

Research



Cite this article: Oskouei SK, Mancini S, Wilde MM. 2019 Union bound for quantum information processing. *Proc. R. Soc. A* **475**: 20180612.
<http://dx.doi.org/10.1098/rspa.2018.0612>

Received: 7 September 2018

Accepted: 3 December 2018

Subject Areas:

quantum computing, quantum physics

Keywords:

quantum union bound, sequential decoding, quantum communication

Author for correspondence:

Mark M. Wilde

e-mail: mwilde@gmail.com

Union bound for quantum information processing

Samad Khabbazi Oskouei¹, Stefano Mancini^{2,3}
 and Mark M. Wilde⁴

¹Department of Mathematics, Islamic Azad University, Varamin-Pishva Branch, 33817-7489 Iran

²School of Science and Technology, University of Camerino, Via M. delle Carceri 9, 62032 Camerino, Italy

³INFN—Sezione Perugia, Via A. Pascoli, 06123 Perugia, Italy

⁴Hearne Institute for Theoretical Physics, Department of Physics and Astronomy, Center for Computation and Technology, Louisiana State University, Baton Rouge, LA 70803, USA

MMW, 0000-0002-3916-4462

In this paper, we prove a quantum union bound that is relevant when performing a sequence of binary-outcome quantum measurements on a quantum state. The quantum union bound proved here involves a tunable parameter that can be optimized, and this tunable parameter plays a similar role to a parameter involved in the Hayashi–Nagaoka inequality (Hayashi & Nagaoka 2003 *IEEE Trans. Inf. Theory* **49**, 1753–1768. (doi:10.1109/TIT.2003.813556)), used often in quantum information theory when analysing the error probability of a square-root measurement. An advantage of the proof delivered here is that it is elementary, relying only on basic properties of projectors, Pythagoras' theorem, and the Cauchy–Schwarz inequality. As a non-trivial application of our quantum union bound, we prove that a sequential decoding strategy for classical communication over a quantum channel achieves a lower bound on the channel's second-order coding rate. This demonstrates the advantage of our quantum union bound in the non-asymptotic regime, in which a communication channel is called a finite number of times. We expect that the bound will find a range of applications in quantum communication theory, quantum algorithms and quantum complexity theory.

1. Introduction

The union bound, alternatively known as Boole's inequality, represents one of the simplest yet non-trivial

methods for bounding the probability that either one event or another occurs, in terms of the probabilities of the individual events (e.g. [1]). By induction, the bound applies to the union of multiple events, and it often provides a good enough bound in a variety of applications whenever the probabilities of the individual events are small relative to the number of events. Concretely, given a finite set $\{A_i\}_{i=1}^L$ of events, the union bound is the following inequality:

$$\Pr \left\{ \bigcup_{i=1}^L A_i \right\} \leq \sum_{i=1}^L \Pr\{A_i\}. \quad (1.1)$$

By applying DeMorgan's Law and basic rules of probability theory, we can rewrite the union bound such that it applies to the probability that an intersection of events does not occur

$$1 - \Pr \left\{ \bigcap_{i=1}^L A_i \right\} \leq \sum_{i=1}^L \Pr\{A_i^c\}, \quad (1.2)$$

and this is the form in which it is typically employed in applications. Recently, the union bound has been listed as the second step to try when attempting to 'upper-bound the probability of something bad', with the first step being to determine if the trivial bound of one is reasonable in a given application [2].

Generalizing the union bound to a quantum-mechanical set-up is non-trivial. A natural setting in which we would consider this generalization is when the goal is to bound the probability that two or more successive measurement outcomes do not occur. Concretely, suppose that the state of a quantum system is given by a density operator ρ . Suppose that there are L projective quantum measurements $\{P_i, I - P_i\}$ for $i \in [1, \dots, L]$, where P_i is a projector, thus satisfying $P_i = P_i^\dagger$ and $P_i = P_i^2$ by definition. Suppose that the first measurement is performed, followed by the second measurement, and so on. If the projectors P_1, \dots, P_L commute, then the probability that the outcomes P_1, \dots, P_L do not occur is calculated by applying the Born rule and can be bounded as

$$1 - \text{Tr}\{P_L P_{L-1} \cdots P_1 \rho P_1 \cdots P_{L-1}\} \leq \sum_{i=1}^L \text{Tr}\{(I - P_i)\rho\}, \quad (1.3)$$

with the bound following essentially from an application of the union bound. However, if the projectors P_1, \dots, P_L do not commute, then classical reasoning does not apply and alternative methods are required.

Recently, Gao proved a quantum union bound [3] that has been useful in a variety of applications, including quantum communication theory [3–7], quantum algorithms [8–10], quantum complexity theory [9,11] and Hamiltonian complexity theory [12,13]. Given an arbitrary set of projectors $\{P_i\}_{i=1}^L$, each corresponding to one outcome of a binary-valued measurement, Gao's quantum union bound is the following inequality [3, Theorem 1]:

$$1 - \text{Tr}\{P_L P_{L-1} \cdots P_1 \rho P_1 \cdots P_{L-1}\} \leq 4 \sum_{i=1}^L \text{Tr}\{(I - P_i)\rho\}. \quad (1.4)$$

By comparing (1.4) with (1.3), we notice that the only difference is the factor of four in (1.4). The factor of four is inconsequential for many applications, but, nevertheless, it is natural to wonder whether this bound can be improved. Furthermore, at least one application in which improving the factor of four does make a difference is in the context of whether a sequential decoding strategy can be used to achieve the second-order coding rate for classical communication—we discuss this application in more detail later.

2. Summary of results

In this paper, we prove the following quantum union bound.

Theorem 2.1 (Quantum union bound). *Let ρ be a density operator acting on a separable Hilbert space \mathcal{H} , let $\{P_i\}_{i=1}^L$ be an arbitrary set of projectors, each acting on \mathcal{H} , and let $c > 0$ be an arbitrary positive constant. Then*

$$1 - \text{Tr}\{P_L P_{L-1} \cdots P_1 \rho P_1 \cdots P_{L-1}\} \leq (1+c) \text{Tr}\{(I - P_L)\rho\} \\ + (2+c+c^{-1}) \sum_{i=2}^{L-1} \text{Tr}\{(I - P_i)\rho\} + (2+c^{-1}) \text{Tr}\{(I - P_1)\rho\}. \quad (2.1)$$

Our proof of the above theorem is elementary, relying only on basic properties of projectors, Pythagoras' theorem and the Cauchy–Schwarz inequality. Furthermore, the theorem directly applies to states of infinite-dimensional quantum systems and can thus be employed to analyse practical situations involving not only qubits but also bosonic quantum systems [14]. Similar to the classical case discussed in the introduction, the quantum union bound of theorem 2.1 provides a useful bound when the individual probabilities $\text{Tr}\{(I - P_i)\rho\}$ are small relative to the number L of them, and this scenario occurs, for example, in the application to communication presented in §5. Furthermore, the tunable parameter $c > 0$ is a significant advantage of our quantum union bound, and it is essential in the application mentioned above, in which it really is necessary for $c > 0$ to be decreasing with the number of channel uses so that the prefactor in front of the term $\text{Tr}\{(I - P_L)\rho\}$ is as close to one as possible. More generally, one could certainly take an infimum over the parameter $c > 0$ in any given application in order to have the upper bound be as tight as possible.

Our quantum union bound represents a strict improvement over that of Gao's in (1.4). Indeed, by setting $c = 1$ and then loosening the above bound further, we recover Gao's. Our quantum union bound can also be compared with the Hayashi–Nagaoka (HN) inequality from [15, Lemma 2], which is often used to analyse the error probability of the square-root measurement. The HN inequality also features a tunable parameter $c > 0$, and this is one of the main reasons why quantum information theory has recently advanced in the direction of characterizing second-order asymptotics for communication tasks [16–24]. Our quantum union bound provides essentially the same trade-off given by the HN inequality, but just slightly improved, in the sense that the prefactor for the term $\text{Tr}\{(I - P_L)\rho\}$ is $1+c$, while the prefactor for $L-2$ other terms is $2+c+c^{-1}$ and the prefactor for the term $\text{Tr}\{(I - P_1)\rho\}$ is $2+c^{-1}$, the last prefactor representing the improvement.

In the previous paragraphs, we focused exclusively on the comparison of theorem 2.1 with Gao's bound in (1.4). However, there are other works that preceded Gao's, which we recall now. Aaronson [25] established a quantum union bound, with applications in quantum complexity theory. Giovannetti *et al.* [26] analysed the error probability of a sequential decoding strategy and proved that it can achieve the Holevo information of a quantum channel for classical communication. The work of [26] then inspired Sen [27], who established another quantum union bound (also called 'non-commutative union bound') of the following form:

$$1 - \text{Tr}\{P_L P_{L-1} \cdots P_1 \rho P_1 \cdots P_{L-1}\} \leq 2 \sqrt{\sum_{i=1}^L \text{Tr}\{(I - P_i)\rho\}}. \quad (2.2)$$

Wilde [28] subsequently generalized the result of [27] beyond projectors, such that it would hold for a set of operators $\{A_i\}_{i=1}^L$, each of which satisfies $0 \leq A_i \leq I$. Then Gao's bound in (1.4) appeared after [28]. Clearly, Gao's bound was a significant improvement over (2.2), eliminating the square root at the cost of a doubling of the prefactor.

