Superadditivity of Quantum Channel Coding Rate with Finite Blocklength Quantum Measurements

Hye Won Chung Saikat Guha Lizhong Zheng

***Abstract*—We investigate superadditivity in the maximum achievable rate of reliable classical communication over a quantum channel. The maximum number of classical information bits extracted per use of the quantum channel strictly increases as the number of channel outputs jointly measured at the receiver increases. This phenomenon is called superadditivity. We provide an explanation of this phenomenon by comparing a quantum channel with a classical discrete memoryless channel (DMC) under concatenated codes. We also give a lower bound on the maximum accessible information per channel use at a finite length of quantum measurements in terms of V, which is the quantum version of channel dispersion, and C, the classical capacity of the quantum channel.**

I. Introduction and Problem Statement

How many classical bits per channel use can be reliably communicated over a quantum channel? This has been a central question in quantum information theory in an effort to understand the intrinsic limit on the classical capacity of quantum channels such as the optical fiber.

The classical capacity of a quantum channel is defined as the maximum number of information bits per channel use that can be modulated into the input quantum states and reliably decoded at the receiver with a set of quantum measurements as the number of transmissions N_c goes to infinity. Consider a pure-state classical-quantum (cq) channel $W : x \to |\psi_x\rangle$, where $x \in \mathcal{X}$ is the classical input, and $\{|\psi_x\rangle\} \in \mathcal{H}$ are corresponding modulation symbols at the output of the channel. For example, the transmission of an ideal laser light over an optical channel can be modeled as a pure-state cq channel $W_c : \boldsymbol{\alpha} \to |\boldsymbol{\alpha}\rangle$ where $\boldsymbol{\alpha} \in \mathbb{C}$ is the complex amplitude of the *coherent state* $|\boldsymbol{\alpha}\rangle$, of which the mean photon number is equal to $|\boldsymbol{\alpha}|^2$. In refs. [1], [2], it was shown that the classical capacity of a pure state cq channel W is given by

$$C = \max_{P_X} \mathrm{Tr}(-\rho \log \rho), \tag{1}$$

where $\rho = \sum_{x \in \mathcal{X}} P_X(x) |\psi_x\rangle\langle\psi_x|$. The states $|\psi_x\rangle$, $x \in \mathcal{X}$, are normalized vectors in a complex Hilbert space $\mathcal{H}$, $\langle\psi_x|$ is the Hermitian conjugate vector of $|\psi_x\rangle$, and ρ is a *density operator*, a linear combination of the outer products $|\psi_x\rangle\langle\psi_x|$ with weights $P_X(x)$. The classical capacity of the cq channel W_c with coherent states under the mean photon number constraint of $\mathbb{E}[|\boldsymbol{\alpha}|^2] \leq \mathcal{E}$ is

$$C(\mathcal{E}) = (1+\mathcal{E}) \log(1+\mathcal{E}) - \mathcal{E} \log \mathcal{E} \text{ [nats/ch.]}, \tag{2}$$

and it is achievable with a circulo-complex Gaussian distribution with variance $\mathcal{E}$, $p(\boldsymbol{\alpha}) = \exp[-|\boldsymbol{\alpha}|^2/\mathcal{E}]/(\pi\mathcal{E})$ [3].

For an input codeword $[x_1, \ldots, x_{N_c}]$, the sequence of outputs of the quantum channel W can be written as a tensor product state, $|\Psi_{x_1,\ldots,x_{N_c}}\rangle := |\psi_{x_1}\rangle \otimes \cdots \otimes |\psi_{x_{N_c}}\rangle \in \mathcal{H}^{\otimes N_c}$. When the received codeword is projected into the orthogonal measurement vectors $\{|\Phi_k\rangle\}$, $k \in \mathcal{K}$, which resolve the identity, i.e., $\sum_k |\Phi_k\rangle\langle\Phi_k| = \mathbb{1}$, in $\mathcal{H}^{\otimes N_c}$, the classical output k is observed with probability equal to the magnitude squared of the inner product between the received codeword state and the measurement vector of the output k, i.e., with $|\langle\Phi_k|\Psi_{x_1,\ldots,x_{N_c}}\rangle|^2$. The orthogonal projective measurement is designed to decode the received codewords with as small average error probability as possible. For any rate $R < C$, there exists a block code of length N_c and rate R that can be decoded with an arbitrarily small probability of error as $N_c \to \infty$ by an appropriate quantum measurement acting jointly on the received codeword in $\mathcal{H}^{\otimes N_c}$ [1], [2].

