I. INTRODUCTION

Transmission of information through noisy quantum communication lines has fascinating properties. Measurements in the basis of entangled states enable extracting more classical information as compared to the individual measurements of each information carrier [1, 2]. Moreover, encoding classical information into entangled states would give even better communication rates [3]. In addition to sending classical information, it is possible to reliably transmit quantum information and create entanglement between the sender and the receiver even if the communication line is noisy, thus opening an avenue for quantum networking [4–6]. A typical kind of noise in quantum communication lines is the loss of information carriers, e.g., photons. For continuous-variable quantum states this effect is intrinsically included in the description (see, e.g., [7]), whereas for discrete-variable quantum states this effect is usually described by an erasure channel [8–9]. For instance, if the information is encoded into polarization degrees of freedom of single photons (that can be potentially entangled among themselves), then the erasure channel accounts for the loss of photons in the line, with the probability to lose a horizontally polarized photon being the same as the probability to lose a vertically polarized photon. Additional effects of decoherence are taken into account by concatenating the erasure channel with the decoherence map, e.g., a combination of the dephasing and the loss results in the so-called “dephrasure” channel [10] and a combination of the amplitude damping channel and the loss is considered in Ref. [11], with the phenomenon of superadditivity of coherent information being observed in the both cases [10, 11]. In particular, the two-letter quantum capacity exceeds the single-letter quantum capacity by about $2.5 \cdot 10^{-3}$ bits per qubit sent in Ref. [10] and by about $5 \cdot 10^{-3}$ bits per qubit sent in Ref. [11]. The difference between the two-letter quantum capacity and the single-letter quantum capacity was experimentally tested for the dephrasure channels in Ref. [12].

However, the physics of photon transmission through optical communication lines is much richer and the losses are polarization-dependent in general [13]. This means the transmission coefficient $p_H$ for a horizontally polarized photon may significantly differ from the transmission coefficient $p_V$ for a vertically polarized photon. Effect of polarization dependent loss on the the quality of transmitted polarization entanglement and the secure quantum communication is discussed in Refs. [14, 15]. The conventional erasure channel is not an adequate description of polarization dependent losses. Similarly, a concatenation of a quantum decoherence channel with the erasure channel is not adequate in description of the transmission of quantum carriers through general lossy communication lines. In this paper, we fill this gap by introducing a generalized erasure channel that covers the above phenomena. Our definition of the generalized erasure channel differs from the generalized erasure channel pair considered in Ref. [11]. In fact, our definition comprises all concatenations of the erasure channel with other channels as partial cases; however, our definition is applicable to a wider class of scenarios with the state-dependent losses, which were not considered before.

The key idea behind the generalized erasure channel is that probabilistic transformations of quantum states are given by quantum operations that are completely positive and trace nonincreasing maps (see, e.g., [16]). Quantum operations are extensively used in description of nondestructive quantum measurements [17] and schemes with sequential measurements [18–21]. Importantly, quantum operations do not generally reduce to attenuated quantum channels. These are biased quantum operations that exhibit the state-dependent probability to lose an information carrier. In this paper, we study physics of the biased quantum operations and relate it with the information transmission through lossy communication lines.

We define the generalized erasure channel as an orthogonal sum of a trace decreasing quantum operation and a map outputting the state-dependent probability to lose the particle. This enables us to treat any type of information capacity for a trace decreasing quantum operation as the same type of capacity for the corresponding generalized erasure channel. We focus on the classical and quantum capacities of the generalized erasure channels and derive lower and upper bounds for them. We elaborate the physical scenario of the polarization dependent losses and discover superadditivity of coherent information for the corresponding generalized erasure channel. For a region of transmission
coefficients for horizontally and vertically polarized photons we analytically prove that the two-letter quantum capacity is strictly greater the single-letter quantum capacity. The maximum difference exceeds 7.197 · 10⁻³ bits per qubit sent, which is the greatest reported difference among qubit-input channels to the best of our knowledge.

The paper is organized as follows. In Sections II A and II B we review subnormalized density operators and trace nonincreasing quantum operations that can be either unbiased or biased depending on the trace of their output. In Section II C we study the ways in which quantum operations can be extended to trace preserving maps and find the minimal extension such that all other extensions are derivatives of the minimal one. In Section II D we study the normalized image of a trace decreasing quantum operation Λ and show that this image coincides with the image of some quantum channel Φ₁, which will be later used in estimation of bounds for capacities. In Section III we give a precise definition of the generalized erasure channel. In Section III A we find lower and upper bounds for the classical capacity and the single-letter classical capacity of the generalized erasure channel. In Section III B we (i) find lower and upper bounds for the quantum capacity and the singe-letter quantum capacity of the generalized erasure channel; (ii) calculate the singe-letter quantum capacity and estimate the two-letter quantum capacity for a generalized erasure channel describing the polarization dependent losses, providing a proof for superadditivity of coherent information within a wide range of polarization transmission coefficients \( p_H \) and \( p_V \). In Section IV brief conclusions are given.

II. TRACE DECREASING QUANTUM OPERATIONS

A. Subnormalized density operators

We consider \( d \)-level quantum systems as information carriers, \( 1 < d < ∞ \). By \( \mathcal{H}_d \) denote a \( d \)-dimensional Hilbert space associated with a single system. \( \mathcal{B}(\mathcal{H}_d) = \{ X : \mathcal{H}_d \to \mathcal{H}_d \} \) is the set of linear operators acting on \( \mathcal{H}_d \). A quantum state of a single information carrier is given by a density operator \( ϱ \in \mathcal{B}(\mathcal{H}_d) \) that is positive-semidefinite and has unit trace. By \( \mathcal{D}(\mathcal{H}_d) \) denote the set of density operators on \( \mathcal{H}_d \), i.e., \( \mathcal{D}(\mathcal{H}_d) = \{ ϱ \in \mathcal{B}(\mathcal{H}_d) | ϱ^\dagger = ϱ ≥ 0, \text{tr}[ϱ] = 1 \} \). Any physical quantity \( f \) associated with the information carrier is mathematically described by a self-adjoint operator \( F ∈ \mathcal{B}(\mathcal{H}_d) \) such that its mean value is given by the Born rule \( \langle f \rangle = \text{tr}[ρF] \).

For instance, if information is encoded into polarization degrees of freedom of a single photon, then \( d = 2 \) and \( \mathcal{H}_2 = \text{Span}(|H⟩, |V⟩) \), where \( |H⟩ \) and \( |V⟩ \) are the orthogonal state vectors describing a photon with horizontal and vertical polarization, respectively. Let \( f = ±1 \) be the values assigned to clicks of detectors located at two outputs of the conventional polarization beam splitter (see Fig. 1). A single photon in the state \( ϱ \) induces a click of one detector only, with the probabilities being \( p(f = −1) = ⟨H|ϱ|H⟩ \) and \( p(f = 1) = ⟨V|ϱ|V⟩ \). The average \( \langle f ⟩ = \sum_{f = ±1} f p(f) = \text{tr}[ρF] \), where \( F = ⟨H⟩|H⟩⟨H| − ⟨V⟩|V⟩⟨V| \).

The need to take a possible loss of the information carrier into account can be satisfied as follows. Extending the Hilbert space by a flag (vacuum) state \( |\text{vac}⟩ \), we can use the extended density operator \( R ∈ \mathcal{B}(\mathcal{H}_{d+1}) \). A measurement of a physical quantity \( f \) associated with the information carrier would give no outcome at all if the carrier is lost, so we assign the value \( f = 0 \) if this situation takes place. For instance, if a photon is lost, then none of the detectors at the outputs of the polarization beam splitter clicks, which we interpret as the outcome 0 of quantity \( f \) (see Fig. 1). The average \( \langle f \rangle = \text{tr}[R(F ⊗ |\text{vac}⟩⟨\text{vac}|)] = \text{tr}[PRP^† F] \), where \( P : \mathcal{H}_{d+1} \to \mathcal{H}_d \) is a projector onto the original Hilbert space associated with the information carrier. The operator \( ϱ = PRP^† \) is Hermitian, positive-semidefinite, and its trace \( \text{tr}[PRP^†] \leq 1 \), so we refer to it as a subnormalized density operator. The trace of the subnormalized density operator, \( \text{tr}[ϱ] \), is nothing else but the probability to detect the information carrier. The probability to lose the carrier equals \( 1 − \text{tr}[ϱ] \). By \( S(\mathcal{H}_d) \) denote the set of subnormalized density operators on \( \mathcal{H}_d \), i.e.,

\[
S(\mathcal{H}_d) = \{ ϱ ∈ \mathcal{B}(\mathcal{H}_d) | ϱ^\dagger = ϱ ≥ 0, \text{tr}[ϱ] ≤ 1 \}.
\]  

The set of subnormalized density operators is a subset of the cone of positive-semidefinite operators.
A quantum operation Λ dual maps the latter condition reads Λ† where Spec(\rho) is trace preserving if and only if Φ⊗Φ trace nonincreasing is equivalent to the relation Λ : \mathcal{B}(\mathcal{H}_d) \rightarrow \mathcal{B}(\mathcal{H}_d) transformation Λ:

- Unbiased operations that are attenuated quantum channels of the form Λ = pΦ, where 0 ≤ p ≤ 1 and Φ ∈ \mathcal{C}(\mathcal{H}_d), for which the output Λ[\rho] is detected with a fixed probability tr[Λ[\rho]] = p regardless of the input density operator \rho;
- Biased operations with the state-dependent probability to detect the outcome, i.e., there exist density operators \rho_1 and \rho_2 such that tr[Λ[\rho_1]] ≠ tr[Λ[\rho_2]].

Biased operations are of primary interest in this paper. A physical example of the biased operation is an optical fiber with polarization dependent losses [13]. Suppose the least attenuated polarization state is either a horizontally polarized state or a vertically polarized state, and let \rho_H and \rho_V be the attenuation factors for the horizontal and vertical polarizations, respectively. Then the effect of the optical fiber with polarization dependent losses, in its simplest form, is described by the following quantum operation with one Kraus operator A [13]:

\[ \Lambda[\rho] = A \rho A^\dagger, \quad A = \sqrt{p_H} |H\rangle\langle H| + \sqrt{p_V} |V\rangle\langle V|. \]

We quantify the bias of a quantum operation Λ ∈ \mathcal{O}(\mathcal{H}_d) by

\[ b(\Lambda) = \sup_{\rho \in \mathcal{D}(\mathcal{H}_d)} \inf_{\rho \in \mathcal{D}(\mathcal{H}_d)} \text{tr}[\Lambda[\rho]] - \inf_{\rho \in \mathcal{D}(\mathcal{H}_d)} \text{tr}[\Lambda[\rho]]. \]

Clearly, the quantity [4] vanishes if and only if the operation Λ is unbiased. Using the formalism of dual maps, we readily get

\[ b(\Lambda) = \max \text{Spec}(\Lambda^\dagger[I]) - \min \text{Spec}(\Lambda^\dagger[I]), \]

where Spec(X) is a spectrum of X. For the operation [3] we have Λ^\dagger[I] = p_H |H\rangle\langle H| + p_V |V\rangle\langle V|, so its bias b(\Lambda) = |p_H - p_V|.

C. Extending a trace decreasing operation to a channel

Any trace decreasing operation can be extended to a quantum channel by adding another trace decreasing operation. A quantum operation Λ' is called an extension for a quantum operation Λ if Λ + Λ' is trace preserving. In terms of the dual maps the latter condition reads Λ^\dagger[I] + (\Lambda')^\dagger[I] = I, which uniquely defines the operator (\Lambda')^\dagger[I] = I - Λ^\dagger[I] but
does not fix the map \( \Lambda' \), so the extension is not unique in general. In fact, since \( \Phi' [I] = I \) for any quantum channel \( \Phi \), then \( (\Lambda')' = \Phi' [I] = \Lambda' [I] \), i.e., a concatenation \( \Phi \circ \Lambda' \) is an extension for \( \Lambda \) provided \( \Lambda' \) is an extension for \( \Lambda \) too. The set \( \{ \Phi \circ \Lambda' | \Phi \in \mathcal{C} (\mathcal{H}_d) \} \) is called an orbit of the operation \( \Lambda' \in \mathcal{O} (\mathcal{H}_d) \). We have just shown that any map from the orbit of some extension \( \Lambda' \) is an extension too, but a natural question arises if all possible extensions can be obtained as an orbit of a single (in a sense, minimal) extension \( \Lambda'_\min \). The following proposition answers this question in affirmative.

**Proposition 1.** The map

\[
\Lambda'_{\min} [\rho] = \sqrt{I - \Lambda' [I]} \rho \sqrt{I - \Lambda' [I]}
\]

is a minimal extension for the quantum operation \( \Lambda \in \mathcal{O} (\mathcal{H}_d) \), i.e., any extension for \( \Lambda \) has the form \( \Phi \circ \Lambda'_{\min} \), \( \Phi \in \mathcal{C} (\mathcal{H}_d) \).

**Proof.** The map \( \Phi \) is an extension for \( \Lambda \) because it is completely positive, trace nonincreasing, and \( (\Lambda'_{\min})' [I] = I - \Lambda' [I] \).