To demonstrate an application in which theorem 2.1 is useful, we show how a sequential decoding strategy achieves a lower bound on the second-order coding rate for classical communication over a quantum channel. We consider the cases in which there is entanglement

assistance as well as no assistance, and our result here also covers the important case when the channel takes input density operators acting on a separable Hilbert space to output density operators acting on a separable Hilbert space. An advantage of our proof is that it is arguably simpler than other approaches that could be taken to solve this problem, with our proof relying on a method called position-based coding [29], as well as sequential decoding [26–28], and an error analysis that uses theorem 2.1. Our proof can be compared with the proof from [30,31], in which it was shown how to achieve the capacity for energy-constrained classical communication (i.e. the first-order coding rate), and we advocate here that our proof is considerably simpler.

We organize the rest of our paper as follows. In §3, we provide a proof of theorem 2.1. In §4, we consider the generalization of theorem 2.1 to positive operator-valued measures (POVMs). Section 5 discusses the application to obtaining a lower bound on the second-order coding rate for classical communication. In §6, we conclude with a summary and discuss some open directions for future research.

3. Proof of theorem 2.1

We prove our main result, theorem 2.1, by establishing the following more general result:

Theorem 3.1. *Let \mathcal{H} be a separable Hilbert space, let $|\psi\rangle \in \mathcal{H}$, let $\{P_i\}_{i=1}^L$ be a finite set of projectors acting on \mathcal{H} , and let $c > 0$. Then*

$$\begin{aligned} \|\psi\|_2^2 - \|P_L P_{L-1} \cdots P_1 |\psi\rangle\|_2^2 &\leq (1+c) \|(I - P_L) |\psi\rangle\|_2^2 \\ &+ (2+c+c^{-1}) \sum_{i=2}^{L-1} \|(I - P_i) |\psi\rangle\|_2^2 + (2+c^{-1}) \|(I - P_1) |\psi\rangle\|_2^2. \end{aligned} \quad (3.1)$$

Theorem 2.1 is a direct consequence of theorem 3.1. Indeed, a density operator ρ acting on a separable Hilbert space has a spectral decomposition as follows:

$$\rho = \sum_{j \in \mathcal{J}} p_j |\psi_j\rangle \langle \psi_j|, \quad (3.2)$$

where the index set \mathcal{J} is countable, $\{p_j\}_{j \in \mathcal{J}}$ is a probability distribution, and $\{|\psi_j\rangle\}_{j \in \mathcal{J}}$ is an orthonormal set of eigenvectors [32]. Applying theorem 3.1, we find that

$$\begin{aligned} 1 - \text{Tr}\{P_L P_{L-1} \cdots P_1 |\psi_j\rangle \langle \psi_j| P_1 \cdots P_{L-1}\} \\ = \|\psi_j\|_2^2 - \|P_L P_{L-1} \cdots P_1 |\psi_j\rangle\|_2^2 \end{aligned} \quad (3.3)$$

$$\begin{aligned} \leq (1+c) \|(I - P_L) |\psi_j\rangle\|_2^2 + (2+c+c^{-1}) \sum_{i=2}^{L-1} \|(I - P_i) |\psi_j\rangle\|_2^2 \\ + (2+c^{-1}) \|(I - P_1) |\psi_j\rangle\|_2^2 \end{aligned} \quad (3.4)$$

$$\begin{aligned} = (1+c) \text{Tr}\{(I - P_L) |\psi_j\rangle \langle \psi_j|\} + (2+c+c^{-1}) \sum_{i=2}^{L-1} \text{Tr}\{(I - P_i) |\psi_j\rangle \langle \psi_j|\} \\ + (2+c^{-1}) \text{Tr}\{(I - P_1) |\psi_j\rangle \langle \psi_j|\}. \end{aligned} \quad (3.5)$$

The reduction from theorem 2.1 to theorem 3.1 follows by averaging over the distribution $\{p_j\}_{j \in \mathcal{J}}$.

So now we shift our focus to proving theorem 3.1, and we do so with the aid of several lemmas. To simplify the notation, hereafter we employ the following shorthand:

$$\|\cdots\| \equiv \|\cdots |\psi\rangle\|_2, \quad (3.6)$$

$$\langle \cdots \rangle \equiv \langle \psi | \cdots | \psi \rangle \quad (3.7)$$

and

$$Q_i \equiv I - P_i. \quad (3.8)$$

The convention we take with the shorthand $\langle A \rangle$ for a non-Hermitian operator A is that $\langle A \rangle = \langle \psi | \varphi \rangle$ where $|\varphi\rangle = A|\psi\rangle$. Furthermore, we also assume without loss of generality that the vector $|\psi\rangle$ in theorem 3.1 is a unit vector. Clearly, this assumption can be easily released by scaling the resulting inequality by an arbitrary positive number.

First recall that, due to the idempotence of projectors, we have the following identities holding for all $i \in \{1, 2, \dots, L\}$:

$$\langle Q_i P_{i-1} \cdots P_1 \rangle = \langle Q_i Q_i P_{i-1} \cdots P_1 \rangle, \quad \langle P_1 \cdots P_i \rangle = \langle P_1 \cdots P_i P_i \rangle, \quad (3.9)$$

under the convention that $P_{i-1} \cdots P_1 = P_1 \cdots P_{i-1} = I$ for $i = 1$.

Lemma 3.2. For a set $\{P_i\}_{i=1}^L$ of projectors acting on a separable Hilbert space \mathcal{H} , a unit vector $|\psi\rangle \in \mathcal{H}$, and employing the shorthand in (3.6)–(3.8), we have the following identities:

$$\sum_{i=1}^L \langle Q_i P_{i-1} \cdots P_1 \rangle = 1 - \langle P_L \cdots P_1 \rangle, \quad (3.10)$$

$$\sum_{i=1}^L \langle P_1 \cdots P_{i-1} Q_i \rangle = 1 - \langle P_1 \cdots P_L \rangle, \quad (3.11)$$

$$\sum_{i=1}^L \langle P_1 \cdots P_{i-1} Q_i P_{i-1} \cdots P_1 \rangle = 1 - \langle P_1 \cdots P_L \cdots P_1 \rangle \quad (3.12)$$

and
$$1 - \sqrt{\langle P_L \rangle} \sqrt{\langle P_1 \cdots P_L \cdots P_1 \rangle} \leq \sum_{i=1}^L \sqrt{\langle Q_i \rangle} \sqrt{\langle P_1 \cdots P_{i-1} Q_i P_{i-1} \cdots P_1 \rangle}, \quad (3.13)$$

under the convention that $P_{i-1} \cdots P_1 = P_1 \cdots P_{i-1} = I$ for $i = 1$.

Proof. The following identities are straightforward:

$$1 = \langle Q_1 \rangle + \langle Q_2 P_1 \rangle + \cdots + \langle Q_{L-1} P_{L-2} \cdots P_1 \rangle + \langle Q_L P_{L-1} \cdots P_1 \rangle + \langle P_L P_{L-1} \cdots P_1 \rangle, \quad (3.14)$$

$$1 = \langle Q_1 \rangle + \langle P_1 Q_2 \rangle + \cdots + \langle P_1 \cdots P_{L-2} Q_{L-1} \rangle + \langle P_1 \cdots P_{L-1} Q_L \rangle + \langle P_1 \cdots P_{L-1} P_L \rangle \quad (3.15)$$

and
$$1 = \langle Q_1 \rangle + \langle P_1 Q_2 P_1 \rangle + \cdots + \langle P_1 \cdots P_{L-2} Q_{L-1} P_{L-2} \cdots P_1 \rangle + \langle P_1 \cdots P_{L-1} Q_L P_{L-1} \cdots P_1 \rangle + \langle P_1 \cdots P_{L-1} P_L P_{L-1} \cdots P_1 \rangle. \quad (3.16)$$

Consequently, from the equalities in (3.14)–(3.16), we obtain (3.10)–(3.12), respectively. The following equality is a direct consequence of (3.14) and (3.9):

$$1 = \langle Q_1 \rangle + \langle Q_2 Q_2 P_1 \rangle + \cdots + \langle Q_{L-1} Q_{L-1} P_{L-2} \cdots P_1 \rangle + \langle Q_L Q_L P_{L-1} \cdots P_1 \rangle + \langle P_L P_L P_{L-1} \cdots P_1 \rangle. \quad (3.17)$$

By applying the Cauchy–Schwarz inequality to (3.17), we find that

$$1 \leq \langle Q_1 \rangle + \sqrt{\langle Q_2 \rangle} \sqrt{\langle P_1 Q_2 P_1 \rangle} + \cdots + \sqrt{\langle Q_L \rangle} \sqrt{\langle P_1 \cdots P_{L-1} Q_L P_{L-1} \cdots P_1 \rangle} + \sqrt{\langle P_L \rangle} \sqrt{\langle P_1 \cdots P_{L-1} P_L P_{L-1} \cdots P_1 \rangle}, \quad (3.18)$$

from which (3.13) immediately follows. ■

Lemma 3.3. For a set $\{P_i\}_{i=1}^L$ of projectors acting on a separable Hilbert space \mathcal{H} , a unit vector $|\psi\rangle \in \mathcal{H}$, and employing the shorthand in (3.6)–(3.8), the following inequality holds for $L \geq 2$:

$$\sum_{i=1}^L \|Q_i(I - P_{i-1} \cdots P_1)\|^2 \leq \sum_{i=1}^{L-1} \|Q_i\|^2, \quad (3.19)$$

under the convention that $P_{i-1} \cdots P_1 = P_1 \cdots P_{i-1} = I$ for $i = 1$. Equivalently,

$$\sum_{i=2}^L \|Q_i(I - P_{i-1} \cdots P_1)\|^2 \leq \sum_{i=1}^{L-1} \|Q_i\|^2, \quad (3.20)$$

due to the aforementioned convention.

