Schemes of transmission of classical information via quantum channels with many senders: Discrete- and continuous-variable cases

L. Czekaj,[1,2,*] J. K. Korbicz,[1,2,3,4] R. W. Chhajlany,[1,2,5] and P. Horodecki[1,2]
[1]*Faculty of Applied Physics and Mathematics, Gdańsk University of Technology, PL-80-952 Gdańsk, Poland*
[2]*National Quantum Information Center of Gdańsk, PL-81-824 Sopot, Poland*
[3]*Institute of Theoretical Physics and Astrophysics, University of Gdańsk, PL-80-952 Gdańsk, Poland*
[4]*ICFO Institut de Ciències Fotòniques, ES-08860 Castelldefels, Barcelona, Spain*
[5]*Faculty of Physics, Adam Mickiewicz University, Umultowska 85, PL-61-614 Poznań, Poland*
(Received 7 November 2011; published 19 January 2012)

Superadditivity effects in the classical capacity of discrete multiaccess channels and continuous variable (CV) Gaussian MACs are analyzed. Several examples of the manifestation of superadditivity in the discrete case are provided, including, in particular, a channel which is fully symmetric with respect to all senders. Furthermore, we consider a class of channels for which *input entanglement across more than two copies of the channels is necessary* to saturate the asymptotic rate of transmission from one of the senders to the receiver. The five-input entanglement of Shor error correction codewords surpass the capacity attainable by using arbitrary two-input entanglement for these channels. In the CV case, we consider the properties of the two channels (a beam-splitter channel and a "nondemolition" quadratures coupling sum (XP) gate channel) analyzed in Czekaj *et al.* [Phys. Rev. A **82**, 020302(R) (2010)] in greater detail and also consider the sensitivity of capacity superadditivity effects to thermal noise. We observe that the estimates of the amount of two-mode squeezing required to achieve capacity superadditivity are more optimistic than previously reported.

DOI: 10.1103/PhysRevA.85.012316 PACS number(s): 03.67.Hk

I. INTRODUCTION

Quantum communication is a dynamically developing branch of quantum information theory [1]. One of its central notions is that of a quantum communication channel [1,2], which models information transfer from senders to receivers using quantum resources. The amount of information which can be encoded in quantum states and reliably sent through a quantum channel is measured, depending on the communication scenario, by various channel capacities: (i) classical capacity (C), defined as the maximal rate at which classical information can be transmitted through the channel; (ii) classical private capacity (P), which is the classical capacity pertaining to the case when the transmitted bits are hidden from an environment; (iii) quantum capacity (Q) characterizing the size of the Hilbert space of states which can be transmitted through the channel. Quantum effects, associated with quantum channels, that have recently attracted much attention are so-called "activations" and "superadditivities." For the quantum capacity Q, various activations were based on bound entanglement, but the most spectacular result was recently obtained in Ref. [3], where an activation of the type $0 \otimes 0 > 0$ was shown. In the case of private capacity P, the corresponding superadditivity was found in Ref. [4] (see also Ref. [5]). Quantum superadditivity of the classical capacity C in the case of multiple access channels (MACs) was shown in the Ref. [6] for discrete variables and in Ref. [7] for continuous variables. The question of additivity of C is still open for the one-sender one-receiver scenario, although a substantial breakthrough on the superadditivity of the Holevo function has recently been achieved in [8].

In the present paper we study a variety of quantum MACs exhibiting superadditivity effects for classical capacity. We do this for both discrete and continuous variable (CV) systems. In particular, for the discrete variable case, we provide a new symmetric scenario where both senders can benefit from capacity superadditivity. This is in contrast to earlier examples studied in Ref. [6], where one of the senders only played a role of an assistant with respect to the other fixed sender. We furthermore go beyond the standard dense coding protocol, which is based on two-particle entanglement and present examples of channels where multipartite entanglement is required to achieve the optimum channel capacity. The use of multipartite entanglement can be seen as the next step in the direction of optimization of the classical capacity of quantum channels. In particular, it is shown that the five-qubit error correction codeword [1] entangled across five inputs beats any two-input-based entanglement encodings for these channels.

In the CV context, we study the examples of Gaussian channels, introduced in Ref. [7] in greater detail. We extend the analysis of nonadditive capacity regions and also study the dependence of the classical capacity of the channels on the choice of the set of input states. We show that for low energies, protocols using two-mode entanglement surpass both coherent state and standard single-mode squeezed-state encodings. Furthermore, we analyze the sensitivity of the superadditivity effects to thermal noise and show that the protocols are relatively sensitive to thermal noise or losses in that 15% of power loss is sufficient to destroy the effect.

The work is organized as follows. All necessary definitions are introduced in Sec. II. Sections III A–III D are devoted to the discrete variable case where we provide a proof of the classical additivity of capacity regions (Sec. III A), an example of a symmetric MAC, exhibiting superadditivity of the classical capacity (Sec. III C), an analysis of the influence of multipartite

*lczekaj@mif.pg.gda.pl

entanglement on the capacity regularization (Sec. III D), and an example of the supperadditivity of regularized capacity regions (Sec. III E). Continuous variable MACs are studied in Secs. IV A–IV C, wherein the locality rule for continuous variable MACs is presented Sec. IV A, the dependence of the classical channel capacities on the choice of input states is studied in Sec. IV B, while the influence of thermal noise is analyzed in Sec. IV C.

II. BASIC DEFINITIONS

The transmission of classical information through a quantum channel corresponds to the following communication sequence [1]:

$$x \mapsto \rho_x \mapsto \Phi(\rho_x) \mapsto \mathrm{tr}[\Phi(\rho_x)E_y] \mapsto y. \tag{1}$$

The sender maps the message x taken from some alphabet, into a state ρ_x of a quantum system, which in turn is sent through a quantum channel Φ to the receiver. The quantum channel models the interaction of ρ_x with the environment. It is assumed that none of the users have access to the environment. The receiver obtains the state $\Phi(\rho_x)$ and performs a measurement $\{E_y\}$, yielding some output result y from which he tries to infer the message sent by the sender. The receiver knows the set of states $\{\rho_x\}$ as well as the respective probabilities p_x with which they are input to the channel. We distinguish two cases: (i) The states $\{\rho_x\}$ belong to a finite-dimensional quantum space and x is a discrete variable (DV); (ii) the states $\{\rho_x\}$ are states of a bosonic system and x is a CV. In the latter situation, a restriction on the average energy sent through the channel must be imposed to obtain a meaningful concept of channel capacity, since cranking up the power of transmission indefinitely allows perfect transfer of information. The restriction usually takes the form of a constraint on the average photon number of the input ensemble $\{p_x,\rho_x\}$: $\mathrm{tr}[\hat{N}\int p_x\rho_x dx] \leqslant N$, where $\hat{N}$ is the photon number operator.

The sender may perform an encoding of his messages into code states to reduce the probability that a message deciphered from the measurement outcome disagrees with the one sent through the channel. Code states belong to the Hilbert space $\mathcal{H}^{\otimes n}$, describing the input of n copies of channel Φ, that is, $\Phi^{\otimes n}$. As $n \to \infty$ the probability of a decoding error can be made arbitrarily small.

The maximal rate at which information can be reliably transmitted through a quantum channel is defined as its *classical capacity* C. By the well-known result [9], the "single-shot" classical capacity $C^{(1)}(\Phi)$ is bounded by the *Holevo quantity*:

$$C^{(1)}(\Phi) \leqslant \chi(\Phi) = \max_{\{p_x,\rho_x\}} \left(S(\Phi(\bar{\rho})) - \sum_x p_x S(\Phi(\rho_x)) \right), \tag{2}$$

where $\bar{\rho} = \sum_x p_x\rho_x$ is the mean input state and $S(\rho) = -\mathrm{tr}[\rho\log\rho]$ is the von Neuman entropy. It can be shown that the above capacity can be achieved by product code states over the copies of $\mathcal{H}$ (Holevo-Schumacher-Westmoreland coding theorem [10]).

However, the input Hilbert space $H^{\otimes n}$ allows also for entangled states, which may be useful for overcoming the above bound. This possibility is quantitatively taken into account by considering the so-called *regularized classical capacity*:

$$C^{(\infty)}(\Phi) = \lim_{n\to\infty} \frac{1}{n}\chi(\Phi^{\otimes n}). \tag{3}$$

The importance of considering entangled encodings is highlighted by Hastings' recent work [8], which showed that there do exist channels for which $C^{(\infty)}(\Phi) > \chi(\Phi)$.

In this paper we consider MACs, where there are at least two senders (we denote them as $A, B, \ldots$, transmitting to one receiver R. Each sender sends his message independently of the other senders; that is, their inputs are completely uncorrelated. They know only the input ensembles and agree upon a set of rules governing the use of the channel: The first n_1 uses of the channel consist of sending states from a fixed first ensemble, the next n_2 uses of the channel consist of states chosen from a second ensemble, and so on. This procedure is called time sharing [11].

For the case of two senders, a MAC acts as a mapping:

$$\rho_{x_A} \otimes \rho_{x_B} \mapsto \Phi(\rho_{x_A} \otimes \rho_{x_B}). \tag{4}$$

Here x_A and x_B are messages pertaining to senders A and B, respectively.

The *capacity region* $\mathcal{R}(\Phi)$ of the classical MAC Φ characterized by the conditional probability distribution of output symbols $p(Y|X_A,X_B)$ is defined as a set of vectors $R = \{R_A,R_B\}$ of rates, simultaneously achievable by adequate coding and time sharing. The capacity region $\mathcal{R}(\Phi)$ of a classical two-sender MAC is given by the convex hull of the rates $\{R_A,R_B\}$ for which there exist probability distributions $p(X_A), p(X_B)$ of transmitted symbols and a joint probability distribution in the form:

$$p(Y,X_A,X_B) = p(Y|X_A,X_B)p(X_A)p(X_B), \tag{5}$$

such that [11]:

$$R_A \leqslant I(X_A : Y|X_B), \tag{6}$$

$$R_B \leqslant I(X_B : Y|X_A), \tag{7}$$

$$R_A + R_B \leqslant I(X_A,X_B : Y), \tag{8}$$

where $I(X_A,X_B : Y)$ denotes the mutual information and $I(X_A : Y|X_B)$ and $I(X_B : Y|X_A)$ are conditional mutual information quantities. These quantities are related to the Shannon entropy $H(X) = -\sum_x p_x \log p_x$ and conditional entropy $H(Y|X) = H(X,Y) - H(X)$ as follows: $I(X : Y) = H(Y) - H(Y|X)$, $I(X : Y|Z) = H(Y|Z) - H(Y|X,Z)$. In the opposite way, for each vector of rates $R \in \mathcal{R}(\Phi)$ there exist joint probability distribution $p(X_A,X_B,Y,Q) = p(Y|X_A,X_B)p(X_A,X_B,Q)$ with input symbols probability distribution in the form:

$$p(X_A,X_B,Q) = p(Q)p(X_A|Q)p(X_B|Q), \tag{9}$$

that following set of inequalities is fulfilled:

$$R_A \leqslant I(X_A : Y|X_B,Q), \tag{10}$$

$$R_B \leqslant I(X_B : Y|X_A,Q), \tag{11}$$

$$R_A + R_B \leqslant I(X_A,X_B : Y|Q). \tag{12}$$

Auxiliary random variable Q refers to time sharing procedure. Special form of the input symbols probability distribution

given by Eq. (9) is connected to the fact that in given moment of time sharing procedure, senders transmit their messages independently.

For the case of a quantum MAC Φ with two senders, a useful notion is that of a "classical-quantum" state: $\rho_{ABR} = \sum_{x_A,x_B} p_{x_A} p_{x_B} e_{x_A} \otimes e_{x_B} \otimes \Phi(\rho_{x_A} \otimes \rho_{x_B})$ where $\{e_{x_A}\}$, $(\{e_{x_B}\})$ are projectors onto the standard basis of the Hilbert space controlled by sender A (B) and $\{p_{x_A},\rho_{x_A}\}$ $(\{p_{x_B},\rho_{x_B}\})$ is the ensemble of code states of A (B). $\Phi(\rho_{x_A} \otimes \rho_{x_B})$ represents state obtained by reciver as the result of transmission.

The single-shot capacity region $\mathcal{R}^{(1)}(\Phi)$ is obtained as a convex closure of all rates (R_A,R_B), for which there exist classical-quantum states ρ fulfilling the following set of inequalities:

$$R_A \leqslant I(X_A : Y|X_B), \tag{13}$$

$$R_B \leqslant I(X_B : Y|X_A), \tag{14}$$

$$R_T = R_A + R_B \leqslant I(X_A,X_B : Y). \tag{15}$$

In distinction to the case of classical channels, the mutual information is now given in terms of the von Neuman entropy $I(X_A,X_B : Y) = S(\rho_{AB}) + S(\rho_R) - S(\rho_{ABR})$ and $I(X_A : Y|X_B) = \sum_{x_B} p_{x_B} I(X_A : Y|X_B = \rho_{x_B})$. Von Neuman entropy is defined as $S(\rho) = -\mathrm{tr}[\rho \log \rho]$. R_T denotes the total capacity and is defined as $R_T = \sum_i R_i$. In the following, we often refer to the notion of the regularized capacity region $\mathcal{R}^{(\infty)}(\Phi) = \lim_{n\to\infty} \mathcal{R}(\Phi^{\otimes n})/n$.

Finally, we use the notion of *parallel composition* of MACs, which we illustrate here by an example of two classical channels (denoted by Φ_I and Φ_{II}) and two senders (A and B). In parallel composition sender A has access to input ports $X_A^I(X_A^{II})$ of the first (second) channel. X_B^I,X_B^{II} denote input ports controlled by sender B. For each input port X_i^j there is a set of possible signals which can be sent through the channel. The channels operate synchronously, which means that the communication process can be divided into steps. In each step, user A sends the vector of symbols $x_A = \{x_A^I,x_A^{II}\}$ while sender B sends symbols $x_B = \{x_B^I,x_B^{II}\}$. In each step a given channel is used by every user exactly once. At the end of the communication step the receiver obtains the output $y = \{y^I,y^{II}\}$.

Let $p(Y^I|X_A^I X_B^I)$,$(p(Y^{II}|X_A^{II} X_B^{II}))$ be the transition probabilities for the MAC's Φ_I (Φ_{II}), then the transition probability for the parallel composition is given by:

$$p(Y|X_A,X_B) = p\left(Y^I\middle|X_A^I,X_B^I\right) p\left(Y^{II}\middle|X_A^{II},X_B^{II}\right). \tag{16}$$

The parallel composition of quantum MACs is defined as the straightforward generalization of the above concept.

