Multiletter codes to boost superadditivity of coherent information in quantum communication lines with polarization dependent losses

Sergey N. Filippov\textsuperscript{1,2,3}

\textsuperscript{1}Steklov Mathematical Institute of Russian Academy of Sciences, Gubkina St. 8, Moscow 119991, Russia
\textsuperscript{2}Valiev Institute of Physics and Technology of Russian Academy of Sciences, Nakhimovskii Pr. 34, Moscow 117218, Russia
\textsuperscript{3}Moscow Institute of Physics and Technology, Institutski Per. 9, Dolgoprudny, Moscow Region 141700, Russia

Coherent information quantifies the achievable rate of the reliable quantum information transmission through a communication channel. Use of the correlated quantum states (multiletter codes) instead of the factorized ones (single-letter codes) may result in an increase in the achievable rate, a phenomenon known as the coherent-information superadditivity. However, even for simple physical models of channels it is rather difficult to detect the superadditivity and find the advantageous multiletter codes. Here we consider the case of polarization dependent losses and propose some physically motivated multiletter codes which outperform all single-letter ones in a wide range of the channel parameters. We show that in the asymptotic limit of the infinite code length the superadditivity phenomenon takes place whenever the communication channel is neither degradable nor antidegradable. Besides the superadditivity identification, we also provide a method how to modify the proposed codes and get a higher quantum communication rate by doubling the code length. The obtained results give a deeper understanding of useful multiletter codes and may serve as a benchmark for quantum capacity estimations and future approaches toward an optimal strategy to transfer quantum information.

I. INTRODUCTION

Quantum information represents quantum states in a variety of forms including superpositions and entanglement. Quantum information significantly differs from classical information because quantum states cannot be deterministically cloned in contrast to classical letters. On the other hand, it is quantum information that should be transferred along physical communication lines to connect quantum computers in a network and manipulate a long-distance entanglement, which potentially has numerous applications \cite{1,2}. A successful transmission of quantum information through a noisy channel implies a perfect transfer (in terms of the fidelity) of any quantum state by arranging appropriate encoding and decoding procedures at the input and the output of the channel, respectively, see Refs. \cite{3,4}. Physical meaning of quantum information transfer is also discussed in Ref. \cite{5} from the viewpoint of creating entanglement between the apart laboratories, provided the channel can be used many times. An \textit{n}-letter encoding scheme implies the use of \textit{n} quantum information carriers, e.g., photons, as a whole. Let $\varrho^{(n)}$ be the average density operator of an ensemble of \textit{n}-partite states used in the quantum communication task. The eigenvectors and eigenvalues of $\varrho^{(n)}$ can be treated as quantum multiletters and their probabilities, respectively, to form typical quantum codewords \cite{3}. In this paper, we report such \textit{n}-letter codes that enable to transmit an increasing amount of quantum information with the increase of \textit{n}.

If each of \textit{n} information carriers propagates through a memoryless noisy quantum channel $\Phi$, then the overall noisy output is $\Phi^{\otimes n}[\varrho^{(n)}]$. The decoder aims at reproducing the encoded state. A figure of merit for this task is the achievable communication rate that quantifies how many qubits per channel use can be reliably transmitted in the sense that the error vanishes in the asymptotic limit of infinitely many channel uses. The quantum capacity $Q(\Phi)$ is defined as the supremum of achievable communication rates among all possible encodings and decodings. The result of the seminal paper \cite{5} generalizes some previous observations \cite{3,4} and shows that

$$Q(\Phi) = \lim_{n \to \infty} Q_n(\Phi),$$

where

$$Q_n(\Phi) = \frac{1}{n} Q_1(\Phi^{\otimes n}),$$

$$Q_1(\Psi) = \sup_{\varrho} I_c(\varrho, \Psi),$$

$$I_c(\varrho, \Psi) = S(\Psi[\varrho]) - S(\Psi[\varrho|\Psi]).$$

$I_c(\varrho, \Psi)$ is a so-called coherent information that quantifies an asymmetry between the von Neumann entropy $S(\Psi[\varrho])$ of the channel output and the von Neumann entropy $S(\Psi[\varrho|\Psi])$ of a complementary channel output. In other words, the coherent information effectively quantifies an asymmetry between the receiver information $S(\Psi[\varrho])$ and the information $S(\Psi[\varrho|\Psi])$ diluted into the environment. To make this description precise, consider a quantum channel $\Psi : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$, where $\mathcal{H}_A$ and $\mathcal{H}_B$ are the Hilbert spaces of input and output, respectively, and $\mathcal{B}(\mathcal{H})$ denotes a set of bounded operators on $\mathcal{H}$. Hereafter, we consider finite-dimensional Hilbert spaces because we will further focus on a finite-dimensional physical model of polarization dependent losses. The Stinespring dilation for $\Psi : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ reads as follows in the Schrödinger picture:

$$\Psi[\varrho] = \text{tr}_E[V \varrho V^\dagger],$$

where $V : \mathcal{H}_A \to \mathcal{H}_B \otimes \mathcal{H}_E$ is an isometry ($V^\dagger V = I_A$), $\mathcal{H}_E$ denotes the Hilbert space of the effective environment, and $\text{tr}_E$ is the partial trace with respect to the effective environment (see, e.g., \cite{7}). The formula

$$\tilde{\Psi}[\varrho] = \text{tr}_B[V \varrho V^\dagger]$$

defines a channel $\tilde{\Psi} : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_E)$ that is complementary to $\Psi : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$. Since the Stinespring...
dilation [1] is not unique for a given channel $\Psi$, neither is the complementary channel $\Psi$; however, all complementary channels are isometrically equivalent (see, e.g., [7]).

