}\end{pmatrix}, \quad Q_{4,4} = \begin{pmatrix} \mathbf{0}_{4\times4} & \tilde{Q}_5^T \\ \tilde{Q}_5 & \mathbf{0}_{2\times2}\end{pmatrix},$$

with $\tilde{B}_5 = (\begin{smallmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1\end{smallmatrix})$ and $\tilde{Q}_5 = (\begin{smallmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0\end{smallmatrix})$. The output state in this case can be diagonalized using the method given in Appendix C.

Example 3. The trivial scenario can occur when $M_\eta = 1\,\forall\,\eta$ where all the off-diagonal terms are proportional to $\mathbb{I}$. However, we consider another scenario where not all the off-diagonal elements are proportional to $\mathbb{I}$ and the output state is given by

$$\rho_{\vec{1}} = \frac{1}{6}\begin{pmatrix} \frac{\mathbb{I}}{d} & \frac{\rho}{d^4} & \frac{\rho}{d^4} & \frac{\rho}{d^4} & \frac{\rho}{d^4} & \frac{\rho}{d^4} \\ \frac{\rho}{d^4} & \frac{\mathbb{I}}{d} & \frac{\mathbb{I}}{d^3} & \frac{\rho}{d^4} & \frac{\mathbb{I}}{d^3} & \frac{\rho}{d^4} \\ \frac{\rho}{d^4} & \frac{\mathbb{I}}{d^3} & \frac{\mathbb{I}}{d} & \frac{\mathbb{I}}{d^3} & \frac{\rho}{d^4} & \frac{\mathbb{I}}{d^3} \\ \frac{\rho}{d^4} & \frac{\rho}{d^4} & \frac{\mathbb{I}}{d^3} & \frac{\mathbb{I}}{d} & \frac{\mathbb{I}}{d^3} & \frac{\rho}{d^4} \\ \frac{\rho}{d^4} & \frac{\mathbb{I}}{d^3} & \frac{\rho}{d^4} & \frac{\mathbb{I}}{d^3} & \frac{\mathbb{I}}{d} & \frac{\mathbb{I}}{d^3} \\ \frac{\rho}{d^4} & \frac{\rho}{d^4} & \frac{\mathbb{I}}{d^3} & \frac{\rho}{d^4} & \frac{\mathbb{I}}{d^3} & \frac{\mathbb{I}}{d} \end{pmatrix}. \tag{B14}$$

This matrix can be diagonalized using the method presented in Appendix C.

APPENDIX C: DETERMINANT OF BLOCK MATRICES

In order to find the eigenvalues of the $dM \times dM$ block matrix, ρ_{out} in the main text, we need to find how its determinant factorizes into the determinant of small matrices as discussed in Refs. [40,41]. The lemma from Refs. [40,41] is stated as follows.

Lemma. Let $\mathbf{A}$ be a $pN \times pN$ complex matrix partitioned into N^2 blocks, each of size $p \times p$, i.e.,

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & \cdots & A_{NN}\end{pmatrix}_{pN\times pN}.$$

Then its determinant is given by

$$\mathrm{Det}[\mathbf{A}] = \prod_{k=1}^{N} \mathrm{Det}\big[\mathbf{X}_{kk}^{(N-k)}\big], \tag{C1}$$

where $\mathbf{X}^{(i)}$ are defined as

$$\mathbf{X}_{ij}^{(0)} = A_{ij},$$
$$\mathbf{X}_{ij}^{(k)} = A_{ij} - \vec{b}_{i,N-k+1}^{T}\tilde{\mathbf{A}}_k^{-1}\vec{a}_{N-k+1,j}, \quad k \geqslant 1,$$

with $\vec{a}_{ij} = (A_{ij}, A_{i+1,j}, \dots, A_{Nj})^T$, $\vec{b}_{ij}^T = (A_{ij}, A_{i,j+1}, \dots, A_{iN})$, and $\tilde{\mathbf{A}}_k$ the $k \times k$ block matrix formed from the lower right corner of $\mathbf{A}$. In Ref. [41] Powell also notices that

$$\mathbf{X}_{ij}^{(k+1)} = \mathbf{X}_{ij}^{(k)} - \mathbf{X}_{i,N-k}^{(k)}\big(\mathbf{X}_{N-k,N-k}^{(k)}\big)^{-1}\mathbf{X}_{N-k,j}^{(k)}. \tag{C2}$$

Equipped with the lemma from [40,41], we should be able to find the eigenvalues of the matrix ρ_M.