III. QUANTUM MACS IN FINITE DIMENSIONAL SPACES

A. Additivity theorem for classical discrete multiaccess channels

We state the additivity theorem for capacity regions of classical discrete MACs in full generality. First recall that the capacity region $\mathcal{R}(\Phi)$ for a classical MAC with arbitrary number of senders is given by the convex hull of the rates $\{R_i\}$ which fulfill

$$R_S \leqslant I(X_S : R|S^C), \tag{17}$$

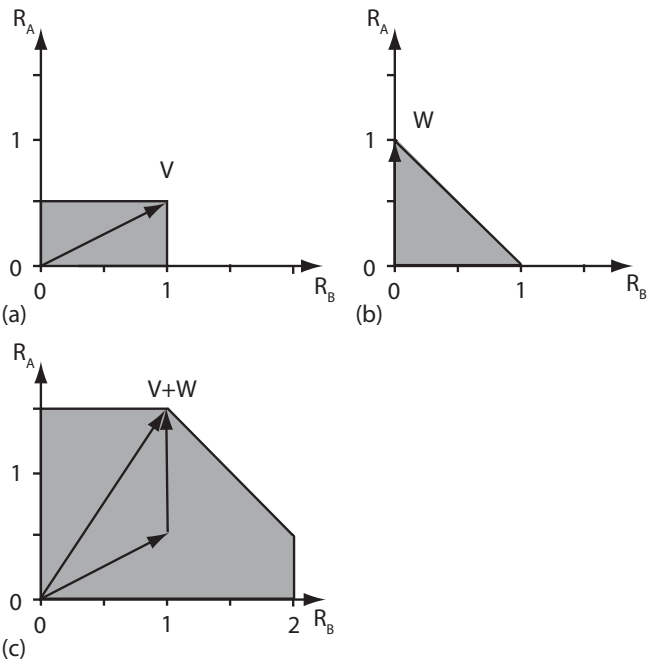

FIG. 1. Additivity of the capacity regions for classical MACs. Capacity regions for channels Φ_I and Φ_{II} are presented, respectively, in panels (a) and (b). The capacity region of the parallel composition $\Phi_I \otimes \Phi_{II}$ of channels Φ_I and Φ_{II} is presented in panel (c) and it is given by the geometrical sum of capacities regions from panels (a) and (b) (see Ref. [6]).

where S enumerates all subsets of senders and $R_S = \sum_{i\in S} R_i$, while S^C is the complement of the set S [11]. For 2-to-1 channels this reduces to the simple form of Eqs. (6)–(8). The capacity region evaluated for fixed probability distribution of input symbols $\tilde{p} = p(Q)\prod_i p(X_i^I,X_i^{II}|Q)$ has a form [cf. Eq. (17)]

$$\tilde{\mathcal{R}} = \{R \in \mathbb{R}^n : \forall_{S\subseteq E} R_S \leqslant I(X_S : Y|X_{S^C},Q),\ \forall_{i\in E} R_i \geqslant 0\}. \tag{18}$$

Pay attention on specific product structure of $\tilde{p}$ which reflect independency of senders.

The *additivity theorem* states that the achievable capacity region $\mathcal{R}$ of a channel being the parallel composition of MACs is the geometrical sum of capacity regions of the constituting channels. More formally, suppose n MACs are used in parallel, with each channel having m senders. Let $\tilde{R} = \{R_1,\ldots,R_m\}$ be the vector of achievable rates for the composite channel; then the capacity additivity theorem states that $\tilde{R}$ can be written as a sum of vectors $\tilde{R}^{(j)}$ describing the capacity region of the jth MAC [12]:

$$\mathcal{R}\left(\bigotimes_i \Phi_i\right) = \sum_i \mathcal{R}(\Phi_i). \tag{19}$$

The additivity theorem for the case of channels with two senders is graphically depicted in Fig. 1.

Here we prove the theorem only for a simple 2-to-1 case $\mathcal{R}(\Phi_I \otimes \Phi_{II}) = \mathcal{R}(\Phi_I) + \mathcal{R}(\Phi_{II})$, a complete proof is postponed to the Appendix. We start with the inclusion $(\subseteq)$. The outline of the proof is as follows: For an arbitrary chosen vector of rates $\tilde{R} = (R_A,R_B) \in \mathcal{R}(\Phi_I \otimes \Phi_{II})$ from

the definition of the capacity region we know that there exists a probability distribution: $\tilde{p} = p(X_A^I, X_A^{II}, X_B^I, X_B^{II}, Q) = p(Q)p(X_A^I, X_A^{II}|Q)p(X_B^I, X_B^{II}|Q)$ and a corresponding fixed probability capacity such region that $\tilde{R} \in \tilde{\mathcal{R}}(\Phi_I \otimes \Phi_{II})$. We will use $\tilde{p}$ to construct probability distribution $\bar{p} = \tilde{p}_I \tilde{p}_{II}$ where $\tilde{p}_I = p(Q^I)p(X_A^I|Q^I)p(X_B^I|Q^I)$ [$\tilde{p}_{II} = p(Q^{II})p(X_A^{II}|Q^{II})p(X_B^{II}|Q^{II})$] is the marginal probability distribution of input symbols of channel Φ_I (Φ_{II}). $\tilde{p}_I$ is obtained from $\tilde{p}$ by tracing out variables X_A^{II}, X_B^{II} and changing the name of time sharing auxiliary variable $Q \mapsto Q^I$. In the same way we obtain $\tilde{p}_{II}$. Notice that in the case of $\bar{p}$ the channels are used independently. Fixed probability capacity region corresponding to $\bar{p}$ will be denoted by $\bar{\mathcal{R}}(\Phi_I \otimes \Phi_{II})$. Since input symbols probability distribution $\bar{p}$ has suitable form [see Eqs. (5) and (9)], it determines the valid fixed probability capacity region $\bar{\mathcal{R}}(\Phi_I \otimes \Phi_{II}) \subseteq \mathcal{R}(\Phi_I \otimes \Phi_{II})$.

We first show that $\tilde{\mathcal{R}}(\Phi_I \otimes \Phi_{II}) \subseteq \bar{\mathcal{R}}(\Phi_I \otimes \Phi_{II}) \subseteq \mathcal{R}(\Phi_I \otimes \Phi_{II})$. Then we show that $\bar{\mathcal{R}}(\Phi_I \otimes \Phi_{II}) = \tilde{\mathcal{R}}(\Phi_I) + \tilde{\mathcal{R}}(\Phi_{II})$, where $\tilde{\mathcal{R}}(\Phi_I)$ is the fixed probability capacity region obtained for channel Φ_I for input symbols probability distribution $\tilde{p}_I$ and $\tilde{\mathcal{R}}(\Phi_{II})$ has a similar meaning for Φ_{II}. From the relation $\bar{\mathcal{R}}(\Phi_I \otimes \Phi_{II}) = \tilde{\mathcal{R}}(\Phi_I) + \tilde{\mathcal{R}}(\Phi_{II})$ we infer that the rate vector $\tilde{R}$ may be presented in the form $\tilde{R} = \tilde{R}_I + \tilde{R}_{II}$, where $\tilde{R}_{I(II)} \in \tilde{\mathcal{R}}(\Phi_{I(II)}) \subseteq \mathcal{R}(\Phi_{I(II)})$, which finishes the proof.

The following facts will be useful in further considerations:

$$H(Y|Q) \leqslant H(Y^I|Q^I) + H(Y^{II}|Q^I), \tag{20}$$

$$H(Y|X_B, Q) \leqslant H\big(Y^I\big|X_B^I, Q^I\big) + H\big(Y^{II}\big|X_B^{II}, Q^{II}\big), \tag{21}$$

$$H(Y|X_A, X_B, Q) = H\big(Y^I\big|X_A^I, X_B^I, Q^I\big) + H\big(Y^{II}\big|X_A^{II}, X_B^{II}, Q^{II}\big). \tag{22}$$

Equation (20) can be proven in the following way:

$$H(Y|Q) = \sum_q p(q)H(Y|Q=q) \tag{23}$$

$$\leqslant \sum_q p(q)(H(Y^I|Q=q) + H(Y^{II}|Q=q)) \tag{24}$$

$$= \sum_q p(q)H(Y^I|Q=q) + \sum_q p(q)H(Y^{II}|Q=q) \tag{25}$$

$$= \sum_{q^I} p(q^I)H(Y^I|Q^I=q^I) + \sum_{q^{II}} p(q^{II})H(Y^{II}|Q^I=q^{II}) \tag{26}$$

$$= H(Y^I|Q^I) + H(Y^{II}|Q^{II}), \tag{27}$$

where in Eq. (24) we used entropy subadditivity and in Eq. (26) we simply renamed auxiliary time sharing variable. This is allowed since q is used as an index in two independent summations. In a similar way one can derive Eq. (21). To prove Eq. (22) it is enough to observe that the conditional transition probability describing the setup $\Phi_I \otimes \Phi_{II}$ factorizes [see Eq. (16)]; hence, we can write

$$H(Y|X_A, X_B, Q) = -\sum_{x_A, x_B, y, q} p \log p(y|x_A, x_B, q) \tag{28}$$

$$= -\sum_{x_A^I, x_B^I, y^I, q^I} p_I \log p\big(y^I\big|x_A^I, x_B^I\big) \tag{29}$$

$$- \sum_{x_A^{II}, x_B^{II}, y^{II}, q^{II}} p_{II} \log p\big(y^{II}\big|x_A^{II}, x_B^{II}\big) \tag{30}$$

$$= H\big(Y^I\big|X_A^I, X_B^I, Q^I\big) + H\big(Y^{II}\big|X_A^{II}, X_B^{II}, Q^{II}\big), \tag{31}$$

where

$$\begin{aligned} p &= p(X_A, X_B, Y, Q) \\ &= p\big(\{x_A^I, x_A^{II}\}, \{x_B^I, x_B^{II}\}, \{y^I, y^{II}\}, q\big) \\ &= p(Y|X_A, X_B)\tilde{p} \\ &= p(Y|X_A, X_B)p(Q)p\big(X_A^I, X_A^{II}\big|Q\big)p\big(X_B^I, X_B^{II}\big|Q\big) \end{aligned} \tag{32}$$

$$\begin{aligned} p_I &= p\big(X_A^I, X_B^I, Y^I, Q^I\big) \\ &= p\big(Y^I\big|X_A^I, X_B^I\big)\tilde{p}_I \\ &= p\big(Y^I\big|X_A^I, X_B^I\big)p(Q^I)p\big(X_A^I\big|Q^I\big)p\big(X_B^I\big|Q^I\big) \end{aligned} \tag{33}$$

$$\begin{aligned} p_{II} &= p\big(X_A^{II}, X_B^{II}, Y^{II}, Q^{II}\big) \\ &= p\big(Y^{II}\big|X_A^{II}, X_B^{II}\big)\tilde{p}_{II} \\ &= p\big(Y^{II}\big|X_A^{II}, X_B^{II}\big)p(Q^{II})p\big(X_A^{II}\big|Q^{II}\big)p\big(X_B^{II}\big|Q^{II}\big) \end{aligned} \tag{34}$$

Now we show that $\tilde{\mathcal{R}}(\Phi_I \otimes \Phi_{II}) \subseteq \bar{\mathcal{R}}(\Phi_I \otimes \Phi_{II})$.

By the definition of the capacity region, there exists the input symbol probability distribution $\tilde{p}$ that the rate vector $\tilde{R}$ satisfies Eqs. (6)–(8). Using Eqs. (20)–(22) we can bound the right-hand side of Eqs. (6)–(8) in the following way:

$$R_A \leqslant I(X_A : Y|X_B, Q) \tag{35}$$

$$= H(Y|X_B, Q) - H(Y|X_A, X_B, Q) \tag{36}$$

$$\leqslant H\big(Y^I\big|X_B^I, Q^I\big) + H\big(Y^{II}\big|X_B^{II}, Q^{II}\big) - H(Y|X_A, X_B, Q) \tag{37}$$

$$= H\big(Y^I\big|X_B^I, Q^I\big) + H\big(Y^{II}\big|X_B^{II}, Q^{II}\big) - H\big(Y^I\big|X_A^I, X_B^I, Q^I\big) - H\big(Y^{II}\big|X_A^{II}, X_B^{II}, Q^{II}\big) \tag{38}$$

$$= I\big(X_A^I : Y^I\big|X_B^I, Q^I\big) + I\big(X_A^{II} : Y^{II}\big|X_B^{II}, Q^{II}\big). \tag{39}$$

An analogous expression can be written for R_B. Then,

$$R_A + R_B \leqslant I(X_A, X_B : Y|Q) \tag{40}$$

$$= H(Y|Q) - H(Y|X_A, X_B, Q) \tag{41}$$

$$\leqslant H(Y^I|Q^I) + H(Y^{II}|Q^{II}) - H(Y|X_A, X_B, Q) \tag{42}$$

$$= H(Y^I|Q^I) + H(Y^{II}|Q^{II}) - H\big(Y^I\big|X_A^I, X_B^I, Q^I\big) - H\big(Y^{II}\big|X_A^{II}, X_B^{II}, Q^{II}\big) \tag{43}$$

$$= I\big(X_A^I, X_B^I : Y^I\big|Q^I\big) + I\big(X_A^{II}, X_B^{II} : Y^{II}\big|Q^{II}\big). \tag{44}$$

$I(X_A^I{:}Y^I|X_B^I,Q^I), I(X_B^I:Y^I|X_A^I,Q^I), I(X_A^I,X_B^I:Y^I|Q^I)$ are calculated for the marginal distribution $\tilde{p}_I$ (analogically for $\tilde{p}_{II}$). Summing up, $\tilde{R}$ belongs to the region given by the set of inequalities:

$$R_A \leqslant I\left(X_A^I:Y^I\middle|X_B^I,Q^I\right)+I\left(X_A^{II}:Y^{II}\middle|X_B^{II},Q^{II}\right), \quad (45)$$

$$R_B \leqslant I\left(X_B^I:Y^I\middle|X_A^I,Q^I\right)+I\left(X_B^{II}:Y^{II}\middle|X_A^{II},Q^{II}\right), \quad (46)$$

$$R_A+R_B \leqslant I\left(X_A^I,X_B^I:Y^I\middle|Q^I\right)+I\left(X_A^{II},X_B^{II}:Y^{II}\middle|Q^{II}\right). \quad (47)$$

However these inequalities define a region $\bar{R}(\Phi_I\otimes\Phi_{II})$, which can be easily checked by evaluation of $I(X_A:Y|X_B,Q)$, $I(X_B:Y|X_A,Q)$ and $I(X_A,X_B:Y|Q)$ on probability distribution $\bar{p}$. Hence $\tilde{R}\in\bar{R}(\Phi_I\otimes\Phi_{II})$ and in this way we have shown the capacity region inclusion.