Proof. Consider the following chain of equalities:

$$\begin{aligned} \sum_{i=1}^L \|Q_i(I - P_{i-1} \cdots P_1)\|^2 &= \sum_{i=1}^L \|Q_i - Q_i P_{i-1} \cdots P_1\|^2 \\ &= \sum_{i=1}^L \left(\|Q_i\|^2 - \langle Q_i P_{i-1} \cdots P_1 \rangle - \langle P_1 \cdots P_{i-1} Q_i \rangle + \langle P_1 \cdots P_{i-1} Q_i P_{i-1} \cdots P_1 \rangle \right) \end{aligned} \quad (3.21)$$

$$= \left(\sum_{i=1}^L \|Q_i\|^2 \right) - 1 + \langle P_L \cdots P_1 \rangle - 1 + \langle P_1 \cdots P_L \rangle + 1 - \langle P_1 \cdots P_L \cdots P_1 \rangle \quad (3.22)$$

$$= \left(\sum_{i=1}^L \|Q_i\|^2 \right) - 1 + \langle P_L P_L P_{L-1} \cdots P_1 \rangle + \langle P_1 \cdots P_{L-1} P_L P_L \rangle - \langle P_1 \cdots P_L \cdots P_1 \rangle. \quad (3.23)$$

To obtain (3.21), we used the identities in (3.9). Next, to get (3.22), the identities in (3.10)–(3.12) of lemma 3.2 were used. Continuing, we have that

$$\text{Equation (3.23)} \leq \left(\sum_{i=1}^L \|Q_i\|^2 \right) - 1 - \langle P_1 \cdots P_L \cdots P_1 \rangle + 2\sqrt{\langle P_L \rangle} \sqrt{\langle P_1 \cdots P_L \cdots P_1 \rangle} \quad (3.24)$$

$$= \left(\sum_{i=1}^L \|Q_i\|^2 \right) - 1 + \langle P_L \rangle - \left(\sqrt{\langle P_L \rangle} - \sqrt{\langle P_1 \cdots P_L \cdots P_1 \rangle} \right)^2 \quad (3.25)$$

$$\leq \left(\sum_{i=1}^L \|Q_i\|^2 \right) - \|Q_L\|^2 = \sum_{i=1}^{L-1} \|Q_i\|^2. \quad (3.26)$$

To obtain (3.24), the Cauchy–Schwarz inequality was employed. ■

We are now in a position to prove theorem 3.1.

Proof of theorem 3.1. Consider that

$$\begin{aligned} 1 - \|P_L \cdots P_1\|^2 &= 1 - \langle P_1 \cdots P_L \cdots P_1 \rangle + 2 \left(1 - \sqrt{\langle P_L \rangle} \sqrt{\langle P_1 \cdots P_L \cdots P_1 \rangle} \right) \\ &\quad - 2 \left(1 - \sqrt{\langle P_L \rangle} \sqrt{\langle P_1 \cdots P_L \cdots P_1 \rangle} \right) \end{aligned} \quad (3.27)$$

$$\begin{aligned} &= 2 \left(1 - \sqrt{\langle P_L \rangle} \sqrt{\langle P_1 \cdots P_L \cdots P_1 \rangle} \right) \\ &\quad - \left(\sqrt{\langle P_L \rangle} - \sqrt{\langle P_1 \cdots P_L \cdots P_1 \rangle} \right)^2 - 1 + \langle P_L \rangle. \end{aligned} \quad (3.28)$$

Continuing, we have that

$$\text{Equation (3.28)} \leq -\|Q_L\|^2 + 2 \left(1 - \sqrt{\langle P_L \rangle} \sqrt{\langle P_1 \cdots P_L \cdots P_1 \rangle} \right) \quad (3.29)$$

$$\leq -\|Q_L\|^2 + 2 \sum_{i=1}^L \sqrt{\langle Q_i \rangle} \sqrt{\langle P_1 \cdots P_{i-1} Q_i P_{i-1} \cdots P_1 \rangle} \quad (3.30)$$

$$\leq -\|Q_L\|^2 + 2 \sum_{i=1}^L \sqrt{\langle Q_i \rangle} (\|Q_i\| + \|Q_i(I - P_{i-1} \cdots P_1)\|). \quad (3.31)$$

First, (3.29) is obtained by observing that

$$-\left(\sqrt{\langle P_L \rangle} - \sqrt{\langle P_1 \cdots P_L \cdots P_1 \rangle}\right)^2 - 1 + \langle P_L \rangle \leq -1 + \langle P_L \rangle = -\|Q_L\|^2. \quad (3.32)$$

Next, (3.30) follows from (3.13) of lemma 3.2. Then, (3.31) is a consequence of the triangle inequality

$$\sqrt{\langle P_1 \cdots P_{i-1} Q_i P_{i-1} \cdots P_1 \rangle} = \|Q_i P_{i-1} \cdots P_1\| \quad (3.33)$$

$$= \|Q_i(-I + I - P_{i-1} \cdots P_1)\| \quad (3.34)$$

$$\leq \|Q_i\| + \|Q_i(I - P_{i-1} \cdots P_1)\|, \quad (3.35)$$

under the convention that $P_{i-1} \cdots P_1 = I$ for $i = 1$. Continuing, we have that

$$\text{Equation (3.31)} = -\|Q_L\|^2 + 2 \sum_{i=1}^L \|Q_i\|^2 + 2 \sum_{i=1}^L (\|Q_i\| \|Q_i(I - P_{i-1} \cdots P_1)\|) \quad (3.36)$$

$$= -\|Q_L\|^2 + 2 \sum_{i=1}^L \|Q_i\|^2 + 2 \sum_{i=2}^L (\|Q_i\| \|Q_i(I - P_{i-1} \cdots P_1)\|) \quad (3.37)$$

$$\leq -\|Q_L\|^2 + 2 \sum_{i=1}^L \|Q_i\|^2 + \sum_{i=2}^L (c \|Q_i\|^2 + c^{-1} \|Q_i(I - P_{i-1} \cdots P_1)\|^2) \quad (3.38)$$

$$\leq -\|Q_L\|^2 + 2 \sum_{i=1}^L \|Q_i\|^2 + c \sum_{i=2}^L \|Q_i\|^2 + c^{-1} \sum_{i=1}^{L-1} \|Q_i\|^2 \quad (3.39)$$

$$\leq (1+c) \|Q_L\|^2 + (2+c^{-1}) \|Q_1\|^2 + (2+c+c^{-1}) \sum_{i=2}^{L-1} \|Q_i\|^2. \quad (3.40)$$

Equation (3.37) follows from the convention that $P_{i-1} \cdots P_1 = I$ for $i = 1$. Equation (3.38) is a consequence of the inequality $2xy \leq cx^2 + c^{-1}y^2$, holding for $x, y \in \mathbb{R}$ and $c > 0$. Finally, (3.39) is obtained by using lemma 3.3. ■

4. Generalization to positive operator-valued measures

Just as the bound from [27] was generalized in [28, Section 3] from projectors to positive semi-definite operators having eigenvalues between zero and one, we can do the same here. This generalization is useful for applications, and we discuss one such application in the next section.

We now give an extension of the quantum union bound in theorem 2.1 that applies for general measurements. The main idea behind it is the well-known Naimark extension theorem, following the approach from [28, Section 3].

Lemma 4.1. *Let ρ be a positive semi-definite operator acting on a separable Hilbert space \mathcal{H}_S , let $\{\Lambda_i\}_{i=1}^L$ denote a set of positive semi-definite operators such that $0 \leq \Lambda_i \leq I$ for all $i \in \{1, \dots, L\}$, and let $c > 0$. Then the following quantum union bound holds*

$$\begin{aligned} \text{Tr}\{\rho\} - \text{Tr}\{\Pi_{\Lambda_L} \cdots \Pi_{\Lambda_1}(\rho \otimes |\bar{0}\rangle\langle\bar{0}|_{p^L})\Pi_{\Lambda_1} \cdots \Pi_{\Lambda_L}\} &\leq (1+c) \text{Tr}\{(I - \Lambda_L)\rho\} \\ &+ (2+c+c^{-1}) \sum_{i=2}^{L-1} \text{Tr}\{(I - \Lambda_i)\rho\} + (2+c^{-1}) \text{Tr}\{(I - \Lambda_1)\rho\}, \end{aligned} \quad (4.1)$$

where $|\bar{0}\rangle_{p^L} \equiv |0\rangle_{p_1} \otimes \cdots \otimes |0\rangle_{p_L}$ is an auxiliary state of L qubit probe systems and Π_{Λ_i} is a projector defined as $\Pi_{\Lambda_i} \equiv U_i^\dagger P_i U_i$, for some unitary U_i and projector P_i such that

$$\text{Tr}\{\Pi_{\Lambda_i}(\rho \otimes |\bar{0}\rangle\langle\bar{0}|_{p^L})\} = \text{Tr}\{\Lambda_i \rho\}. \quad (4.2)$$