Let \( P_+ \in \mathcal{B} (\mathcal{H}_d) \) denote the projector onto the support of \( I - \Lambda' [I] \). By \( P_0 \in \mathcal{B} (\mathcal{H}_d) \) we denote the projector onto the kernel of \( I - \Lambda' [I] \). Let \( X^{-1} \) be the Moore–Penrose inverse of \( X \), then the operator \( (I - \Lambda' [I])^{-1/2} \) is well defined and its support coincides with the support of \( P_+ \).

Consider any other extension \( \Lambda' \) for \( \Lambda \) and its Kraus decomposition \( \Lambda' [\rho] = \sum_i B_i \rho B_i^\dagger \). Then \( (\Lambda')' [I] = \sum_i B_i^\dagger B_i = I - \Lambda' [I] \) and \( \text{supp} B_i = \text{supp} B_i^\dagger B_i \subset \text{supp} P_+ \). Define the completely positive map \( \Phi \) by the Kraus sum \( \Phi [\rho] = P_0 \rho P_0 + \sum_i B_i (I - \Lambda' [I])^{-1/2} g (I - \Lambda' [I])^{-1/2} B_i^\dagger \), then \( \Phi' [I] = P_0 + (I - \Lambda' [I])^{-1/2} \sum_i B_i (I - \Lambda' [I])^{-1/2} = P_0 + P_+ = I \) and \( \Phi \) is trace preserving, which implies \( \Phi \in \mathcal{C} (\mathcal{H}_d) \). On the other hand, \( \Phi [\Lambda'_{\min} [\rho]] = \sum_i B_i P_0 \rho P_0 B_i^\dagger \). Recalling the relation \( \text{supp} B_i = \text{supp} B_i^\dagger B_i \subset \text{supp} P_+ \), we conclude that \( \Phi \circ \Lambda'_{\min} = \Lambda' \).

Note that the minimal extension \( \Lambda'_{\min} \) has the Kraus rank 1. If \( \Lambda [\rho] = A \rho A^\dagger \), \( \Lambda'_{\min} [\rho] = B_0 B_0^\dagger \), where \( B = \sqrt{I - \Lambda' [I]} \). In particular, for an optical fiber with polarization dependent losses given by Eq. (3) we have \( B = \sqrt{I - \Lambda' [I]} \) and \( \sqrt{I - \Lambda' [I]} | V \rangle \langle V | \).

**D. Normalized image of a trace decreasing operation**

Consider a trace decreasing operation \( \Lambda \in \mathcal{O} (\mathcal{H}_d) \) and its outcome \( \Lambda [\rho] \) for some density operator \( \rho \in \mathcal{D} (\mathcal{H}_d) \). Suppose the probability to detect the outcome particle is nonzero, i.e., \( \text{tr} [\rho \Lambda' [I]] \neq 0 \). While measuring a physical quantity \( f \) in a statistical experiment described in section II A, we can exclude the outcomes when none of the detectors clicks and get a conditional distribution for values of \( f \). This is equivalent to normalizing the output operator \( \Lambda [\rho] \), which leads to a map \( \Lambda_D : \mathcal{D} (\mathcal{H}_d) \to \mathcal{D} (\mathcal{H}_d) \) defined by

\[
\Lambda_D [\rho] = \frac{\Lambda [\rho]}{\text{tr} [\Lambda [\rho]]} \quad \text{tr} [\Lambda [\rho]] \neq 0.
\]

Eq. (7) describes a conditional output density operator that is commonly reconstructed in quantum optics experiments via postselection (see, e.g., [25]). Eq. (7) is also used to describe a conditional output state of a quantum measuring apparatus [17–21]. If \( \Lambda \) is unbiased, i.e., \( \Lambda = p \Phi \) for some \( 0 < p \leq 1 \) and some channel \( \Phi \), then \( \Lambda_D [\sum_i \lambda_i \rho_i] = \sum_i \lambda_i \Lambda_D [\rho_i] \) for any ensemble \( \{ \lambda_i, \rho_i \} \) of density operators, \( \{ \lambda_i \} \), being the probability distribution. In other words, for an unbiased operation \( \Lambda \) the map \( \Lambda_D \) is quasi-linear on convex sums of density operators; however, \( \Lambda_D \) is nonlinear in general because \( \Lambda_D [c \rho] = \Lambda_D [\rho] \) for all \( c \neq 0 \). If \( \Lambda \) is biased, then quasi-linearity does not hold and \( \Lambda_D [\sum_i \lambda_i \rho_i] \neq \sum_i \lambda_i \Lambda_D [\rho_i] \).

Our main interest in this section is the image \( \Lambda_D [\mathcal{D} (\mathcal{H}_d)] = \{ \Lambda_D [\rho] | \rho \in \mathcal{D} (\mathcal{H}_d) \} \) of all conditional output density operators. As \( \Lambda_D \) is nonlinear and does not exhibit quasi-linearity in general, one may expect that \( \Lambda_D [\mathcal{D} (\mathcal{H}_d)] \) significantly differs from the image \( \Phi [\mathcal{D} (\mathcal{H}_d)] \) of any quantum channel \( \Phi \in \mathcal{C} (\mathcal{H}_d) \). The following example shows that this is not the case.

**Example 1.** Consider a qubit operation \( \Lambda : \mathcal{B} (\mathcal{H}_2) \to \mathcal{B} (\mathcal{H}_2) \) of the form

\[
\Lambda [\rho] = \frac{1}{2} \left( 1 + a | c \rangle \langle d | + \text{tr} [| c \rangle \langle d |] \sigma_z + b \text{tr} [| c \rangle \langle d |] \sigma_y + \text{tr} [\rho \sigma_z] (c \sigma_z + d I) \right),
\]

where \( \sigma_z, \sigma_y, \sigma_x \) is the conventional set of Pauli operators and real parameters \( a, b, c, d \) satisfy the relations \( a \geq |c| + |d|, (a + c)^2 \geq 4b^2 + d^2 \), and \( a + |d| \leq 1 \), which make \( \Lambda \) be completely positive and trace nonincreasing. Since \( \Lambda' [I] = a I + d \sigma_z \), the bias \( b (\Lambda) = 2 |d| \). If \( a = |d| \), then \( b = c = 0 \) and the image \( \Lambda_D [\mathcal{D} (\mathcal{H}_d)] \) becomes highly degenerate, namely, \( \Lambda_D [\mathcal{D} (\mathcal{H}_d) \setminus \{ \rho_0 \}] = \{ \frac{1}{2} I \} \), where \( \rho_0 = \frac{1}{2} (I - \text{sgn}(d) \sigma_z) \) has vanishing detection probability, \( \Lambda [\rho_0] = 0 \). In the following, we consider the case \( a > |d| \).
The Bloch vector parametrization for $\rho \in D(\mathcal{H})$ reads $\rho = \frac{1}{2}(I + r \cdot \sigma)$, $r \in \mathbb{R}^3$, $|r| \leq 1$. The conditional output density operator $\Lambda_D[\rho]$ has the Bloch vector $q$ with components $q_x = br_x/(a + dr_x)$, $q_y = br_y/(a + dr_z)$, $q_z = cr_z/(a + dr_z)$. Rewriting the inequality $r \cdot r \leq 1$ in terms of $q$, we get

$$
\frac{q_x^2}{(b/a^2 - d^2)} + \frac{q_y^2}{(b/a^2 - d^2)} + \frac{q_z^2}{(ac/a^2 - d^2)} \leq 1,
$$

which defines an ellipsoid of revolution in $\mathbb{R}^3$. Not any ellipsoid within a unit ball can be associated with the image of a quantum channel [23]. In our case, the image $\Lambda_D[D(\mathcal{H})]$ coincides with the image $\Phi_D[D(\mathcal{H})]$ of the phase covariant map [26]

$$
\Phi_D[\rho] = \frac{1}{2} \left( \text{tr}[\rho](I + t_z \sigma_z) + \lambda \text{tr}[\rho \sigma_x] \sigma_x + \lambda \text{tr}[\rho \sigma_y] \sigma_y + \lambda \text{tr}[\rho \sigma_z] \sigma_z \right), \quad \lambda = \frac{b}{\sqrt{a^2 - d^2}}, t_z = -\frac{cd}{a^2 - d^2}. \tag{8}
$$

The map (8) is clearly trace preserving, whereas it is completely positive if and only if $|\lambda_z| + |t_z| \leq 1$ and $4\lambda^2 + t_z^2 \leq (1 + \lambda_z)^2$ (see Ref. [29]), with the both conditions being automatically fulfilled if $\Lambda$ is a valid quantum operation and $a > |d|$. If $a = |d|$, then we put $\lambda = \lambda_z = t_z = 0$. Finally, $\Lambda_D[D(\mathcal{H})] = \Phi_D[D(\mathcal{H})]$, where $\Phi_D$ is a quantum channel. △

Generalizing the above example to arbitrary qubit operations, one can readily see that the image of a nonlinear qubit map (7) is the same as the image of some linear, completely positive, and trace preserving qubit map $\Phi$. The following result generalizes this relation further for an arbitrary finite dimension $d$ of the underlying Hilbert space $\mathcal{H}$ and specifies the explicit form of the channel $\Phi_D$.

**Proposition 2.** For a quantum operation $\Lambda \in O(\mathcal{H})$, $\Lambda^\dagger[I] \neq 0$, the image $\Lambda_D[D(\mathcal{H})]$ coincides with the image $\Phi_D[D(\mathcal{H})]$ of the quantum channel

$$
\Phi_D[\rho] = \Lambda \left( [\Lambda^\dagger[I]]^{-1/2} \rho [\Lambda^\dagger[I]]^{-1/2} \right) + \text{tr}[\rho \Pi_0] \xi, \tag{9}
$$

where $X^{-1}$ is the Moore–Penrose inverse of $X \in B(\mathcal{H})$, $\Pi_0 \in B(\mathcal{H})$ is a projector onto the kernel of operator $\Lambda^\dagger[I]$, and $\xi$ is an arbitrary density operator from the image $\Lambda_D[D(\mathcal{H})]$.

Proof. We note that $\Phi_D$ is completely positive as the maps $X \rightarrow (\Lambda^\dagger[I])^{-1/2} X (\Lambda^\dagger[I])^{-1/2}$, $X \rightarrow \Pi_+ X \Pi_0$, $X \rightarrow \text{tr}[X] \xi$, and $\Lambda$ are all completely positive. Denoting by $\Pi_+ \in B(\mathcal{H})$ the projector onto the support of operator $\Lambda^\dagger[I]$, we calculate $\Phi^\dagger[I] = (\Lambda^\dagger[I])^{-1/2} \Lambda^\dagger[I] (\Lambda^\dagger[I])^{-1/2} + \text{tr}[\xi] \Pi_0 = \Pi_+ + \Pi_0 = I$. Therefore, $\Phi$ is trace preserving and, consequently, $\Phi$ is a quantum channel.

Given the Kraus decomposition $\Lambda[\rho] = \sum_k A_k \rho A_k^\dagger$, we have $\sum_k A_k^\dagger A_k = \Lambda^\dagger[I]$, which implies $\text{supp} A_k = \text{supp} A_k^\dagger \subset \text{supp} \Lambda^\dagger[I] = \text{supp} \Pi_+$. Then $\Lambda[\rho] = \lambda \mathbb{1} + \rho \Pi_+$ for any $\rho \in D(\mathcal{H})$. Relying on the latter observation, we divide $D(\mathcal{H})$ into three subsets and consider them separately.

(i) Consider a subset $D_+(\mathcal{H}) \subset D(\mathcal{H})$ of the density operators whose support belongs to the support of $\Lambda^\dagger[I]$, i.e., $D_+(\mathcal{H}) = \{ \rho \in D(\mathcal{H}) \mid \text{supp} \rho \subset \text{supp} \Pi_+ \}$. For any $\rho \in D_+(\mathcal{H})$ we have $\text{tr}[\rho \Pi_0] = 0$ and

$$
\Phi_D[\rho] = \frac{\Lambda[\rho]}{\text{tr}[\Lambda[\rho]]}, \quad \rho' = \frac{(\Lambda^\dagger[I])^{-1/2} \rho [\Lambda^\dagger[I]]^{-1/2}}{\text{tr}[\Lambda^\dagger[I]]} \in \mathcal{D}_+(\mathcal{H}), \tag{10}
$$

i.e., for any $\rho \in D_+(\mathcal{H})$ there exists $\rho' \in D_+(\mathcal{H})$ such that $\Phi_D[\rho] = \Lambda_D[\rho']$. Conversely, for any $\rho' \in D_+(\mathcal{H})$ we have

$$
\Lambda_D[\rho'] = \frac{\Lambda[\rho']}{\text{tr}[\Lambda[\rho']]} = \Phi_D[\rho], \quad \rho = \frac{\sqrt{\Lambda^\dagger[I]} \rho' \sqrt{\Lambda^\dagger[I]}}{\text{tr}[\Lambda[\rho']]} \in \mathcal{D}_+(\mathcal{H}), \tag{11}
$$

\[\Phi_D[\rho] = \frac{\Lambda[\rho]}{\text{tr}[\Lambda[\rho]]}, \quad \rho' = \frac{(\Lambda^\dagger[I])^{-1/2} \rho [\Lambda^\dagger[I]]^{-1/2}}{\text{tr}[\Lambda^\dagger[I]]} \in \mathcal{D}_+(\mathcal{H}), \tag{10}\]
i.e., for any $\rho' \in \mathcal{D}_+(\mathcal{H}_d)$ there exists $\rho \in \mathcal{D}_+(\mathcal{H}_d)$ such that $\Lambda_D[\rho'] = \Phi_\Lambda[\rho]$. Recalling the fact that $\Lambda[\rho] = \Lambda[\Pi_+ \rho \Pi_+]$ for any $\rho \in \mathcal{D}(\mathcal{H}_d)$ and combining it with Eqs. (10) and (11), we conclude

$$\Lambda_D[\mathcal{D}(\mathcal{H}_d)] = \Lambda_D[\mathcal{D}_+(\mathcal{H}_d)] = \Phi_\Lambda[\mathcal{D}_+(\mathcal{H}_d)].$$

(12)

(ii) Suppose $\text{supp}\rho \subset \text{supp}\Pi_0$, then $\text{tr}[\rho \Pi_0] = 1$ and $\Phi_\Lambda[\rho] = \xi \in \mathcal{D}[\mathcal{D}(\mathcal{H}_d)]$ by the statement of proposition. Hence, $\Phi_\Lambda[\rho] \in \Phi_\Lambda[\mathcal{D}_+(\mathcal{H}_d)]$.