To achieve this capacity, however, a joint detection receiver (JDR) needs to be implemented, which can measure the length-N_c sequence of states jointly and decode it reliably among $e^{N_c C}$ possible messages. The number of measurement outcomes thus scales exponentially with the length of the codeword N_c, and the measurement vectors stay in the high dimensional Hilbert space $\mathcal{H}^{\otimes N_c}$. Hence, the complexity of physical implementation of the receiver in general also grows exponentially in N_c. Considering this exponential growth in complexity, one might want to limit the maximum length $N \leq N_c$ of the sequence of symbols to be jointly detected at the receiver, independent of the length of the codeword N_c. However, such quantum measurements of fixed blocklengths cannot achieve the capacity of the quantum channel [4]. Moreover, it was shown that for some examples of quantum channels, as the number of channel outputs N jointly measured increases, the maximum number of information bits extracted per use of the quantum channel increases [5], [6]. This phenomenon is called *superadditivity* of the maximum achievable information rate over a quantum channel.

In this paper, we will study the trade-off between information rate and receiver complexity, for classical communication over a quantum channel. We will investigate the maximum number of classical bits that can be reliably decoded per use of the quantum channel, when quantum

Hye Won Chung (hwchung@mit.edu) and Lizhong Zheng (lizhong@mit.edu) are with the EECS department at MIT; Saikat Guha (sguha@bbn.com) is with the Quantum Information Processing (QuIP) group at Raytheon BBN Technologies. This research was supported by the DARPA Information in a Photon (InPho) program under contract number HR0011-10-C-0159.

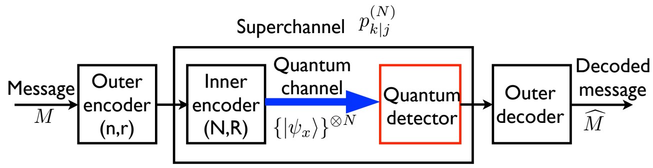

Fig. 1. Concatenated coding over a classical-quantum channel

states of a finite blocklength $N \leq N_c$ are jointly measured by a quantum receiver. After the receiver detects the quantum states, it can collect all the classical information extracted from each block of length N, and then apply any classical decoding algorithm over the collected information to decode the transmitted message reliably. To explain how it works, we introduce the architecture of concatenation over a quantum channel in Fig. 1.

In the communication system depicted in Fig. 1, a *concatenated code* is used to transmit the message M over the quantum channel. For an inner code of length N and rate R, there can be a total of e^{NR} inputs to the inner encoder, $J \in \{1, \ldots, e^{NR}\}$. The inner encoder maps each input J to a length-N classical codeword, which maps to a length-N sequence of quantum states at the output of the quantum channel. The quantum joint-detection receiver measures the length-N quantum codeword and generates an estimate $K \in \{1, \ldots, e^{NR}\}$ of the encoded message J. For a good inner code and joint measurement with large N, the estimate would generally match the input message. But for a fixed N, the error probability may not be close to 0. The inner encoder, the classical-quantum channel, and the inner decoder (the joint detection receiver), collectively form a discrete memoryless *superchannel*, with transition probabilities $p_{k|j}^{(N)} := \mathsf{Pr}(K = k | J = j)$. We define the maximum mutual information of this superchannel, over all choices of inner codes of blocklength N, and over all choices of inner-decoder joint measurements of length N as:

$$C_N := \max_{p_j} \max_{\{N\text{-symbol inner code-measurement pairs}\}} I(p_j, p_{k|j}^{(N)}). \tag{3}$$