Proof. This extension of theorem 2.1 follows easily by employing the Naimark extension theorem. Concretely, to each operator Λ_i , we associate the following unitary:

$$U_{SP_i} \equiv \sqrt{I_S - (\Lambda_i)_S} \otimes [|0\rangle\langle 0|_{P_i} + |1\rangle\langle 1|_{P_i}] + \sqrt{(\Lambda_i)_S} \otimes [|1\rangle\langle 0|_{P_i} - |0\rangle\langle 1|_{P_i}]. \quad (4.3)$$

Then defining the projectors $\Pi_{\Lambda_i} \equiv U_{SP_i}^\dagger (I_S \otimes |1\rangle\langle 1|_{P_i}) U_{SP_i}$, a straightforward calculation gives that $\text{Tr}\{\Pi_{\Lambda_i}(\rho \otimes |\bar{0}\rangle\langle \bar{0}|_{P^L})\} = \text{Tr}\{\Lambda_i \rho_S\}$. Observe that the operator Π_{Λ_i} is an orthogonal projector (because it is Hermitian and idempotent), so that theorem 2.1 applies to each of these operators. Then (4.1) follows. ■

5. Lower bound on the second-order coding rate for classical communication

One application of our main result, theorem 2.1, is in achieving the second-order coding rate for classical communication. As we stated earlier, this area of quantum information theory has advanced in recent years [16–24], with one of the main reasons being the availability of the tunable parameter $c > 0$ in the Hayashi–Nagaoka inequality [15, Lemma 2]. That is, one can let $c > 0$ vary, to become closer to zero, as the number of channel uses increases.

An advantage of our theorem 2.1 is that it applies directly to the case of states and projectors that act on an infinite-dimensional, separable Hilbert space. Thus, the theorem can be applied directly in order to achieve a lower bound on the second-order coding rate for classical communication. To our knowledge, prior to our work here, Wilde *et al.* [20] presented the only case in which lower bounds on the second-order coding rates have been considered in this general case, and there the analysis was limited to channels that accept a classical input and output a pure quantum state. The situation that we analyse here is thus more general.

(a) Information quantities

Before we begin with the application, let us recall some information quantities that are essential in the analysis. Let \mathcal{H} denote a separable Hilbert space, and let $\mathcal{D}(\mathcal{H})$ denote the set of density operators acting on \mathcal{H} (positive, semi-definite operators with trace equal to one). Let spectral decompositions of $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ be given as

$$\rho = \sum_{x \in \mathcal{X}} \lambda_x P_x \quad \text{and} \quad \sigma = \sum_{y \in \mathcal{Y}} \mu_y Q_y, \quad (5.1)$$

where \mathcal{X} and \mathcal{Y} are countable index sets, $\{\lambda_x\}_{x \in \mathcal{X}}$ and $\{\mu_y\}_{y \in \mathcal{Y}}$ are probability distributions with $\lambda_x, \mu_y \geq 0$ for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ and $\sum_{x \in \mathcal{X}} \lambda_x = \sum_{y \in \mathcal{Y}} \mu_y = 1$, and $\{P_x\}_{x \in \mathcal{X}}$ and $\{Q_y\}_{y \in \mathcal{Y}}$ are sets of projections such that $\sum_{x \in \mathcal{X}} P_x = \sum_{y \in \mathcal{Y}} Q_y = I$.

The hypothesis testing relative entropy $D_H^\varepsilon(\rho \parallel \sigma)$ is defined for $\varepsilon \in [0, 1]$ as [33,34]

$$D_H^\varepsilon(\rho \parallel \sigma) \equiv -\log_2 \inf_{\Lambda} \{\text{Tr}\{\Lambda \sigma\} : \text{Tr}\{\Lambda \rho\} \geq 1 - \varepsilon \wedge 0 \leq \Lambda \leq I\}. \quad (5.2)$$

The quantum relative entropy [35], the quantum relative entropy variance [16,36,37], and the T quantity [16,36,37] are defined as

$$D(\rho \parallel \sigma) \equiv \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} \lambda_x \text{Tr}\{P_x Q_y\} \log_2 \left(\frac{\lambda_x}{\mu_y} \right), \quad (5.3)$$

$$V(\rho \parallel \sigma) \equiv \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} \lambda_x \text{Tr}\{P_x Q_y\} \left[\log_2 \left(\frac{\lambda_x}{\mu_y} \right) - D(\rho \parallel \sigma) \right]^2, \quad (5.4)$$

$$\text{and} \quad T(\rho \parallel \sigma) \equiv \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} \lambda_x \text{Tr}\{P_x Q_y\} \left| \log_2 \left(\frac{\lambda_x}{\mu_y} \right) - D(\rho \parallel \sigma) \right|^3. \quad (5.5)$$

For states ρ and σ satisfying

$$D(\rho \parallel \sigma), V(\rho \parallel \sigma), T(\rho \parallel \sigma) < \infty, \quad V(\rho \parallel \sigma) > 0, \quad (5.6)$$

the following expansion holds for the hypothesis testing relative entropy for $\varepsilon \in (0, 1)$ and a sufficiently large positive integer n :

$$D_H^\varepsilon(\rho^{\otimes n} \parallel \sigma^{\otimes n}) = nD(\rho \parallel \sigma) + \sqrt{nV(\rho \parallel \sigma)}\Phi^{-1}(\varepsilon) + O(\log n), \quad (5.7)$$

where

$$\Phi(a) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^a dx \exp\left(-\frac{x^2}{2}\right), \quad \Phi^{-1}(\varepsilon) \equiv \sup\{a \in \mathbb{R} \mid \Phi(a) \leq \varepsilon\}. \quad (5.8)$$

The equality in (5.7) was proven for finite-dimensional states ρ and σ in [16,36]. For the case of states acting on infinite-dimensional, separable Hilbert spaces, the inequality \leq in (5.7) was proven in [38] and [37, Appendix C]. In appendix A, we prove the inequality \geq in (5.7). The proof that we detail follows the development in [38, Appendix C] very closely, which is in turn based on [36, Section 3.2].

(b) Communication codes

We now recall what we mean by a code for classical communication and one for entanglement-assisted classical communication, starting with the former. Note that classical communication was considered for the asymptotic case in [39,40]. Suppose that a channel $\mathcal{N}_{A \rightarrow B}$ connects a sender Alice to a receiver Bob. For positive integers n and M , and $\varepsilon \in [0, 1]$, an (n, M, ε) code for classical communication consists of a set $\{\rho_{A^n}^m\}_{m \in \mathcal{M}}$ of quantum states, which are called quantum codewords, and where $|\mathcal{M}| = M$. It also consists of a decoding POVM $\{\Lambda_{B^n}^m\}_{m \in \mathcal{M}}$ and satisfies the following condition:

$$\frac{1}{M} \sum_{m \in \mathcal{M}} \text{Tr}\{(I_{B^n} - \Lambda_{B^n}^m) \mathcal{N}_{A \rightarrow B}^{\otimes n}(\rho_{A^n}^m)\} \leq \varepsilon, \quad (5.9)$$

which we interpret as saying that the average error probability is no larger than ε , when using the quantum codewords and decoding POVM described above. The non-asymptotic classical capacity of $\mathcal{N}_{A \rightarrow B}$, denoted by $C(\mathcal{N}_{A \rightarrow B}, n, \varepsilon)$ is equal to the largest value of $\frac{1}{n} \log_2 M$ (bits per channel use) for which there exists an (n, M, ε) code as described above.

Entanglement-assisted classical communication is defined similarly, but one allows for Alice and Bob to share an arbitrary quantum state $\Psi_{A'B'}$ before communication begins. Note that entanglement-assisted classical communication was considered for the asymptotic case in [41–43]. For positive integers n and M , and $\varepsilon \in [0, 1]$, an (n, M, ε) code for entanglement-assisted classical communication consists of the resource state $\Psi_{A'B'}$, a set $\{\mathcal{E}_{A' \rightarrow A^n}^m\}_{m \in \mathcal{M}}$ of encoding channels, where $|\mathcal{M}| = M$. It also consists of a decoding POVM $\{\Lambda_{B^n B'}^m\}_{m \in \mathcal{M}}$ and satisfies the following condition:

$$\frac{1}{M} \sum_{m \in \mathcal{M}} \text{Tr}\{(I_{B^n B'} - \Lambda_{B^n B'}^m) \mathcal{N}_{A \rightarrow B}^{\otimes n}(\mathcal{E}_{A' \rightarrow A^n}^m(\Psi_{A'B'}))\} \leq \varepsilon, \quad (5.10)$$

which we interpret as saying that the average error probability is no larger than ε , when using the entanglement-assisted code described above. The non-asymptotic entanglement-assisted classical capacity of $\mathcal{N}_{A \rightarrow B}$, denoted by $C_{\text{EA}}(\mathcal{N}_{A \rightarrow B}, n, \varepsilon)$ is equal to the largest value of $\frac{1}{n} \log_2 M$ (bits per channel use) for which there exists an (n, M, ε) entanglement-assisted code as described above.