(iii) Consider the case when $\rho \in \mathcal{D}(\mathcal{H}_d) \setminus \mathcal{D}_+(\mathcal{H}_d)$ but $\text{supp}\rho \not\subset \text{supp}\Pi_0$, then $p_+ := \text{tr}[\Pi_+ \rho \Pi_+] > 0$, $p_0 := \text{tr}[\Pi_0 \rho \Pi_0] > 0$, $p_+ + p_0 = 1$, and

$$\Phi_\Lambda[\rho] = \Phi_\Lambda[\Pi_+ \rho \Pi_+] + \text{tr}[\rho \Pi_0] \xi = p_+ \Lambda_D[\rho'] + p_0 \xi \in \Phi_\Lambda[\mathcal{D}_+(\mathcal{H}_d)]$$

because $\Phi_\Lambda[\mathcal{D}_+(\mathcal{H}_d)]$ is a convex set.

Therefore, in all the considered cases we have $\Phi_\Lambda[\mathcal{D}(\mathcal{H}_d)] = \Phi_\Lambda[\mathcal{D}_+(\mathcal{H}_d)]$. Recalling Eq. (12), we obtain the equality $\Lambda_D[\mathcal{D}(\mathcal{H}_d)] = \Phi_\Lambda[\mathcal{D}(\mathcal{H}_d)]$.

Though the images $\Lambda_D[\mathcal{D}(\mathcal{H}_d)]$ and $\Phi_\Lambda[\mathcal{D}(\mathcal{H}_d)]$ coincide, the distributions of the output states for the maps $\Lambda_D$ and $\Phi_\Lambda$ are different in general. Provided the distribution of the input density operators is uniform with respect to the Hilbert-Schmidt measure [27], the outcome distribution for $\Lambda$ would be uniform only if $\Lambda$ is unbiased. Fig. 2 illustrates the geometric meaning of the fact the higher the detection probability $\text{tr}[\Lambda[\rho]]$ the greater the density of the output states for the map $\Lambda_D$. Fig. 2 explains the geometric meaning of Proposition 2 too.

As a quantum operation $\Lambda$ and the corresponding quantum channel $\Phi_\Lambda$ are intimately related, a natural question arises if $\Lambda$ can be extended to $\Phi_\Lambda$. The following example shows this is not the case in general.

Example 2. Let $\Lambda$ be the qubit operation (3) describing polarization dependent losses with $p_H > 0$ and $p_V > 0$. Then $\Lambda[I] = p_H (|H\rangle\langle H| + |V\rangle\langle V|)$ is strictly positive and $\Phi_\Lambda = \text{Id}$. Consider the density operator $\rho = (1/2(|H\rangle\langle H| + |V\rangle\langle V|)$, then the operator $(\Phi_\Lambda - \Lambda)[\rho] = (1/2)(1 - p_H)|H\rangle\langle H| + (1 - \sqrt{p_H p_V})(|H\rangle\langle H| + |V\rangle\langle V|)) + (1 - p_V)|V\rangle\langle V|$ is not positive semidefinite whenever $p_H \neq p_V$, which implies that the map $\Phi_\Lambda - \Lambda$ is not positive, so $\Phi_\Lambda - \Lambda$ is not a quantum operation and cannot be an extension for $\Lambda$.

\[\square\]

### III. GENERALIZED ERASURE CHANNEL

A trace decreasing quantum operation $\Lambda \in \mathcal{O}(\mathcal{H}_d) \setminus \mathcal{C}(\mathcal{H}_d)$ probabilistically describes the information transmission through a lossy quantum communication line. The probabilistic nature of that transmission is due to a finite detection probability $\text{tr}[\Lambda[\rho]] \leq 1$ of a single information carrier initially prepared in the state $\rho \in \mathcal{D}(\mathcal{H}_d)$. If the information carriers enter the communication line within predefined time bins, then the loss of a carrier is detected by recording no measurement outcome within a given time bin, see Fig. 3. Operationally, we treat the absence of a measurement carriers enter the communication line within predefined time bins, then the loss of a carrier is detected by recording no measurement outcome within a given time bin, see Fig. 3. Recalling the minimal extension (6), we see that $\Gamma = \Lambda \odot (\text{Tr} \circ \Lambda_{\min}'$, where $\text{Tr}[\rho] = \text{tr}[\rho]|e\rangle\langle e|$ is a so-called trace-and-prepare quantum channel [21]. Since both $\Lambda$ and $\Lambda_{\min}'$ are completely positive, so is $\Gamma$. The map (13) is trace preserving because the map $\Lambda \odot \Lambda_{\min}'$ is trace preserving. Therefore, $\Gamma$ is a quantum channel. It is worth mentioning that our definition is very similar to a map considered in Ref. [10], section VI.B, where the authors discuss the existence of a trace nonincreasing map $\Lambda$ such that $\Lambda[\rho_0] = p_0 \rho_0$, $0 \leq p_0 \leq 1$, for two given sets of density operators $\{\rho_i\}_{i=1}^N$ and $\{\rho'_i\}_{i=1}^N$.

If $\Lambda = \text{Id}$, $0 \leq p \leq 1$, then $\Gamma_{\text{Id}}$ is nothing else but the conventional erasure channel [8, 9]. If $\Lambda = \rho \rho_0$, where $\Phi : \mathcal{B}(\mathcal{H}_d) = \mathcal{B}(\mathcal{H}_d)$ is a dephasing channel, then $\Gamma_{\rho \rho_0}$ is a so-called dephasing channel [10]. The authors of Ref. [11] consider $\Lambda = \rho \Phi$, where $\Phi : \mathcal{B}(\mathcal{H}_d) = \mathcal{B}(\mathcal{H}_d)$ is a general channel or an amplitude damping channel, in particular. In contrast to these specific cases, $\Lambda$ does not have to be unbiased, so the erasure probability $\text{tr}[\rho(I - \Lambda[I])]$ is state-dependent in general. This is the reason we refer to the channel (13) as a generalized erasure channel. Note that our definition differs from the concept of the generalized erasure channel pair introduced in Ref. [11]. As the information transmission through a lossy communication line has operational meaning only in the described scenario with predefined time bins, we associate an information transmission capacity of the quantum operation $\Lambda$ with the corresponding capacity of the generalized erasure channel $\Gamma$. In the following sections, we consider specific scenarios of classical and quantum information transmission through a lossy quantum communication line.
channel whose classical capacity is well known\cite{9,31}, namely, $C(\Lambda)$
valid quantum operation because $\Theta$ is completely positive and $\Theta$ implies $C(\Lambda)$ may vanish for unbiased quantum operations $\Lambda$, we need to establish a reasonable lower bound for $C(\Gamma)$ in the case of biased quantum operations $\Lambda$.

A. Classical capacity

Encoding classical information into $d$-dimensional quantum systems, sending all the systems through the same memoryless quantum channel $\Phi$, and measuring the outcome, it becomes possible to transmit classical information via quantum communication lines. The maximum rate of reliable information transmission per system used is called classical capacity and reads\cite{11,22}

$$C(\Phi) = \lim_{n \to \infty} \frac{1}{n} C^\chi(\Phi^\otimes n), \quad C^\chi(\Psi) = \sup_{\{\pi_i, \rho_i\}} \left\{ S\left(\sum_i \pi_i \rho_i\right) - \sum_i \pi_i S(\rho_i) \right\},$$  \hspace{1cm} (14)

where $C^\chi(\Psi)$ is the Holevo capacity of a channel $\Psi : B(H_d) \to B(H_{d^2})$, $S(\rho) = -\text{tr}[\rho \log \rho]$ is the von Neumann entropy, and $\{\pi_i, \rho_i\}$ is an ensemble of density operators ($\pi_i \geq 0, \sum_i \pi_i = 1, \rho_i \in D(H_d)$). Hereafter, the base of the log can be chosen at wish depending on the preferred units of information; the base equals 2 if the information is quantified in bits. The regularized capacity $C(\Phi)$ may exceed the Holevo capacity $C^\chi(\Phi)$\cite{3}; however, it is hard to evaluate $C(\Phi)$ explicitly for a given channel $\Phi$, so many recent studies are devoted to the search of lower and upper bounds for $C(\Phi)$ (see, e.g., Refs.\cite{28-30}). Below in this section, we find the lower and upper bounds for classical capacity of the generalized erasure channel.

As the concatenation $\Phi \circ \Psi$ of quantum channels $\Phi$ and $\Psi$ satisfies $C(\Phi \circ \Psi) \leq C(\Psi)$ (see, e.g.,\cite{22}), we first establish an analogous relation for generalized erasure channels.

**Proposition 3.** Suppose the quantum operations $\Lambda_1, \Lambda_2, \Theta \in \mathcal{O}(H_d)$ satisfy the relation $\Lambda_1 = \Theta \circ \Lambda_2$, then $C(\Gamma_{\Lambda_1}) \leq C(\Gamma_{\Lambda_2})$.

**Proof.** Define the map $\Xi : B(H_{d+1}) \to B(H_{d+1})$ by its action on matrices in the basis $\{\ldots, |\rangle\rangle\}$:

$$\Xi \left[\begin{array}{c|c}
\rho & \vdots \\
\hline
\ldots & c
\end{array}\right] = \begin{pmatrix}
\Theta[\rho] & 0 \\
0 & c + \text{tr}[\rho(I - \Theta[\Gamma])] \\
\end{pmatrix},$$  \hspace{1cm} (15)

then $\Xi$ is completely positive and trace preserving, i.e., $\Xi \in C(H_{d+1})$. It is not hard to see that $\Gamma_{\Lambda_1} = \Xi \circ \Gamma_{\Lambda_2}$, which implies $C(\Gamma_{\Lambda_1}) \leq C(\Gamma_{\Lambda_2})$ by the concatenation property for quantum channels. \hfill $\square$

Using the result of Proposition\cite{3} we can find an upper bound for $C(\Gamma)$ in terms of the quantum operation $\Lambda$.

**Proposition 4.** Let $\Lambda \in \mathcal{O}(H_d)$, then $C(\Gamma_{\Lambda}) \leq (\log d) \max \text{Spec} (\Lambda^\dagger[I])$.

**Proof.** If $\Lambda = 0$, then apparently $C(\Gamma_{\Lambda}) = 0$. Suppose $\Lambda \neq 0$, then $p_{\max} := \max \text{Spec} (\Lambda^\dagger[I]) > 0$ and $\Theta = \Lambda^{-1}\Lambda$ is a valid quantum operation because $\Theta$ is completely positive and $\Theta^\dagger[I] \leq p_{\max}^{-1}\Lambda^\dagger[I] \leq I$. Therefore, $\Lambda = \Theta \circ \Lambda_2$, where $\Lambda_2 = p_{\max}\text{Id}$. By Proposition\cite{3} we have $C(\Gamma_{\Lambda}) \leq C(\Gamma_{\Lambda_2})$. On the other hand, $\Gamma_{\Lambda_2} \equiv \Gamma_{p_{\max}\text{Id}}$ is the conventional erasure channel whose classical capacity is well known\cite{21,21}, namely, $C(\Gamma_{p_{\max}\text{Id}}) = p_{\max}\log d$. \hfill $\square$

It is tempting to treat $(\log d) \min \text{Spec} (\Lambda^\dagger[I])$ as a lower bound for $C(\Gamma_{\Lambda})$, but a simple counterexample is the unbiased operation $\Lambda[\rho] = p\text{tr}[\rho]/d I$, for which $C(\Gamma_{\Lambda}) = 0 < p\log d = (\log d) \min \text{Spec} (\Lambda^\dagger[I])$ if $0 < p \leq 1$. As the classical capacity $C(\Gamma_{\Lambda})$ may vanish for unbiased quantum operations $\Lambda$, we need to establish a reasonable lower bound for $C(\Gamma_{\Lambda})$ in the case of biased quantum operations $\Lambda$.\hfill $\square$
Proposition 5. Let $\Lambda \in \mathcal{O}(\mathcal{H}_d)$, then $C(\Gamma_\Lambda) \geq F(p_{\text{min}},p_{\text{max}})$, where
\[
F(p_{\text{min}},p_{\text{max}}) = \begin{cases} 
\log \left(1 + \exp \left[-\frac{h(p_{\text{max}}) - h(p_{\text{min}})}{p_{\text{max}} - p_{\text{min}}}\right]\right) - \frac{p_{\text{max}}h(p_{\text{min}}) - p_{\text{min}}h(p_{\text{max}})}{p_{\text{max}} - p_{\text{min}}} & \text{if } p_{\text{min}} < p_{\text{max}}, \\
0 & \text{if } p_{\text{min}} = p_{\text{max}}, 
\end{cases}
\] (16)

Exp is the inverse function to log, $p_{\text{max}} = \text{max Spec}(\Lambda^\dagger[I])$, $p_{\text{min}} = \text{min Spec}(\Lambda^\dagger[I])$, and $h(x) = -x \log x - (1 - x) \log(1 - x)$.