Suppose two quantum channels $\Phi : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ and $\Phi' : \mathcal{B}(\mathcal{H}_{A'}) \to \mathcal{B}(\mathcal{H}_{B'})$ are both degradable, i.e., there exist quantum channels $D$ and $D'$ such that $\Phi = D \circ \Phi$ and $\Phi' = D' \circ \Phi'$; the symbol $\circ$ denotes a concatenation of maps. Then the coherent information is subadditive [5] in the sense that

$$I_c(\rho_{AA'}, \Phi \otimes \Phi') \leq I_c(\rho_A, \Phi) + I_c(\rho_{A'}, \Phi').$$

(2)

An immediate consequence of Eq. (2) is the additivity of the single-letter capacity, $Q_1(\Phi \otimes \Phi') = Q_1(\Phi) + Q_1(\Phi')$. If $\Phi' = \Phi^{\otimes (n-1)}$, then we get $Q_1(\Phi^{\otimes n}) = nQ_1(\Phi)$ by mathematical induction. Hence, if the channel $\Phi$ is degradable, then the quantum capacity $Q(\Phi)$ coincides with the single-letter quantum capacity $Q_1(\Phi)$. Subadditivity of coherent information for degradable channels significantly simplifies calculations of the quantum capacity and shows that some single-letter encodings have the same efficacy as the best multiletter ones.

If the channel $\Phi$ is antidegradable, i.e., there exists a quantum channel $A$ such that $\Phi = A \circ \Phi$, then $I_c(\rho, \Phi)$ is nonpositive and vanishes for pure states $\rho = \langle \psi | \psi \rangle$. Similarly, $I_c(\rho, \Phi^{\otimes n}) \leq 0$. This implies the trivial equality $Q(\Phi) = Q_1(\Phi) = 0$, i.e., all single-letter encodings and all multiletter encodings are equally useless for quantum information transmission.

If $\Phi : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ is neither degradable nor antidegradable, then it may happen that there exists an $n$-partite quantum state $\rho^{\otimes n} = \rho_{A_1 \ldots A_n}$ such that

$$I_c(\rho_{A_1 \ldots A_n}, \Phi^{\otimes n}) > \sum_k I_c(\rho_{A_k}, \Phi)$$

and $Q_1(\Phi) > Q_1(\Phi')$. This case corresponds to superadditivity of coherent information, which implies that some multiletter encodings outperform all single-letter ones so that the achievable quantum communication rate increases with the use of correlated states. The superadditivity phenomenon is predicted for qubit depolarizing channels if $n \geq 3$ [9, 10], so-called dephrasure qubit channels if $n \geq 2$ [11] (for which superadditivity was also analyzed experimentally [12]), a concatenation of an erasure qubit channel with an amplitude damping qubit channel [13], some qutrit channels and their higher-dimensional generalizations [14], and a collection of specific channels if $n \geq n_0$, where $n_0 \geq 2$ can be arbitrary [15]. In this paper, we focus on quantum communication lines with polarization dependent losses [16][20], which also exhibit the coherent information superadditivity for some values of attenuation factors [21].

Consider a lossy quantum communication line such that the transmission coefficient for horizontally polarized photons, $p_H$, differs from that for vertically polarized photons, $p_V$. The simplest example is a horizontally oriented linear polarizer for which $p_H = 1$ and $p_V = 0$. In practice, however, all values $0 \leq p_H \leq 1$ and $0 \leq p_V \leq 1$ are attainable (see, e.g., [22]), which leads to a two-parameter family of qubit-to-qutrit channels

$$\Gamma = \left( \begin{array}{ccc} \rho_{HH} & \sqrt{\rho_{HV}} & 0 \\ \sqrt{\rho_{VH}} & \rho_{VV} & 0 \\ 0 & 0 & (1-p_H)\rho_{HH} + (1-p_V)\rho_{VV} \end{array} \right),$$

(3)

with $p_H$ and $p_V$ being the parameters. The extra (third) dimension in Eq. (3) corresponds to the vacuum contribution $|\text{vac}\rangle$ that leads to no detector clicks. If $p_H = p_V$, then we get the standard erasure channel [23, 24]. If $p_H \neq p_V$, then Eq. (3) defines a generalized erasure channel [21] (cf. a similar but different concept in Ref. [13]) induced by the trace decreasing operation $\rho \to \Lambda_F[\rho] := F\rho F^\dagger$, where

$$F = \sqrt{p_H}|H\rangle\langle H| + \sqrt{p_V}|V\rangle\langle V|,$$

$|H\rangle$ and $|V\rangle$ are the single-photon states with horizontal and vertical polarization, respectively. The brief version of Eq. (3) is

$$\Gamma[\rho] = F\rho F^\dagger \oplus \text{tr}(I - F F^\dagger)\rho |\text{vac}\rangle\langle \text{vac}|.$$

The term $\text{tr}(I - F F^\dagger)\rho$ is the state dependent erasure probability. Denoting $G = \sqrt{1 - FF^\dagger}$ and recalling the notation $\Lambda_G[\rho] := G\rho G^\dagger$, the channel [4] takes the form

$$\Gamma = \Lambda_F \oplus (\text{Tr} \circ \Lambda_G),$$

(4)

where $\text{Tr}$ denotes the trash-and-prepare map $\rho \to \text{tr}[\rho]|\text{vac}\rangle\langle \text{vac}|$. Interestingly, a complementary channel

![FIG. 1: Attenuation factors $p_H$ and $p_V$ for horizontally and vertically polarized photons for which the quantum channel (3) is degradable (red area), antidegradable (green area), both degradable and antidegradable (black points). Yellow regions correspond to the coherent-information superadditivity detected with the use of two-letter encodings [21].](Image)
can be expressed as 21
\[ \Gamma = \Lambda_G \oplus (\text{Tr} \circ \Lambda_F), \]
which is equivalent to the change \( p_H \to 1 - p_H \) and \( p_V \to 1 - p_V \) in Eq. (3).