(c) Lower bound on second-order coding rate

Defining the ε -mutual information of a bipartite state τ_{CD} as

$$I_H^\varepsilon(C; D)_\tau \equiv D_H^\varepsilon(\tau_{CD} \parallel \tau_C \otimes \tau_D), \quad (5.11)$$

the following inequality was proven recently in [44, Theorem 8] for the finite-dimensional case, improving upon a prior result from [19]:

$$C_{\text{EA}}(\mathcal{N}_{A \rightarrow B}, 1, \varepsilon) \geq I_H^{\varepsilon-\eta}(R; B)_\xi - \log_2 \left(\frac{4\varepsilon}{\eta^2} \right), \quad (5.12)$$

where $\varepsilon \in (0, 1)$, $\eta \in (0, \varepsilon)$, $\zeta_{RB} \equiv \mathcal{N}_{A \rightarrow B}(\rho_{RA})$, and ρ_{RA} is a bipartite state. The techniques employed in the proof of [44, Theorem 8] were position-based coding [29] and the Hayashi–Nagaoka inequality [15, Lemma 2]. Note that the position-based coding method can be understood as a variation of the well known and studied coding technique called pulse position modulation [45,46]. We now generalize the inequality in (5.12) to the infinite-dimensional case by applying theorem 2.1, along with position-based coding [29] and the sequential decoding strategy from [28].

Theorem 5.1. *Let \mathcal{H}_A , \mathcal{H}_B and \mathcal{H}_R be separable Hilbert spaces. Let $\mathcal{N}_{A \rightarrow B}$ be a quantum channel, taking $\mathcal{D}(\mathcal{H}_A)$ to $\mathcal{D}(\mathcal{H}_B)$. Then the following bound holds:*

$$C_{\text{EA}}(\mathcal{N}_{A \rightarrow B}, 1, \varepsilon) \geq I_H^{\varepsilon-\eta}(R; B)_\zeta - \log_2 \left(\frac{4\varepsilon}{\eta^2} \right), \quad (5.13)$$

where $\varepsilon \in (0, 1)$, $\eta \in (0, \varepsilon)$, $\zeta_{RB} \equiv \mathcal{N}_{A \rightarrow B}(\rho_{RA})$, and $\rho_{RA} \in \mathcal{D}(\mathcal{H}_R \otimes \mathcal{H}_A)$ is a bipartite state.

Proof. Let Λ_{RB} be a measurement operator (i.e., $0 \leq \Lambda_{RB} \leq I_{RB}$) satisfying

$$\text{Tr}\{(I_{RB} - \Lambda_{RB})\mathcal{N}_{A \rightarrow B}(\rho_{RA})\} \leq \varepsilon - \eta. \quad (5.14)$$

To this operator Λ_{RB} is associated a unitary U_{RBP} , defined as

$$U_{RBP} \equiv \sqrt{I_{RB} - \Lambda_{RB}} \otimes [|0\rangle\langle 0|_P + |1\rangle\langle 1|_P] + \sqrt{\Lambda_{RB}} \otimes [|1\rangle\langle 0|_P - |0\rangle\langle 1|_P]. \quad (5.15)$$

Then defining the projectors

$$\Pi_{RBP} \equiv U_{RBP}^\dagger (I_{RB} \otimes |1\rangle\langle 1|_P) U_{RBP} \quad (5.16)$$

and

$$\hat{\Pi}_{RBP} \equiv I_{RBP} - \Pi_{RBP} = U_{RBP}^\dagger (I_{RB} \otimes |0\rangle\langle 0|_P) U_{RBP}, \quad (5.17)$$

the inequality in (5.14) and a simple calculation imply that

$$\text{Tr}\{(I_{RBP} - \Pi_{RBP})\mathcal{N}_{A \rightarrow B}(\rho_{RA}) \otimes |0\rangle\langle 0|_P\} \leq \varepsilon - \eta. \quad (5.18)$$

This kind of construction and equality is known as the Naimark extension theorem.

The position-based coding strategy then proceeds as follows. We let Alice and Bob share M copies of the resource state ρ_{RA} , where Bob has the R systems and Alice the A systems. If Alice would like to transmit message $m \in \mathcal{M}$, then she simply selects the m th A system, and sends it through the channel $\mathcal{N}_{A \rightarrow B}$. The marginal state of Bob's systems is then as follows:

$$\rho_{R_1} \otimes \cdots \otimes \rho_{R_{m-1}} \otimes \mathcal{N}_{A_m \rightarrow B}(\rho_{R_m A_m}) \otimes \rho_{R_{m+1}} \otimes \cdots \otimes \rho_{R_M}. \quad (5.19)$$

Bob then uses a sequential decoding strategy to determine which message Alice transmitted. He introduces M auxiliary probe systems in the state $|0\rangle\langle 0|$, so that Bob's overall state is now

$$\begin{aligned} \omega_{RMBP}^m \equiv & \rho_{R_1} \otimes \cdots \otimes \rho_{R_{m-1}} \otimes \mathcal{N}_{A_m \rightarrow B}(\rho_{R_m A_m}) \otimes \rho_{R_{m+1}} \\ & \otimes \cdots \otimes \rho_{R_M} \otimes |0\rangle\langle 0|_{P_1} \otimes \cdots \otimes |0\rangle\langle 0|_{P_M}. \end{aligned} \quad (5.20)$$

He then performs the binary measurements $\{\Pi_{R_i BP_i}, \hat{\Pi}_{R_i BP_i}\}$ sequentially, in the order $i = 1, i = 2$, etc. With this strategy, the probability that he decodes the m th message correctly is given by

$$\text{Tr}\{\Pi_{R_m BP_m} \hat{\Pi}_{R_{m-1} BP_{m-1}} \cdots \hat{\Pi}_{R_1 BP_1} \omega_{RMBP}^m \hat{\Pi}_{R_1 BP_1} \cdots \hat{\Pi}_{R_{m-1} BP_{m-1}}\}. \quad (5.21)$$

Applying theorem 2.1, we can bound the complementary (error) probability as

$$p_e(m) \equiv 1 - \text{Tr} \left\{ \Pi_{R_m B P_m} \hat{\Pi}_{R_{m-1} B P_{m-1}} \cdots \hat{\Pi}_{R_1 B P_1} \omega_{R^M B P^M}^m \hat{\Pi}_{R_1 B P_1} \cdots \hat{\Pi}_{R_{m-1} B P_{m-1}} \right\} \quad (5.22)$$

$$\leq (1+c) \text{Tr} \{ \hat{\Pi}_{R_m B P_m} \omega_{R^M B P^M}^m \} + (2+c+c^{-1}) \sum_{i=1}^{m-1} \text{Tr} \{ \Pi_{R_i B P_i} \omega_{R^M B P^M}^m \} \quad (5.23)$$

$$= (1+c) \text{Tr} \{ (\Lambda_{RB} - \Lambda_{RB}) \mathcal{N}_{A \rightarrow B}(\rho_{RA}) \} + (2+c+c^{-1})(m-1) \text{Tr} \{ \Lambda_{RB} [\rho_R \otimes \mathcal{N}_{A \rightarrow B}(\rho_A)] \} \quad (5.24)$$

$$\leq (1+c)(\varepsilon - \eta) + (2+c+c^{-1})M \text{Tr} \{ \Lambda_{RB} [\rho_R \otimes \mathcal{N}_{A \rightarrow B}(\rho_A)] \}, \quad (5.25)$$

where $c > 0$. Since the whole development above holds for all measurement operators Λ_{RB} satisfying (5.14), we can take an infimum over all of them to obtain the following uniform bound on the error probability when sending an arbitrary message $m \in \mathcal{M}$:

$$p_e(m) \leq (1+c)(\varepsilon - \eta) + (2+c+c^{-1})M 2^{-I_H^{\varepsilon-\eta}(R;B)_\zeta}. \quad (5.26)$$

Picking $c = \eta/(2\varepsilon - \eta)$ and taking M such that

$$\log_2 M = I_H^{\varepsilon-\eta}(R;B)_\zeta - \log_2 \left(\frac{4\varepsilon}{\eta^2} \right) \quad (5.27)$$

then implies that $p_e(m) \leq \varepsilon$ for all $m \in \mathcal{M}$. Since we have shown the existence of a $(1, M, \varepsilon)$ entanglement-assisted code, where M satisfies (5.27), this concludes the proof. ■

Remark 5.2. It is worthwhile to note that the code above has an error probability less than ε for every message $m \in \mathcal{M}$, and so the error criterion is maximal error probability and not just average error probability.

The above result also implies rates that are achievable for unassisted classical communication, by combining theorem 5.1 with an analysis nearly identical to that given in [24, Section 3.3]. In particular, we could allow Alice and Bob to share many copies of the following classical–quantum state before communication begins:

$$\rho_{XA} \equiv \sum_{x \in \mathcal{X}} p(x) |x\rangle \langle x|_X \otimes \rho_A^x, \quad (5.28)$$

where \mathcal{H}_X and \mathcal{H}_A are separable Hilbert spaces, \mathcal{X} is a countable index set, $\{p(x)\}_{x \in \mathcal{X}}$ is a probability distribution, $\{|x\rangle_X\}_{x \in \mathcal{X}}$ is a set of orthonormal states, and $\{\rho_A^x\}_{x \in \mathcal{X}}$ is a set of states. This state then plays the role of the resource state ρ_{RA} in the proof of theorem 5.1. However, the above state is a classical–quantum state, and as such, the code can be derandomized. Specifically, to do so, we can employ the analysis given in [24, Section 3.3], but replacing the square-root measurement there with the sequential decoding strategy. This leads to the following result, which generalizes one of the main results of [34] to the infinite-dimensional case:

Corollary 5.3. Let \mathcal{H}_A , \mathcal{H}_B , and \mathcal{H}_X be separable Hilbert spaces. Let $\mathcal{N}_{A \rightarrow B}$ be a quantum channel, taking $\mathcal{D}(\mathcal{H}_A)$ to $\mathcal{D}(\mathcal{H}_B)$. Then the following bound holds:

$$C(\mathcal{N}_{A \rightarrow B}, 1, \varepsilon) \geq I_H^{\varepsilon-\eta}(X;B)_\zeta - \log_2 \left(\frac{4\varepsilon}{\eta^2} \right), \quad (5.29)$$

where $\varepsilon \in (0, 1)$, $\eta \in (0, \varepsilon)$, $\zeta_{XB} \equiv \mathcal{N}_{A \rightarrow B}(\rho_{XA})$, and $\rho_{XA} \in \mathcal{D}(\mathcal{H}_X \otimes \mathcal{H}_A)$ is a bipartite, classical–quantum state of the form in (5.28).