We move to $\bar{\mathcal{R}}(\Phi_I\otimes\Phi_{II})=\tilde{\mathcal{R}}(\Phi_I)+\tilde{\mathcal{R}}(\Phi_{II})$. Fixed probability capacity region $\tilde{\mathcal{R}}_I$ obtained for input symbol probability distribution $\tilde{p}_I$ is given by

$$R_A \leqslant I\left(X_A^I:Y^I\middle|X_B^I,Q^I\right), \quad (48)$$

$$R_B \leqslant I\left(X_B^I:Y^I\middle|X_A^I,Q^I\right), \quad (49)$$

$$R_A+R_B \leqslant I\left(X_A^I,X_B^I:Y^I\middle|Q^I\right). \quad (50)$$

Geometrical sum $\tilde{\mathcal{R}}_I+\tilde{\mathcal{R}}_{II}$ can be easy obtained as a convex hull of sums of vertices of the fixed probability capacity regions $\tilde{\mathcal{R}}_I,\tilde{\mathcal{R}}_{II}$ and is equal to the region $\bar{\mathcal{R}}$. Because $\tilde{R}$ was chosen arbitrary, we have proven that $\mathcal{R}(\Phi_I\otimes\Phi_{II})\subseteq\mathcal{R}(\Phi_I)+\mathcal{R}(\Phi_{II})$.

We move to the opposite inclusion ($\supseteq$). Let $\tilde{R}_I\in\mathcal{R}(\Phi_I)$ belong to the fixed probability capacity region associated with the input distribution $\tilde{p}_I$ and similarly for $\tilde{R}_{II}$. It is easy to check by direct evaluation of Eqs. (10)–(12) that the rate vector $\tilde{R}_I+\tilde{R}_{II}$ belongs to the fixed probability capacity region of $\Phi_I\otimes\Phi_{II}$ obtained for the input distribution $\tilde{p}=\tilde{p}_I\tilde{p}_{II}$. That proves $\tilde{R}_I+\tilde{R}_{II}\in\mathcal{R}(\Phi_I\otimes\Phi_{II})$.

B. Supperadditivity

Supperadditivity is defined as the situation when for a certain type of capacity $\tilde{C}$ and two channels Φ_I,Φ_{II}, the following holds:

$$\tilde{C}(\Phi_I\otimes\Phi_{II})>\tilde{C}(\Phi_I)+\tilde{C}(\Phi_{II}). \quad (51)$$

One may distinguish the following types of superadditivities: (a) superaddivity of channel capacity, when $\tilde{C}=C^{(\infty)}$ [see Eq. (3)]; (b) superadditivity of Holevo capacity, when $\tilde{C}=\chi$; (c) self-superadditivity, if $\tilde{C}=\chi$ and $\Phi_I=\Phi_{II}$. For self-superadditivity, $C^{(\infty)}>C^{(1)}$. Note that the right-hand side of (51) expresses the capacity achieved with product inputs on Φ_I and Φ_{II}. Superadditivity means that using encoded states that are correlated (entangled) across uses of channels is advantageous.

In the context of MACs superadditivity effects are identified in terms of the capacity regions: $\mathcal{R}(\Phi_I\otimes\Phi_{II})\mathcal{R}\not\supseteq(\Phi_I)+\mathcal{R}(\Phi_{II})$, where $+$ denotes the geometrical sum of two regions. Superadditivity occurs if there exists a vector in the region $\mathcal{R}(\Phi_I\otimes\Phi_{II})$ which cannot be expressed as a sum of two vectors from $\mathcal{R}(\Phi_I),\mathcal{R}(\Phi_{II})$. To prove supperadditivty effects in terms of the capacity regions it is enough to show that the maximal rate achieved by one of the senders (say sender A) exhibits superadditivity. This means that we may concentrate only on the rate of a single sender or, in other words, show the effect only by analysis of its "coordinate" (or "dimension") in the multidimensional geometric regions $\mathcal{C}(\Phi_I\otimes\Phi_{II})$, $\mathcal{C}(\Phi_I)$, and $\mathcal{C}(\Phi_{II})$.

C. Superadditivity effect in symmetric channels

Examples of channels presented in [6,7], which exhibit superadditivity effects, are highly unsymmetrical. One of the senders performs a "remote" dense coding on the part of an entangled state transmitted by the other. In the described communication schemes one sender is a true sender who transmits messages, while the role of the others is only to help in the communication process since their transmission rates are zero. It might suggest that in the channels based on the dense coding scheme there is only a single supersender who takes advantage of the entangled state transmission. This is not the case as shown here. A channel can be constructed that is symmetric with respect to the exchange of senders, facilitating a superadditivity effect for all of them.

Here we consider a channel Φ (see Fig. 2) with two senders: A and B. Each of the senders controls two one-qubit lines. The channel operates in two modes: F and S. Each occurs with a probability $1/2$. In the first mode, the operation of the channel is depicted in Fig. 2(b). In the second mode, A and B are swapped; that is, lines A_1 and A_2 now belong to B while B_1 and B_2 belong to A. The channel is explicitly symmetric with respect to the senders. Information that the first (second) case occurred is sent to the receiver as a label $|F\rangle(|S\rangle)$. The crosses at the ends of the lines denote replacement of the transmitted state by a completely mixed state. The action

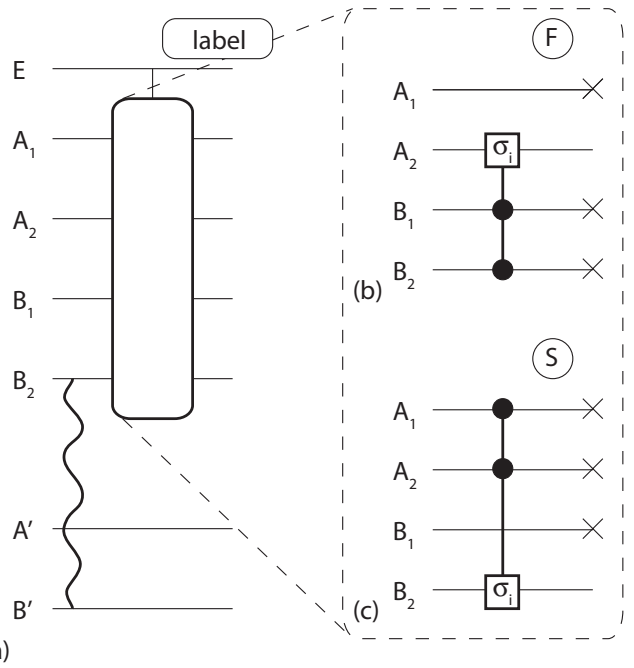

FIG. 2. Channel Φ from Sec. III C. (a) Channel Φ working in parallel with identity channel $\mathcal{I}$; waved line denotes entangled state. (b),(c) Two modes of work of channel Φ.

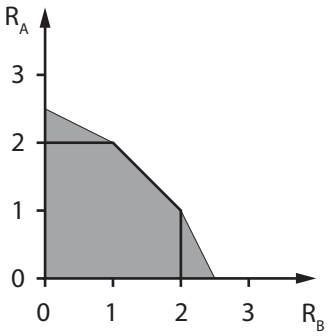

FIG. 3. Lower bound for achieved capacity region for channel Φ from Sec. III C working in parallel with the identity channel. Thick lines refer to the upper bound for geometrical sum of capacity regions of component channels.

of the controlled σ_i gate is $|00\rangle\langle 00| \otimes I + |01\rangle\langle 01| \otimes \sigma_x + |10\rangle\langle 10| \otimes \sigma_z + |11\rangle\langle 11| \otimes \sigma_y$. The capacity region $\mathcal{R}(\Phi)$ is upper bounded by the inequalities $R_A \leqslant 1, R_B \leqslant 1, R_A + R_B \leqslant 1$ as a direct consequence of the dimensionality of the output space (one-qubit space).

Each user is supplied with an additional one-qubit identity connection with the receiver. These two channels are jointly referred to as $\mathcal{I}$. Note that its capacity region $\mathcal{R}(\mathcal{I})$ is given by $R_A \leqslant 1, R_B \leqslant 1, R_A + R_B \leqslant 2$.

The upper bound for $\mathcal{R}(\Phi) + \mathcal{R}(\mathcal{I})$ thus becomes $R_A \leqslant 2, R_B \leqslant 2, R_A + R_B \leqslant 3$. On the other hand, the lower bound for the achievable capacity region of the composite channel $\mathcal{R}(\Phi \otimes \mathcal{I})$ can be seen in Fig. 3. To see this, we present a protocol which achieves the capacity (2.5,0). Due to symmetry of the channel, it follows that the rates (0,2.5) are also achievable. Notice immediately that the rates (1,2) and (2,1) can be obtained by product code states. All the other rates presented in Fig. 3 are obtained by time sharing.

Consider the following protocol: Sender A sends states $|i\rangle|i'\rangle$ with a probability 1/8, where $|i\rangle \in \{|00\rangle, \ldots, |11\rangle\}$ are all possible standard basis states of two qubits, while $|i'\rangle \in \{|0\rangle, |1\rangle\}$. The two-qubit states $|i\rangle$ are input to Φ while $|i'\rangle$ are input to the supporting identity channel $\mathcal{I}$. B sends a fixed state $1/\sqrt{2}|0\rangle(|00\rangle + |11\rangle)$ with one qubit of the Bell state sent through line B_2 and the other through the supporting channel.

For a given $\{i,i'\}$ the receiver obtains

$$\rho_{i,i'} = \tfrac{1}{2}|F\rangle\langle F| \otimes \tfrac{1}{2}\mathrm{I} \otimes |i_{A_2}\rangle\langle i_{A_2}| \otimes \tfrac{1}{8}\mathrm{I}^{\otimes 3} \otimes |i'\rangle\langle i'| + \tfrac{1}{2}|S\rangle\langle S| \otimes \tfrac{1}{8}\mathrm{I}^{\otimes 3} \otimes |\psi_i\rangle\langle\psi_i| \otimes |i'\rangle\langle i'|,$$

where $|F\rangle, |S\rangle$ denote the mode of operation of channel Φ. The output state consists of the mode label and six qubits. The first four qubits are output by channel Φ, while the fifth and sixth qubits are outputs pertaining to $\mathcal{I}$. If channel Φ works in mode F, either the identity operation I or σ_x is performed on line A_2. However, states sent by sender A ($|0\rangle$ and $|1\rangle$) are invariant under the mentioned operations since the receiver obtains an unchanged state from line A_2. If channel Φ operates in the mode S, the controlled σ_i gate, triggered by the state $|i\rangle$ from sender A, is performed on a half of the Bell state input by sender B. The result of this operation is denoted by $|\psi_i\rangle$. The entropy of the conditional output state $\rho_{i,i'}$ is equal to 4.5. Note that the entropy has the same value for each input state $|i\rangle|i'\rangle$.

The mean output state is $\bar{\rho} = \frac{1}{8}\sum_{i,i'} \rho_{i,i'}$ and can be written as

$$\rho = \tfrac{1}{2}\left(|F\rangle\langle F| \tfrac{1}{64}\mathrm{I}^{\otimes 6} + |S\rangle\langle S| \tfrac{1}{64}\mathrm{I}^{\otimes 6}\right) \quad (52)$$

$$= \tfrac{1}{128}\mathrm{I}^{\otimes 7}. \quad (53)$$

Its entropy is $S(\rho) = 7$. In the presented scheme, sender B always transmits the same state and, hence, his rate is zero. Since the setup $\Phi \otimes \mathcal{I}$ can be viewed as a channel with the single sender A, while the state from the helper-sender B is formally included to the environment, by Holevo's theorem [cf. Eq. (2)], we obtain that the rate that sender A can attain is 2.5 bits.

Although rates (2.5,0) and (0,2.5) are achieved in the protocol where there is still one true sender, while the other is a helper-sender and there is no superadditivity of the total rate $R_T = R_A + R_B$, potentially both of senders can take advantage of entangled state transmission.

D. Multiparticle entanglement and the regularized capacity region

In this section we give an example of a channel where senders must use multiparticle entangled states to achieve the regularized capacity region.

We start by describing the class of channels $\Phi_{n,n'}$ that will be used in the search for supperadditivity effects. The channels have one distinguished sender A and n helper-senders B_i. Sender A controls n' of two-qubit lines which are measured in the standard basis by the channel (alternatively, it can be seen as a control of n' 2-bit lines), while senders B_i control only one-qubit lines. Each time the channel is used, one of the helper-senders is attributed to each 2-bit line of sender A. One helper-sender can be attributed to only one line of sender A. Selected helper-senders become active helper-senders. It means that they participate in a transmission of messages from sender A. The state from the active helper-sender is modified by the unitary operation from the set $I, \sigma_x, \sigma_y, \sigma_z$, which is triggered by the state of the appropriate line of sender A. The states of the other helper-senders remain unchanged. The described selection of active helper-senders is performed in a random way. Each selection can be chosen with equal probability. States transmitted by A are absorbed (i.e., the output degrees of freedom of A are traced out). The receiver obtains only the states coming from senders B_i and a label w with information about attribution of active helper-senders to the lines of A. For example, if $n = 3, n' = 2$, the label $w = \{2,3\}$ tells the receiver that states from senders B_2 and B_3 were chosen as the targets of the unitaries controlled by the first and the second two-qubit lines of sender A, respectively. This channel is schematically depicted in Fig. 4. Note that the message included in the label w may be represented as a $n'\lceil \log_2 n \rceil$-qubit state $|w\rangle = |(i_1)_b, \ldots, (i_{n'})_b\rangle$, where i_k is the number of the helper-senders chosen to be the target of the unitary operation controlled by kth line of sender A and $(\cdot)_b$ denotes the binary representation of the value i_k. For example, in the above-mentioned case of $n = 3$ the label $w = 2,3$ corresponds to $|w\rangle = |10,11\rangle$. We use this notation in the analysis of a specific example.

Here we study a parallel setup of m copies of the channel $\Phi_{n,n'=1}$ from the class described above. For simplicity we

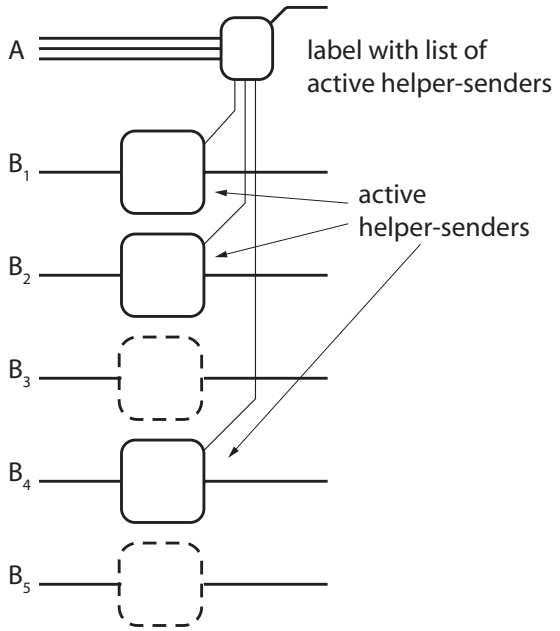

FIG. 4. The channel described in Sec. III D with $n = 5$ helper-senders and $n' = 3$ lines belonging to sender A. The message represented by the label $w = 1,2,4$ is additionally sent to the receiver, which may be represented as a "flag" state $|w\rangle = |001,010,100\rangle$.

denote the channel by Φ. Note that in what follows we have chosen $n' = 1$. In the setup $\Phi^{\otimes m}$, senders B_i can send at most m-particle entangled states through their inputs. We assume that entanglement cannot be used throughout inputs of two different senders (see Fig. 5).