The fact that \( \Gamma \) and \( \tilde{\Gamma} \) have the same structure was used in Ref. 21 to prove that \( Q(\Gamma) = 0 \) if and only if \( \max(p_H, p_V) \leq \frac{1}{2} \) or \( p_H = 0 \) or \( p_V = 0 \), see the green region in Fig. 1. It was also shown in Ref. 21 that \( Q(\Gamma) > 0 \) beyond the antidegradability region, \( \Gamma \) is degradable so that \( Q(\Gamma) = Q_1(\Gamma) \) if and only if \( \min(p_H, p_V) \geq \frac{1}{2} \) or \( p_H = 1 \) or \( p_V = 1 \), see the red region in Fig. 1. The final result of Ref. 21 is the analytical proof of superadditivity region. \( \Gamma \) can be expressed as 
\[ \Gamma \] 

II. SUPERADDITIVITY IDENTIFICATION

Technically, it is quite difficult to maximize the coherent information \( I_c(\hat{\rho}^{(n)}), \Gamma^{(n)} \) with respect to  \( n \)-qubit density operators \( \hat{\rho}^{(n)} \) even if \( p_H, p_V \), and \( n \) are all fixed. In the case of the single-letter encoding \( (n = 1) \), the optimal state \( \hat{\rho}_{\text{opt}}^{(1)} \) is shown to be diagonal in the basis \( \{|H\},\{|V\} \) for all \( p_H \) and \( p_V \), i.e., 
\[ \hat{\rho}_{\text{opt}}^{(1)} = \rho_{HH}|H\rangle\langle H| + \rho_{VV}|V\rangle\langle V|; \]
however, a closed-form expression for the coefficients \( \rho_{HH} \) and \( \rho_{VV} \) is still missing so they appear as a solution of some equation that can be readily solved numerically 211. If \( \Gamma \) is not antidegradable, then both \( \rho_{HH} > 0 \) and \( \rho_{VV} > 0 \) so that \( Q_1(\Gamma) = I_c(\hat{\rho}_{\text{opt}}^{(1)}, \Gamma) > 0 \). Therefore, the optimal single-letter codes to transmit quantum information exploit no superpositions of horizontally and vertically polarized photons. If the degradability property holds for \( \Gamma \) (see the red region in Fig. 1), there is no benefit in using multiletter codes either. If \( \Gamma \) is neither degradable nor antidegradable, then there is a potential for improvement. In Section II A we review in detail an approach of Ref. 21 to find the two-letter encodings outperforming the single-letter ones for some parameters \( p_H \) and \( p_V \). In Section II B we generalize that approach to \( n \)-qubit encodings.

A. Two-letter encoding

Suppose \( n = 2 \). Consider the state 
\[ \hat{\rho}^{(2)} = (\hat{\rho}_{\text{opt}}^{(1)})^{\otimes 2} + p_H p_V \rho_{HH} \rho_{VV}(|HV\rangle\langle HV| + |VH\rangle\langle VH|) \]
\[ = 2p_H p_V \rho_{HH} \rho_{VV} \frac{|HV\rangle\langle HV| + |VH\rangle\langle VH|}{\sqrt{2^2}} \]
\[ + 2p_H p_V \rho_{HH} \rho_{VV} \frac{|HV\rangle\langle HV| + |VH\rangle\langle VH|}{\sqrt{2^2}}. \] 

Clearly, the diagonals of density matrices \( \hat{\rho}^{(2)} \) and \( (\hat{\rho}_{\text{opt}}^{(1)})^{\otimes 2} \) coincide in the standard basis \( (|HH\rangle, |HV\rangle, |VH\rangle, |VV\rangle) \). The two photon states \( |HV\rangle \) and \( |VH\rangle \) experience the same attenuation even if \( p_H \neq p_V \) due to the obvious symmetry. In fact, all vectors from the subspace \( \mathcal{H}_{1,1} := \text{Span}(|HV\rangle, |VH\rangle) \) are equally attenuated, which makes it easy to calculate the output state
\[ \Gamma^{(2)} [\hat{\rho}^{(2)}] \]
\[ = (\Gamma [\hat{\rho}_{\text{opt}}^{(1)}])^{\otimes 2} + p_H p_V \rho_{HH} \rho_{VV} (|HV\rangle\langle HV| + |VH\rangle\langle VH|) \]
\[ = 2p_H p_V \rho_{HH} \rho_{VV} \frac{|HV\rangle\langle HV| + |VH\rangle\langle VH|}{\sqrt{2^2}} \]
\[ + (1 - p_H) \rho_{HH} (1 - p_v) \rho_{VV} \]
\[ \times \left( \hat{\rho}_{\text{opt}}^{(1)} \otimes |vac\rangle\langle vac| + |vac\rangle\langle vac| \otimes \hat{\rho}_{\text{opt}}^{(1)} \right) \]
\[ + (1 - p_H) \rho_{HH} (1 - p_v) \rho_{VV} [|vac\rangle\langle vac| \otimes |vac\rangle\langle vac|]. \]