(d) Energy constraints

It is common in the theory of communication over infinite-dimensional channels [30,31,47] to impose energy constraints on the space of inputs. If we do not so, then the capacities can be infinite. The definitions of these energy-constrained non-asymptotic capacities are the same as we discussed previously, except that we impose energy constraints on the channel input states.

Before defining them, let us first recall the notion of an energy observable [47,48]:

Definition 5.4 (Energy observable). For a Hilbert space \mathcal{H} , let $G \in \mathcal{L}_+(\mathcal{H})$ denote a positive semi-definite operator, defined in terms of its action on a vector $|\psi\rangle$ as

$$G|\psi\rangle = \sum_{j=1}^{\infty} g_j |e_j\rangle \langle e_j| \psi\rangle, \quad (5.30)$$

for $|\psi\rangle$ such that $\sum_{j=1}^{\infty} g_j |\langle e_j|\psi\rangle|^2 < \infty$. In the above, $\{|e_j\rangle\}_j$ is an orthonormal basis and $\{g_j\}_j$ is a sequence of non-negative, real numbers. Then $\{|e_j\rangle\}_j$ is an eigenbasis for G with corresponding eigenvalues $\{g_j\}_j$.

Definition 5.5. The n th extension \bar{G}_n of an energy observable G is defined as

$$\bar{G}_n \equiv \frac{1}{n} [G \otimes I \otimes \cdots \otimes I + \cdots + I \otimes \cdots \otimes I \otimes G], \quad (5.31)$$

where n is the number of factors in each tensor product above.

Then the non-asymptotic, energy-constrained classical capacity $C(\mathcal{N}_{A \rightarrow B}, G, P, n, \varepsilon)$ is defined exactly as it was previously in §5b, except that we demand that

$$\frac{1}{M} \sum_{m \in \mathcal{M}} \text{Tr}[\bar{G}_n \rho_{A^n}^m] \leq P, \quad (5.32)$$

for a real $P \in [0, \infty)$. Similarly, the non-asymptotic, energy-constrained entanglement-assisted classical capacity $C_{\text{EA}}(\mathcal{N}_{A \rightarrow B}, G, P, n, \varepsilon)$ is defined exactly as it was previously in §5b, except that we demand that

$$\frac{1}{M} \sum_{m \in \mathcal{M}} \text{Tr}[(\bar{G}_n \otimes I_{B'}) \mathcal{E}_{A' \rightarrow A^n}^m(\Psi_{A'B'})] \leq P. \quad (5.33)$$

One could alternatively demand that the energy constraint hold for every codeword, not just on average. Note that we recover the usual notion of capacity (unconstrained) by taking $G = I$ and setting $P = 1$.

An advantage of the approach given in the proof of theorem 5.1 is that we easily obtain a lower bound on the second-order coding rate for energy-constrained entanglement-assisted classical communication over a quantum channel:

Theorem 5.6. Let \mathcal{H}_A , \mathcal{H}_B , and \mathcal{H}_R be separable Hilbert spaces. Let $\varepsilon \in (0, 1)$. Let G be an energy observable, and let $P \in [0, \infty)$. Let $\mathcal{N}_{A \rightarrow B}$ be a quantum channel, taking $\mathcal{D}(\mathcal{H}_A)$ to $\mathcal{D}(\mathcal{H}_B)$. Then the following bound holds:

$$C_{\text{EA}}(\mathcal{N}_{A \rightarrow B}, G, P, n, \varepsilon) \geq I(R; B)_{\zeta} + \sqrt{\frac{1}{n} V(R; B)_{\zeta}} \Phi^{-1}(\varepsilon) + O\left(\frac{1}{n} \log n\right), \quad (5.34)$$

where $\zeta_{RB} \equiv \mathcal{N}_{A \rightarrow B}(\rho_{RA})$ and $\rho_{RA} \in \mathcal{D}(\mathcal{H}_R \otimes \mathcal{H}_A)$ is a bipartite state such that

$$I(R; B)_{\zeta}, V(R; B)_{\zeta}, T(R; B)_{\zeta} < \infty, \quad V(R; B)_{\zeta} > 0, \quad (5.35)$$

and $\text{Tr}[G \rho_A] \leq P$. In the above, we have the mutual information, mutual information variance and another quantity:

$$I(R; B)_{\zeta} \equiv D(\zeta_{RB} \| \zeta_R \otimes \zeta_B), \quad V(R; B)_{\zeta} \equiv V(\zeta_{RB} \| \zeta_R \otimes \zeta_B), \quad T(R; B)_{\zeta} \equiv T(\zeta_{RB} \| \zeta_R \otimes \zeta_B). \quad (5.36)$$

Proof. Let ζ_{RA} be a state satisfying the conditions stated above. Then the claim follows by applying theorem 5.1, picking $\eta = 1/\sqrt{n}$, and invoking the expansion in (5.7). ■

The proof given above is quite simple once all of the relevant components are in place (namely, the quantum union bound from theorem 2.1, position-based coding [29], and the expansion in (5.7)). This is to be contrasted with the approach taken in [30,31], in which the energy-constrained entanglement-assisted classical capacity was identified. Not only can we argue to have a simpler

approach for the achievability part, but our method also delivers a lower bound on the second-order coding rate. An important open question remaining however is to determine whether this lower bound on the second-order coding rate is tight. To our knowledge, this tightness has only been shown for finite-dimensional channels that are covariant [19].

We note here that the bound in theorem 5.6 applies to the practically relevant case of bosonic Gaussian channels [14]. The energy-constrained entanglement-assisted classical capacity of these channels was identified in [30,31,49,50]. The additive-noise, thermal and amplifier channels are of major interest for applications, as stressed in [47,51]. It is known that the energy-constrained, entanglement-assisted capacity formula for these channels is achieved by a two-mode squeezed vacuum state, whose reduction to the channel input system has an average photon number meeting the desired photon number constraint. Thus, we could evaluate the lower bound in theorem 5.6 by taking ρ_{RA} therein to be the two-mode squeezed vacuum and then applying the formula from [52] to evaluate the mutual information variance $V(R;B)_\zeta$, while noting that the quantity $T(R;B)_\zeta$ is finite for any finite-energy state, as proven in [37, Appendix D].

We end this section on a different note, by remarking that the same argument as above gives a non-trivial lower bound on the second-order coding rate for energy-constrained classical communication with randomness assistance. However, it remains open to determine whether this rate is achievable without the assistance of randomness. It is also open to extend the result to a continuous (uncountable) index set \mathcal{X} . We suspect that these extensions should be possible but leave it for future work.

6. Conclusion

In this paper, we proved theorem 2.1, which improves Gao's quantum union bound to include a tunable parameter $c > 0$ that plays a role similar to the tunable parameter available in the Hayashi–Nagaoka inequality from [15, Lemma 2]. An advantage of the proof of theorem 2.1 is that it is elementary, relying only on basic properties of projectors, Pythagoras' theorem, and the Cauchy–Schwarz inequality. Owing to our improvement, the quantum union bound can now be employed in a wider variety of scenarios. As an example application, we showed how to achieve a lower bound on the second-order coding rate for classical communication over a quantum channel by employing a sequential decoding strategy.

For future directions, we think it would be interesting to determine whether the improved bound in theorem 2.1 would find application in areas including quantum algorithms [8–10], quantum complexity theory [9,11] and Hamiltonian complexity theory [12,13]. We also wonder whether theorem 2.1 could be useful outside of quantum information, for example, in the analysis of projection algorithms [55].

Data accessibility. This work does not have any experimental data.

Authors' contributions. All authors contributed equally toward proving the main results and writing the paper.

Competing interests. We have no competing interests.

Funding. M.M.W. acknowledges support from the US National Science Foundation under grant no. 1714215.

Acknowledgements. M.M.W. is grateful to S.M. for hosting him for a research visit to University of Camerino in June and July, 2016, during which this research project was initiated. S.M. is grateful to MMW for hosting him for a research visit to Louisiana State University at Baton Rouge in March, 2018, during which this research project was completed.

Appendix A. Proof of the inequality \geq in equation (5.7)

The goal of this appendix is to prove the inequality \geq in (5.7). The proof follows the development in appendix C of [38] very closely, which is in turn based on [36, Section 3.2].