Proof. Consider a quantum channel $\Xi : \mathcal{B}(\mathcal{H}_{d+1}) \to \mathcal{B}(\mathcal{H}_2)$, which affects matrices in the basis $\{\ldots, |c\rangle\}$ as follows:
\[
\Xi \left( \begin{array}{ccc} \rho & & \\ & \ddots & \\ & & c \end{array} \right) = \left( \begin{array}{cc} \text{tr}[\rho] & 0 \\ 0 & c \end{array} \right).
\]

Then $\Xi \circ \Gamma_\Lambda$ is a quantum channel too and
\[
C(\Gamma_\Lambda) \geq C(\Xi \circ \Gamma_\Lambda) \geq C(\Xi \circ \Gamma_\Lambda) - \sum_{i=1,2} \pi_i S(\Xi \circ \Gamma_\Lambda[\rho_i])
\] (17)
for some ensemble $\{\pi_i, \rho_i\}_{i=1,2}$ consisting of two density matrices $\rho_1, \rho_2 \in \mathcal{D}(\mathcal{H}_d)$ emerging with probabilities $\pi_1$ and $\pi_2 = 1 - \pi_1$, respectively. Let $\rho_1 = |f_{\text{max}}\rangle \langle f_{\text{max}}|$ and $\rho_2 = |f_{\text{min}}\rangle \langle f_{\text{min}}|$, where $|f_{\text{max}}\rangle \in \mathcal{H}_d$ and $|f_{\text{min}}\rangle \in \mathcal{H}_d$ are normalized vectors such that $\Lambda^\dagger[I]|f_{\text{max}}\rangle = p_{\text{max}}|f_{\text{max}}\rangle$ and $\Lambda^\dagger[I]|f_{\text{min}}\rangle = p_{\text{min}}|f_{\text{min}}\rangle$. Then $\Xi \circ \Gamma_\Lambda[\rho_1] = \text{diag}(p_{\text{max}}, 1 - p_{\text{max}})$ and $\Xi \circ \Gamma_\Lambda[\rho_2] = \text{diag}(p_{\text{min}}, 1 - p_{\text{min}})$. The rightmost side of Eq. (17) equals $h(p_{\text{max}} + p_{\text{min}} - \pi_1 b(p_{\text{max}}) - \pi_2 b(p_{\text{min}}))$. Maximizing the latter expression with respect to a binary probability distribution $(\pi_1, \pi_2)$, we get the right hand side of Eq. (16).

The lower bound (16) is nonvanishing whenever $p_{\text{max}} > p_{\text{min}}$, i.e., for any biased operation $\Lambda$. The following result provides a lower bound for the classical capacity of $\Gamma_\Lambda$ only in terms of the bias $b(\Lambda)$.

Corollary 1. Let $\Lambda \in \mathcal{O}(\mathcal{H}_d)$, then $C(\Gamma_\Lambda) \geq \log 2 - h(\frac{1+b(\Lambda)}{2})$, where $b(\Lambda) = p_{\text{max}} - p_{\text{min}}$, $p_{\text{max}} = \text{Spec}(\Lambda^\dagger[I])$, $p_{\text{min}} = \text{Spec}(\Lambda^\dagger[I])$.

Proof. If $b(\Lambda) = 0$, then we get a trivial bound $C(\Gamma_\Lambda) \geq 0$. Suppose $b(\Lambda) > 0$ is fixed. Let us use the expressions $p_{\text{min}} = \frac{1}{2} + x - \frac{1}{2} b(\Lambda)$ and $p_{\text{max}} = \frac{1}{2} + x + \frac{1}{2} b(\Lambda)$, where $x := \frac{1}{2}(p_{\text{max}} + p_{\text{min}} - 1) \in [\frac{1}{2} b(\Lambda) - \frac{1}{2}, \frac{1}{2} b(\Lambda)].$ Then the lower bound $F(p_{\text{min}},p_{\text{max}})$ for $C(\Gamma_\Lambda)$ can be rewritten as $f(x) := F(\frac{1}{2} + x - \frac{1}{2} b(\Lambda), \frac{1}{2} + x + \frac{1}{2} b(\Lambda))$. Note that $f(x) = f(-x)$ because the replacement $x \to -x$ leads to $p_{\text{min}} \to 1 - p_{\text{max}}$ and $p_{\text{max}} \to 1 - p_{\text{min}}$, whereas $\log(1 + \exp[-y]) = \log(1 + \exp[y]) - y$. This means $f(x)$ is an even function and $\frac{df}{dx} = 0$. Moreover, $\frac{d^2 f}{dx^2} = \left[\frac{4 \pi^2(\Lambda)}{1 - b(\Lambda)^2} + \frac{1}{b(\Lambda)} \ln \frac{1 - b(\Lambda)^2}{1 - b(\Lambda)^2} + 2 b(\Lambda)^2\right] \log e > 0$, $\frac{df}{dx} < 0$ if $x < 0$, and $\frac{df}{dx} > 0$ if $x > 0$. Therefore, $C(\Gamma_\Lambda) \geq F(p_{\text{min}},p_{\text{max}}) \geq f(0) = \log 2 - h\left(\frac{1+b(\Lambda)}{2}\right)$. 

Example 3. Consider the reflection of photons from a dielectric surface, where the angle of incidence equals Brewster’s angle. In this case, we deal with the quantum operation $\Lambda \in \mathcal{O}(\mathcal{H}_2)$ given by Eq. (3), where $p_{\text{P}} > p_{\text{V}} = 0$. Then $p_{\text{max}} = p_{\text{P}}$, $p_{\text{min}} = 0$, and Propositions 4 and 5 yield $\log (1 + p_{\text{P}}(1 - p_{\text{H}})/(1 - p_{\text{V}})) \leq C(\Gamma_\Lambda) \leq C(\Gamma_\Lambda) \leq \log 2$. If $p_{\text{H}} = 1$, then $C(\Gamma_\Lambda) = \log 2$.

The disadvantage of Propositions 4 and 5 is that they exploit only two quantities, $p_{\text{max}}$ and $p_{\text{min}}$, leaving the structure of the quantum operation $\Lambda$ beyond the scope. As we know from Proposition 2, the normalized image of $\Lambda$ coincides with the image of the channel $\Phi_\Lambda$ given by Eq. (9), which enables us to relate the Holevo capacity $C(\Phi_\Lambda)$ with the Holevo capacity $C(\Gamma_\Lambda)$.

Proposition 6. Let $\Lambda \in \mathcal{O}(\mathcal{H}_d)$, then $p_{\text{min}} C(\Phi_\Lambda) \leq C(\Gamma_\Lambda) \leq p_{\text{max}} C(\Phi_\Lambda) + F(p_{\text{min}},p_{\text{max}})$, where $p_{\text{max}} = \text{max Spec}(\Lambda^\dagger[I])$, $p_{\text{min}} = \text{min Spec}(\Lambda^\dagger[I])$, $\Phi_\Lambda$ is given by Eq. (9), and $F(p_{\text{min}},p_{\text{max}})$ is given by Eq. (16).

Proof. Consider an ensemble $\{\pi_k, \xi_k\}$, where $\{\pi_k\}$ is a nondegenerate probability distribution and $\xi_k \in \mathcal{D}(\mathcal{H}_d)$. The Holevo capacity of a channel $\Psi$ reads $C(\Psi) = \text{sup} \{\pi_k, \xi_k\} \chi(\{\xi_k, \Psi[\xi_k]\})$, where $\chi(\{\pi_k, \Psi[\xi_k]\}) = S(\Psi[\sum_k \pi_k \xi_k]) = -\sum_k \pi_k S(\Psi[\xi_k])$ is the so-called Holevo quantity. Let the ensemble $\{\pi_k, \xi_k\}$ pass through the generalized erasure channel $\Gamma_\Lambda$, then the output ensemble is $\{\pi_k, \Gamma_\Lambda[\xi_k]\}$. Since $\Lambda$ is trace nonincreasing, we have $\Lambda[\xi_k] = \text{tr}[\Lambda[\xi_k]] \in [0,1]$ and $\rho_k \in \mathcal{D}(\mathcal{H}_d)$. We have $S(\pi_k \rho_k) = S(\rho_k) - \pi_k \log p_k$, and a straightforward calculation yields
\[ S \left( \begin{pmatrix} p_k & p_k \\ 0 & 1 - p_k \end{pmatrix} \right) = p_k S(\rho_k) + h(p_k). \] Denote \( \bar{\rho} = \sum_k \pi_k p_k > 0 \) and introduce the renormalized probabilities \( q_k = \pi_k p_k / \bar{\rho} \) and the average state \( \bar{\rho} = \sum_k q_k p_k \), then \( \Lambda(\sum_k \pi_k \xi_k) = \sum_k \pi_k p_k \rho_k = \bar{\rho} \bar{\rho}. \) We obtain

\[
\chi(\{\pi_k, \Lambda(\xi_k)\}) = S(\Gamma_\lambda \left( \sum_k \pi_k \xi_k \right)) - \sum_k \pi_k S(\Gamma_\lambda(\xi_k)) = \bar{\rho} S(\rho_k) + h(\bar{\rho}) - \sum_k \pi_k (p_k S(\rho_k) + h(p_k))
\]

\[ = \bar{\rho} \left( S(\bar{\rho}) - \sum_k q_k S(\rho_k) \right) + h \left( \sum_k \pi_k p_k \right) - \sum_k \pi_k h(p_k).
\]

As for any \( q_k > 0 \) we have \( \rho_k = \Phi_\lambda(\rho'_k) \) for some \( \rho'_k \in \mathcal{D}(\mathcal{H}_d) \) due to Proposition 2 we obtain

\[
\chi(\{\pi_k, \Lambda(\xi_k)\}) = \bar{\rho} \chi(\{q_k, \Phi_\lambda(\rho'_k)\}) + h \left( \sum_k \pi_k p_k \right) - \sum_k \pi_k h(p_k).
\]

(18)

Let us consider two cases.

(i) Suppose \( \{\pi_k, \xi_k\} \) is an optimal ensemble such that \( C_\lambda(\Gamma_\lambda) = \chi(\{\pi_k, \Lambda(\xi_k)\}) \), then \( \{q_k, \rho'_k\} \) is some (generally nonoptimal) ensemble and \( \chi(\{q_k, \Phi_\lambda(\rho'_k)\}) \leq C_\lambda(\Phi_\lambda) \). As \( p_{\text{min}} \leq p_k \leq p_{\text{max}} \) for all \( k \), we have \( \bar{\rho} \leq p_{\text{max}} \). Also, for any \( k \) there exists \( \mu_k \in [0, 1] \) such that \( p_k = \mu_k p_{\text{min}} + (1 - \mu_k) p_{\text{max}} \). Since the binary entropy \( h(x) \) is a concave function, \( h(p_k) \geq \mu_k h(p_{\text{min}}) + (1 - \mu_k) h(p_{\text{max}}) \). Denote \( \tilde{\pi}_1 = \sum_k \pi_k \mu_k \) and \( \tilde{\pi}_2 = \sum_k \pi_k (1 - \mu_k) \), then \( \tilde{\pi}_1, \tilde{\pi}_2 \) is a binary probability distribution and \( \sum_k \pi_k h(p_k) \geq \tilde{\pi}_1 h(p_{\text{min}}) + \tilde{\pi}_2 h(p_{\text{max}}) \), which implies \( h(\sum_k \pi_k p_k - \sum_k \pi_k h(p_k) \leq h(\tilde{\pi}_1 p_{\text{min}} + \tilde{\pi}_2 p_{\text{max}}) - \tilde{\pi}_1 h(p_{\text{min}}) - \tilde{\pi}_2 h(p_{\text{max}}) \leq F(p_{\text{max}}, p_{\text{max}}) \). Finally, we get the upper bound \( C_\lambda(\Gamma_\lambda) \leq p_{\text{max}} C_\lambda(\Phi_\lambda) + F(p_{\text{max}}, p_{\text{max}}) \).