We focus on the upper bound for the achievable rate for sender A. We restrict ourselves to the scheme where helper-senders B_i send one state at all times. Vectors of rates for such schemes take the form $(R_A,0)$. Formally, we can consider the channel $\Phi^{\otimes m}$ in the setup as a 1-to-1 channel and determine the capacity $C_A(\Phi^{\otimes m})$ of sender A. Now we prove that the upper bound for the capacity $C_A(\Phi^{\otimes m})$ has the form

$$C_A(\Phi^{\otimes m}) \leqslant n \sum_{i=0}^{m} p^i (1-p)^{m-i} \binom{m}{i} \min(2i,m), \quad (54)$$

where n is the number of helper-senders, m is the number of channels used for transmission that is equivalent to the number of parties in the entangled state pertaining to B_i, and $p = 1/n$.

Proof. First we find an upper bound for the Holevo capacity of the setup $\Phi^{\otimes m}$ in the case when the helper-sender B_i was active l_i times. Then we use these results to calculate the upper bound for the capacity of $\Phi^{\otimes m}$.

The orthogonal label $|w_j\rangle$ describes which sender B_i was active in the jth copy of Φ. Label $|w\rangle = |w_1,\ldots,w_m\rangle = |w_1\rangle \otimes \cdots \otimes |w_m\rangle$ is the complete list of the active helper-senders in the setup. Given the label we know that sender B_i was active l_i times. The probability of occurrence of the situation described in $|w\rangle$ is given by $p_w = p^m$.

Suppose that $|w\rangle$ is obtained as the result of the action of $\Phi^{\otimes m}$. This fixes the attribution of senders B_i to the lines of A. We denote this case as $\Phi_w^{\otimes m}$. Now, the m uses of the channel Φ can be thought as n separate channels Γ_i. The input of each channel Γ_i consists of the subset of lines from A and all lines

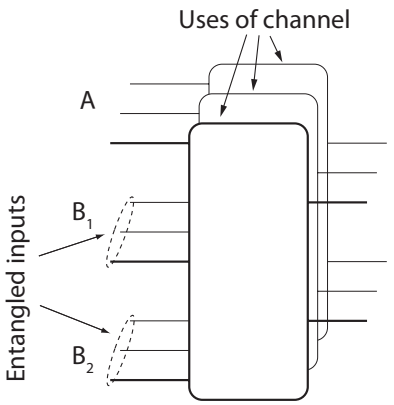

FIG. 5. The parallel setup of channels described in Sec. III D. The inputs of the channels used for transmission of entangled states are shown. The presented case consists of channels with two helper-senders $n = 2$, each of which can send three particle entangled states $m = 3$.

from B_i. None of the Γ_i share input lines with any other Γ_j. Each channel Γ_i has a $2l_i$-qubit input from sender A, a m-qubit input from sender B_i, and a m-qubit output. The equivalence $\Phi_w^{\otimes m} = \otimes_i \Gamma_i$ is depicted in Fig. 6.

Taking into account the dimensionality, one can infer that A can transmit at most $\min(2l_i,m)$ classical bits of information through Γ_i. Given $|w\rangle$, the channels Γ_i work independently. There cannot be entanglement shared between Γ_i and Γ_j because sender A transmits only classical states and users B_i and B_j cannot share entanglement due to the definition of a MAC. This leads to the total conditional capacity:

$$C_A\left(\Phi_w^{\otimes m}\right) = \sum_{i=1}^{n} C(\Gamma_i) = \sum_i \min(2l_i,m). \quad (55)$$

The following observation is helpful for the calculation of $C_A(\Phi^{\otimes m})$. Consider a channel $\Delta(\rho) = \sum_w p_w \Delta_w(\rho)|w\rangle\langle w|$, which acts with probability p_w as some channel Δ_w. Assume

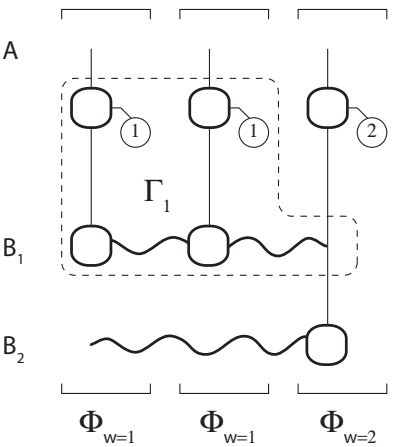

FIG. 6. Equivalence $\Phi_w^{\otimes m} = \otimes_i \Gamma_i$ on example of channel $\Phi^{\otimes 3}_{w=\{1,1,2\}}$. The dashed line delimits channel Γ_1; waved lines denote entangled states from by senders B_1 and B_2; labels with active helper-sender are in the circles.

again that the label $|w\rangle$ is sent to the receiver which identifies the case that occurs. For this channel we have

$$C(\Delta) = \max_{\{p_x,\rho_x\}} S\left[\Delta\left(\sum p_x\rho_x\right)\right] - \sum_x p_x S(\Delta(\rho_x)) \quad (56)$$

$$= \max_{\{p_x,\rho_x\}} S\left(\sum_w p_w \Delta_w\left(\sum p_x\rho_x\right)|w\rangle\langle w|\right) - \sum_x p_x S\left(\sum_w p_w \Delta_w(\rho_x)|w\rangle\langle w|\right) \quad (57)$$

$$= \max_{\{p_x,\rho_x\}} \sum_w p_w \left\{S\left[\Delta_w\left(\sum p_x\rho_x\right)\right] + H(\{p_w\}) - \sum_x p_x S(\Delta_w(\rho_x)) - H(\{p_w\})\right\} \quad (58)$$

$$\leqslant \sum_w p_w \max_{\{p_x^w,\rho_x^w\}} \left\{S\left[\Delta_w\left(\sum p_x^w\rho_x^w\right)\right] - \sum_x p_x^w S\left(\Delta_w\left(\rho_x^w\right)\right)\right\} \quad (59)$$

$$= \sum_w p_w C(\Delta_w), \quad (60)$$

where equality occurs if the same ensemble achieves the capacity of each channel Δ_w. A similar argument can be used to show that the rates achieved for the channel Δ in a certain protocol obey

$$R(\Delta) = \sum_w p_w R(\Delta_w), \quad (61)$$

where $R(\Delta_w)$ are the rates achieved by this protocol for Δ_w's.

Using the above observation with $\Delta = \Phi_w^{\otimes n}$, $p_w = p^m$, and $C(\Delta_w)$ substituted by $C_A(\Phi_w^{\otimes m})$ from Eq. (55), we obtain

$$C_A(\Phi^{\otimes m}) \leqslant p^m \sum_w [\min(2l_1(w),m) \quad (62)$$

$$+ \cdots + \min(2l_n(w),m)], \quad (63)$$

where $l_i(w)$ denotes value of l_i encoded in label w. After some rearrangement:

$$C_A(\Phi^{\otimes m}) \leqslant p^m \sum_w [\min(2l_1(w),m) + \cdots] \quad (64)$$

$$= p^m \sum_{l_1+\ldots+l_n=m} \frac{m!}{l_1!\cdot\cdots\cdot l_n!} \times [\min(2l_1,m) + \cdots + \min(2l_n,m)] \quad (65)$$

$$= np^m \sum_{l=0}^{m} \binom{m}{l} \min(2l,m)\alpha_l \quad (66)$$

$$= n\sum_{l=0}^{m} \binom{m}{l} \min(2l,m)p^m(n-1)^{m-l} \quad (67)$$

$$= n\sum_{l=0}^{m} p^l(1-p)^{m-l}\binom{m}{l}\min(2l,m). \quad (68)$$

In Eq. (65) we collected in the common factor all w with the same $\{l_1,\ldots,l_n\}$. Because formulas with $\min(l_2,m),\ldots,\min(l_n,m)$ in Eq. (65) have the same form as the one with $\min(l_1,m)$, we omitted them and introduced in Eq. (66) a factor n. Moreover, we introduced $\alpha_l = \sum_{l_2+\ldots+l_n=m-l} \binom{m-l}{l_2,\ldots,l_n}$. In Eq. (67) we used the relation

$$\sum_{k_1+\cdots+k_n=m} \binom{m}{k_1,\ldots,k_n} = n^m. \quad (69)$$

Recalling that $p = 1/n$, one has $(n-1)p = (n-1)/n = 1-p$, which was used in Eq. (68). Upper bound for the regularized capacity $C_A^{(m)} = \frac{1}{m}C_A^{(m)}(\Phi^{\otimes m})$ is presented in Fig. 7.

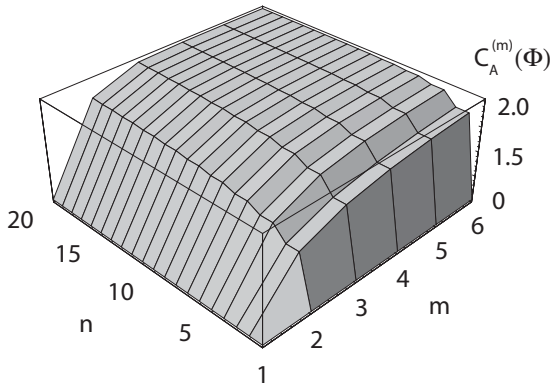

FIG. 7. Upper bound for the regularized capacity $C_A^{(m)} = \frac{1}{m}C_A(\Phi^{\otimes m})$ as a function of m—the number of channel copies the capacity is evaluated on—and the number of helper senders n in the channel Φ.

The upper bound given by (54) is achieved in the case of $m \in \{1,2,5\}$ by the protocols which run as follows: A transmits with an equal probability all states from the standard basis of his $2m$-qubit input space of $\Phi^{\otimes m}$, while all B_i's transmit either the state $|0\rangle$ from the standard base, one of the Bell states $|\Phi^+\rangle$ or $|0_L\rangle$, a five-qubit correction codeword (see [1]) for $m = 1$, $m = 2$ or $m = 5$, respectively:

$$\begin{aligned}|0_L\rangle = \tfrac{1}{4}[&|00000\rangle + |10010\rangle + |01001\rangle + |10100\rangle \\ &+ |01010\rangle - |11011\rangle - |00110\rangle - |11000\rangle \\ &- |11101\rangle - |00011\rangle - |11110\rangle - |01111\rangle \\ &- |10001\rangle - |01100\rangle - |10111\rangle + |00101\rangle].\end{aligned} \quad (70)$$

To prove this, we show that for each w, an ensemble used in the protocol gives the capacity (55). Equality in Eq. (54) then becomes a simple consequence of relation Eq. (61).

Assuming the knowledge of w, the output entropy of the channel is equal to 0 for each state transmitted in the described protocol. Hence, we have to check if, under the condition of w, the mean output state entropy $S(\rho_w)$ reaches $\sum_i \min[2\,l_i(w),m]$. Since senders B_i are uncorrelated, we can focus only on one sender B_i and consider only the value l_i. Let the set e contain all positions where the state coming from sender B_i was affected by the channel Φ ($|e| = l_i$). We denote by $\mathcal{E}^k(\rho)$ the completely depolarizing channel acting on the kth qubit of the state ρ. For a given e, part of the mean output state coming from B_i has the form $\rho_e = (\bigotimes_{j=1}^{l_i} \mathcal{E}^{(e_j)})[|\phi\rangle\langle\phi|]$ where e_j denotes the jth element of e and $|\phi\rangle$ is $|0\rangle$, $|\Phi^+\rangle$ or $|0_L\rangle$. The condition of whether $S(\rho_e) = \min(2l_i,m)$ occurs for all e was checked numerically. The program enumerated all

e's, then for each e it computed the state ρ_e and its entropy $S(\rho_e)$. Obtained results confirmed that $S(\rho_e) = \min(2l_i, m)$ for $|0\rangle$, $|\Phi^+\rangle$ and $|0_L\rangle$.

In the presented protocol, entanglement increases diversity of the mean output state. An important feature of the code state from the five-qubit correction code is that the increase of entropy of the output state depends only on the number of qubits affected by the unitary. It does not depend on a localization of the affected qubits. We cannot exceed m bits of entropy per state; hence, the closer l_i is to m the smaller is the entropy increase. Due to the asymptotical equipartition property, for $n > 1$ the larger the entanglement in the state, the smaller is the chance that l_i will be close to the m.

The above analysis opens a possibility of further study of a type of entanglement which is the best in a case of various channels. The possible classification of noise with respect to classes of entanglement seems especially interesting. For instance, one can ask whether there are any channels for which cluster-type entanglement is the best in saturating the asymptotic rates of the channel. We leave those questions for further research.

E. Supperadditivity of the regularized capacity

We now turn to the study of the supperadditivity effect for the regularized capacity. We investigate a setup which consists of two channels of the type already described in Sec. III D. For the channel Φ_I we choose $n = 10, n' = 9$ and for the channel Φ_{II} we choose $n = 10, n' = 1$. We are interested in maximal transmission rate from sender A, that is, the case when all senders B_i help sender A by transmitting the same state all the time. Senders B_i transmits with rates equal zero. Formally we can include senders B to the environment and view channels Φ_I and Φ_{II} as 1-to-1 channels.

First, we show that for channels Φ_I and Φ_{II}, the upper bounds C_A for the rates achievable by sender A fulfill $C_A^{(\infty)} > C_A^{(1)}$. For this, we consider a protocol where senders B_i transmit one of the Bell states and show that this protocol achieves a regularized rate strictly greater than $C_A^{(1)}$. Calculation is performed for general sizes of the set of selected helper-senders equal to n'. The single-shot capacity is given by the joint dimensionality of the states of the selected helper-senders and it reads $C_A^{(1)} = n'$. In case of two uses of the channel and Bell states transmission, the probability that the same set of selected helper-senders was chosen twice is $p = 1/\binom{n}{n'}$. With probability $1 - p$ sets of the selected helper-senders in the first and second uses of the channel differ in at least one sender B_i. It means that the input Bell state of two senders, say B_1 and B_2, was affected by the channel only once and, due to the dense coding, the states carry full information from appropriate two-qubit input lines of A (lines 1 and 2). In this case sender A takes advantage of a transmission of an additional 2 bits of information. Under condition of output label w, the output entropy of the channel is 0; therefore, the rate achievable by the protocol for given w is equal to entropy of the mean output state (strictly speaking, entropy of the part coming from senders B). As usual, sender A transmits with equal probability all states from the standard basis. From Eq. (61), the rate achievable by sender A is at least $R_A = p2n' + (1-p)(2n'+2) = 2[n' + (1-p)]$. This leads to $C_A^{(\infty)} \geqslant (1/2)R_A = n' + (1-p) > n' = n' \geqslant C^{(1)}$.