A classical Shannon-capacity-achieving outer code can be used to reliably communicate information through the superchannel of the maximum mutual information C_N. By Shannon's coding theorem, for any rate $r < C_N$, there exists an outer code of length n and rate r that can be decoded by *classical data processing* with arbitrarily small decoding error as $n \to \infty$. Since the overall length of the concatenated code, which is composed of the inner code and the outer code, is $N_c = nN$, the total rate of the concatenated code is $R_c = r/N$. Therefore, the maximum information rate achievable by the concatenated code *per use of the quantum channel* can approach C_N/N.

From the definition of C_N, superadditivity of the quantity, i.e., $C_{N_1} + C_{N_2} \leq C_{N_1+N_2}$, can be shown. Holevo [7] showed that the limit $\lim_{N\to\infty} C_N/N$ is equal to the ultimate capacity of the quantum channel, $\lim_{N\to\infty} C_N/N = C = \max_{P_X} \mathrm{Tr}(-\rho \log \rho)$. Therefore, C_N/N is an increasing sequence in N with its limit equal to the capacity.

The question we want to answer is: *How does the maximum achievable information rate C_N/N increase as the length of the quantum measurement, N, increases?*

On the receiver side, since quantum processing occurs only at the quantum decoder for the inner code of a finite length N, the complexity of the quantum processing only depends on N, but not on the outer code length n. Therefore, the trade-off between the (rate) performance and the (quantum) complexity can be captured by how fast C_N/N increases with N. It is known that for some examples of input states, strict superadditivity of C_N can be demonstrated [4], [5]. However, the calculation of C_N, even for a pure-state binary alphabet, is extremely hard for $N > 1$ because the complexity of optimization increases exponentially with N.

Instead of aiming to calculate the exact C_N, in this paper, a lower bound of C_N/N, which becomes tight for large enough N, will be derived. From this bound, it will be possible to calculate the inner code blocklength N at which a given fraction of the ultimate capacity is achievable. We also provide a new framework for understanding the strict superadditivity of C_N in quantum channels, which is different from the previous explanation of the phenomenon by *entangled measurements* and the resulting memory in the quantum channel [6]. The detailed proofs of the results in this paper can be found in [8].

II. Strict Superadditivity of C_N

Before investigating how C_N/N increases with N, we will show examples where strict superadditivity of C_N can be shown, i.e., $C_1 < C$. For binary pure states $\{|\psi_0\rangle, |\psi_1\rangle\}$, C_1 and C can be calculated as simple functions of the inner product $\gamma = |\langle\psi_0|\psi_1\rangle|$, as summarized below [4]:

$$\begin{aligned} C_1 =& \frac{1-\sqrt{1-\gamma^2}}{2} \log\left(1-\sqrt{1-\gamma^2}\right) \\ &+ \frac{1+\sqrt{1-\gamma^2}}{2} \log\left(1+\sqrt{1-\gamma^2}\right), \end{aligned} \tag{4}$$

$$C = -\frac{1-\gamma}{2} \log \frac{1-\gamma}{2} - \frac{1+\gamma}{2} \log \frac{1+\gamma}{2}. \tag{5}$$

The capacity C is strictly greater than C_1 for all $0 < \gamma < 1$, which demonstrates the strict superadditivity of C_N for all binary input quantum channels.

Now let us consider the superadditivity of C_N in quantum channels *with an input constraint*, in the context of optical communication. The constraint will be the mean photon number of input states. A *coherent state* $|\boldsymbol{\alpha}\rangle$ is the quantum description of a single spatio-temporal-polarization mode of an ideal laser-light field, where $\boldsymbol{\alpha} \in \mathbb{C}$ is the complex amplitude, and $|\boldsymbol{\alpha}|^2$ is the mean photon number of the mode.

The important question of how many bits can be reliably communicated through a quantum channel per transmission under the average photon number constraint has been answered in [3], as stated in (2).