[1] C. E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948).
[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).
[3] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, Phys. Rev. Lett. **83**, 3081 (1999).
[4] N. Gisin, N. Linden, S. Massar, and S. Popescu, Phys. Rev. A **72**, 012338 (2005).
[5] L.-P. Lamoureux, E. Brainis, N. J. Cerf, P. Emplit, M. Haelterman, and S. Massar, Phys. Rev. Lett. **94**, 230501 (2005).
[6] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2002).
[7] M. M. Wilde, *Quantum Information Theory*, 2nd ed. (Cambridge University Press, Cambridge, 2017).
[8] Y. Aharonov, J. Anandan, S. Popescu, and L. Vaidman, Phys. Rev. Lett. **64**, 2965 (1990).
[9] D. K. L. Oi, Phys. Rev. Lett. **91**, 067902 (2003).
[10] L. Hardy, J. Phys. A: Math. Theor. **40**, 3081 (2007).
[11] O. Oreshkov, F. Costa, and Č. Brukner, Nat. Commun. **3**, 1092 (2012).
[12] I. Ibnouhsein and A. Grinbaum, Phys. Rev. A **92**, 042124 (2015).
[13] K. Goswami, C. Giarmatzi, M. Kewming, F. Costa, C. Branciard, J. Romero, and A. G. White, Phys. Rev. Lett. **121**, 090503 (2018).
[14] O. Oreshkov, Quantum **3**, 206 (2019).
[15] P. A. Guérin, A. Feix, M. Araújo, and Č. Brukner, Phys. Rev. Lett. **117**, 100502 (2016).
[16] 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).
[17] G. Chiribella, Phys. Rev. A **86**, 040301(R) (2012).
[18] M. Frey, Quantum Inf. Process. **18**, 96 (2019).
[19] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, Phys. Rev. A **88**, 022318 (2013).
[20] 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).
[21] 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).
[22] D. Ebler, S. Salek, and G. Chiribella, Phys. Rev. Lett. **120**, 120502 (2018).
[23] K. Goswami, Y. Cao, G. A. Paz-Silva, J. Romero, and A. G. White, Phys. Rev. Res. **2**, 033292 (2020).
[24] G. Chiribella, M. Banik, S. S. Bhattacharya, T. Guha, M. Alimuddin, A. Roy, S. Saha, S. Agrawal, and G. Kar, New J. of Phys. **23**, 033039 (2021).
[25] S. Salek, D. Ebler, and G. Chiribella, Quantum communication in a superposition of causal orders, arXiv:1809.06655.
[26] 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).
[27] C. Mukhopadhyay and A. K. Pati, J. Phys. Commun. **4**, 105003 (2020).
[28] C. Mukhopadhyay, M. K. Gupta, and A. K. Pati, Superposition of causal order as a metrological resource for quantum thermometry, arXiv:1812.07508.
[29] X. Zhao, Y. Yang, and G. Chiribella, Phys. Rev. Lett. **124**, 190503 (2020).
[30] M. Caleffi and A. S. Cacciapuoti, IEEE J. Sel. Areas Commun., **38**, 575 (2020).
[31] M. K. Gupta and U. Sen, Transmitting quantum information by superposing causal order of mutually unbiased measurements, arXiv:1909.13125.
[32] N. Loizeau and A. Grinbaum, Phys. Rev. A **101**, 012340 (2020).
[33] P. A. Guérin, G. Rubino, and Č. Brukner, Phys. Rev. A **99**, 062317 (2019).
[34] A. A. Abbott, J. Wechs, D. Horsman, M. Mhalla, and C. Branciard, Quantum **4**, 333 (2020).
[35] L. M. Procopio, F. Delgado, M. Enríquez, N. Belabas, and J. A. Levenson, Phys. Rev. A **101**, 012346 (2020).
[36] L. M. Procopio, F. Delgado, M. Enríquez, N. Belabas, and J. A. Levenson, Entropy **21**, 1012 (2019).

[37] T. Colnaghi, G. M. D'Ariano, S. Facchini, and P. Perinotti, Phys. Lett. A **376**, 2940 (2012).
[38] A. S. Holevo, IEEE Trans. Inf. Theory **44**, 269 (1998).
[39] B. Schumacher and M. D. Westmoreland, Phys. Rev. A **56**, 131 (1997).
[40] J. R. Silvester, Math. Gaz. **84**, 460 (2000).
[41] P. D. Powell, Calculating determinants of block matrices, arXiv:1112.4379.
[42] G. Chiribella, M. Wilson, and H.-F. Chau, Quantum and classical data transmission through completely depolarising channels in a superposition of cyclic orders, arXiv:2005.00618.
[43] A. Kaveh and H. Rahami, Acta Mech. **217**, 51 (2011).