Now we pass to the supperadditivity of the regularized capacities [cf. Eq. (51)]. We show that $C_A^{(1)}(\Phi_I \otimes \Phi_{II}) > C_A^{(\infty)}(\Phi_I) + C_A^{(\infty)}(\Phi_{II})$. We first provide upper bounds for $C_A^{(\infty)}(\Phi_I)$ and $C_A^{(\infty)}(\Phi_{II})$.

Recall that entangled states can be transmitted only through the inputs controlled by the same users. Channel capacity is upper bounded by the minimum value of the logarithm of its input and output spaces. Therefore, for the channel $\Phi_{n,n'}$, we have $C_A^{(m)} \leqslant 1/m \min(2n'm, nm) = \min(2n', n)$ and it leads to $C_A^{(\infty)}(\Phi_I) \leqslant \min(2 \times 9, 10) = 10$ and $C_A^{(\infty)}(\Phi_{II}) \leqslant \min(2 \times 1, 10) = 2$.

Now we move to the case $\Phi_I \otimes \Phi_{II}$, that is, the case where entanglement between the inputs of the channels Φ_I and Φ_{II} controlled by the same user is allowed. One can use the following protocol to provide a lower bound for $C_A^{(1)}(\Phi_I \otimes \Phi_{II})$: Sender A only uses the inputs of Φ_I and sends each state from the standard basis of the input space of Φ_I with the same probability; through channel Φ_{II}, he sends only one chosen state $|00\rangle$ all the time. It is easy to see that the channel Φ_{II} does not change the states coming from senders B_i and in fact it can be seen as an identity channel. Senders B_i send one chosen Bell state $|\Phi^+\rangle$. The first qubit of the Bell state goes through channel Φ_I while the second through channel Φ_{II}. If the qubit is affected by the channel Φ_I, the dense coding scheme is reproduced. Each time the setup $\Phi_I \otimes \Phi_{II}$ is used, all the lines controlled by A find as a target different Bell states. Therefore, the rate achieved by the protocol is given by the dimensionality of input space of the channel Φ_I controlled by sender A and reads 18 bits. It is a lower bound for $C_A^{(1)}(\Phi_I \otimes \Phi_{II})$ and shows that $C_A^{(\infty)}(\Phi_I \otimes \Phi_{II}) \geqslant C_A^{(1)}(\Phi_I \otimes \Phi_{II}) \geqslant 18 > 12 \geqslant C_A^{(\infty)}(\Phi_I) + C_A^{(\infty)}(\Phi_{II})$ and proves that the supperadditivity effect indeed occurs.

IV. QUANTUM GAUSSIAN MACS

We now consider the capacity properties of Gaussian multiaccess channels. Before going further, we first collect certain basic notions and definitions that will be subsequently useful.

Recall first the concept of classical Gaussian multiaccess channels [11]. Inputs and outputs of classical CV Gaussian MACs are real numbers. A Gaussian MAC models the influence of an additive Gaussian noise Z (with a variance N) on the total input signal; that is, the output is

$$Y = \sum_i X_i + Z. \tag{71}$$

To prevent unphysical infinite capacities, power constraints are imposed on the input signals $\langle X_i^2 \rangle \leqslant P_i$. Under these constraints, the capacity region for a classical Gaussian MAC is given by [11]

$$\sum_i R_i \leqslant C\left(\sum_i P_i/N\right), \tag{72}$$

where $C(x) = 1/2 \log(1+x)$.

For a *quantum* Gaussian MAC, the input and the output spaces are described by infinite dimensional Hilbert spaces, isomorphic to those describing a finite number of bosonic modes [13]. The latter are equipped with the "position" and "momentum" canonical observables $\{\hat{x}_1,\ldots,\hat{x}_n,\hat{p}_1,\ldots,\hat{p}_n\}$ fulfilling the standard commutation relations $[\hat{x}_i,\hat{p}_j]=\mathrm{i}\delta_{i,j}$, where i,j enumerate the modes of the system. States of a bosonic system can be expressed in terms of characteristic functions $\chi_\rho(\xi)=\mathrm{Tr}[\rho W_\xi]$, where $W_\xi=\exp(-\mathrm{i}\xi^T R)$ is the so-called Weyl operator and $\hat{R}=(\hat{x}_1,\hat{p}_1,\ldots,\hat{x}_n,\hat{p}_n)^T$ is the vector of canonical observables [13,14]. Gaussian states are the states whose characteristic functions are Gaussian:

$$\chi(\xi)=\exp\left[-\tfrac{1}{4}\xi^T\gamma\xi+id^T\xi\right], \tag{73}$$

where d is the displacement vector [with $d_j=\mathrm{tr}(\rho\hat{R}_j)$] and γ is the covariance matrix with entries $\gamma_{jk}=2\mathrm{tr}[\rho(\hat{R}_j-d_j)(\hat{R}_k-d_k)]-iJ^{(n)}_{jk}$ that completely define the Gaussian state. $J^{(n)}$ is the symplectic form for the multimode system:

$$J^{(n)}=\bigoplus_{i=1}^{n}J,\quad J=\begin{pmatrix}0&1\\-1&0\end{pmatrix}. \tag{74}$$

Gaussian channels are defined as mappings that transform Gaussian states into Gaussian states. They can be expressed as transformations of γ and d:

$$\gamma\mapsto X\gamma X^T+Y, \tag{75}$$

$$d\mapsto Xd. \tag{76}$$

Complete positivity of the channel is guaranteed by the condition

$$Y+iJ-iX^TJX\geqslant 0. \tag{77}$$

We now show how to determine X,Y for an arbitrary Gaussian channel Φ. Recall that the action of any general channel is given by: $\Phi(\rho_s)=\mathrm{tr}_e[\hat{U}(\rho_s\otimes\rho_e)\hat{U}^\dagger]$, where $\hat{U}=\exp(-\mathrm{i}\hat{H})$ is a unitary operation generated by a Hamiltonian $\hat{H}$. Gaussian channels are generated by Hamiltonians $\hat{H}$ that are quadratic in the canonical operators: $\hat{H}=\mathrm{i}\hat{R}^T h\hat{R}$, where h is a $2n\times 2n$ Hermitian matrix [15]. Here ρ_s is the input state and ρ_e is a state of the environment. Now, for Gaussian channels, both ρ_s and ρ_e are Gaussian states with covariance matrices γ_s, γ_e and displacement vectors d_s, d_e, respectively. The displacement of the output state depends linearly on d_e. As any displacement of output states by a constant vector is a unitary operation and as such it does not influence the channel capacity, we assume that $d_e=0$. The action of $\hat{U}$ on the canonical observables can be identified with the linear transformation $\hat{U}^\dagger\hat{R}^T\hat{U}=M\hat{R}^T$. Now we express M in block form with respect to the system-environment partition:

$$M=\begin{pmatrix}M_{ss}&M_{se}\\M_{es}&M_{ee}\end{pmatrix}.$$

From the latter one obtains $X=M_{ss}$ and $Y=M_{se}\gamma_e M_{se}^T$.

Finally, note that in the context of quantum Gaussian channels, power constraints are usually expressed as a limitation on a mean number of photons transmitted per channel use.

Squeezed states represent an important class of Gaussian states for communication tasks. A one-mode squeezed state saturates the Heisenberg uncertainty principle, with lower quantum noise (variance) in one of the quadratures as compared with a coherent state. In the photon number basis a one-mode vacuum squeezed state has the following form:

$$|\zeta;0\rangle=\sqrt{\mathrm{sech}\,r}\sum_{n=0}^{\infty}\frac{\sqrt{(2n)!}}{n!}\left[-\frac{1}{2}e^{i\phi}\tanh r\right]^n|2n\rangle, \tag{78}$$

where r is the squeezing parameter. In terms of the covariance matrix formalism, the $\phi=0$ squeezed vacuum state is described by

$$\gamma=\begin{pmatrix}e^{-2r}&0\\0&e^{2r}\end{pmatrix} \tag{79}$$

and the displacement vector $d=0$. Displacing a squeezed vacuum state using the displacement operator $D_{\tilde{d}}$ leads to a state with an unchanged covariance matrix but with the displacement vector $d=\tilde{d}$. In the two-mode case, we utilize the two-mode squeezed vacuum state, with squeezing of the relative position x_1-x_2 and the total momentum p_1+p_2. The covariance matrix of this state takes the form [15]

$$\gamma=H^T\mathrm{diag}(e^{-2r},e^{2r},e^{2r},e^{-2r})H, \tag{80}$$

where

$$H=\frac{1}{\sqrt{2}}\begin{pmatrix}1&0&-1&0\\1&0&1&0\\0&1&0&-1\\0&1&0&1\end{pmatrix}, \tag{81}$$

while the displacement vector $d=0$.

Last, for a calculation of channel capacities we need the entropy of n-mode Gaussian states ρ. This is given by the formula [15]

$$S(\rho)=\sum_{j=1}^{n}g\left(\frac{v_j-1}{2}\right) \tag{82}$$

in terms of normal modes of the system. Here $g(x)=(x+1)\ln(x+1)-x\ln x$ is the entropy of a normal mode with average occupation number x. The v_j's are the symplectic eigenvalues of the covariance matrix γ corresponding to the state ρ, that is, the square roots of the eigenvalues of the matrix $-J^{(n)}\gamma J^{(n)}\gamma$. (Note that the symplectic spectrum for each mode is doubly degenerate and that in the entropy formula each value is taken only once).

A. Locality rule for classical Gaussian MACs

The analysis of capacity regions is more intricate in the CV case than in the DV case, already for classical channels. This is intimately related to the fact that the capacities are dependent on power constraints which may lead to various scenarios. To see this, consider an example of two 1-to-1 classical channels Φ_1 and Φ_2 with noise levels N_1 and N_2 and the same power constraints $\tilde{P}$. We assume that $N_1<N_2$. Suppose each channel works separately; then Φ_i achieves the capacity $\tilde{C}_i=\frac{1}{2}\log[1+\tilde{P}/N_i]$ [11]. Now suppose the channels work in a parallel setup. The sender aims to maximize the total capacity $C_T=C_1+C_2=\frac{1}{2}(\log[1+P_1/N_1]+\log[1+P_2/N_2])$, where P_i is the power allocated to the channel Φ_i. One demands that the total power available to the user in this case is identical to the total power used when the channels were utilized separately; that is, P_1 and P_2 obey the constraint $P_1+P_2\leqslant 2\tilde{P}$. Now since the noise levels N_1,N_2 are different,

the senders can increase the total capacity by allocating more power in the transmission through the channel with the lower level of noise. When $N_1 + 2\tilde{P} < N_2$, the optimal choice is to put $P_1 = 2\tilde{P}$. In the other case, the optimal allocation is determined from the relation $N_1 + P_1 = N_2 + P_2$. Using this power redistribution, the sender can achieve capacity $\tilde{C} > C_1 + C_2$. This process of optimization is the so-called water-filling scheme (see, e.g., [11]).

Thus, we see that for Gasussian channels the additivity theorem of Sec. III A cannot be stated as such. However, observe that the local rates depend only on the local power constraints [cf. Eq. (72)]. This means that, in a multiuser scenario, adding a resource (channel or energy) to one sender never helps the others beat their maximal achievable rates (power constraints pertaining to different users are not allowed to be combined; hence, no interuser water-filling effect can take place). We call this observation the locality rule for classical Gaussian MACs [7] and treat it as the appropriate analog of the additivity rule for classical capacities for discrete channels.

B. A comparison of strategies which lead to increased transmission rates for Gaussian state encoding and homodyne detection

We first study the setup presented in Fig. 8(a). It consists of a channel Φ_θ, which is an asymmetrical beam spliter with two one-mode input lines and a one-mode output. The channel Φ_θ works in parallel with a one-mode ideal (identity) channel $\mathcal{I}$. Sender A has access to the input A of channel Φ_θ and sender B to the input B of channel Φ_θ and B' of channel $\mathcal{I}$. The signals from the input lines A and B are mixed in channel Φ_θ on the beam splitter, which has transitivity $T = \cos^2\theta$. The receiver has access to only one output mode of the beam splitter, the second being blocked. The channel Φ_θ is thus characterized by a loss $N_A \cos^2\theta$ of the power input N_A through line A.

We place the following power constraints, expressed in terms of mean number of photons used by the senders:

$$\{\text{sender } A \text{ average number of photons }\} = N_A, \quad (83)$$

$$\{\text{sender } B \text{ average number of photons}\} = N_B. \quad (84)$$

We now compare how the choice of quantum states used in the communication protocol influences the transmission rate $R_A(\Phi_\theta \otimes \mathcal{I})$ for sender A, while sender B is a helper-sender all the time and his rate is $R_B(\Phi_\theta \otimes \mathcal{I}) = 0$. We also point out cases where the locality rule is broken. We consider only transmission of Gaussian states. We focus on the following protocols.

(1) Senders A and B transmit coherent states. A encodes messages as displacements of both canonical variables of the vacuum state while the probability of displacement is chosen, as is standard, to be a Gaussian distribution $p(x,p) = \frac{1}{2\pi\sigma^2}\exp(-\frac{x^2+p^2}{2\sigma^2})$ with $\sigma^2 = 2N_1$. Sender B transmits a fixed chosen coherent state all the time. The receiver performs homodyne detection on both quadratures to decode the message. This is a typical setup for transmission of information through optical fibers [16]. The achievable rate depends only on the output power corresponding to user A and reads

$$R_A^{\text{coherent}} \leqslant \log(1 + N_A \sin^2\theta). \quad (85)$$

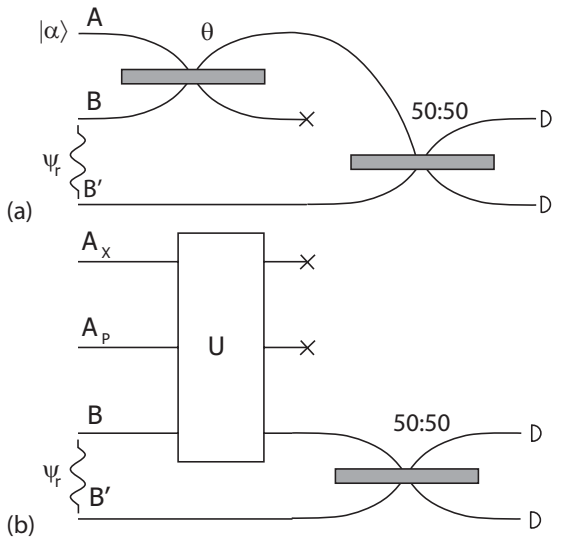

FIG. 8. Gaussian channels exhibiting violation of locality rule on individual transmission rates. (a) Beam-splitter channel Φ_θ; (b) triple QND sum gate channel Φ. The channels were first presented in Ref. [7].