The number of information bits that can be reliably communicated *per photon*—the photon information efficiency (PIE)—under a mean photon number constraint per mode,

$\mathcal{E}$, is given by $C(\mathcal{E})/\mathcal{E}$ (nats/photon). From (2), it can be shown that in order to achieve high PIE, $\mathcal{E}$ must be small. In the $\mathcal{E} \to 0$ regime, the capacity (2) can be approximated as

$$C(\mathcal{E}) = \mathcal{E}\log\frac{1}{\mathcal{E}} + \mathcal{E} + o(\mathcal{E}), \tag{6}$$

which shows that PIE $\sim -\log\mathcal{E}$ for $\mathcal{E} \ll 1$. Thus there is no upper limit in principle to the photon information efficiency.

We will now show that in the high-PIE (low photon number) regime, this ultimate capacity can be closely approached even with a simple Binary Phase Shift Keying (BPSK) coherent state constellation $\{|\sqrt{\mathcal{E}}\rangle, |-\sqrt{\mathcal{E}}\rangle\}$, which satisfies the energy constraint with any prior distribution. The inner product between the two states, $\gamma = |\langle\boldsymbol{\alpha}|\boldsymbol{\beta}\rangle| = \exp[-|\boldsymbol{\alpha}-\boldsymbol{\beta}|^2/2]$. Therefore, $|\langle\sqrt{\mathcal{E}}|-\sqrt{\mathcal{E}}\rangle| = \exp[-2\mathcal{E}]$. By plugging this γ into (5), we obtain the capacity of the BPSK input constellation,

$$C_{\mathsf{BPSK}}(\mathcal{E}) = \mathcal{E}\log\frac{1}{\mathcal{E}} + \mathcal{E} + o(\mathcal{E}), \tag{7}$$

which is equal to $C(\mathcal{E})$ for the first- and second-order terms in the limit $\mathcal{E} \to 0$. For BPSK input states, by using (4), the maximum achievable rate at $N = 1$ is

$$C_{1,\mathsf{BPSK}}(\mathcal{E}) = 2\mathcal{E} + o(\mathcal{E}) \tag{8}$$

Thus, the PIE of the BPSK channel caps off at 2 nats/photon for $N = 1$, while for large N, achievable PIE $\to \infty$ as $\mathcal{E} \to 0$. It would therefore be interesting to ask how large a JDR length N is needed to bridge the gap between (8) and (7) with the BPSK channel. To answer this question, we will need the general lower bound on C_N that we develop in the following section.

III. Lower Bound on C_N

In this section, a lower bound is derived for the maximum achievable information rate at a finite blocklength N of quantum measurements.

Theorem 1: For a given set of pure input states $\{|\psi_x\rangle\}$, $x \in \mathcal{X}$ the maximum achievable information rate using quantum measurements of blocklength N, which is C_N/N as defined in (3), is lower bounded by

$$\frac{C_N}{N} \geq \max_R \left(\left(1 - 2e^{-NE(R)}\right)R - \frac{\log 2}{N}\right), \tag{9}$$

where

$$E(R) = \max_{0\leq s\leq 1}\left(\max_{P_X}\left(-\log \mathrm{Tr}\left(\rho^{1+s}\right)\right) - sR\right), \tag{10}$$

with $\rho = \sum_{x\in\mathcal{X}} P_X(x)|\psi_x\rangle\langle\psi_x|$.

By using this theorem, for the BPSK input channel, a blocklength N can be calculated at which the lower bound of (9) exceeds certain targeted rates below capacity. In the previous section, it was shown that there is a gap between $C_{1,\mathsf{BPSK}}(\mathcal{E})/\mathcal{E}$ in (8) and $C_{\mathsf{BPSK}}(\mathcal{E})/\mathcal{E}$ in (7) as $\mathcal{E} \to 0$. The performance of the BPSK channel depends significantly on the regime of N. We will now find how much quantum processing is sufficient in order to communicate using the BPSK alphabet at rates close to its capacity.