(ii) Suppose \( \{q_k, \rho'_k\} \) is an optimal ensemble such that \( C_\lambda(\Phi_\lambda) = \chi(\{q_k, \Phi_\lambda(\rho'_k)\}) \). For any \( \rho'_k \in \mathcal{D}(\mathcal{H}_d) \) such that \( \Lambda(\xi_k)/p_k = \Phi_\lambda(\rho'_k) \) and \( p_k = \text{tr}[\Lambda(\xi_k)] > 0 \) due to Proposition 3. Define \( 1 / \bar{\rho} = \sum_k q_k / p_k \), then formula \( \pi_k = \bar{\rho} q_k / p_k \) defines a probability distribution \( \{\pi_k\} \) and Eq. (18) is valid. Since \( \{\pi_k, \xi_k\} \) is some (generally nonoptimal) ensemble for the channel \( \Gamma_\lambda \), we have \( C_\lambda(\Phi_\lambda) \geq \chi(\{\pi_k, \Lambda(\xi_k)\}) = p C_\lambda(\Phi_\lambda) + h(\sum_k \pi_k p_k) - \sum_k \pi_k h(p_k) = p_{\text{min}} C_\lambda(\Phi_\lambda) \).

Note that the inequality \( p_{\text{min}} C_\lambda(\Phi_\lambda) \leq C_\lambda(\Gamma_\lambda) \) cannot be derived in a way similar to the proof of Proposition 3 because, in general, there exists no quantum operation \( \Theta \) such that \( \Theta \circ \Lambda = p_{\text{min}} \Phi_\lambda \). For instance, for the operation \( \Lambda \) in Example 1 we explicitly find the unique \( \Theta = p_{\text{min}} \Phi_\lambda \circ \Lambda^{-1} \) if \( abc \neq 0 \); however, the obtained map \( \Theta \) turns out to be nonpositive.

For unbiased operations \( \Lambda \) we have \( p_{\text{min}} = p_{\text{max}} = 0 \), so we readily get the following result.

Corollary 2. Let \( \Lambda \in \mathcal{O}(\mathcal{H}_d) \) be an unbiased quantum operation, i.e., \( \Lambda = \rho \Phi \) for some \( 0 \leq \rho \leq 1 \) and a quantum channel \( \Phi \in \mathcal{O}(\mathcal{H}_d) \), then \( C_\lambda(\Gamma_\lambda) = \rho C_\lambda(\Phi) \).

To conclude this section, we establish the relation between tensor products \( \Lambda^{\otimes n} \) and \( \Lambda^{\otimes n} \). Using the definition and the expression \( \Gamma_\lambda^{\otimes 2} = \Lambda^{\otimes 2} \oplus \{\Lambda \otimes (\text{Tr} \circ \Lambda^{\otimes n}_\min)\} \oplus \{(\text{Tr} \circ \Lambda^{\otimes n}_\min) \otimes \Lambda\} \oplus (\text{Tr} \circ \Lambda^{\otimes n}_\min)^{\otimes 2} \).

(19)

Similarly, we have \( \Gamma_\lambda^{\otimes n} = \Lambda^{\otimes n} \oplus \mathcal{T} \), where the image of the map \( \mathcal{T} \) is orthogonal to the image of the map \( \Lambda^{\otimes n} \) with respect to the Hilbert–Schmidt scalar product. For any \( \rho \in \mathcal{B}(\mathcal{H}_d^{\otimes n}) \) the support of \( \mathcal{T}[\rho] \) belongs to a linear subspace of dimension \( (d+1)^n - d^n \); we denote this subspace by \( \mathcal{H}^{\otimes n}_{d+1} \setminus \mathcal{H}^{\otimes n}_d \). Let \( \text{Id}^{\otimes n} : \mathcal{B}(\mathcal{H}_d^{\otimes n}) \rightarrow \mathcal{B}(\mathcal{H}_d^{\otimes n}) \) be the identity transformation and \( \text{Tr}_{\mathcal{H}_d^{\otimes n_{d+1}} \setminus \mathcal{H}_d^{\otimes n}} \) be a trash-and-prepare quantum channel that maps any \( \rho \in \mathcal{B}(\mathcal{H}_d^{\otimes n_{d+1}} \setminus \mathcal{H}_d^{\otimes n}) \) to \( \text{Tr}[\rho] |e\rangle \langle e| \). Since both \( \Gamma_\lambda^{\otimes n} \) and \( \text{Tr}_{\mathcal{H}_d^{\otimes n_{d+1}} \setminus \mathcal{H}_d^{\otimes n}} \) are trace preserving, we have \( (\text{Id}^{\otimes n} \oplus \text{Tr}_{\mathcal{H}_d^{\otimes n_{d+1}} \setminus \mathcal{H}_d^{\otimes n}}) \circ \Gamma_\lambda^{\otimes n} = \Lambda^{\otimes n} \otimes \text{tr}[\rho - \Lambda^{\otimes n}[\rho]] |e\rangle \langle e| = \Lambda^{\otimes n} \).

Therefore, the channel \( \Gamma_\lambda^{\otimes n} \) is a concatenation of channels \( \Gamma_\lambda^{\otimes n} \) and \( \text{Id}^{\otimes n} \otimes \text{Tr}_{\mathcal{H}_d^{\otimes n_{d+1}} \setminus \mathcal{H}_d^{\otimes n}} \), and we immediately get the following result.

Proposition 7. Let \( \Lambda \in \mathcal{O}(\mathcal{H}_d) \), then \( C(\Gamma_\lambda) \geq \frac{1}{n} C_\lambda(\Gamma_\lambda^{\otimes n}) \geq \frac{1}{n} C_\lambda(\Gamma_\lambda^{\otimes n}) \).

B. Quantum capacity

Encoding quantum states into higher-dimensional multipartite quantum systems via an isometric map, sending all the systems through the same memoryless quantum channel \( \Phi \), and decoding the outcome via a dimension-reducing quantum channel, one can asymptotically achieve the perfect transfer of any initial quantum state provided the noise in the communication line is not too intense [10]. The rate of this quantum communication is quantified by the logarithm
of the transferred state dimension per channel use. The maximum reliable communication rate is called quantum capacity of the channel \( \Phi \) and reads [6]

\[
Q(\Phi) = \lim_{n \to \infty} \frac{1}{n} Q_1(\Phi^{\otimes_n}), \quad Q_1(\Psi) = \sup_{\rho \in D(\mathcal{H}_\text{in})} \{ S(\Psi[\rho]) - S(\tilde{\Psi}[\rho]) \},
\]

(20)

where \( \Psi : \mathcal{B}(\mathcal{H}_\text{in}) \to \mathcal{B}(\mathcal{H}_\text{out}) \) is a complementary channel to the channel \( \Psi : \mathcal{B}(\mathcal{H}_\text{in}) \to \mathcal{B}(\mathcal{H}_\text{out}) \) with the Kraus rank \( k \). To be precise, \( \Psi[\rho] = \text{tr}_{\mathcal{H}_\text{in}}[W \rho W^\dagger] \), where \( W : \mathcal{H}_\text{in} \to \mathcal{H}_\text{out} \otimes \mathcal{H}_k \) is an isometry \( (W^\dagger W = I) \) in the Stinespring dilation \( \Psi[\rho] = \text{tr}_{\mathcal{H}_\text{in}}[W \rho W^\dagger] \). Physically, the complementary channel output \( \Psi[\rho] \in D(\mathcal{H}_\text{out}) \) shows an effective state of the environment after a density operator \( \rho \in D(\mathcal{H}_\text{in}) \) has passed through a quantum channel \( \Psi : \mathcal{B}(\mathcal{H}_\text{in}) \to \mathcal{B}(\mathcal{H}_\text{out}) \) with the Kraus rank \( k \). The quantity \( S(\Psi[\rho]) - S(\tilde{\Psi}[\rho]) \) is known as the coherent information, whereas \( \frac{1}{n} Q_1(\Phi^{\otimes_n}) \) is usually referred to as an \( n \)-letter quantum capacity. Useful conditions for strict positivity of \( Q_1(\Phi) \) are given in Ref. [32].

The quantum capacity is known to satisfy the additivity property \( Q(\Phi) = Q_1(\Phi) \) if \( \Phi \) is degradable, i.e., if there exists a quantum channel \( \Xi \) such that \( \Phi = \Xi \circ \Phi \) [33]. If \( \Phi \) is antidegradable, i.e., there exists a quantum channel \( \Xi \) such that \( \Phi = \Xi \circ \tilde{\Phi} \), then \( Q(\Phi) = 0 \) and the additivity property is trivially fulfilled (see, e.g., [34]). The superadditivity of coherent information, i.e., the strict inequality \( Q_1(\Phi^{\otimes_n}) > n Q_1(\Phi) \), is known to hold for some depolarizing channels if \( n \geq 3 \) [35, 36], some depenuishment channels if \( n \geq 2 \) [10], concatenations of the erasure channel with the amplitude damping channel [11], the state-of-the-art channels \( \Phi : \mathcal{B}(\mathcal{H}_d) \to \mathcal{B}(\mathcal{H}_d) \) with \( \frac{1}{n} Q_1(\Phi^{\otimes_n}) \approx 4.4 \cdot 10^{-2} \) and their higher-dimensional generalizations [37], and for a collection of peculiar channels if \( n \geq n_0 \), where \( n_0 \geq 2 \) specifies the channel and can be arbitrary [38]. In this section, we find lower and upper bounds for the quantum capacity of a generalized erasure channel. Then we study degradability and antidegradability for a class of generalized erasure channels. For a 2-parameter map of that class, we reveal the superadditivity property \( Q_1(\Gamma^{\otimes 2}) > 2 Q_1(\Gamma) \) within a wide range of parameters.

**Proposition 8.** Let \( \Lambda \in \mathcal{O}(\mathcal{H}_d) \), then \( Q(\Gamma_\Lambda) \geq \frac{1}{n} Q_1(\Gamma^{\otimes_n}_\Lambda) \geq \frac{1}{n} Q_1(\Gamma^{\otimes_n}_\Lambda) \) for all \( n \in \mathbb{N} \).

**Proof.** The proof readily follows from the relation \( \Gamma^{\otimes_n}_\Lambda = \text{tr}_{\mathcal{H}_d} \cdots \text{tr}_{\mathcal{H}_d} \Phi^{\otimes_n} \otimes \Gamma^{\otimes_n}_\Lambda \) and the property \( Q(\Psi_2 \circ \Psi_1) \leq Q(\Psi_1) \) for concatenated quantum channels (see, e.g., [22]).

**Proposition 9.** Suppose the quantum operations \( \Lambda_1, \Lambda_2, \Theta \in \mathcal{O}(\mathcal{H}_d) \) satisfy the relation \( \Lambda_1 = \Theta \circ \Lambda_2 \), then \( Q(\Gamma_{\Lambda_1}) \leq Q(\Gamma_{\Lambda_2}) \).

**Proof.** Following the lines of Proposition 9 we get \( \Gamma_{\Lambda_1} = \Xi \circ \Gamma_{\Lambda_2} \), where the channel \( \Xi \) is given by Eq. (15). By the concatenation property for quantum channels we have \( Q(\Gamma_{\Lambda_1}) \leq Q(\Gamma_{\Lambda_2}) \).

**Proposition 10.** Let \( \Lambda \in \mathcal{O}(\mathcal{H}_d) \), then \( Q(\Gamma_\Lambda) \leq \max(0, 2p_{\text{max}} - 1) \log d, \) where \( p_{\text{max}} = \text{max Spec } (\Lambda^\dagger[I]). \)

**Proof.** If \( \Lambda = 0 \), then apparently \( Q(\Gamma_\Lambda) = 0 \). Suppose \( \Lambda \neq 0 \), then \( p_{\text{max}} > 0 \) and \( \Theta = \frac{1}{p_{\text{max}}} \Lambda \) is a valid quantum operation because \( \Theta \) is completely positive and \( \Theta[I] \leq p_{\text{max}} \Lambda^\dagger[I] \leq I \). Therefore, \( \Lambda = \Theta \circ \Lambda_2 \), where \( \Lambda_2 = p_{\text{max}} \Lambda_2. \) By Proposition 9 we have \( Q(\Gamma_\Lambda) \leq Q(\Gamma_{\Lambda_2}) \). On the other hand, \( Q(\Gamma_{\Lambda_2}) \leq \max(0, 2p_{\text{max}} - 1) \log d, \) see Ref. [22].

The relation between the operation \( \Lambda \) and the channel \( \Phi_\Lambda \) in Eq. (9) enables us to find both the lower and upper bounds for \( Q_1(\Gamma_{\Lambda}) \) in terms of \( Q_1(\Phi_\Lambda) \).

**Proposition 11.** Let \( \Lambda \in \mathcal{O}(\mathcal{H}_d) \) and suppose \( \Lambda^\dagger[I] > 0 \), then

\[
p_{\text{min}} Q_1(\Phi_\Lambda) - (1 - p_{\text{min}}) \log d \leq Q_1(\Gamma_\Lambda) \leq p_{\text{max}} Q_1(\Phi_\Lambda),
\]

(21)

where \( p_{\text{max}} = \text{max Spec } (\Lambda^\dagger[I]), p_{\text{min}} = \text{min Spec } (\Lambda^\dagger[I]), \) and \( \Phi_\Lambda \) given by Eq. (9).