This rate refers to the case of a lossy channel with transmitivity $T = \sin^2\theta$ in case when sender performs encoding in coherent states and receiver performs homodyne detection on the output [17]. It depends only on the power constraints for A and manifestly obeys the locality rule.

(2) Senders A and B use single-mode squeezed vacuum states. Both users transmit states which are squeezed in the same canonical variable, say x. A encodes his message in the displacement of the variable x of his state, whose value is Gaussian distributed with variance σ_x^2. The receiver performs homodyne detection only on x. This setup was studied in [18]. The rate is given by

$$R_A^{\text{squeezed}} = \frac{1}{2}\log\left[1 + \frac{\sigma_x^2 \sin^2\theta}{e^{-2R}\sin^2\theta + e^{-2r}\cos^2\theta}\right], \quad (86)$$

where R and r denote the squeezing parameters of the x quadrature for senders A and B, respectively. The energy constraints can be written as $\sigma_x^2 \leqslant 4(N_A - \sinh^2 R)$, $\sinh^2 r \leqslant N_B$. User A performs optimization of the parameter R; that is, he optimizes the power allocation between the squeezing and the mean-square displacement. For fixed θ, in the limit $N_A \to \infty, N_B \to \infty$ we get asymptotically

$$R_A^{\text{squeezed}} = \log[1 + N_A]. \quad (87)$$

(3) Sender A again sends coherent states, encoding his message in the displacement of both canonical variables. The displacement has a probability density distribution as in the case (IV B). Sender B transmits a two-mode squeezed state, one mode through Φ_θ and the second one through the extra resource $\mathcal{I}$. The receiver has access to the output of Φ_θ and $\mathcal{I}$. The decoding consists of a joint measurement of the canonical variables $x_{\Phi_\theta} - x_I$ and $p_{\Phi_\theta} + p_I$ on the output modes of the setup. To achieve this, the output modes of Φ_θ and $\mathcal{I}$ are mixed on a 50:50 beam splitter followed by homodyne measurements of x_1 and p_2 on the output modes of the 50:50 beam splitter.

In this setup, sender B is assumed to make use of entangled states. The formula for the transmission rate is now given by

$$R_A^{\text{ent}} = \log\left[1 + \frac{\sigma^2 \sin^2\theta}{2(\cosh r - \cos\theta \sinh r)^2}\right]. \quad (88)$$

Here $\sigma^2 = 2N_A$ is the variance in the displacement of canonical variables in A's mode and r denotes the squeezing parameter of the two-mode squeezed state sent by B. The imposed power constraints imply that $\sinh^2 r = N_B/2$. For a given N_A, the optimal values of (88) lie on a curve,

$$\cos\theta = \tanh r, \quad (89)$$

which leads to the following maximal rate formula:

$$R_A^{\text{ent-opt}} = \log[1 + N_A]. \quad (90)$$

$R_A^{\text{ent-opt}}$ is, in fact, equal to the rate achievable by a one-mode ideal channel when the sender performs encoding in coherent states and the receiver performs homodyne detection on the output. Thus, entanglement can be used to completely overcome the power loss in the case of coherent state encoding. The same effect is also obtained in the second case, without entanglement, as described above, but only in the asymptotic regime of infinite power [see Eq. (87)].

It is indeed worth noting, for comparison, that the limit $N_A \to \infty, N_B \to \infty$ of Eq. (86) under the constraint Eq. (89) leads to

$$R_A^{\text{squeezed}} \leqslant \tfrac{1}{2}\log[1 + 16N_A] \approx \tfrac{1}{2}\log[1 + N_A] = \tfrac{1}{2}R_A^{\text{ent-opt}}. \quad (91)$$

Comparison of this result with Eq. (87) shows that one-mode squeezed states transmission requires much higher squeezing to reach the rates achievable by the two-mode squeezed-state transmission.

We can also calculate two upper bounds $R_A(\Phi_\theta)$ for the transmission rates only through the channel Φ_θ.

(1) A bound based on the maximal entropy of a state with a mean number of photons equal to the mean number of photons in the output mode of the channel Φ_θ. We refer to it as to the *output entropy bound*. This tells us how large a rate is achievable if no entanglement is allowed in the communication protocol and is given by

$$R_A^{\text{prod-bound}} = g(N_{\text{out}}) \quad (92)$$

$$= g(\sqrt{N_A}\sin^2\theta + \sqrt{N_B}\cos^2\theta). \quad (93)$$

(2) A bound based on the maximal entropy of a state with a mean number of photons equal to the mean number of photons in the input mode A of the channel Φ_θ. This may be referred to as the *input entropy bound*. This bound cannot be violated by any type of communication protocol, entanglement-free or entanglement-aided, and it tells us how much information can be transmitted with given energy constraints if sender A is connected to the receiver by a one-mode ideal line. We check how close the protocols described above approach this bound, which is given by

$$R_A^{\max} \leqslant g(N_A). \quad (94)$$

The bounds $R_A^{\text{prod-bound}}$ and $R_A^{\max}$ allow us to express the theoretical maximum rate for sender A in the form $\min(R_A^{\text{prod-bound}}, R_A^{\max})$.

Figure 9 shows the behavior of the rates achievable by the different encoding schemes and parameter regimes as a function of the energy constraint N_B for sender B [the schemes and bound correspond directly to the points (1–5) in the main text]. Figure 9(a) represents the situation where using entangled states quickly becomes more efficient than using any product state encoding, while on the other hand Fig. 9(b) represents a situation where entanglement cannot beat the upper bound for the rates achievable by the product states encoding. In Fig. 9(c) we consider the behavior of rates for a large range of values of N_B. We can observe that in the low N_B range, the strategy using entanglement states is the best among the three considered approaches. However, increasing N_B leads to a maximal value of R_A^{ent} after which further growth leads to a diminishing rate. This can be explained by increasing of the entanglement of the output state with the erased mode. In the case of R_A^{squeezed} the situation looks different. It was shown in Ref. [18] that in the limit $N_A \to \infty, N_B \to \infty$ the rate R_A^{squeezed} asymptotically approaches the upper bound for the transmission rate for sender A expressed by $R_A^{\max}$. It has to be stated again here that this strategy requires extremely high squeezing to approach the maximal rate achieved by the protocol using a two-mode squeezing scenario.

In case of the protocol using two-mode squeezed states it is interesting to ask about the lower limit of squeezing for which the rate achieved starts to be higher than the rate achieved by any protocol based only on product-state encoding. Figure 10(a) presents demarcation curves $R_A^{\text{ent}}/R_A^{\text{prod-bound}} = 1$ in the $\theta - N_B$ parameter plane for three different values of the parameter N_A. For a fixed N_A, with increasing N_B, we move above the demarcation curve and fall into the area where $R_A^{\text{ent}} > R_A^{\text{prod-bound}}$. The minimal mean photon number in the entangled state, required to approach this area, amounts to around $N_B = 1, 0.6, 0.55$ for $N_A = 10^3, 10^6, 10^9$. These values of N_B refer to the following squeezing levels, which are experimentally realistic: 5.72 dB, 4.55 dB, 4.37 dB. The demarcation curve is crossed as θ equals 0.28, 0.1, 0.02 or transmissivity 0.077, 0.01, 0.0004. For a large N_A, the locality rule is broken for $\theta \approx 0$. In this regime the setup reproduces the CV dense coding scheme. $N_B = 1$ means that we use a two-mode squeezed state with squeezing equal to 5.72 dB, which is a reasonable value for experimental setup. In Fig. 10(b). we change the scale of observation and show that breaking of the locality rule occurs for quite a large range of the parameter N_B and θ.

C. Realization of *XP* gate by linear optics and one-mode squeezed states: Influence of noise on the superadditivity effect

In this section, we start with details of realization of the three input quantum nondemolition channel Φ presented in Fig. 8(b). This leads naturally to a discussion of the interplay between noise (or imperfections) and superadditivity.

The channel $\Phi : A_X A_P B \mapsto R$ acts as follows: $\Phi(\rho_{A_X A_P} \otimes \rho_B) = \text{tr}_{A_X A_P}[\hat{U}(\rho_{A_X A_P} \otimes \rho_B)\hat{U}^\dagger]$. Sender A holds lines A_X

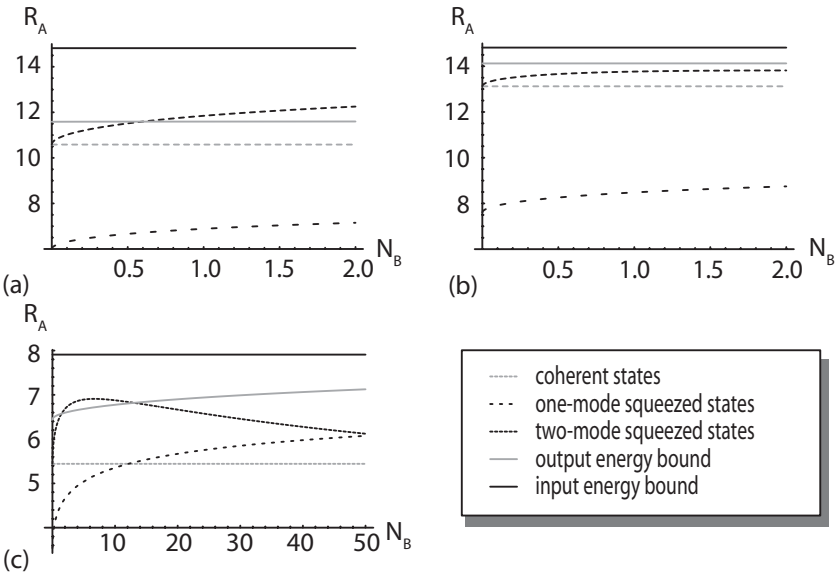

FIG. 9. Rates which are achievable by the different encoding schemes for the beam-splitter channel Φ_θ as a function of the energy constraint N_B for sender B. The rates are evaluated for the following values of θ and the energy constraints N_A for sender A: (a) $\theta = \pi/4, N_A = 10^6$; (b) $\theta = 0.2, N_A = 10^6$; (c) $\theta = 0.5, N_A = 10^3$.

and A_P, while sender B holds line B. $\hat{U}$ is a unitary operator of the form $\hat{U} = \exp[-i(\hat{x}_X \hat{p}_B - \hat{p}_P \hat{x}_B)]$, which can be factorized as follows:

$$\hat{U} = \exp[-i(\hat{x}_X \hat{p}_B - \hat{p}_P \hat{x}_B)] \tag{95}$$

$$= \exp\left[\frac{i}{2}\hat{x}_X \hat{p}_P\right] \exp[-i\hat{x}_X \hat{p}_B] \exp[i\hat{p}_P \hat{x}_B]. \tag{96}$$

The XP interaction, appearing here, can be obtained by measurement-induced CV quantum interactions as described in Ref. [19]. An experimental proof of concept has been presented in Ref. [20].

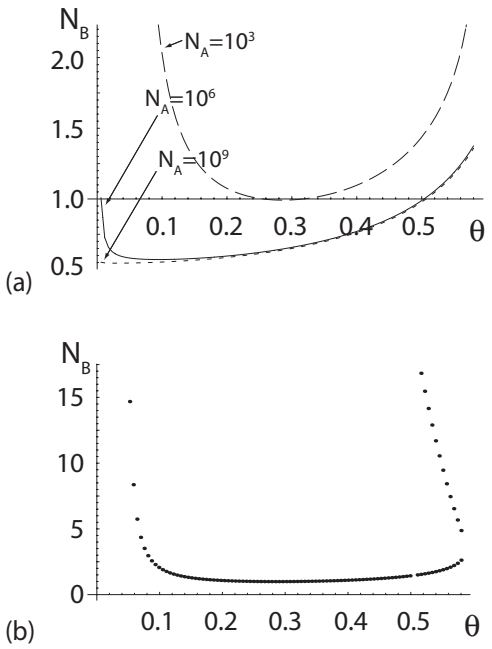

FIG. 10. (a) Lines present bounds of the areas where $R_A^{\text{ent}} > R_A^{\text{prod-bound}}$. They are given by the condition $R_A^{\text{ent}}/R_A^{\text{prod-bound}} = 1$. Lines refer to the cases of $N_A = 10^3, 10^6, 10^9$. For a given N_A the superadditive region lies above the corresponding line. (b) The dotted line delimits the area where $R_A^{\text{ent}} > R_A^{\text{prod-bound}}$ for $N_A = 10^3$. Notice that the superadditive region is considered in a large scale of N_B. The figure shows that there is large window of parameters N_B, θ for which $R_A^{\text{ent}} > R_A^{\text{prod-bound}}$ can be observed.

The superadditive effect of the single user capacity (breaking of the locality rule) for this channel was considered by us in [7], where the XP interaction was assumed to be ideally implemented. However, the method of measurement-induced CV quantum interactions is, in practice, imperfect and introduces errors in the output states. To study such errors it is useful to write down how canonical observables are transformed by the realization of the XP gate [19,20]:

$$\hat{x}_1^{\text{out}} = \hat{x}_1^{\text{in}} - \sqrt{\alpha}\hat{x}_0 - \sqrt{\beta}\hat{x}_{S_1}, \tag{97}$$

$$\hat{p}_1^{\text{out}} = \hat{p}_1^{\text{in}} - \frac{1-T}{\sqrt{T}}\hat{p}_2^{\text{in}} + \sqrt{\alpha/T}\,\hat{p}_0 + \sqrt{T\beta}\,\hat{p}_{S_2}, \tag{98}$$

$$\hat{x}_2^{\text{out}} = \hat{x}_2^{\text{in}} + \frac{1-T}{\sqrt{T}}\hat{x}_1^{\text{in}} - \sqrt{\alpha/T}\,\hat{p}_0 + \sqrt{T\beta}\hat{x}_{S_1}, \tag{99}$$

$$\hat{p}_2^{\text{out}} = \hat{p}_2^{\text{in}} - \sqrt{\alpha}\hat{p}_0 + \sqrt{\beta}\,\hat{p}_{S_2}, \tag{100}$$

where $\alpha = (1-T)(1-\eta)/(1+T)\eta, \beta = (1-T)/(1+T)$, $\hat{x}_{S_1}, \hat{p}_{S_2}$ are canonical observables of two different modes in squeezed states, η is efficiency of the homodyne detectors inside the XP gate realization, and $\hat{x}_0, \hat{p}_0$ are canonical observables of two different modes in the coherent states used that homodyne detectors. Parameter T depends on the configuration of the XP gate realization and can be manipulated. Choosing $T = \frac{1}{2}(3-\sqrt{5})$, the coefficients of $\hat{p}_2^{\text{in}}$ and $\hat{x}_1^{\text{in}}$ in Eqs. (98) and (99) become -1 and 1, respectively. We hereby represent this XP gate realization as a quantum noisy channel described by transformation matrices X, Y, using the method outlined in Sec. IV. Here we assume that errors introduced by linear optical elements can be neglected in comparison with natural noise due to the physical generation of highly squeezed states. Then

$$X = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad Y = \begin{pmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_2^2 & 0 \\ 0 & 0 & 0 & \sigma_1^2 \end{pmatrix}, \tag{101}$$

with $\sigma_1^2 = \alpha + \beta e^{-2s}, \sigma_2^2 = \alpha/T + \beta T e^{-2s}$. We assumed that squeezed states in both modes have the same squeezing level. This gate reproduces the ideal XP gate in the limit of infinite squeezing $s \to \infty$ and ideal homodyne detection $\eta \to 1$. If all XP gates used in the implementation of the considered channel have the same parameters, we can collect all noise components under the common factor $\sigma^2_{\text{noise}} = \sigma_1^2 + \sigma_2^2$.