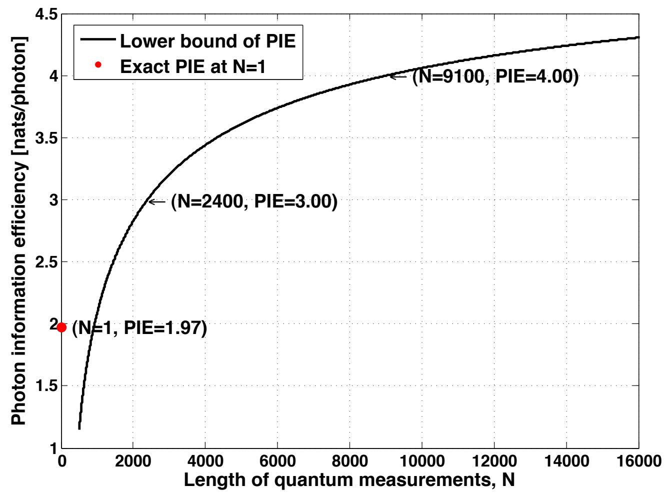

Fig. 2. A lower bound of photon information efficiency of the BPSK channel, $C_N/(N\mathcal{E})$, at $\mathcal{E} = 0.01$ for the finite blocklength N.

The following corollary summarizes an answer for the question. Note that for the BPSK inputs $\{|\sqrt{\mathcal{E}}\rangle, |-\sqrt{\mathcal{E}}\rangle\}$, any input distribution satisfies the energy constraint of $\mathcal{E}$. Consequently, we can directly apply Theorem 1 to the BPSK channel—while automatically satisfying the energy constraint—even though the theorem itself does not assume any energy constraint.

Corollary 1: For the coherent-state BPSK channel, at the length of the joint measurement

$$N = \mathcal{E}^{-1}\left(\log(1/\mathcal{E})\right)^2\left(\log\log(1/\mathcal{E})\right)^2, \tag{11}$$

we can achieve

$$\frac{C_{N,\mathsf{BPSK}}}{N} \geq \mathcal{E}\log\frac{1}{\mathcal{E}} + \mathcal{E} + o(\mathcal{E}). \tag{12}$$

Using the result of Theorem 1 applied for the BPSK channel, the photon information efficiency achievable by the BPSK channel is plotted as a function of N in Fig. 2. When the average photon number transmitted per symbol, $\mathcal{E}$, is 0.01, the PIE at an arbitrarily large N is 5.55 nats/photon from (5), and at $N = 1$, is 1.97 nats/photon from (4). Therefore, as N increases from 1 to ∞, the PIE of the BPSK channel should strictly increase from 1.97 to 5.55 nats/photon. From the lower bound of PIE in Fig. 2, it can be seen that at $N = 2400$, a PIE of 3.0 nats/photon can be achieved, and at $N = 9100$, 4.0 nats/photon is achievable. The lower bound is not tight in the regime of very small N, but it gets tighter as N increases, and approaches the ultimate limit of PIE as $N \to \infty$.

IV. Proof of Theorem 1

The proof of Theorem 1 is based on two ideas: First, instead of tracking the exact superchannel distribution $p_{k|j}^{(N)}$, which depends on the detailed structure of the length-N inner code and joint measurement, we focus on one representative quantity derived from $p_{k|j}^{(N)}$ that can be easily analyzed and optimized. Second, among superchannels that have the same value of this representative quantity, we find a superchannel whose mutual information is the smallest. The representative quantity is the average decoding error probability of the inner code, under a uniform distribution over the inner codewords, defined as

$$p_e = e^{-NR} \sum_{j=1}^{e^{NR}} \sum_{k \neq j} p_{k|j}^{(N)}, \tag{13}$$

where R is the rate of the inner code.

In ref. [4], Holevo showed the following upper bound on p_e for a code of length N and rate R.