**Proof.** Since \( \Lambda^\dagger[I] > 0 \), the operator \( \Pi_0 \) in Eq. (9) is the zero operator and \( \Phi_\Lambda[\rho] = \Lambda \left[ (\Lambda^\dagger[I])^{-1/2} \rho (\Lambda^\dagger[I])^{-1/2} \right] \). We can rewrite the generalized erasure channel \( \Gamma_\Lambda[\rho] \) in the form

\[
\Gamma_\Lambda[\rho] = \begin{pmatrix}
\Phi_\Lambda \left[ \sqrt{\Lambda^\dagger[I]} \rho \sqrt{\Lambda^\dagger[I]} \right] & 0 \\
0^T & \text{tr} \left[ \sqrt{I - \Lambda^\dagger[I]} \rho \sqrt{I - \Lambda^\dagger[I]} \right]
\end{pmatrix}.
\]

(22)

Let \( \Phi_\Lambda[\rho] = \sum_\alpha V_\alpha \rho V_\alpha^\dagger \), then the Kraus operators \( V_\alpha \) of the complementary channel \( \tilde{\Phi}_\Lambda \) satisfy \( \langle \alpha | \tilde{V}_j | j \rangle = \langle j | V_\alpha \rangle \), where \( \{ | j \rangle \} \) is an orthonormal basis for the output Hilbert space and \( \{ | \alpha \rangle \} \) is an orthonormal basis for the effective environment [39]. Hence, \( \tilde{V}_j = \sum_\alpha | \alpha \rangle \langle j | V_\alpha \) and \( \sqrt{\Lambda^\dagger[I]} = \sum_\alpha | \alpha \rangle \langle \alpha | \Lambda^\dagger[I] \) and \( \sqrt{I - \Lambda^\dagger[I]} = \sum_\alpha | \alpha \rangle \langle \alpha | I - \Lambda^\dagger[I] \), i.e., the map \( \rho \to \tilde{\Phi}_\Lambda \left[ \sqrt{\Lambda^\dagger[I]} \rho \sqrt{\Lambda^\dagger[I]} \right] \)
is complementary to the map \( \rho \to \Phi_\Lambda \left[ \sqrt{\Lambda^{\dagger}} \rho \sqrt{\Lambda^{\dagger}} \right] \). Since the identity channel Id is known to be complementary to the trash-and-prepare channel Tr (see, e.g., [22]), we conclude that the map \( \rho \to \sqrt{1 - \Lambda^{\dagger}} \rho \sqrt{1 - \Lambda^{\dagger}} \) is complementary to the map \( \rho \to \text{tr} \left[ \sqrt{1 - \Lambda^{\dagger}} \rho \sqrt{1 - \Lambda^{\dagger}} \right] \). Therefore,

\[
\Gamma_\Lambda[\rho] = \begin{pmatrix} \tilde{\Phi}_\Lambda \left[ \sqrt{\Lambda^{\dagger}} \rho \sqrt{\Lambda^{\dagger}} \right] & O \\ O & \sqrt{1 - \Lambda^{\dagger}} \rho \sqrt{1 - \Lambda^{\dagger}} \end{pmatrix}.
\]  

(23)

Let \( \rho \in \mathcal{D}(\mathcal{H}_d) \). Denoting \( p = \text{tr} \left[ \rho \Lambda^{\dagger} \right] \in (0, 1] \), \( \xi = p^{-1} \sqrt{\Lambda^{\dagger} \rho \sqrt{\Lambda^{\dagger}}} \in \mathcal{D}(\mathcal{H}_d) \), and \( \omega = (1 - p)^{-1} \sqrt{1 - \Lambda^{\dagger}} \rho \sqrt{1 - \Lambda^{\dagger}} \in \mathcal{D}(\mathcal{H}_d) \) if \( p \neq 1 \), with \( \omega \in \mathcal{D}(\mathcal{H}_d) \) being arbitrary if \( p = 1 \), we get

\[
S(\Gamma_\Lambda[\rho]) - S(\widetilde{\Gamma}_\Lambda[\rho]) = S \left( \begin{pmatrix} p\Phi_\Lambda[\xi] & 0 \\ 0 & (1 - p)\omega \end{pmatrix} \right) - S \left( \begin{pmatrix} \rho \Phi_\Lambda[\xi] & O \\ O & (1 - p)\omega \end{pmatrix} \right) = pS(\Phi_\Lambda[\xi]) - pS(\tilde{\Phi}_\Lambda[\xi]) - (1 - p)S(\omega).
\]  

(24)

Let us consider two cases.

(i) Suppose \( \rho \) is optimal in the sense that \( Q_1(\Gamma_\Lambda) = S(\Gamma_\Lambda[\rho]) \), then Eq. (24) implies \( Q_1(\Gamma_\Lambda) \leq pS(\Phi_\Lambda[\xi]) - pS(\tilde{\Phi}_\Lambda[\xi]) \leq p\max Q_1(\Phi_\Lambda) \).

(ii) Suppose \( \xi \) is optimal in the sense that \( Q_1(\Phi_\Lambda) = S(\Phi_\Lambda[\xi]) \), then Eq. (24) implies \( Q_1(\Gamma_\Lambda) \geq S(\Gamma_\Lambda[\rho]) - S(\widetilde{\Gamma}_\Lambda[\rho]) = pQ_1(\Phi_\Lambda) - (1 - p)S(\omega) \geq p\min Q_1(\Phi_\Lambda) - (1 - p)\min \log d \).

Example 4. Let \( \mathcal{H} \in \mathcal{O}(\mathcal{H}_d) \) be a quantum operation describing polarization dependent losses, Eq. (3). If \( p_{HV}p_V \neq 0 \), then \( \Phi_\Lambda = \text{Id} \) and Proposition [11] yields \( 2 \min(p_{HV}, p_V) \log 2 \leq Q_1(\Gamma_\Lambda) \leq \max(p_{HV}, p_V) \log 2 \). Proposition [10] gives a tighter upper bound, namely, \( Q_1(\Gamma_\Lambda) \leq 2 \max(p_{HV}, p_V) \log 2 \). Hence, \( Q_1(\Gamma_\Lambda) = 0 \) if \( \max(p_{HV}, p_V) \leq \frac{1}{2} \). Suppose \( \max(p_{HV}, p_V) > \frac{1}{2} \) and \( p_{HV}p_V \neq 0 \). We fix the orthonormal basis \( \{|H\}, |V\rangle \) and consider a general input density matrix \( \rho = \begin{pmatrix} \rho_{HH} & \rho_{HV} \\ \rho_{VH} & \rho_{VV} \end{pmatrix} \), \( \rho_{HH} + \rho_{VV} = 1 \). Then \( \Lambda[\rho] = \begin{pmatrix} \rho_{HH} & \rho_{HV} & \sqrt{\rho_{HH}\rho_{HV}} & \sqrt{\rho_{HV}\rho_{VV}} \\ \rho_{VH} & \rho_{VV} & \sqrt{\rho_{VH}\rho_{VV}} & \sqrt{\rho_{HH}\rho_{VV}} \\ 0 & 0 & \sqrt{1 - \rho_{HH}} & (1 - \rho_{VV}) \\ 0 & 0 & (1 - \rho_{VH}) & \rho_{VV} \end{pmatrix} \) and the probability to detect a photon at the output equals \( \text{tr} \left[ \Lambda[\rho] \right] = \rho_{HH} + \rho_{VV} > 0 \). Since \( \Phi_\Lambda = \text{Id} \) and \( \Phi_\Lambda = \text{Tr} \), Eqs. (22) and (23) take the form

\[
\Gamma_\Lambda[\rho] = \begin{pmatrix} \rho_{HH} & \rho_{HV} & \sqrt{\rho_{HH}\rho_{HV}} & \sqrt{\rho_{HV}\rho_{VV}} \\ \rho_{VH} & \rho_{VV} & \sqrt{\rho_{VH}\rho_{VV}} & \sqrt{\rho_{HH}\rho_{VV}} \\ 0 & 0 & \sqrt{1 - \rho_{HH}} & (1 - \rho_{VV}) \\ 0 & 0 & (1 - \rho_{VH}) & \rho_{VV} \end{pmatrix}.
\]

Consider the function \( q(\rho) := S(\Gamma_\Lambda[\rho]) - S(\widetilde{\Gamma}_\Lambda[\rho]) \), then \( Q_1(\Gamma_\Lambda) = \max_{\rho\in\mathcal{D}(\mathcal{H}_d)} q(\rho) \). To find the maximum of \( q(\rho) \), we first notice that the entropies \( S(\Gamma_\Lambda[\rho]) \) and \( S(\widetilde{\Gamma}_\Lambda[\rho]) \) do not depend on the phase of \( \rho_{HV} \), so we put \( \rho_{HV} = \rho_{VV} = \frac{1}{2} \geq 0 \) and use the Bloch ball parametrization \( \rho = \frac{1}{2} (I + x_\sigma z z^* \sigma) \), where \( 0 \leq x \leq \sqrt{1 - z^2} \). At the boundary \( x = \sqrt{1 - z^2} \) the function \( q \) vanishes because this boundary corresponds to pure states for which \( S(\Gamma_\Lambda[|\psi\rangle\langle\psi|]) = S(\widetilde{\Gamma}_\Lambda[|\psi\rangle\langle\psi|]) \) (see, e.g., [22]). Suppose \( \max_{\rho\in\mathcal{D}(\mathcal{H}_d)} q(\rho) > 0 \). Then this maximum is attained at some point \((x_*, z_*)\) satisfying \( 0 \leq x_* < \sqrt{1 - z_*^2} \). Note that \( z_* \in ( - \sqrt{1 - x_*^2}, \sqrt{1 - x_*^2}) \), with \( \frac{\partial q}{\partial z} \bigg|_{x=x_*,z=z_*} = 0 \). Consider an interior point \((x', z')\), where \( 0 < x' < \sqrt{1 - z_*^2} \), \( |q|_{x=x',z=z_*} < 0 \), and \( \frac{\partial q}{\partial z} \bigg|_{x=x',z=z_*} = 0 \). The direct calculation yields \( \frac{\partial q}{\partial z} \bigg|_{x=x',z=z_*} < 0 \), which means \( x_* \neq x' \) and the maximum cannot be attained at the interior point, so with necessity \( x_* = 0 \). To find \( z_* \) we need to solve the equation \( \frac{\partial q}{\partial z} \bigg| (\frac{1}{2} (I + z_* \sigma z)) \bigg) = 0 \). We simplify

\[
q \left( \frac{1}{2} (I + z_* \sigma z) \right) = H \left( \frac{1}{2} p_{HH} (1 + z), \frac{1}{2} p_{HV} (1 - z), 1 - \frac{1}{2} p_{HH} (1 + z) - \frac{1}{2} p_{VV} (1 - z) \right) - H \left( \frac{1}{2} p_{HH} (1 + z) + \frac{1}{2} p_{HV} (1 - z), \frac{1}{2} (1 - p_{HH}) (1 + z), \frac{1}{2} (1 - p_{VV}) (1 - z) \right),
\]

(25)

where \( H(\{\lambda_i\}) \) is the Shannon entropy of the probability distribution \( \{\lambda_i\} \). It is not hard to see that the equation \( \frac{\partial q}{\partial z} \bigg| (\frac{1}{2} (I + z_* \sigma z)) = 0 \) is equivalent to the equation \( G(p_{HV}, p_{VV}, z) = G(p_{VV}, p_{HH}, -z) \), where

\[
G(p_1, p_2, z) = -p_1 \log \frac{p_1 (1 + z)}{p_1 (1 + z) + p_2 (1 - z)} + (1 - p_1) \log \frac{(1 - p_1) (1 + z)}{(1 - p_1) (1 + z) + (1 - p_2) (1 - z)}.
\]
The analysis of derivative \( \frac{d}{dz} q \left( \frac{1}{2} (I + z \sigma_z) \right) \) shows that the maximum of \( q \) corresponds to such a solution \( z = z_* \) of the equation \( G(p_H, p_V, z) = G(p_V, p_H, -z) \) for which \( \text{sgn}(z_*) = \text{sgn}(p_V - p_H) \). Substituting this solution into Eq. (25), it can be readily checked that \( q \left( \frac{1}{2} (I + z_* \sigma_z) \right) = G(p_H, p_V, z_*) = G(p_V, p_H, -z_*) \). On the other hand, \( Q_1(\Gamma_\Lambda) = q \left( \frac{1}{2} (I + z_\sigma_z) \right) \), which enables us to find \( Q_1(\Gamma_\Lambda) \) by numerically solving the equation \( G(p_H, p_V, z) = G(p_V, p_H, -z) \) and selecting a solution of a proper sign. If \( \max(p_H, p_V) > \frac{1}{2} \) and \( p_H p_V \neq 0 \), then \( -1 < z_* < 1 \) and \( Q_1(\Gamma_\Lambda) > 0 \), which justifies our assumption that \( \max_{\rho \in \mathcal{D}(\mathcal{H}_A)} q(\rho) > 0 \). For the sake of completeness, we also provide an approximate solution \( z' \approx z_* \), for which \( Q_1(\Gamma_\Lambda) = q \left( \frac{1}{2} (I + z_\sigma_z) \right) \geq q \left( \frac{1}{2} (I + z' \sigma_z) \right) > 0 \), namely,

\[
\frac{1 - z'}{1 + z'} = \left( \frac{1 - p_V}{1 - p_H} \right) \frac{1 - p_H}{2p_H - 1} \left( \frac{p_H}{p_V} \right) \frac{p_H}{2p_H - 1} = \frac{(p_H - p_V)(p_H + p_V - 2p_H p_V)}{(2p_H - 1)(1 - p_V)p_V} \quad \text{if} \quad p_H \geq p_V, \\
\frac{1 + z'}{1 - z'} = \left( \frac{1 - p_H}{1 - p_V} \right) \frac{1 - p_V}{2p_V - 1} \left( \frac{p_V}{p_H} \right) \frac{p_V}{2p_V - 1} = \frac{(p_V - p_H)(p_V + p_V - 2p_H p_V)}{(2p_V - 1)(1 - p_H)p_H} \quad \text{if} \quad p_V > p_H.
\]

The heat map of \( Q_1(\Gamma_\Lambda) \) as function of \( p_H \) and \( p_V \) is depicted in Fig. 4. △

In what follows, we study degradability and antidegradability for a specific class of generalized erasure channels \( \Gamma_\Lambda \), where \( \Lambda \) has the Kraus rank 1.