Consider quantum states ρ and σ acting on a separable Hilbert space \mathcal{H} , with spectral decompositions as given in (5.1). Observe that each P_x is finite-dimensional, as a consequence of the fact that ρ is trace class. Indeed, were it not the case, then $\text{Tr}\{P_x\}$ would be infinite, and ρ could thus not be trace class. By the same reasoning, each Q_y is finite-dimensional.

By defining a random variable Z taking values $\log_2(\lambda_x/\mu_y)$ with probability $p(x, y) \equiv \lambda_x \text{Tr}\{P_x Q_y\}$, observe that

$$D(\rho \parallel \sigma) = \mathbb{E}\{Z\}, \quad V(\rho \parallel \sigma) = \text{Var}\{Z\}, \quad T(\rho \parallel \sigma) = \mathbb{E}\left\{|Z - \mathbb{E}\{Z\}|^3\right\}, \quad (\text{A } 1)$$

where $D(\rho \parallel \sigma)$, $V(\rho \parallel \sigma)$, and $T(\rho \parallel \sigma)$ are defined in (5.3)–(5.5).

Lemma A.1. *Let ρ and σ denote states acting on a separable Hilbert space \mathcal{H} . Let $L > 0$. Then there exists a measurement operator T_L (i.e. $0 \leq T_L \leq I$) such that*

$$\text{Tr}\{T_L \rho\} \geq \Pr\{Z \geq \log_2 L\}, \quad \text{Tr}\{T_L \sigma\} \leq \frac{1}{L}, \quad (\text{A } 2)$$

where Z is the random variable defined just before (A 1).

Proof. Let us define the positive semi-definite operator \tilde{T}_L as

$$\tilde{T}_L \equiv \sum_{x, y: L \leq \lambda_x / \mu_y} Q_y P_x Q_y. \quad (\text{A } 3)$$

By inspection, this operator is positive semi-definite. The measurement operator T_L is then defined to be the projection onto the support of \tilde{T}_L . Let $|\psi\rangle$ be a unit vector such that $P_x |\psi\rangle = |\psi\rangle$ for some x . It follows that $|\psi\rangle\langle\psi| \leq P_x$. Then, for any μ_y such that $L \leq \lambda_x / \mu_y$, we have, from the definition of \tilde{T}_L , that $|\psi\rangle\langle\psi| \leq \tilde{T}_L$. This in turn implies that $Q_y |\psi\rangle \in \text{supp}(\tilde{T}_L)$. From this, we then conclude that

$$\frac{Q_y |\psi\rangle\langle\psi| Q_y}{\|Q_y |\psi\rangle\|^2} \leq T_L. \quad (\text{A } 4)$$

Furthermore, we have that $\{Q_y |\psi\rangle\}_y$ forms a family of orthogonal vectors. Then the following inequality holds

$$\sum_{y: L \leq \lambda_x / \mu_y} \frac{Q_y |\psi\rangle\langle\psi| Q_y}{\|Q_y |\psi\rangle\|^2} \leq T_L. \quad (\text{A } 5)$$

From this, we conclude that

$$\text{Tr}\{T_L |\psi\rangle\langle\psi|\} \geq \text{Tr} \left\{ \sum_{y: L \leq \lambda_x / \mu_y} \frac{Q_y |\psi\rangle\langle\psi| Q_y}{\|Q_y |\psi\rangle\|^2} |\psi\rangle\langle\psi| \right\} = \sum_{y: L \leq \lambda_x / \mu_y} \text{Tr} \left\{ \frac{Q_y |\psi\rangle\langle\psi| Q_y}{\|Q_y |\psi\rangle\|^2} |\psi\rangle\langle\psi| \right\} \quad (\text{A } 6)$$

$$= \sum_{y: L \leq \lambda_x / \mu_y} \frac{\|Q_y |\psi\rangle\|^4}{\|Q_y |\psi\rangle\|^2} = \sum_{y: L \leq \lambda_x / \mu_y} \|Q_y |\psi\rangle\|^2 = \sum_{y: L \leq \lambda_x / \mu_y} \text{Tr}\{Q_y |\psi\rangle\langle\psi| Q_y\}. \quad (\text{A } 7)$$

Now, recall that P_x is a finite-dimensional projector for each x . (As stated above, if P_x were not finite-dimensional, then this would contradict the assumption that ρ is trace class.) Furthermore, we can write it as $P_x = \sum_{l=1}^{\text{Tr}\{P_x\}} |\psi_{x,l}\rangle\langle\psi_{x,l}|$, for some orthonormal set $\{|\psi_{x,l}\rangle\}_{l=1}^{\text{Tr}\{P_x\}}$. Then the development in (A 6) and (A 7) implies that

$$\text{Tr}\{T_L P_x\} \geq \sum_{y: L \leq \lambda_x / \mu_y} \text{Tr}\{Q_y P_x Q_y\} = \sum_{y: L \leq \lambda_x / \mu_y} \text{Tr}\{Q_y P_x\}. \quad (\text{A } 8)$$

We can then use this to conclude that

$$\text{Tr}\{T_L \rho\} = \sum_x \lambda_x \text{Tr}\{T_L P_x\} \geq \sum_{x, y: L \leq \lambda_x / \mu_y} \lambda_x \text{Tr}\{Q_y P_x\} \quad (\text{A } 9)$$

$$= \sum_{x, y: \log_2 L \leq \log_2(\lambda_x / \mu_y)} \lambda_x \text{Tr}\{Q_y P_x\} = \Pr\{Z \geq \log_2 L\}, \quad (\text{A } 10)$$

where the second equality uses the fact that $\log_2 : (0, \infty) \rightarrow (-\infty, \infty)$ is invertible, and the last line follows from the definition of the random variable Z , given just before (A 1).

What remains is to place an upper bound on $\text{Tr}\{T_L\sigma\}$. Observe that for all x and y , the following equivalence holds:

$$\text{ran}(Q_y P_x) = \text{supp}(Q_y P_x Q_y), \quad (\text{A } 11)$$

where ran denotes the range of an operator. For some x , define the following subspace:

$$\tilde{S}_x \equiv \bigvee_{y: L \leq \lambda_x / \mu_y} \text{ran}(Q_y P_x) = \bigvee_{y: L \leq \lambda_x / \mu_y} \text{supp}(Q_y P_x Q_y), \quad (\text{A } 12)$$

where the operation \vee realizes a space formed as the union of subspaces. Due to the fact that P_x is finite-dimensional, it follows that the subspace \tilde{S}_x is finite-dimensional. We now employ a Gram–Schmidt orthogonalization procedure for these subspaces. First order the eigenvalues of ρ as $\lambda_1 > \lambda_2 > \dots$. Now define a family $\{S_x\}_x$ of subspaces of the whole Hilbert space \mathcal{H} as

$$S_1 \equiv \tilde{S}_1, \quad S_x \equiv \left(\bigvee_{i=1}^x \tilde{S}_i \right) \wedge \left(\bigvee_{i=1}^{x-1} \tilde{S}_i \right)^\perp \quad \text{for } x \geq 2, \quad (\text{A } 13)$$

where the operation \wedge corresponds to the intersection of subspaces. The subspaces S_x are mutually orthogonal by construction. Furthermore, by the procedure given above, the following holds for any positive integer $w \geq 1$:

$$\bigvee_{x=1}^w \tilde{S}_x = \bigvee_{x=1}^w S_x. \quad (\text{A } 14)$$

By definition, T_L is the projection onto the following subspace:

$$\bigvee_{x, y: L \leq \lambda_x / \mu_y} \text{supp}(Q_y P_x Q_y) = \bigvee_{y: L \leq \lambda_x / \mu_y} \text{ran}(Q_y P_x) = \bigvee_x \tilde{S}_x = \bigvee_x S_x = \bigoplus_x S_x. \quad (\text{A } 15)$$

Thus, it follows that T_L can be written as $T_L = \sum_x P_{S_x}$, where P_{S_x} is the projection onto S_x . By the procedure given above, we have that $S_x \subseteq \tilde{S}_x$, and from the definition of \tilde{S}_x , we find that

$$\text{Tr}\{P_{S_x}\} \leq \text{Tr}\{P_{\tilde{S}_x}\} \leq \text{Tr}\{P_x\}. \quad (\text{A } 16)$$

We then find that

$$\text{Tr}\{T_L\sigma\} = \sum_{y, x} \mu_y \text{Tr}\{Q_y P_{S_x}\} = \sum_{y, x: L \leq \lambda_x / \mu_y} \mu_y \text{Tr}\{Q_y P_{S_x}\} \quad (\text{A } 17)$$

$$\leq \frac{1}{L} \sum_{y, x: L \leq \lambda_x / \mu_y} \lambda_x \text{Tr}\{Q_y P_{S_x}\} \leq \frac{1}{L} \sum_{y, x} \lambda_x \text{Tr}\{Q_y P_{S_x}\} \quad (\text{A } 18)$$

$$= \frac{1}{L} \sum_x \lambda_x \text{Tr}\{P_{S_x}\} \leq \frac{1}{L} \sum_x \lambda_x \text{Tr}\{P_x\} = \frac{1}{L} \text{Tr}\{\rho\} = \frac{1}{L}. \quad (\text{A } 19)$$

In the above, the second equality follows because $Q_y P_{S_x} = 0$ unless $L \leq \lambda_x / \mu_y$ (from the definition of the space S_x and the fact that $S_x \subseteq \tilde{S}_x$). The third equality follows from $\sum_y Q_y = I$, and the third inequality follows from (A 16). ■