Lemma 1: [Holevo] *For a set of input states $\{|\psi_x\rangle\}$, there exists a block code of length N and rate R that can be decoded by a set of measurements with*

$$p_e \leq 2\exp[-NE(R)], \tag{14}$$

for $E(R)$ in (10).

Now, among superchannels that have the same value of p_e, we find a superchannel whose mutual information is the smallest. An *equierror superchannel*, which was first introduced in [9], is defined with the following distribution:

$$\overline{p}_{k|j}^{(N)} := \begin{cases} 1 - p_e, & k = j\ ; \\ \left(e^{NR} - 1\right)^{-1} p_e, & k \neq j. \end{cases} \tag{15}$$

This channel assumes that the probability of making an error is equal for every input, and when an error occurs, all wrong estimates $k \neq j$ are equally likely. Due to the symmetry, the input distribution that maximizes the mutual information of this channel is uniform. The resulting maximum mutual information of the equierror superchannel,

$$\begin{aligned} \max_{p_j} I\left(p_j, \overline{p}_{k|j}^{(N)}\right) &= NR - p_e \log\left(e^{NR} - 1\right) - H_{\mathsf{B}}(p_e) \\ &> (1 - p_e)NR - \log 2, \end{aligned}$$

where $H_{\mathsf{B}}(p) = -p\log p - (1-p)\log(1-p)$.

We can show that the mutual information of the equierror channel is smaller than that of any other superchannel with the same average probability of error, p_e.

Lemma 2: For any $p_{k|j}^{(N)}$ with a fixed p_e defined in (13),

$$\max_{p_j} I(p_j, p_{k|j}^{(N)}) \geq \max_{p_j} I(p_j, \overline{p}_{k|j}^{(N)}) \tag{16}$$

for the equierror superchannel, $\overline{p}_{k|j}^{(N)}$ with the same p_e.

Then, by the definition of C_N and Lemma 2, when there exists an inner code of length N and rate R that can be decoded by a set of length N measurements with an average error probability p_e,

$$\frac{C_N}{N} \geq \max_{p_j} \frac{I(p_j, p_{k|j}^{(N)})}{N} > (1 - p_e)R - \frac{\log 2}{N}. \tag{17}$$

By (17) and Lemma 1, Theorem 1 can be proven.

V. Interpretation of Superadditivity: Classical DMC vs. Quantum Channel

Previously, the superadditivity of C_N has been thought of as a unique property that can be observed only in quantum channels, but not in classical DMCs. One popular interpretation of this phenomenon is that a set of length-N *entangled* quantum measurements can induce a classical superchannel that has memory over the N channel uses, which results in the higher capacity as N increases. Despite the fact that the above intuition of why superadditivity appears in the capacity of quantum channels is somewhat satisfying, this viewpoint does not provide enough quantitative insight to fully understand the phenomenon. In this section, we will introduce a new aspect of understanding strict superadditivity of C_N by comparing the performance of concatenated coding over quantum channels and classical DMCs, for a fixed inner code length N.

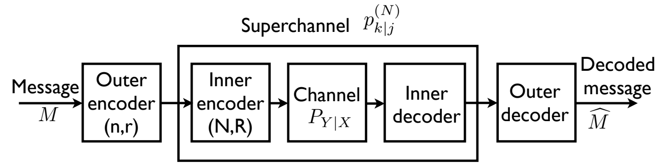

Fig. 3. Concatenated coding over a classical DMC

Fig. 3 illustrates a concatenated coding architecture over a classical DMC. In [9], Forney introduced this concatenated coding architecture for a classical DMC to analyze the trade-off between (rate) performance and (coding) complexity [9]. It is obvious that when the inner decoder generates a sufficient statistic of the channel output and forwards it to the outer decoder, there is no loss of information, so that the performance of the concatenated code can be as good as an optimum code, even within the restricted structure of code concatenation. Despite the fact that the performance remains intact, the decoding complexity increases exponentially with the overall length of the code. On the other hand, it was shown in [9] that even if there is some loss of information at the inner decoder by making a hard-decision on the message of the inner code, *as the inner code blocklength N goes to infinity,* the capacity of the underlying classical DMC can be achieved with the concatenated code. Moreover, the overall complexity of the decoding algorithm is significantly reduced to be almost linear in the length of the concatenated code. The loss of information at the inner decoder, however, degrades the achievable error exponent over all rates below capacity.