**Proposition 12.** Suppose the quantum operation \( \Lambda \in \mathcal{O}(\mathcal{H}_A) \) has a single Kraus operator \( A \), i.e., \( \Lambda[\rho] = A\rho A^\dagger \), then \( \Gamma_\Lambda \) is degradable if and only if \( AA^\dagger \geq \frac{1}{2} I \) or \( I - A^\dagger A \) is a rank-1 operator.

**Proof.** Using the relation between the Kraus operators for a channel and a complementary channel [39], we readily get

\[
\Gamma_\Lambda[\rho] = \begin{pmatrix} A\rho A^\dagger \\ A^\dagger \end{pmatrix} = \begin{pmatrix} 0 \\ \text{tr}[A\rho A^\dagger] \end{pmatrix}, \quad \tilde{\Gamma}_\Lambda[\rho] = \begin{pmatrix} 0 \\ \text{tr}[A\rho A^\dagger] \end{pmatrix} = \begin{pmatrix} 0 \\ \text{tr}[A\rho A^\dagger] \end{pmatrix}.
\]

(26)

Up to a proper change of the output basis, \( \tilde{\Gamma}_\Lambda \) coincides with \( \Gamma_{\Lambda_{\text{min}}} \). Let us consider three distinctive cases.
(i) Suppose $I - A^\dagger A$ is a rank-1 operator, i.e., $I - A^\dagger A = p|\psi\rangle\langle\psi|$ for some normalized vector $|\psi\rangle \in \mathcal{H}_d$ and $p > 0$. Then we have

$$\sqrt{I - A^\dagger A} \rho \sqrt{I - A^\dagger A} = p|\psi\rangle\langle\psi| \rho|\psi\rangle\langle\psi| = \text{tr}\left[\sqrt{I - A^\dagger A} \rho \sqrt{I - A^\dagger A}\right]|\psi\rangle\langle\psi|$$

and $\Gamma_\Lambda$ is degradable because $\Gamma_\Lambda = \Xi \circ \Gamma_\Lambda$, where the quantum channel $\Xi$ reads $\Xi \left[\left(\begin{array}{c} \rho \\ \cdots \\ c \end{array}\right)\right] = \left(\begin{array}{c} \text{tr}[\rho] \\ 0^T \\ c|\psi\rangle\langle\psi|\right)$. 

(ii) Suppose $I - A^\dagger A$ is not a rank-1 operator. Then $\sqrt{I - A^\dagger A} \rho \sqrt{I - A^\dagger A}$ generally has the rank greater than or equal to 2, so this operator cannot be obtained from a linear map acting on $\sqrt{I - A^\dagger A} \rho \sqrt{I - A^\dagger A}$. Therefore, the operator $\sqrt{I - A^\dagger A} \rho \sqrt{I - A^\dagger A}$ should be obtained from a linear map acting on $A \rho A^\dagger$. Suppose $\det A \neq 0$, then there exists a unique linear map that for all $\rho \in \mathcal{B}(\mathcal{H}_d)$ maps the operator $A \rho A^\dagger$ to the operator $\sqrt{I - A^\dagger A} \rho \sqrt{I - A^\dagger A}$. Therefore, if $\det A \neq 0$, then there exists a unique linear map $\Xi$ such that $\Gamma_\Lambda = \Xi \circ \Gamma_\Lambda$. It reads

$$\Xi \left[\left(\begin{array}{c} \rho \\ \cdots \\ c \end{array}\right)\right] = \left(\begin{array}{c} c + \text{tr}[2I - (AA^\dagger)^{-1}] \\ 0^T \\ \sqrt{I - A^\dagger A} A^{-1}(A^\dagger)^{-1} \sqrt{I - A^\dagger A}\right).$$

It is not hard to see that $\Xi$ is trace preserving; however, $\Xi$ is completely positive if and only if $2I - (AA^\dagger)^{-1} \geq 0$, which is equivalent to $AA^\dagger \geq \frac{1}{2} I$. On the other hand, if an operator $A$ satisfies $AA^\dagger \geq \frac{1}{2} I$ then $\det A \neq 0$ automatically.

(iii) Suppose $I - A^\dagger A$ is not a rank-1 operator and $\det A = 0$. If $A = 0$, then $\Gamma_\Lambda$ is obviously not degradable, so in what follows we additionally assume $\text{supp} \rho \neq \emptyset$. Let $|f\rangle \in \text{ker}A$ and $|g\rangle \in \text{supp}A$, then $\Gamma_\Lambda(|f\rangle\langle g|) = 0$ but $\Gamma_\Lambda(|f\rangle\langle f|) = 0 \neq |f\rangle\langle f| \sqrt{I - A^\dagger A}$, so $\Gamma_\Lambda(|f\rangle\langle f|) = 0$ if and only if $A^\dagger A |g\rangle = |g\rangle$. Therefore, the degradability of $\Gamma_\Lambda$ implies $A^\dagger A |g\rangle = |g\rangle$ for all $|g\rangle \in \text{supp}A$, i.e., $A^\dagger A$ is to be a projector. Since $\det A = 0$, the rank of the projector $A^\dagger A$ is bounded from above by $d - 1$. If rank$A^\dagger A \leq d - 2$, then there exist two orthonormal vectors $|f_1\rangle, |f_2\rangle \in \text{ker}A$ and $\Gamma_\Lambda(|f_1\rangle\langle f_2|) = 0$ whereas $\Gamma_\Lambda(|f_1\rangle\langle f_2|) = 0 \neq |f_1\rangle\langle f_2|$. Therefore, the degradability of $\Gamma_\Lambda$ implies $A^\dagger A$ is a projector of rank $d - 1$. This contradicts the assumption that $I - A^\dagger A$ is not a rank-1 operator. \hfill $\Box$

**Proposition 13.** Suppose the quantum operation $\Lambda \in \mathcal{O}(\mathcal{H}_d)$ has a single Kraus operator $A$, i.e., $\Lambda[\rho] = A \rho A^\dagger$, then $\Gamma_\Lambda$ is antidegradable if and only if $A^\dagger A \leq \frac{1}{2} I$ or $A^\dagger A$ is a rank-1 operator.

Proof. $\Gamma_\Lambda$ and $\widetilde{\Gamma_\Lambda}$ are given by Eq. (26). Antidegradability of $\Gamma_\Lambda$ is equivalent to degradability of $\widetilde{\Gamma_\Lambda}$. The change $B = \sqrt{I - A^\dagger A}$ leads to the relation

$$\widetilde{\Gamma_\Lambda}[\rho] = \left(\text{tr} \left[\sqrt{I - B^\dagger B} \rho \sqrt{I - B^\dagger B}\right] \right)_{B^\dagger B^\dagger}.$$

Therefore, $\Gamma_\Lambda$ is unitarily equivalent to the generalized erasure channel $\Gamma_\Lambda$, where $\Gamma[\rho] = B \rho B^\dagger$. By Proposition 12 $\Gamma_\Lambda$ is degradable if and only if $BB^\dagger \geq \frac{1}{2} I$ or $I - B^\dagger B$ is a rank-1 operator. Substituting $B = \sqrt{I - A^\dagger A}$ into these relations, we get that $\Gamma_\Lambda$ is degradable if and only if $A^\dagger A \leq \frac{1}{2} I$ or $A^\dagger A$ is a rank-1 operator. \hfill $\Box$

**Example 5.** Let $\Lambda \in \mathcal{O}(\mathcal{H}_d)$ be a quantum operation describing polarization dependent losses, Eq. (3). As the Kraus rank of $\Lambda$ equals 1, we apply Propositions 12 and 13 and obtain the following results:

(i) $\Gamma_\Lambda$ is degradable if and only if $\min(\rho_{HH}, \rho_{VV}) \geq \frac{1}{2}$ or $\rho_{H} = 1$ or $\rho_{V} = 1$;

(ii) $\Gamma_\Lambda$ is antidegradable if and only if $\max(\rho_{HH}, \rho_{VV}) \leq \frac{1}{2}$ or $\rho_{H} = 0$ or $\rho_{V} = 0$.

Therefore, $Q(\Gamma_\Lambda) = Q_1(\Gamma_\Lambda)$ if $\min(\rho_{HH}, \rho_{VV}) \geq \frac{1}{2}$ or $\rho_{H} = 1$ or $\rho_{V} = 1$ and, moreover, we exactly know $Q(\Gamma_\Lambda)$ thanks to the result of Example 4. Additionally, we know that $Q(\Gamma_\Lambda) = 0$ if $\max(\rho_{HH}, \rho_{VV}) \leq \frac{1}{2}$ or $\rho_{H} = 0$ or $\rho_{V} = 0$. \hfill $\triangle$

For $\Lambda$ in Eq. (3), Example 4 leaves $Q(\Gamma_\Lambda)$ uncertain in two regions of parameters, where either $\frac{1}{2} < \rho_{H} < 1$ and $0 < \rho_{V} < \frac{1}{2}$ or $0 < \rho_{V} < \frac{1}{2}$ and $\frac{1}{2} < \rho_{V} < 1$. The following result shows that in half of this region the superadditivity of coherent information takes place.

**Proposition 14.** Let $\Lambda \in \mathcal{O}(\mathcal{H}_d)$ describe polarization dependent losses with parameters $\rho_{HH}$ and $\rho_{VV}$. The strict inequality $\frac{1}{2} Q_1(\Gamma_\Lambda) > Q_1(\Gamma_\Lambda)$ holds if either $\frac{1}{2} < \rho_{H} < 1$ and $0 < \rho_{V} < \frac{1}{2}$ or $0 < \rho_{V} < \frac{1}{2}$ and $\frac{1}{2} < \rho_{V} < 1$.

Proof. Let $\rho_2 = \rho_{H} H H H H + 2 \rho_{H} \rho_{VV} V |\varphi_+\rangle\langle\varphi_+| + \rho_{VV} |V V \rangle\langle V V |$, $|\varphi_+\rangle = \frac{1}{\sqrt{2}} (H V - V H)$.

$$\rho_2 = \rho_{H} H H H H + 2 \rho_{H} \rho_{VV} V |\varphi_+\rangle\langle\varphi_+| + \rho_{VV} |V V \rangle\langle V V |,$$
where $|\varphi-\rangle\langle\varphi-|$ is an entangled pure state. In the basis \{\(\{HH\}, |HV\rangle, |VH\rangle, |VV\rangle\)\} the diagonal of $\rho_2$ is exactly the diagonal of the diagonal matrix $\rho_1^{\otimes 2}$. Since both $\Lambda$ and $\Lambda'=\Lambda$ have a single diagonal Kraus operator, the application of maps $\Lambda^{\otimes 2}$, $\Lambda\otimes\Lambda'$, $\Lambda'\otimes\Lambda$, and $(\Lambda'\otimes\Lambda)$ preserves the positions of non-zero elements in the matrices $\rho_2$ and $\rho_1^{\otimes 2}$.