The density operators \( \Gamma^{(2)} [\hat{\rho}^{(2)}] \) and \( (\Gamma [\hat{\rho}_{\text{opt}}^{(1)}])^{\otimes 2} \) differ by their action in the subspace \( \mathcal{H}_{1,1} \), namely, \( \Gamma^{(2)} [\hat{\rho}^{(2)}] \) acts as a coherent operator
\[ 2p_H p_V \rho_{HH} \rho_{VV} \frac{|HV\rangle\langle HV| + |VH\rangle\langle VH|}{\sqrt{2^2}} \]
\[ \frac{|HV\rangle\langle HV| + |VH\rangle\langle VH|}{\sqrt{2^2}}, \]
whereas \( (\Gamma [\hat{\rho}_{\text{opt}}^{(1)}])^{\otimes 2} \) acts as an incoherent operator
\[ p_H p_V \rho_{HH} \rho_{VV} (|HV\rangle\langle HV| + |VH\rangle\langle VH|). \]

This leads to a readily accountable difference in spectra of the two states. Spectrum of (3) is \( (2p_H p_V \rho_{HH} \rho_{VV},0) \) and that of (7) is \( (p_H p_V \rho_{HH} \rho_{VV}, p_H p_V \rho_{HH} \rho_{VV}) \). We have
\[ S(\Gamma^{(2)} [\hat{\rho}^{(2)}]) = S \left( (\Gamma [\hat{\rho}_{\text{opt}}^{(1)}])^{\otimes 2} \right) \]
\[ - (2 \log 2) p_H p_V \rho_{HH} \rho_{VV}. \]

As the complementary channel \( \tilde{\Gamma} \) is obtained from the direct channel \( \Gamma \) by the change \( p_H \to 1 - p_H \) and \( p_V \to 1 - p_V \), we readily have
\[ S(\Gamma^{(2)} [\hat{\rho}^{(2)}]) = S \left( (\tilde{\Gamma} [\hat{\rho}_{\text{opt}}^{(1)}])^{\otimes 2} \right) \]
\[ - (2 \log 2) (1 - p_H) (1 - p_V) \rho_{HH} \rho_{VV}. \]
Finally, we get
\[
I_c(\varrho^{(2)}, \Gamma^{\otimes 2}) = S(\Gamma^{\otimes 2}[\varrho^{(2)}]) - S(\Gamma^{\otimes 2}[\varrho^{(2)}]) \\
= S\left(\left(\Gamma^{(1)}_{\text{opt}}\right)^{\otimes 2}\right) - S\left(\left(\Gamma^{(1)}_{\text{opt}}\right)^{\otimes 2}\right) \\
+ 2I_c(\varrho^{(1)}, \Gamma) + (2\log 2)(1 - p_H - p_V) \varrho_{HH} \varrho_{VV} \\
= 2Q_2(\Gamma) + (2\log 2)(1 - p_H - p_V) \varrho_{HH} \varrho_{VV}. \tag{8}
\]

The coherent information is superadditive if \((1 - p_H - p_V) \varrho_{HH} \varrho_{VV} > 0\), i.e., if \(p_H + p_V < 1\) and the state \(\varrho_{\text{opt}}^{(1)}\) is nondegenerate. Combining these conditions we get two yellow regions in Fig. 1, where
\[
Q_2(\Gamma) \geq \frac{1}{2} I_c(\varrho^{(2)}, \Gamma^{\otimes 2}) > Q_1(\Gamma).
\]

B. Multiletter encoding

Suppose \(n > 2\). A generalization of the approach in Section 2A would be to consider a state \((\varrho_{\text{opt}}^{(1)})^{\otimes n}\) and modify it to a state \(\varrho^{(n)}\), which would differ from \((\varrho_{\text{opt}}^{(1)})^{\otimes n}\) when acting on some subspace that is symmetric with respect to permutations of photons. Physically, the subspace is to be chosen in such a way as to ensure a high enough detection probability for all states from the subspace. Suppose \(p_H > p_V\), then the state \(|H\rangle^{\otimes n}\) has the highest detection probability, but the corresponding subspace \(\mathcal{H}_{n,0} := \text{Span}(|H\rangle^{\otimes n})\) is trivial (has dimension 1). So we consider the subspace \(\mathcal{H}_{n,1}\) spanned by the vector \(|H\rangle^{\otimes (n-1)} \otimes |V\rangle\) and all its photon-permutated versions. The detection probability for all states from this subspace equals \(p_H^{n-1} p_V\). The following \(n\)-qubit generalization of the \(W\)-state belongs to \(\mathcal{H}_{n,1,1}\):

\[
|W^{(n)}\rangle = \frac{1}{\sqrt{n}} \left( |HH\ldots HHV\rangle_{n-1} + |HH\ldots HVV\rangle_{n-2} + \ldots + |VH\ldots HHH\rangle_{n-1} \right) \in \mathcal{H}_{n,1,1}. \tag{9}
\]