We now ask a new question for the concatenated code over the classical DMC, similar to the one we asked for the quantum channel: When the inner decoder makes a hard estimate of messages of the inner code *at a finite blocklength N*, how does the maximum achievable information rate (error-free bits per use of the underlying DMC) with the concatenated code increase as N increases (with no restriction on the complexity of the outer code)?

The maximum achievable information rate at a finite blocklength N of the inner code is C_N/N, where

$$C_N = \max_{p_j} \max_{\{\text{N-symbol inner code-decoder pairs}\}} I(p_j, p_{k|j}^{(N)}) \tag{18}$$

for the superchannel distribution $p_{k|j}^{(N)}$, which is determined by the decoding algorithm, given an inner code. When the inner decoder makes hard-decisions at a finite blocklength, we can observe a phenomenon similar to the superadditivity of C_N in the quantum channel, even in the classical DMC.

The following theorem summarizes the lower bound for the maximum achievable information rate of the concatenated codes over the classical DMC at a finite N.

Theorem 2: With a fixed inner code blocklength N,

$$\frac{C_N}{N} \geq \max_R \left(\left(1 - e^{-NE(R)}\right) R - \frac{\log 2}{N} \right), \quad (19)$$

where

$$E(R) = \max_{0 \leq s \leq 1} \left(\max_{P_X} (E_0(s, P_X)) - sR \right) \quad (20)$$

with

$$E_0(s, P_X) := -\log \sum_y \left[\sum_x P_X(x) P_{Y|X}(y|x)^{1/(1+s)} \right]^{1+s}. \quad (21)$$

Note that the lower bound on C_N/N in (19) strictly increases with N, and it has the same form as that for the quantum channel in (9) except for the difference in $E(R)$ and a constant multiplying $e^{-NE(R)}$. C_N is below C, the capacity of the channel, for a finite inner code blocklength N because the hard-decision at the inner decoder results in a significant amount of loss of information, which hurts the rate of the communication. Therefore, the superadditivity of C_N can now be interpreted as a degradation of the performance by the loss of information at the inner decoder that makes the hard-decision at a finite blocklength. This new understanding can also be applied to explain the same phenomenon observed in the quantum channel by replacing the role of inner decoder with a quantum joint-detection receiver that makes hard-decisions on finite blocks of quantum states. However, it is important to note that unlike the classical inner decoder, which has an option to maintain sufficient statistics of the channel outputs with the cost of complexity, a loss of information at the quantum detector, which results in the superadditivity, is not avoidable

We can further simplify the lower bound of C_N by finding an approximation of the error exponent $E(R)$ for the quantum channel and for the classical DMC. Both the error exponent of the classical DMC, $E(R)$ in (20), and that of the quantum channel, $E(R)$ in (10), can be approximated by the Taylor expansion at the rate R close to C as

$$E(R) = \frac{1}{2V}(R - C)^2 + O\left((R - C)^3\right), \quad (22)$$

with a parameter V. For the classical channel, $V = V^{(c)}$ where $V^{(c)}$ is the variance of $\log(P_{Y|X}/P_Y^*)$ where P_Y^* is the capacity achieving output distribution. $V^{(c)}$ was defined and termed *channel dispersion* in [10]. Similarly, the quantum version of the channel dispersion $V^{(q)}$ is defined as

$$V^{(q)} = \sum_{i=1}^{J} \sigma_i (-\log \sigma_i)^2 - \left(\sum_{i=1}^{J} \sigma_i \left(-\log \sigma_i\right) \right)^2, \quad (23)$$

where $\{\sigma_i\}$ are the eigenvalues of the density operator $\rho = \sum_x P_X(x) |\psi_x\rangle\langle\psi_x|$ at P_X^* that attains the maximum in Eq. (1). For the quantum channel, $V = V^{(q)}$ in (22).