Recalling the definition of the trash-and-prepare channel, $\text{Tr}[\rho] = \text{tr}[\rho|\psi\rangle\langle\psi|]$, we have

$$
\Lambda \otimes (\text{Tr} \circ \Lambda'_{\text{min}}) [\rho_2] = \Lambda \otimes (\text{Tr} \circ \Lambda'_{\text{min}}) [\rho_1^{\otimes 2}],
$$

$$
(\text{Tr} \circ \Lambda'_{\text{min}}) \otimes \Lambda [\rho_2] = (\text{Tr} \circ \Lambda'_{\text{min}}) \otimes \Lambda [\rho_1^{\otimes 2}],
$$

$$
(\text{Tr} \circ \Lambda'_{\text{min}}) \otimes \Lambda [\rho_2] = (\text{Tr} \circ \Lambda'_{\text{min}}) (\otimes 2) [\rho_1^{\otimes 2}].
$$

It follows from Eq. \(19\) that the only difference between $\Gamma_2^{\otimes 2}[\rho_2]$ and $\Gamma_1^{\otimes 2} [\rho_1^{\otimes 2}]$ is in the blocks $\Lambda^{\otimes 2} [\rho_2]$ and $\Lambda^{\otimes 2} [\rho_1^{\otimes 2}]$. Moreover, within these blocks the difference is present only in $2 \times 2$ submatrices, namely, the submatrix $2\rho_1 p_\rho p_v p_H p_H p_V p_V \psi\varphi-\rangle\langle\varphi-\mid$ and the submatrix $\rho_1 p_\rho p_v p_H p_H p_V p_V (|HV\rangle\langle HV| + |VH\rangle\langle VH|)$ for $\Lambda^{\otimes 2} [\rho_2]$ and $\Lambda^{\otimes 2} [\rho_1^{\otimes 2}]$, respectively. Since $\text{Spec}(\{\varphi-\rangle\langle\varphi-\mid\}) = \{2, 0\}$ and $\text{Spec}(\{HV\rangle\langle HV| + |VH\rangle\langle VH|) = \{1, 1\},$ we explicitly relate the entropies as follows:

$$
S(\Gamma_2^{\otimes 2}[\rho_2]) = S(\Gamma_1^{\otimes 2} [\rho_1^{\otimes 2}]) - 2 p_\rho p_v p_H p_H p_V p_V \log 2. \quad (29)
$$

Analogous consideration for the complementary channel yields

$$
S(\Gamma_2^{\otimes 2}[\rho_2]) = S(\Gamma_1^{\otimes 2} [\rho_1^{\otimes 2}]) - 2 (1 - p_\rho) (1 - p_V) p_H p_H p_V p_V \log 2. \quad (30)
$$

Since $S(\Gamma_1^{\otimes 2}[\rho_1]) = 2S(\Gamma_1[\rho_1])$ and $S(\Gamma_2^{\otimes 2}[\rho_1^{\otimes 2}]) = 2S(\Gamma_2[\rho_1])$, we readily obtain the following lower bound for the two-letter quantum capacity:

$$
\frac{1}{2} Q_1(\Gamma_2^{\otimes 2}) \geq \frac{1}{2} \left[ S(\Gamma_2^{\otimes 2}[\rho_2]) - S(\Gamma_2^{\otimes 2}[\rho_2]) \right] = S(\Gamma_2[\rho_1]) - S(\Gamma_2[\rho_1]) + (1 - p_\rho - p_V) p_H p_H p_V p_V \log 2 \geq (1 - p_\rho) (1 - p_V) p_H p_H p_V p_V \log 2.
$$

If $p_H$ and $p_V$ satisfy the requirements in the statement of Proposition \(14\) then $1 - p_\rho - p_V > 0$ and $-1 < z_\ast < 1$, which implies $(1 - p_\rho - p_V) p_H p_H p_V p_V > 0$. 

\[\square\]
In Fig. 5 we depict the derived lower bound \( (1 - p_H - p_V) \rho_{HH} \rho_{VV} \) bits for the difference \( \frac{1}{2} Q_1(\Gamma_{\Lambda}^{\otimes 2}) - Q_1(\Gamma_{\Lambda}) \) in the region of parameters \( \frac{1}{2} < p_H < 1 \) and \( 0 < p_V < \frac{1}{2} \). Numerics show that the actual difference \( \frac{1}{2} Q_1(\Gamma_{\Lambda}^{\otimes 2}) - Q_1(\Gamma_{\Lambda}) \) has a similar shape within the specified region and vanishes (up to a machine precision) if \( p_H + p_V \geq 1 \). The maximum achievable difference \( \frac{1}{2} Q_1(\Gamma_{\Lambda}^{\otimes 2}) - Q_1(\Gamma_{\Lambda}) \) approximately equals \( 7.197 \cdot 10^{-3} \) and is achieved in the vicinity of parameters \( p_H = 0.7 \) and \( p_V = 0.19 \) (or vice versa).

Physical meaning of Eqs. (29) and (30) is that the use of \( \rho_2 \) instead of \( \rho_1^{\otimes 2} \) in the two-letter scenario diminishes both the entropy of the channel output and the entropy of the complementary channel output. However, the decrement in Eq. (29) is less than the decrement in Eq. (30), i.e., less information is dissolved into environment and more information reaches the receiver as compared to the single-letter case. Despite the fact that the losses are asymmetric, i.e., \( p_H \neq p_V \), the contribution \( |\varphi^-\rangle \langle \varphi^-| \) in \( \rho_2 \) preserves its form in the output states \( \Gamma_{\Lambda}^{\otimes 2}[\rho_2] \) and \( \Gamma_{\Lambda}^{\otimes 2}[\rho_1] \) because the same product \( p_H p_V \) characterizes the transmission of both \( HV \) and \( VH \) pairs of photons.

IV. CONCLUSIONS

We reviewed physical properties of trace decreasing quantum operations and clarified a distinction between biased and unbiased quantum operations. We emphasized the importance of biased quantum operations and motivated the introduction of the generalized erasure channel. We identified information capacities of a trace decreasing quantum operation with the corresponding capacities of the generalized erasure channel.

As to general mathematical results, we proved some simple yet fruitful characterizations for extensions of a quantum operation to a channel (Proposition 1) and the normalized image of a trace decreasing operation (Proposition 2). The channel \( \Phi_{\Lambda} \) found in Proposition 2 was subsequently used in finding lower and upper bounds for the single-letter classical and quantum capacities of the generalized erasure channel (Propositions 6 and 11). Bound on the regularized classical and quantum capacities of the generalized erasure channel were expressed through the minimal and maximal detection probabilities (Propositions 4, 5, and 10). We showed that the biasedness of a quantum operation automatically guarantees nonzero classical capacity of the generalized erasure channel (Proposition 5). For quantum operations with Kraus rank 1 we fully characterized necessary and sufficient conditions for degradability and antidegradability of the corresponding generalized erasure channel (Propositions 12 and 13).

As a prominent physical example of a biased quantum operation we considered polarization dependent losses. In addition to the calculation of the single-letter quantum capacity for that physical situation in Example 4 we managed to provide an analytical proof for the superadditivity of coherent information, i.e., a strict separation between the single-letter quantum capacity and the two-letter quantum capacity (Proposition 14). Importantly, the observed difference \( \frac{1}{2} Q_1(\Gamma_{\Lambda}^{\otimes 2}) - Q_1(\Gamma_{\Lambda}) \) was shown to achieve \( 7.197 \cdot 10^{-3} \) bits per qubit sent, which is the maximum reported value for superadditivity of coherent information among qubit-input channels. These results show that the polarization dependent losses may serve as a testbed for exploring other interesting effects, for instance, checking the superadditivity of private information.

Acknowledgments

The author thanks Vikesh Siddhu for useful comments and the anonymous referee for valuable comments to improve the quality of the manuscript. The study was supported by the Russian Science Foundation, project no. 19-11-00086.

[1] A. S. Holevo, The capacity of quantum channel with general signal states, IEEE Trans. Inf. Theory 44, 269 (1998).
[2] B. Schumacher and M. D. Westmoreland, Sending classical information via noisy quantum channels, Phys. Rev. A 56, 131 (1997).
[3] M. Hastings, Superadditivity of communication capacity using entangled inputs, Nature Phys. 5, 255 (2009).
[4] S. Lloyd, Capacity of the noisy quantum channel, Phys. Rev. A 55, 1613 (1997).
[5] H. Barnum, M. A. Nielsen, and B. Schumacher, Information transmission through a noisy quantum channel, Phys. Rev. A 57, 4153 (1998).
[6] I. Devetak, The private classical capacity and quantum capacity of a quantum channel, IEEE Transactions on Information Theory 51, 44 (2005).
[7] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, Gaussian quantum information, Rev. Mod. Phys. 84, 621 (2012).
[8] M. Grassl, T. Beth, and T. Pellizzari, Codes for the quantum erasure channel, Phys. Rev. A 56, 33 (1997).
[9] C. H. Bennett, D. P. DiVincenzo, and J. A. Smolin, Capacities of quantum erasure channels, Phys. Rev. Lett. 78, 3217 (1997).
[10] F. Leditzky, D. Leung, and G. Smith, Dephrasure channel and superadditivity of coherent information, Phys. Rev. Lett. 121, 160501 (2018).
16

[11] V. Siddhu and R. B. Griffiths, Positivity and nonadditivity of quantum capacities using generalized erasure channels, arXiv:2003.00583.

[12] S. Yu, Y. Meng, R. B. Patel, Y.-T. Wang, Z.-J. Ke, W. Liu, Z.-P. Li, Y.-Z. Yang, W.-H. Zhang, J.-S. Tang, C.-F. Li, and G.-C. Guo, Experimental observation of coherent-information superadditivity in a dephrasure channel, Phys. Rev. Lett. 125, 060502 (2020).

[13] N. Gisin and B. Huttner, Combined effects of polarization mode dispersion and polarization dependent losses in optical fibers, Optics Communications 142, 119 (1997).

[14] B. T. Kirby, D. E. Jones, and M. Brodsky, Effect of polarization dependent loss on the quality of transmitted polarization entanglement, Journal of Lightwave Technology 37, 95 (2019).

[15] C. Li, M. Curty, F. Xu, O. Bedroya, and H.-K. Lo, Secure quantum communication in the presence of phase- and polarization-dependent loss, Phys. Rev. A 98, 042324 (2018).

[16] T. Heinosaari, M. A. Jivulescu, D. Reeb, and M. M. Wolf, Extending quantum operations, J. Math. Phys. 53, 102208 (2012).

[17] E. B. Davies and J. T. Lewis, An operational approach to quantum probability, Commun. Math. Phys. 17, 239 (1970).

[18] C. Carmeli, T. Heinosaari, and A. Toigo, Sequential measurements of conjugate observables, J. Phys. A: Math. Theor. 44, 285304 (2011).

[19] I. A. Luchnikov and S. N. Filippov, Quantum evolution in the stroboscopic limit of repeated measurements, Phys. Rev. A 95, 022113 (2017).

[20] V. A. Zhuravlev, S. N. Filippov, Quantum state tomography via sequential uses of the same informationally incomplete measuring apparatus, Lobachevskii J. Math. 41, 2005 (2020).

[21] L. Leppäjärvi and M. Sedlák, Post-processing of quantum instruments, arXiv:2010.15816 [quant-ph].

[22] A. S. Holevo, Quantum Systems, Channels, Information. A Mathematical Introduction (de Gruyter, Berlin, Boston, 2012).

[23] M. M. Wilde, Quantum Information Theory (Cambridge University Press, Cambridge, 2013).

[24] T. Heinosaari and M. Ziman, The Mathematical Language of Quantum Theory (Cambridge University Press, Cambridge, 2012).

[25] Yu. I. Bogdanov, E. V. Moreva, G. A. Maslennikov, R. F. Galeev, S. S. Straupe, and S. P. Kulik, Polarization states of four-dimensional systems based on biphotons, Phys. Rev. A 73, 063810 (2006).

[26] S. N. Filippov, A. N. Glinov, and L. Leppäjärvi, Phase covariant qubit dynamics and divisibility, Lobachevskii J. Math. 41, 617 (2020).

[27] I. Bengtsson and K. Życzkowski, Geometry of Quantum States. An Introduction to Quantum Entanglement (Cambridge University Press, New York, 2006).

[28] F. Leditzky, E. Kaur, N. Datta, and M. M. Wilde, Approaches for approximate additivity of the Holevo information of quantum channels, Phys. Rev. A 97, 012332 (2018).

[29] S. N. Filippov, Lower and upper bounds on nonunital qubit channel capacities, Reports on Mathematical Physics 82, 100 (2018).

[30] S. N. Filippov and K. V. Kuzhamuratova, Quantum informational properties of the Landau-Streater channel, J. Math. Phys. 60, 042202 (2019).

[31] G. G. Amosov and S. Mancini, The decreasing property of relative entropy and the strong superadditivity of quantum channels, Quantum Information and Computation 7, 594 (2009).

[32] V. Siddhu, Log-singularities for studying capacities of quantum channels, arXiv:2003.10367.

[33] I. Devetak and P. Shor, The capacity of a quantum channel for simultaneous transmission of classical and quantum information, Commun. Math. Phys. 256, 287 (2005).

[34] T. S. Cubitt, M. B. Ruskai, and G. Smith, The structure of degradable quantum channels, J. Math. Phys. 49, 102104 (2008).

[35] D. P. DiVincenzo, P. W. Shor, and J. A. Smolin, Quantum-channel capacity of very noisy channels, Phys. Rev. A 57, 830 (1998).

[36] J. Fern and K. B. Whaley, Lower bounds on the nonzero capacity of Pauli channels, Phys. Rev. A 78, 062335 (2008).

[37] V. Siddhu, Leaking information to gain entanglement, arXiv:2011.15116.

[38] T. Cubitt, D. Elkouss, W. Matthews, M. Ozols, D. Pérez-García, and S. Strelchuk, Unbounded number of channel uses may be required to detect quantum capacity, Nature Commun. 6, 6739 (2015).

[39] S. Holevo, Complementary channels and the additivity problem, Theory Probab. Appl. 51, 92 (2007).