Consider the \(n\)-qubit density operator \(\varrho^{(n)}\) defined through
\[
\varrho^{(n)}\langle \varphi | = \begin{cases} 
(\varrho_{\text{opt}}^{(1)})^{\otimes n}|\varphi\rangle & \text{if } |\varphi\rangle \perp \mathcal{H}_{n,1,1}, \\
n\varrho_{HH}^{n-1} \varrho_{VV}^{n-1} |W^{(n)}\rangle \langle W^{(n)}| & \text{if } |\varphi\rangle \in \mathcal{H}_{n,1,1}.
\end{cases}
\]

The restriction of \(\varrho^{(n)}\) to the subspace \(\mathcal{H}_{n,1,1}\) is a coherent (rank-1) operator
\[
[\varrho^{(n)}]_{\mathcal{H}_{n,1,1}} := n\varrho_{HH}^{n-1} \varrho_{VV}^{n-1} |W^{(n)}\rangle \langle W^{(n)}|, \tag{10}
\]
whereas the restriction of \((\varrho_{\text{opt}}^{(1)})^{\otimes n}\) to the subspace \(\mathcal{H}_{n,1,1}\) is a mixed (rank-\(n\)) operator
\[
[(\varrho_{\text{opt}}^{(1)})^{\otimes n}]_{\mathcal{H}_{n-1,1}} \\
:= \varrho_{HH}^{n-1} \varrho_{VV}^{n-1} \left( |HH\ldots HHV\rangle_{n-1} \langle HH\ldots HHV|_{n-1} + |HH\ldots HVV\rangle_{n-2} \langle HH\ldots HVV|_{n-2} + \ldots + |VH\ldots HHH\rangle_{n-1} \langle VH\ldots HHH|_{n-1} \right), \tag{11}
\]
but beyond that restriction
\[
[\varrho^{(n)}]_{\mathcal{H}_{n,1,1}} - [\varrho^{(n)}]_{\mathcal{H}_{n-1,1}} = (\varrho_{\text{opt}}^{(1)})^{\otimes n} - [(\varrho_{\text{opt}}^{(1)})^{\otimes n}]_{\mathcal{H}_{n-1,1}}.
\]

Using the direct sum representation of the channel \(\Gamma^{\otimes n}\), we explicitly find its tensor power
\[
\Gamma^{\otimes n} = \varLambda_F^{\otimes n} \oplus \ldots \oplus \varLambda_F^{\otimes (n-k)} \otimes (\text{Tr} \circ \varLambda_G)_{\otimes k} \oplus \ldots, \tag{12}
\]
where the brace denotes a direct sum of \(\binom{n}{k}\) different terms, with each term being a permuted tensor product of \(n - k\) maps \(\varLambda_F\) and \(k\) maps \(\varLambda_G\). Let us consider how the term \(\varLambda_F^{\otimes (n-k)} \otimes (\text{Tr} \circ \varLambda_G)_{\otimes k}\) affects the operators \([\varrho^{(n)}]_{\mathcal{H}_{n,1,1}}\) and \([[(\varrho_{\text{opt}}^{(1)})^{\otimes n}]_{\mathcal{H}_{n-1,1}}\). Recalling the effect of the partial trace on \(W\)-states, we see that the coherent component of \(\varLambda_F^{\otimes (n-k)} \otimes (\text{Tr} \circ \varLambda_G)_{\otimes k}\) \([[\varrho^{(n)}]_{\mathcal{H}_{n-1,1}}\]
reads
\[
\varrho_{HH}^{n-1} \varrho_{VV}^{n-1} p_H^{n-k-1} p_V (1 - p_H)^k \\
\times \langle n - k | W_{n-k} \rangle \langle W_{n-k} | \otimes (|\text{vac}\rangle \langle \text{vac}|)^{\otimes k}, \tag{13}
\]
whereas \(\varLambda_A^{\otimes (n-k)} \otimes (\text{Tr} \circ \varLambda_G)_{\otimes k}\) \([[(\varrho_{\text{opt}}^{(1)})^{\otimes n}]_{\mathcal{H}_{n-1,1}}\]
has the completely incoherent component
\[
\varrho_{HH}^{n-1} \varrho_{VV}^{n-1} p_H^{n-k-1} p_V (1 - p_H)^k \\
\times \langle HH\ldots HHV\rangle_{n-k-1} \langle HH\ldots HHV|_{n-k-2} + \ldots + |VH\ldots HHH\rangle_{n-k-1} \langle VH\ldots HHH|_{n-k-2} \rangle_{n-k-1} \langle VH\ldots HHH|_{n-k-1} \langle VH\ldots HHH|_{n-k-1} \otimes (|\text{vac}\rangle \langle \text{vac}|)^{\otimes k}. \tag{14}
\]

The operator \([13]\) has the only nonzero eigenvalue, whereas the operator \([14]\) has \(n - k\) coincident nonzero eigenvalues, with traces of the two operators being the same. Therefore, the only nonzero eigenvalue of the operator \([13]\) is \((n - k)\) multiplied by any nonzero eigenvalue of the operator \([14]\). This leads to a simple expression for the difference in entropies, namely,
\[
S\left(\varLambda_F^{\otimes (n-k)} \otimes (\text{Tr} \circ \varLambda_G)_{\otimes k}[\varrho^{(n)}]\right) \\
= S\left(\varLambda_F^{\otimes (n-k)} \otimes (\text{Tr} \circ \varLambda_G)_{\otimes k}[[(\varrho_{\text{opt}}^{(1)})^{\otimes n}]_{\mathcal{H}_{n-1,1}}\right) \\
- \varrho_{HH}^{n-1} \varrho_{VV}^{n-1} p_H^{n-k-1} (1 - p_H)^k (n - k) \log(n - k). \tag{15}
\]
Since the operators $|\psi\rangle$ and $|\phi\rangle$ are invariant with respect to permutations of photons, each term in the brace in Eq. (12) results in the same entropy decrement as in Eq. (15). Summing all the decrements, we get

$$S(\Gamma \otimes n | \rho^{(n)}) = S(\rho^{(n)} | \otimes n) - \frac{n-1}{2} e_{HH}^V$$

$$\sum_{k=0}^{n-1} \binom{n}{k} p_H p_{V}^{n-k-1} (1 - p_H) k (n - k) \log(n - k).$$