Now suppose that sender B also has access to the input B' of a one mode ideal channel $\mathcal{I}$ [see Fig. 8(b)]. We are interested in the maximal rate $R_A^{(1)}(\Phi \otimes \mathcal{I})$ for sender A under the following protocol [7] when sender A transmits displaced squeezed one-mode vacuum states. States transmitted through line A_X (A_P) are squeezed in the canonical observable $\hat{x}$ ($\hat{p}$), where the squeezing parameter is R. Sender A encodes his message in the displacement of the canonical observables $\hat{x}$ ($\hat{p}$) for line A_X (A_P). As usual, the displacement has a Gaussian distribution with variance σ^2. Sender B continuously transmits a constant fixed two-mode squeezed vacuum state, with squeezing parameter r. One mode is transmitted through the B and the other through line B'. The receiver performs joint homodyne detection of $x_B - x_{B'}$ and $p_B + p_{B'}$ on the output of the channels Φ and $\mathcal{I}$ to decode the message. The rate in this case is now calculated to be

$$R_A^{(1)} = \log\left(1 + \frac{\sigma^2}{e^{-2R} + \frac{e^{-2r}}{2} + \frac{\sigma^2_{\text{noise}}}{2}}\right). \tag{102}$$

The imperfections in the implementation of the desired unitary evolution appear in the form of the extra noise term $\sigma^2_{\text{noise}}/2$ in the expression for maximal rate, as compared to the ideal case described in [7].

For a more realistic description, we model the influence of various unavoidable imperfections, associated, for example, with the implementation of displacement during encoding, the measurement process realized by the receiver for decoding, and the interaction with the environment at a nonzero temperature, by thermal noise channels. This type of a channel is a 1-to-1 lossy channel mixing an input state with a thermal state $\rho_{N_{Th}}$ containing an average of N_T photons, at a beam splitter with transmissivity $T = \cos^2\omega$. The receiver receives only one of the output modes of this mixing beam splitter. The covariance matrix γ of an input state is transformed by this channel as follows:

$$\gamma \mapsto T\gamma + (1-T)\gamma_{N_{\text{Th}}}, \tag{103}$$

where $\gamma_{N_{\text{Th}}} = N_{\text{Th}}\mathbb{I}$. Below we assume that the XP gate is perfect and put the thermal noise channels parameterized by ω and N_{Th} at two places: between the output of the nondemolition channel Φ and the receiver and between the output of the supporting channel $\mathcal{I}$ and the receiver. Now the transmission rate for the upper sender, using the same protocol as described earlier in the noiseless case, is modified and is calculated here to be

$$R_A^{(1)} = \log\left(1 + \frac{\sigma^2\cos^2\omega}{\left(e^{-2R} + \frac{e^{-2r}}{2}\right)\cos^2\omega + (1+N_{\text{Th}})\sin^2\omega}\right). \tag{104}$$

In a similar way, we now also model effects of noise on the beam-splitter MAC channel Φ_θ discussed earlier [see Fig. 8(a)]) of Sec. IV B. We again place the thermal noise channels (parameterized by ω, N_{Th}) between the output of the Φ_θ channel and the receiver and between the output of the supporting channel $\mathcal{I}$ and the receiver. In this case, we obtain the following rate of the upper sender:

$$R_{A_1} = \log\left(1 + \frac{T\sigma^2\sin^2\theta}{T(\cosh r - \cos\theta\sinh r)^2 + (1-T)(1+2N_{Th})}\right). \tag{105}$$

Here θ is the parameter of the BS channel Φ_θ, $T = \cos^2\omega$ is the transmissivity of the thermal noise channel, and N_{Th} is the mean photon number of the environment. In Fig. 11 we use this result to illustrate how the capacity $R_A^{(1)}(\Phi_\theta \otimes \mathcal{I})$ changes with the parameters of the thermal noise channel. Even if the effect of thermal noise channel is small—its transmissivity is large and $N_{\text{Th}} = 0$—the capacity gain becomes negligible and from $T = 0.85$ no enhancement over the upper bound for rates obtained using product codes is observed. This scenario corresponds to the case where there are only losses in the thermal channel.

Now we use the presented results to discuss the possibility of an experimental verification of the considered superadditive effects in the context of available technological resources

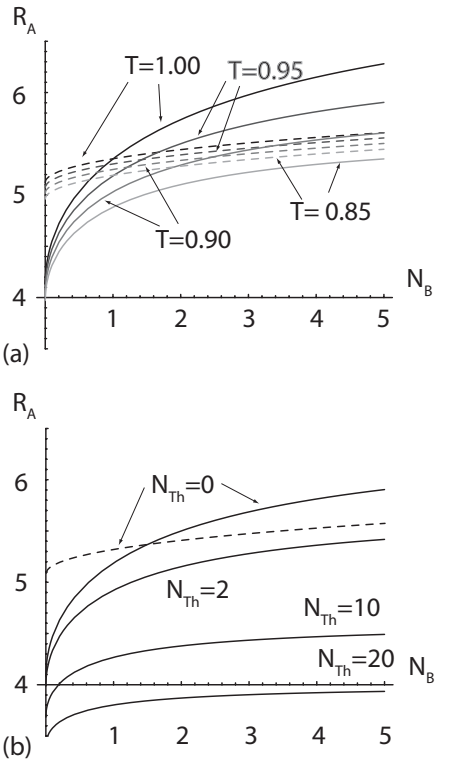

FIG. 11. Influence of thermal noise on capacity $R_A^{(1)}(\Phi_\theta \otimes \mathcal{I})$ of the BS channel Φ_θ working with parameters $\theta = 0.25, N_A = 10^3$. (a) Dependence on the transitivity T of the thermal noise channel, here $N_{\text{Th}} = 0$; rates $R_A^{(1)}(\Phi_\theta \otimes \mathcal{I})$ (solid) are compared with output entropy bound $R_A^{\text{prod-bound}}$ (dashed). (b) Dependency on the mean photon number N_{Th} in the thermal state for $N_{\text{Th}} = 0, 2, 10, 20$ and an exemplary value $T = 0.95$; here we provide output entropy bound $R_A^{\text{prod-bound}}$ only for $N_{\text{Th}} = 0$ since additional thermal noise increases output entropy and makes the bound less tight.

[21,22]. For homodyne detection we assume quantum efficiency at the level $\eta = 99\%$ as in [23] and the dark noise level at 20 dB below the shot noise of the local oscillator. For the power constraint of sender A we assume $N_A = 1000$ and we remain in the regime of linear approximation of homodyne detection (mean number of photons of local oscillator is on the level 4×10^6). Note finally that the highest observed value of a single-mode squeezing [21] is at the level of 10 dB, which corresponds to the mean photon number of 2.025.

We start with a discussion of the setup $\Phi \otimes \mathcal{I}$ in the context of the implementation of the XP gate presented in [20]. In that experiment, squeezing of 5.6 dB and detectors with quantum efficiency $\eta = 98\%$ were used. A quadrature transfer coefficient of $T_P = \mathrm{SNR}_{\mathrm{out}}/\mathrm{SNR}_{\mathrm{in}} = 0.4$ was reported, where $\mathrm{SNR}_{\mathrm{in(out)}}$ is the signal-to-noise ratio for the signal input (output) of the gate. This coefficient T_P can be translated to $\sigma^2_{\mathrm{noise}} = 5$ in the model of the Φ gate. Assuming this value in Eq. (102) and comparing the rate with the output entropy bound leads to the conclusion that the superadditivity effect described here cannot be achieved for these parameters. On the other hand, implementation of the XP gate with the use of 10 dB squeezing leads to $\sigma^2_{\mathrm{noise}} = 0.098$ and moves the transmission into the supperadditivity regime.

For the setup $\Phi_\theta \otimes \mathcal{I}$ the situation is more optimistic. With a realistic loss level of 5% on the optical elements and homodyne detection efficiency as described above one gets $\cos^2 \omega = 0.94, N_T = 0.09$. Our results show that the superadditivity effect can be observed for $\theta = 0.25$ for squeezing upward of 7.8 dB (the mean photon number of 2.1). Thus, one can draw a conclusion that a loop-hole free verification of the superadditivity effect can be done with the present state-of-the-art quantum optical experimental techniques.

Formulas (102), (104), and (105) can be understood in a generic signal-to-noise "phenomenological" scheme as $R_A = \log(1 + \sigma^2_{\mathrm{signal}}/\sigma^2_{\mathrm{noise}})$. The variance $\sigma^2_{\mathrm{signal}}$ describes how spread out are the input states of the sender in phase space and $\sigma^2_{\mathrm{noise}}$ describes the effective noise level associated with measurement of the displacement, which here is the carrier of classical information. Senders can manipulate $\sigma^2_{\mathrm{signal}}$ and $\sigma^2_{\mathrm{noise}}$ by changing energy allocation used for displacement and squeezing, and in this way approach the bound for the channel capacity. Noise introduced by imperfections of elements of the communication setup plays the role of a lower bound for $\sigma^2_{\mathrm{noise}}$ and the user cannot decrease measurement errors below this level.

V. CONCLUSIONS

The superadditivity of classical capacity regions has been previously reported in the case of discrete [6] and CV (Gaussian) [7] cases. Here we have analyzed these problems in more detail. We have been able to show that channel asymmetry is not crucial for the occurrence of the superadditivity effect. Even more interestingly, we have proven explicitly that two-input entanglement is not sufficient in some cases and have provided an example analytically showing that at least five-input entanglement is required. It is interesting that here the five-qubit code error correction code has been used (see [1]) to beat the C-type (classical) multiaccess capacity, which so far was a tool related to Q-type (quantum) bipartite capacity. Moreover, we do not know of any example related to bipartite classical capacity where more than two-input entanglement is needed to achieve the asymptotic bound. In fact, the celebrated effect of breaking of additivity of Holevo function [8] needs two copies of the channel. We believe that our result will inspire the search for requirement of multipartite entanglement for achieving the asymptotic Holevo capacity in the bipartite case. In both bipartite and multiuser cases, this opens the intriguing question concerning which types of multipartite entanglement (bipartite quantum codewords, cluster, Dicke-type, etc.) are the best for achieving asymptotic classical capacities. We leave these types of questions for further research.

In the CV case, we carefully compared different communication scenarios including the one described by Yen and Shapiro [18]. We explicitly incorporated imperfections of the schemes into the calculations. The success of superadditivity depends on the power of light used and may be destroyed by the thermal noise or even by large-enough losses. On the other hand, we found that the condition for two-mode squeezing used for the effect may not be very demanding (4.55 dB). This opens the possibility of an experimental confirmation of the effect in the near future. Again the question of channels that require multipartite CV-type entanglement for gaining quantum advantage in communication protocols appears quite naturally and is an interesting area for further research.

ACKNOWLEDGMENTS

The work was supported by the EU Commission through the QESSENCE project and by the Polish Ministry of Science and Higher Education through Grant No. NN202231937. Part of the work was done in Quantum Information Centre of Gdansk. J.K.K. was also supported by EU project NAMEQUAM.

APPENDIX

Here we prove that capacity regions of discrete classical n-to-1 channels are additive:

$$\mathcal{R}(\Phi_I \otimes \Phi_{II}) = \mathcal{R}(\Phi_I) + \mathcal{R}(\Phi_{II}), \tag{A1}$$

where $\mathcal{R}(\Phi_I) + \mathcal{R}(\Phi_{II}) = \{u_I + u_{II} : u_I \in \mathcal{R}(\Phi_I), u_{II} \in \mathcal{R}(\Phi_{II})\}$. For simplicity it is assumed that both channels, Φ_I, Φ_{II}, have the same number of senders. This situation is easy to obtain by formal extension of the set of senders for one of the channels. The messages transmitted by these additional senders are then always lost. We use short-hand notation $R_S = \sum_{i \in S} R_i$ for vector of rates $R \in \mathbb{R}^n$, where R_i is ith element of R and $S \subseteq E$ is a subset of senders E. The proof for the n-to-1 case follows the same principles as for 2-to-1 channels. Here we provide only the parts that are distinct from the latter case.

We start with $\tilde{\mathcal{R}}(\Phi_I \otimes \Phi_{II}) \subseteq \bar{\mathcal{R}}(\Phi_I \otimes \Phi_{II}) \subseteq \mathcal{R}(\Phi_I \otimes \Phi_{II})$. Let us assume that the senders transmit with rates given by a vector $\tilde{R}$. As in the 2-to-1 case, these vectors have to belong to the fixed probability capacity region for input symbol probability distribution $\tilde{p} = p(Q) \prod_i p(X_i^I, X_i^{II} | Q)$. On the basis of $\tilde{p}$, we again construct $\bar{p} = \tilde{p}_I \tilde{p}_{II}$, where $\tilde{p}_I = p(Q^I) \prod_i p(X_i^I | Q^I)$ and $\tilde{p}_{II} = p(Q^{II}) \prod_i p(X_i^{II} | Q^{II})$ are marginal distributions obtained from $\tilde{p}$ by tracing out

proper variables and renaming $Q \mapsto Q^I, Q \mapsto Q^{II}$. Below we provide an upper bound for this region:

$$\begin{aligned}
\tilde{R}_S &\leqslant I(X_S : Y|X_{S^c}, Q) \\
&= H(Y|X_{S^c}, Q) - H(Y|X_S, X_{S^c}, Q) && \text{(A2a)} \\
&= H(Y|X_{S^c}, Q) - H\left(Y^I \middle| X_S^I, X_{S^c}^I, Q^I\right) \\
&\quad - H\left(Y^{II} \middle| X_S^{II}, X_{S^c}^{II}, Q^{II}\right) && \text{(A2b)} \\
&\leqslant H\left(Y^I \middle| X_{S^c}^I, Q^I\right) + H\left(Y^{II} \middle| X_{S^c}^{II}, Q^{II}\right) \\
&\quad - H\left(Y^I \middle| X_S^I, X_{S^c}^I, Q^I\right) - H\left(Y^{II} \middle| X_S^{II}, X_{S^c}^{II}, Q^{II}\right) && \text{(A2c)} \\
&= I\left(X_S^I : Y^I \middle| X_{S^c}^I, Q^I\right) + I\left(X_S^{II} : Y^{II} \middle| X_{S^c}^{II}, Q^{II}\right), && \text{(A2d)}
\end{aligned}$$

where Eq. (A2b) is based on the factorization of conditional probabilities defining the channel action for the product channel, while in Eq. (A2c) we use entropy subadditivity. On the other hand, evaluation of Eq. (18) for input symbols probability distribution $\bar{p} = \tilde{p}_I \tilde{p}_{II} = [p(Q^I) \prod_i p(X_i^I|Q^I)][p(Q^{II}) \prod_i p(X_i^{II}|Q^{II})]$ leads to the region

$$\begin{aligned}
\bar{\mathcal{R}} = \big\{ & R \in \mathbb{R}^n : \forall_{i \in E} R_i \geqslant 0, \\
& \forall_{S \subseteq E}\ R_S \leqslant I\left(X_S^I : Y^I \middle| X_{S^c}^I, Q^I\right) \\
& + I\left(X_S^{II} : Y^{II} \middle| X_{S^c}^{II}, Q^{II}\right) \big\}. \qquad \text{(A3)}
\end{aligned}$$

Combining this result with the bound from Eq. (A2d), it is easy to see that $\tilde{\mathcal{R}}(\Phi_I \otimes \Phi_{II}) \subseteq \bar{\mathcal{R}}(\Phi_I \otimes \Phi_{II})$ holds.