Since the lower bound on C_N as well as the approximated error exponent $E(R)$ have similar forms for the classical DMC and for the quantum channel, it is possible to compare the classical DMC and the quantum channel by a common simplified lower bound on C_N, which can be written with the parameter V and C as follows.

Theorem 3: For both a classical DMC and a pure-state classical-quantum channel, when the channel dispersion V and the capacity C satisfy i) $\sqrt{\frac{V}{NC^2}} \to 0$ as $N \to \infty$ and ii) $V \cdot C$ is finite, the maximum achievable information rate at the inner code blocklength N is lower bounded by

$$\begin{aligned} \frac{C_N}{N} \geq & C \cdot \left(1 - \sqrt{\frac{V}{NC^2} \log\left(\frac{NC^2}{V}\right)}\right) - \frac{\log 2}{N} \\ & + O\left(\sqrt{\frac{V}{N \log \frac{NC^2}{V}}} \log\log\left(\frac{NC^2}{V}\right)\right). \end{aligned} \quad (24)$$

Remark 1: From the lower bound of Theorem 3, we can see that the inner code blocklength N at which the lower bound is equal to a given fraction of the capacity is proportional to V/C^2.

Since the same bound on C_N/N as in (24) holds both for the quantum and the classical channels, using the parameter V/C^2, we can compare the behaviors of the quantum channel and the classical DMC. For the BPSK quantum channel, $V_{\mathsf{BPSK}}/C_{\mathsf{BPSK}}^2 \approx 1/\mathcal{E}$ for the low photon number regime where $\mathcal{E} \to 0$. For the classical additive white Gaussian noise (AWGN) channel in the low-power regime where $\mathsf{SNR} \to 0$, $V_{\mathsf{AWGN}}/C_{\mathsf{AWGN}}^2$ can be calculated by using the result of [10], and it is $4/\mathsf{SNR}$. For both channels, V/C^2 is inversely proportional to the energy to transmit the information per channel use. This means that as the energy decreases, in order to make the lower bound meet a targeted fraction of capacity, it is necessary to adopt a longer inner code.

References

[1] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, and W. K. Wootters, "Classical information capacity of a quantum channel," *Phys. Rev. A*, vol. 54, pp. 1869–1876, Sep 1996.

[2] A. Holevo, "The capacity of the quantum channel with general signal states," *Information Theory, IEEE Transactions on*, vol. 44, no. 1, pp. 269–273, 1998.

[3] V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, J. H. Shapiro, and H. P. Yuen, "Classical capacity of the lossy bosonic channel: The exact solution," *Physical review letters*, vol. 92, no. 2, p. 027902, 2004.

[4] A. S. Holevo, "Coding theorems for quantum channels," *arXiv preprint quant-ph/9809023*, 1998.

[5] A. Peres and W. K. Wootters, "Optimal detection of quantum information," *Physical Review Letters*, vol. 66, no. 9, pp. 1119–1122, 1991.

[6] M. Sasaki, K. Kato, M. Izutsu, and O. Hirota, "A demonstration of superadditivity in the classical capacity of a quantum channel," *Physics Letters A*, vol. 236, no. 1, pp. 1–4, 1997.

[7] A. Holevo, "Problems in the mathematical theory of quantum communication channels," *Reports on Mathematical Physics*, vol. 12, no. 2, pp. 273–278, 1977.

[8] H. W. Chung, S. Guha, and L. Zheng, "Superadditivity of quantum channel coding rate with finite blocklength quantum measurements," *arXiv preprint arXiv:1310.3793*, 2013.

[9] G. D. Forney, *Concatenated codes*. Citeseer, 1966, vol. 11.

[10] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *Information Theory, IEEE Transactions on*, vol. 56, no. 5, pp. 2307–2359, 2010.