Similarly, for the complementary channel we have

$$S(\tilde{\Gamma} \otimes n | \rho^{(n)}) = S(\rho^{(n)} | \otimes n) - \frac{n-1}{2} e_{HH}^V$$

$$\sum_{k=0}^{n-1} \binom{n}{k} (1 - p_V)(1 - p_H) k (n - k) (n - k) \log(n - k).$$

Finally, we get

$$Q_n(\Gamma) - Q_1(\Gamma) \geq \frac{1}{n} \left[ I_r(\rho^{(n)}, \Gamma^{\otimes n}) - I_r(\rho^{(1)}_{opt}, \Gamma^{\otimes n}) \right]$$

$$= \frac{1}{n} e_{HH}^V$$

$$\sum_{k=0}^{n-1} \binom{n}{k} (1 - p_V)(1 - p_H) k (n - k) \log(n - k)$$

$$\sum_{k=0}^{n-1} \binom{n-1}{k} (n - k) \log(n - k)$$

$$\sum_{k=0}^{n-1} \binom{n-1}{k} (1 - p_H) k (n - k) \log(k + 1).$$

If the obtained expression (16) is positive, then we successfully identify the coherent-information superadditivity in the form $Q_n(\Gamma) > Q_1(\Gamma)$. Suppose $\Gamma$ is not antidegradable, then $e_{HH}^V > 0$, $e_{HH}^V > 0$, and $Q_n(\Gamma) > Q_1(\Gamma)$ if the sum in Eq. (16) is positive.

In the above analysis, we assumed $p_H > p_V$. The converse case $p_V > p_H$ obviously reduces to the considered one if we replace $|H\rangle \leftrightarrow |V\rangle$ in Eq. (9). Therefore, we make the following conclusion: $Q_n(\Gamma) > Q_1(\Gamma)$ if $\Gamma$ is not antidegradable and $w_n(p_H, p_V) > 0$, where

$$w_n(p_H, p_V) := \sum_{k=0}^{n-1} \binom{n-1}{k} (1 - p_H) k (n - k) \log(k + 1)$$

if $p_H > p_V$, then

$$w_n(p_H, p_V) := \sum_{k=0}^{n-1} \binom{n-1}{k} (1 - p_V) k (n - k) \log(k + 1)$$

if $p_V > p_H$.

In the case $n = 2$, the condition $w_2(p_H, p_V) > 0$ is equivalent to $p_H + p_V < 1$, i.e., we reproduce the results of Section II A. If $n \geq 3$, then the region of parameters $p_H$ and $p_V$, where $Q_n(\Gamma) > Q_1(\Gamma)$, is strictly larger than the region, where $Q_2(\Gamma) > Q_1(\Gamma)$, see Fig. 2. Interestingly, the greater the region, where $Q_n(\Gamma) > Q_1(\Gamma)$.

FIG. 2: Superadditivity regions $Q_n(\Gamma) > Q_1(\Gamma)$ for $n = 2$ (dotted line), $n = 3, 10, 10^2, 10^3, 10^4$ (dashed lines from left to right).

The binomial distribution $\left\{ \binom{n-1}{k} (1 - p)^n - k - 1 \right\}_{k=0}^{n-1}$ tends to the normal distribution $N(n p, n p (1 - p))$ with the mean value $n p$ and the standard deviation $\sqrt{n p (1 - p)}$ when $0 < p < 1$ and $n$ tends to infinity. Therefore, the terms with $k \approx n p$ contribute the most to Eq. (17), and the terms with $k \approx n (1 - p)$ contribute the most to Eq. (18). In the asymptotic limit $n \rightarrow \infty$ we have

$$w_n(p_H, p_V) \approx (1 - 2 p_V) n + (1 - p_V) \log(1 - p_H)$$

$$- p_V \log p_H$$

if $p_H > p_V$,

$$w_n(p_H, p_V) \approx (1 - 2 p_H) n + (1 - p_H) \log(1 - p_V)$$

$$- p_H \log p_V$$

if $p_V > p_H$.

Hence, $w_n(p_H, p_V) > 0$ in the asymptotic limit $n \rightarrow \infty$ if $0 < p_V < \frac{1}{2}$, $p_H < 1$, and $0 < p_H < \frac{1}{2}$, which is exactly the region, where $\Gamma$ is neither degradable nor antidegradable (see Fig. 2).