It remains to be shown now that

$$\bar{\mathcal{R}}(\Phi_I \otimes \Phi_{II}) = \tilde{\mathcal{R}}(\Phi_I) + \tilde{\mathcal{R}}(\Phi_{II}), \qquad \text{(A4)}$$

where again $\tilde{\mathcal{R}}(\Phi_I)$ $[\tilde{\mathcal{R}}(\Phi_{II})]$ is evaluated for the marginal probability distribution $\tilde{p}_I$ $[\tilde{p}_{II}]$. The following argument is based on a fact that the fixed probability capacity region [cf. Eq. (16)] is a polymatroid [24].

Definition 1. Let $E = \{1, \ldots, n\}$ and $f : 2^E \mapsto \mathbb{R}_+$ be a set function (i.e., a function that maps subsets of E into $\mathbb{R}_+$. The polyhedron,

$$B(f) = \{x \in \mathbb{R}^n : \forall_{S \subseteq E} x_S \leqslant f(s), \forall_{i \in E} x_i \geqslant 0\}, \qquad \text{(A5)}$$

is a polymatroid if the set function f satisfies (i) $f(\emptyset) = 0$, (ii) $S \subseteq T \Rightarrow f(S) \leqslant f(T)$, (iii) $f(S) + f(T) \geqslant f(S \cap T) + f(S \cup T)$.

Lemma 1. The fixed probability capacity region [cf. Eq. (16)] is a polymatroid.

Proof. Observe that the conditional mutual information $I(X_S : Y|X_{S^c}, Q)$ plays a role of the set function $f(S)$ in the above definition. All we have to do now is to check conditions (i)–(iii) defining the polymatroid. By definition, if there is no sender, mutual information is equal to 0, which proves (i). Now let us write

$$\begin{aligned}
f(T) &= I(X_T : Y|X_{T^c}, Q) && \text{(A6a)} \\
&= H(Y|X_{T^c}, Q) - H(Y|X_T, X_{T^c}, Q) && \text{(A6b)} \\
&\geqslant H(Y|X_{S^c}, Q) - H(Y|X_S, X_{S^c}, Q) && \text{(A6c)} \\
&= I(X_S : Y|X_{S^c}, Q) && \text{(A6d)} \\
&= f(S), && \text{(A6e)}
\end{aligned}$$

where in Eq. (A6c) we use the fact that additional information reduces entropy $(S \subseteq T \Rightarrow T^C \subseteq S^C)$ and $S \cup S^C = T \cup T^C = E$. Since subsets S, T are arbitrary, the condition (ii) is satisfied. Similarly, can check the condition (iii):

$$\begin{aligned}
f(S) + f(T) &= I(X_S : Y|X_{S^c}, Q) + I(X_T : Y|X_{T^c}, Q) \\
&= H(Y|X_{S^c}, Q) - H(Y|X_S, X_{S^c}, Q) + H(Y|X_{T^c}, Q) - H(Y|X_T, X_{T^c}, Q) \\
&= H(Y, X_{S^c}, Q) - H(X_{S^c}, Q) - H(Y|X_S, X_{S^c}, Q) + H(Y, X_{T^c}, Q) - H(X_{T^c}, Q) - H(Y|X_T, X_{T^c}, Q) && \text{(A7a)} \\
&= H(Y, X_{S^c}, Q) + H(Y, X_{T^c}, Q) - \sum_{i \in S^C} H(X_i|Q) - H(Q) - \sum_{i \in T^C} H(X_i|Q) - H(Q) - 2H(Y|X_E, Q) && \text{(A7b)} \\
&\geqslant H(Y, X_{S^c \cup T^c}, Q) + H(Y, X_{S^c \cap T^c}, Q) - \sum_{i \in S^C} H(X_i|Q) - H(Q) \\
&\quad - \sum_{i \in T^C} H(X_i|Q) - H(Q) - 2H(Y|X_E, Q) && \text{(A7c)} \\
&= H(Y, X_{S^c \cup T^c}, Q) + H(Y, X_{S^c \cap T^c}, Q) - \sum_{i \in S^C \cap T^C} H(X_i|Q) - H(Q) \\
&\quad - \sum_{i \in S^C \cup T^C} H(X_i|Q) - H(Q) - 2H(Y|X_E, Q) && \text{(A7d)} \\
&= H(Y, X_{S^c \cup T^c}, Q) + H(Y, X_{S^c \cap T^c}, Q) - H(X_{S^c \cap T^c}, Q) - H(X_{S^c \cup T^c}, Q) - 2H(Y|X_E, Q) && \text{(A7e)} \\
&= I(X_{S \cap T} : Y|X_{(S \cap T)^c}, Q) + I(X_{S \cup T} : Y|X_{(S \cup T)^c}, Q) && \text{(A7f)} \\
&= f(S \cap T) + f(S \cup T), && \text{(A7g)}
\end{aligned}$$

where in Eqs. (A7b), (A7e) we used the chanin rule [i.e., $H(A,B,C) = H(A|B,C) + H(B|C) + H(C)$] together with the independency of the senders [i.e., $H(A|B) = H(A)$] and in Eq. (A7c) we used the strong subadditivity of entropy [i.e., $H(A,B) + H(A,C) \geqslant H(A,B,C) + H(A)$]. In Eq. (A7f) we use de Morgan's laws $S^C \cup T^C = (S \cap T)^C, S^C \cap T^C = (S \cup T)^C$. ■

We go back to Eq. (A4) and verify that $\bar{\mathcal{R}}(\Phi_I \otimes \Phi_{II}) \supseteq \tilde{\mathcal{R}}(\Phi_I) + \tilde{\mathcal{R}}(\Phi_{II})$. This can be done by a direct coordinate sum, that is,

$$\tilde{R}_S^I + \tilde{R}_S^{II} \leqslant I\big(X_S^I : Y^I \big| X_{S^C}^I, Q^I\big) + I\big(X_S^{II} : Y^{II} \big| X_{S^C}^{II}, Q^{II}\big), \tag{A8}$$

where $\tilde{R}_S^I \in \tilde{\mathcal{R}}(\Phi_I), \tilde{R}_S^{II} \in \tilde{\mathcal{R}}(\Phi_{II})$ and by definition obey Eq. (16).

Finally, we show that $\bar{\mathcal{R}}(\Phi_I \otimes \Phi_{II}) \subseteq \tilde{\mathcal{R}}(\Phi_I) + \tilde{\mathcal{R}}(\Phi_{II})$. Since there is an equivalence between the vertex and the half space representation [25] of a convex polyhedron, we only have to show that each vertex $v \in \bar{\mathcal{R}}(\Phi_I \otimes \Phi_{II})$ can be expressed as $v = u + w$, where u and w are suitable vertices of $\tilde{\mathcal{R}}_I$ and $\tilde{\mathcal{R}}_{II}$, respectively.

As we have seen, the fixed probability capacity region is a polymatroid. This leads to a key property of the set of its vertices [26]. Let π be an ordered choice from the set of senders E. For each ordered choice π, there is a vertex v with entries $v_{\pi_1} = f(\pi_1)$, $v_{\pi_i} = f(\{\pi_1, \ldots, \pi_i\}) - f(\{1, \ldots, \pi_{i-1}\})$, and $\forall_{i \notin \pi} v_i = 0$. On the other hand, we can always find an ordered choice π which defines a given vertex. It may happen that more than one ordered choice gives the vertex with the same entries. For example, in the 2-to-1 case, fixed probability capacity region is given by the vertices:

$$\pi = \emptyset : \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{A9}$$

$$\pi = \{1\} : \begin{pmatrix} I\big(X_1^I : Y^I \big| X_2^I, Q^I\big) + I\big(X_1^{II} : Y^{II} \big| X_2^{II}, Q^{II}\big) \\ 0 \end{pmatrix}, \tag{A10}$$

$$\pi = \{2\} : \begin{pmatrix} 0 \\ I\big(X_2^I : Y^I \big| X_1^I, Q^I\big) + I\big(X_2^{II} : Y^{II} \big| X_1^{II}, Q^{II}\big) \end{pmatrix}, \tag{A11}$$

$$\pi = \{1,2\} : \begin{pmatrix} I\big(X_1^I : Y^I \big| X_2^I, Q^I\big) + I\big(X_1^{II} : Y^{II} \big| X_2^{II}, Q^{II}\big) \\ I\big(X_2^I : Y^I \big| Q^I\big) + I\big(X_2^{II} : Y^{II} \big| Q^{II}\big) \end{pmatrix}, \tag{A12}$$

$$\pi = \{2,1\} : \begin{pmatrix} I\big(X_1^I : Y^I \big| Q^I\big) + I\big(X_1^{II} : Y^{II} \big| Q^{II}\big) \\ I\big(X_2^I : Y^I \big| X_1^I, Q^I\big) + I\big(X_2^{II} : Y^{II} \big| X_1^{II}, Q^{II}\big) \end{pmatrix}. \tag{A13}$$

Using the chain rule, we obtain that for a given ordered choice π, rates achieved in vertex $v(\pi)$ are

$$\begin{aligned} R_{\pi_i} = {} & I\big(X_{\pi_i}^I : Y^I \big| X_{\pi_{i+1}}^I, \ldots, X_{\pi_n}^I, Q^I\big) \\ & + I\big(X_{\pi_i}^{II} : Y^{II} \big| X_{\pi_{i+1}}^{II}, \ldots, X_{\pi_n}^{II}, Q^{II}\big) \end{aligned} \tag{A14}$$

and can be viewed as a sum of the vectors of rates $u(\pi)$ and $w(\pi)$ with entries

$$\begin{aligned} R_{\pi_i}^I &= I\big(X_{\pi_i}^I : Y^I \big| X_{\pi_{i+1}}^I, \ldots, X_{\pi_n}^I, Q^I\big), \\ R_{\pi_i}^{II} &= I\big(X_{\pi_i}^{II} : Y^{II} \big| X_{\pi_{i+1}}^{II}, \ldots, X_{\pi_n}^{II}, Q^{II}\big), \end{aligned} \tag{A15}$$

which in an obvious way belong to fixed probability capacity regions $\tilde{\mathcal{R}}_I$ and $\tilde{\mathcal{R}}_{II}$, respectively. This completes the proof.

The proof for n MACs can be obtained through an induction of the above proof. Indeed, it suffices to divide the set of n MACs into two MACs—one composite MAC consisting of $n-1$ MACs and a second channel consisting of the remaining MAC—and apply the above procedure to them.

[1] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[2] C. H. Bennett and P. W. Shor, Science **303**, 1784 (2004).

[3] G. Yard and J. Smith, Science **321**, 1812 (2008).

[4] K. Li, A. Winter, X.-B. Zou, and G.-C. Guo, Phys. Rev. Lett. **103**, 120501 (2009).

[5] G. Smith and J. A. Smolin, Phys. Rev. Lett. **103**, 120503 (2009).

[6] L. Czekaj and P. Horodecki, Phys. Rev. Lett. **102**, 110505 (2009).

[7] L. Czekaj, J. K. Korbicz, R. W. Chhajlany, and P. Horodecki, Phys. Rev. A **82**, 020302(R) (2010).

[8] M. B. Hastings, Nat. Phys. **5**, 255 (2009).

[9] A. S. Holevo, IEEE Trans. Info. Theory **44**, 269 (1998).

[10] B. Schumacher and M. Westmoreland, Phys. Rev. A **56**, 131 (1997).

[11] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley & Sons, New York, 1991).

[12] Note that the assumption that each channel has the same number of inputs does not in any way influence the generality of the theorem, since m can be chosen to be the maximal number of senders allowed among all considered channels and the definition of the other channels can be extended simply in the following manner $p(yx_A, x_B, \ldots, x_m) = p(yx_A, x_B, \ldots, x_w)$ if $w < m$.

[13] X. B. Wang, T. Hiroshima, A. Tomita, and M. Hayashi, Phys. Rep. **448**, 1 (2007).

[14] Ch. C. Gerry and P. L. Knight, *Introductory Quantum Optics* (Cambridge University Press, Cambridge, 2005).

[15] X.-B. Wang, T. Hiroshima, A. Tomita, and M. Hayashi, Phys. Rep. **448**, 1 (2007).

[16] Rajiv Ramaswami, Kumar N. Sivarajan, *Optical Networks: A Practical Perspective* (Morgan Kaufmann, San Francisco, 2002).

[17] C. M. Caves and P. D. Drummond, Rev. Mod. Phys. **66**, 2 (1994).

[18] B. J. Yen, J. H. Shapiro,, Phys. Rev. A **72**, 062312 (2005).

[19] R. Filip, P. Marek, and U. L. Andersen, Phys. Rev. A **71**, 042308 (2005).

[20] Jun-ichi Yoshikawa, Y. Miwa, A. Huck, U. L. Andersen, P. vanLoock, and A. Furusawa, Phys. Rev. Lett. **101**, 250501 (2008).

[21] H. Vahlbruch *et al.*, Phys. Rev. Lett. **100**, 033602 (2008).
[22] H. A. Bachor and T. C. Ralph, *A Guide To Experiments In Quantum Optics* (Wiley-VCH, Weinheim, 2004).
[23] Note that we are working in the bright-states regime. We do not need single-photon resolution. Under those conditions quantum efficiency on the level $\eta = 99\%$ is avaible, for example, PIN photodiodes.
[24] S. Hanly and P. Whiting, in *Proceedings of the IEEE International Symposium on Information Theory* (Trondheim, Norway, 1994), p. 54.
[25] Gunter M. Ziegler, *Lectures on Polytopes* (Springer, Berlin, 1995).
[26] J. Edmonds, in *Proc. Calgary International Conference on Combinatorial Structures and Applications* (Calgary, Alberta, 1969), p. 69.