Suppose the parameters $p_H$ and $p_V$ are fixed. Exploiting the asymptotic formulas (19)–(20) and solving the inequality $w_n(p_H, p_V) \geq 0$, we estimate the multiletter length $n$ needed to observe the superadditivity phenomenon $Q_n > Q_1$:

$$n \geq n_0 := \left\{ \begin{array}{ll}
\left( \frac{p_H}{(1 - p_H)^2 - p_V} \right)^{\frac{1}{1 - 2 p_V}} & \text{if } 0 < p_V < \frac{1}{2} < p_H < 1, \\
\left( \frac{p_H}{(1 - p_H)^2 + p_V} \right)^{\frac{1}{1 - 2 p_H}} & \text{if } 0 < p_H < \frac{1}{2} < p_V < 1.
\end{array} \right.$$

If $n > n_0$, then the proposed code yields the following
benefit in the quantum communication rate:
\[
\frac{1}{n} I_n(\rho^{(n)}, \Gamma^{\otimes n}) \approx Q_1(\Gamma)
\]
\[+
\begin{cases}
(1 - 2p_H)\rho_{HH}^{2} \rho_{VV} \log n & \text{if } 0 < p_H < \frac{1}{2} < p_V < 1, \\
(1 - 2p_V)\rho_{HH}^{2} \rho_{VV} \log n & \text{if } 0 < p_H < \frac{1}{2} < p_V < 1.
\end{cases}
\]

III. SUPERADDITIVITY IMPROVEMENT

The goal of the previous section was to detect the coherent-information superadditivity in the widest region of parameters \(p_H\) and \(p_V\). In this section we discuss how to get a higher quantum communication rate (for fixed values of \(p_H\) and \(p_V\)) by using longer codes.

Our approach is to combine two \(n\)-letter encodings from Section II and slightly modify the result to get a better \(2n\)-letter encoding. To illustrate this approach, consider the region \(0 < p_V < 1 - p_H < \frac{1}{2}\), where \(Q_2 > Q_1\) (see Section II A). Let \(\rho^{(2)}\) be a partially coherent state given by Eq. \((5)\).

The four-qubit state \(\rho^{(2)} \otimes \rho^{(2)}\) inherits some superpositions in the subspace spanned by 12 vectors: \(|HHHV\rangle, |HHHV\rangle, |HHHV\rangle, |HVHV\rangle, |VVVV\rangle, |VVHV\rangle, |VHHV\rangle, |VHHV\rangle, |VHVV\rangle, |VHVV\rangle, |VVHV\rangle, |VVHV\rangle\). On the other hand, the states \(|HHHV\rangle|HVHV\rangle\) and \(|VHVV\rangle|VHHV\rangle\) incoherently contribute to \(\rho^{(2)} \otimes \rho^{(2)}\) though they have the same detection probability \(p_Hp_V^2\). We use the latter fact to construct a more coherent version of the state \(\rho^{(2)} \otimes \rho^{(2)}\) as follows:

\[
\xi^{(4)} := \rho^{(2)} \otimes \rho^{(2)} + 2\rho_{HH}^2 \rho_{VV} \rho_{HH}^2 |HHHV\rangle \langle VVVH| + |VVHH\rangle \langle HHVV|.
\]

The states \(\rho^{(2)} \otimes \rho^{(2)}\) and \(\xi^{(4)}\) have the almost identical spectra, with the difference being in the eigenspace spanned by \(|HHHV\rangle\) and \(|VVHH\rangle\). That difference is translated into the operators \(\Lambda^{\otimes 4}_F[\xi^{(4)}]\) and \(\Lambda^{\otimes 4}_F[\rho^{(2)} \otimes \rho^{(2)}]\), which results in

\[
S(\Lambda^{\otimes 4}_F[\xi^{(4)}]) = S(\Lambda^{\otimes 4}_F[\rho^{(2)} \otimes \rho^{(2)}]) - (2\log 2) p_H^2 p_V^2 \rho_{HH}^2 \rho_{VV}.
\]

Since the partial trace of the operator \(|HHHV\rangle \langle VVVH| + |VVHH\rangle \langle HHVV|\) with respect to any photon vanishes, this means that \(\Lambda^{\otimes 3}_F \otimes (\text{Tr} \circ \Lambda_G)[\xi^{(4)}] = \Lambda^{\otimes 3}_F \otimes (\text{Tr} \circ \Lambda_G)[\rho^{(2)} \otimes \rho^{(2)}]\), etc., so that the density operators \(\xi^{(4)}\) and \(\rho^{(2)} \otimes \rho^{(2)}\) are both mapped to the same operator when affected by any map involving the trivial-and-prepare operation \(\text{Tr}\) for at least one of the qubits. Recalling the fact that \(\Gamma^{\otimes 4} = [\Lambda_F \otimes (\text{Tr} \circ \Lambda_G)]^{\otimes 4}\), we get

\[
S(\Gamma^{\otimes 4}[\xi^{(4)}]) = S(\Gamma^{\otimes 4}[\rho^{(2)} \otimes \rho^{(2)}]) - (2\log 2) p_H^2 p_V^2 \rho_{HH}^2 \rho_{VV}.
\]

Similarly,

\[
S(\tilde{\Gamma}^{\otimes 4}[\xi^{(4)}]) = S(\tilde{\Gamma}^{\otimes 4}[\rho^{(2)} \otimes \rho^{(2)}]) - (2\log 2)(1 - p_H)^2(1 - p_V)^2 \rho_{HH}^2 \rho_{VV}.
\]

These relations lead to a greater coherent information as compared to twice the expression \((8)\), namely,

\[
\begin{align*}
I_n(\xi^{(4)}, \Gamma^{\otimes n}) &= S(\Gamma^{\otimes 4}[\xi^{(4)}]) - S(\tilde{\Gamma}^{\otimes 4}[\xi^{(4)}]) \\
&= S(\Gamma^{\otimes 4}[\rho^{(2)} \otimes \rho^{(2)}]) - S(\tilde{\Gamma}^{\otimes 4}[\rho^{(2)} \otimes \rho^{(2)}]) \\
&+ (2\log 2) p_H^2 \rho_{HH}^2 \rho_{VV}^2 (|1 - p_H|^2(1 - p_V)^2 - p_H^2 p_V^2) \\
&= 4Q_1(\Gamma) + (4\log 2) p_H^2 \rho_{HH}^2 \rho_{VV}^2 (1 - p_H)(1 - p_V).
\end{align*}
\]

Dividing Eq. \((21)\) by 4, we get a better lower bound

\[
Q_4(\Gamma) - Q_1(\Gamma) \geq \frac{1}{4} I_n(\xi^{(4)}, \Gamma^{\otimes 4}) - Q_1(\Gamma)
\]

\[= \left[1 + \frac{1}{2} \rho_{HH}^2 \rho_{VV}^2 (1 - p_H - p_V + 2p_Hp_V)\right] \times \log 2(1 - p_H - p_V) \rho_{HH}^2 \rho_{VV}.
\]

The lower bound \((22)\) significantly outperforms the lower bound \((16)\) for \(n = 4\) in a wide range of parameters \(p_H\) and \(p_V\). For instance, if \(p_H = 0.7\) and \(p_V = 0.2\), then Eq. \((22)\) yields \(Q_4(\Gamma) - Q_1(\Gamma) \geq 6.3 \times 10^{-3}\) bits, whereas Eq. \((16)\) yields \(Q_4(\Gamma) - Q_1(\Gamma) \geq 9.1 \times 10^{-5}\) bits.

Clearly, the presented approach works well to extend the \(n\)-letter code from Section III B two a \(2n\)-letter code by modifying the state \((\rho^{(n)})^{\otimes 2}\) in the subspace spanned by \(|HHHV\rangle \langle VVVH| + |VVHH\rangle \langle HHVV|\). Similarly, the modified \(2n\)-letter code \(\xi^{(2n)}\) further be improved to a \(4n\)-letter code and so on ad infinitum. Starting with the two-letter code in Section I A, we get the following result:

\[
Q(\Gamma) - Q_1(\Gamma) \geq (2\log 2)(1 - p_H - p_V) \sum_{m=0}^{\infty} \frac{2^m \rho_{HH}^m \rho_{VV}^m}{2^m} \times \prod_{k=0}^{2n-1} (1 - p_H)^{2n-k-1}(1 - p_V)^{2n-k-1} p_H^k p_V^k.
\]

IV. CONCLUSIONS

A phenomenon of the coherent-information superadditivity makes it possible to enhance the quantum communication rate by using clever multiletter codes instead of single-letter ones. In this paper, we have studied the superadditivity phenomenon in physically relevant quantum communication lines with polarization dependent losses. Such lines represent a two-parameter family of generalized erasure channels \(\Gamma\), with the attenuation factors \(p_H\) and \(p_V\) for horizontally and vertically polarized photons being the parameters. In prior research, two-letter codes were shown to outperform all single-letter ones for some values of \(p_H\) and \(p_V\) within the region \(p_H + p_V < 1\). Interestingly, if \(p_H + p_V > 1\), then \(\Gamma\) is input-degradable in the sense that there exists a quantum channel \(Y\) such that \(\tilde{\Gamma} = \Gamma \circ Y\). Making an analogy with the case of standard degradable channels, it is tempting to conjecture that the input-degradability implies \(Q_4(\Gamma) = Q_1(\Gamma)\). Our study shows that this conjecture is false: the 3-letter code in Section III B insures \(Q_4(\Gamma) > Q_1(\Gamma)\) if \(p_H + p_V = 1\) and \(0 \neq p_H \neq p_V \neq 0\), see Fig. 2.
The longer the code in Section II B the wider the region of parameters $p_H$ and $p_V$, where the superadditivity phenomenon takes place. In the limit of the infinitely long encoding, $Q(\Gamma) > Q_1(\Gamma)$ for all $p_H$ and $p_V$ satisfying $0 < p_V < \frac{1}{2} < p_H < 1$ or $0 < p_H < \frac{1}{2} < p_V < 1$, i.e., $Q(\Gamma) > Q_1(\Gamma)$ whenever $\Gamma$ is neither degradable nor antidegradable. A feature of the code proposed in Section II B is that it has a clear physical meaning: $\rho^{(n)}$ has an entangled component proportional to $[W^{(n)}]_{\langle W^{(n)} \rangle}$, which in turn has a high detection probability and whose structure is preserved by polarization dependent losses due to the permutation symmetry. Clearly, one could alternatively use another Dicke state [26, 27] instead of $[W^{(n)}]$; however, the detection probability would be less in that case.

In this work, we were interested not only in the superadditivity identification but also in its improvement with the increase of the code length. In Section II B we proposed a method how to get a higher quantum communication rate by doubling the code length. We believe that the scheme is far from being optimal, which necessitates a further search of better codes, e.g., by using a neural network state ansatz [23]. Nonetheless, our analytically derived codes with known asymptotic behavior may serve as a benchmark for future codes generated by numerical optimization.

Acknowledgments

The study in Sections II, III, and IV was supported by the Russian Science Foundation under Project No. 19-11-00086 and performed in Steklov Mathematical Institute of Russian Academy of Sciences. The review in Section II was written in Moscow Institute of Physics and Technology as well as in Valiev Institute of Physics and Technology of Russian Academy of Sciences.