We now apply the above lemma to the i.i.d. states $\rho^{\otimes n}$ and $\sigma^{\otimes n}$, with spectral decompositions

$$\rho^{\otimes n} = \sum_{x^n} \lambda_{x^n} P_{x^n}, \quad \sigma^{\otimes n} = \sum_{y^n} \mu_{y^n} Q_{y^n}, \quad (\text{A } 20)$$

where $x^n = (x_1, \dots, x_n)$, $y^n = (y_1, \dots, y_n)$, $\lambda_{x^n} = \lambda_{x_1} \times \dots \times \lambda_{x_n}$, $\mu_{y^n} = \mu_{y_1} \times \dots \times \mu_{y_n}$, $P_{x^n} = P_{x_1} \otimes \dots \otimes P_{x_n}$, and $Q_{y^n} = Q_{y_1} \otimes \dots \otimes Q_{y_n}$. Then the i.i.d. random sequence $Z^n \equiv (Z_1, \dots, Z_n)$ takes the

values

$$\log_2 \left(\frac{\lambda_{x^n}}{\mu_{y^n}} \right) = \sum_{i=1}^n \log_2 \left(\frac{\lambda_{x_i}}{\mu_{y_i}} \right), \quad (\text{A } 21)$$

with probability

$$p(x^n, y^n) = \lambda_{x^n} \text{Tr}\{P_{x^n} Q_{y^n}\} = \prod_{i=1}^n \lambda_{x_i} \text{Tr}\{P_{x_i} Q_{y_i}\}. \quad (\text{A } 22)$$

The Berry–Essen theorem [53,54] states that if A_1, \dots, A_n are i.i.d. random variables such that $\mathbb{E}\{A_1\} = 0$, $\mathbb{E}\{|A_1|^2\} = \tau^2 \in (0, \infty)$, and $\mathbb{E}\{|A_1|^3\} = \omega < \infty$, then

$$\left| \Pr \left\{ \frac{B_n \sqrt{n}}{\tau} \leq x \right\} - \Phi(x) \right| \leq \frac{C\omega}{\tau^3 \sqrt{n}}, \quad (\text{A } 23)$$

where $x \in \mathbb{R}$, $\Phi(x) \equiv [2\pi]^{-1/2} \int_{-\infty}^x dy \exp(-y^2/2)$, $B_n \equiv (1/n) \sum_{i=1}^n A_i$, and C is the Berry–Esseen constant satisfying $0.40973 \leq C \leq 0.4784$.

Proposition A.2. Let ρ and σ denote states acting on a separable Hilbert space \mathcal{H} . Suppose that $D(\rho \parallel \sigma), V(\rho \parallel \sigma), T(\rho \parallel \sigma) < \infty$ and $V(\rho \parallel \sigma) > 0$. Suppose n is sufficiently large such that $\varepsilon - C \cdot T(\rho \parallel \sigma) / \sqrt{n[V(\rho \parallel \sigma)]^3} > 0$. Then

$$D_H^\varepsilon(\rho^{\otimes n} \parallel \sigma^{\otimes n}) \geq nD(\rho \parallel \sigma) + \sqrt{nV(\rho \parallel \sigma)} \Phi^{-1} \left(\frac{\varepsilon - C \cdot T(\rho \parallel \sigma)}{\sqrt{n[V(\rho \parallel \sigma)]^3}} \right) \quad (\text{A } 24)$$

$$= nD(\rho \parallel \sigma) + \sqrt{nV(\rho \parallel \sigma)} \Phi^{-1}(\varepsilon) + O(1). \quad (\text{A } 25)$$

Proof. Applying the Berry–Esseen theorem to the random sequence $Z_1 - D(\rho \parallel \sigma), \dots, Z_n - D(\rho \parallel \sigma)$, with Z_i defined in (A 21) and (A 22), we find that

$$\left| \Pr \left\{ \bar{Z}^n \sqrt{\frac{n}{V(\rho \parallel \sigma)}} \leq x \right\} - \Phi(x) \right| \leq \frac{C \cdot T(\rho \parallel \sigma)}{\sqrt{n[V(\rho \parallel \sigma)]^3}}, \quad (\text{A } 26)$$

where $\bar{Z}^n \equiv (1/n) \sum_{i=1}^n [Z_i - D(\rho \parallel \sigma)]$, which implies that

$$\Pr \left\{ \sum_{i=1}^n Z_i \leq nD(\rho \parallel \sigma) + x \sqrt{nV(\rho \parallel \sigma)} \right\} \leq \Phi(x) + \frac{C \cdot T(\rho \parallel \sigma)}{\sqrt{n[V(\rho \parallel \sigma)]^3}}. \quad (\text{A } 27)$$

Picking $x = \Phi^{-1}(\varepsilon - C \cdot T(\rho \parallel \sigma) / \sqrt{n[V(\rho \parallel \sigma)]^3})$, this becomes

$$\Pr \left\{ \sum_{i=1}^n Z_i \leq nD(\rho \parallel \sigma) + \sqrt{nV(\rho \parallel \sigma)} \Phi^{-1} \left(\varepsilon - \frac{C \cdot T(\rho \parallel \sigma)}{\sqrt{n[V(\rho \parallel \sigma)]^3}} \right) \right\} \leq \varepsilon. \quad (\text{A } 28)$$

Choosing L such that

$$\log_2 L = nD(\rho \parallel \sigma) + \sqrt{nV(\rho \parallel \sigma)} \Phi^{-1} \left(\varepsilon - \frac{C \cdot T(\rho \parallel \sigma)}{\sqrt{n[V(\rho \parallel \sigma)]^3}} \right) \quad (\text{A } 29)$$

and applying lemma A.1, we find that

$$\text{Tr}\{T^n \rho^{\otimes n}\} \geq \Pr \left\{ \sum_{i=1}^n Z_i \geq \log_2 L \right\} = 1 - \Pr \left\{ \sum_{i=1}^n Z_i \leq \log_2 L \right\} \geq 1 - \varepsilon,$$

while

$$\text{Tr}\{T^n \sigma^{\otimes n}\} \leq \frac{1}{L} = e^{-[nD(\rho \parallel \sigma) + \sqrt{nV(\rho \parallel \sigma)} \Phi^{-1}(\varepsilon - C \cdot T(\rho \parallel \sigma) / \sqrt{n[V(\rho \parallel \sigma)]^3})]}. \quad (\text{A } 30)$$

This implies that

$$-\log_2 \text{Tr}\{T^n \sigma^{\otimes n}\} \geq nD(\rho \parallel \sigma) + \sqrt{nV(\rho \parallel \sigma)} \Phi^{-1} \left(\varepsilon - \frac{C \cdot T(\rho \parallel \sigma)}{\sqrt{n[V(\rho \parallel \sigma)]^3}} \right). \quad (\text{A } 31)$$

Since $D_H^\varepsilon(\rho^{\otimes n} \parallel \sigma^{\otimes n})$ involves an optimization over all possible measurement operators T^n satisfying $\text{Tr}\{T^n \rho^{\otimes n}\} \geq 1 - \varepsilon$, we conclude the bound in (A 24). The equality after (A 24) follows from expanding Φ^{-1} at the point ε using Lagrange's mean value theorem. ■

References

1. Rozanov YA. 1977 *Probability theory: a concise course*. Dover Books on Mathematics. New York, NY: Dover Publications.
2. Aaronson S. 2018 How to upper-bound the probability of something bad. See <https://www.scottaaronson.com/blog/?p=3712>.
3. Gao J. 2015 Quantum union bounds for sequential projective measurements. *Physical Review A* **92**, 052331. (doi:10.1103/PhysRevA.92.052331)
4. Hirche C. 2015 Polar codes in quantum information theory. Master's thesis, Leibniz Universität Hannover. (<http://arxiv.org/abs/1501.03737>)
5. Boche H, Cai N, Nötzel J. 2016 The classical-quantum channel with random state parameters known to the sender. *J. Phys. A: Math. Theor.* **49**, 195302. (doi:10.1088/1751-8113/49/19/195302)
6. Renes JM. 2017 Belief propagation decoding of quantum channels by passing quantum messages. *New J. Phys.* **19**, 072001. (doi:10.1088/1367-2630/aa7c78)
7. Yamasaki H, Soeda A, Murao M. 2017 Graph-associated entanglement cost of a multipartite state in exact and finite-block-length approximate constructions. *Phys. Rev. A* **96**, 032330. (doi:10.1103/PhysRevA.96.032330)
8. Montanaro A, deWolf R. 2016 *A survey of quantum property testing*. Graduate Surveys, no. 7. Theory of Computing Library. (<http://arxiv.org/abs/1310.2035>)
9. Aaronson S. 2016 The complexity of quantum states and transformations: from quantum money to black holes. (<http://arxiv.org/abs/1607.05256>)
10. Harrow AW, Yen-Yu Lin C, Montanaro A. 2017 Sequential measurements, disturbance and property testing. In *Proc. of the Twenty-Eighth Annual ACM-SIAM Symp. on Discrete Algorithms*, pp. 1598–1611. (<http://arxiv.org/abs/1607.03236>)
11. BeneWatts A, Harrow AW, Kanwar G, Natarajan A. 2018 Algorithms, bounds, and strategies for entangled XOR games. (<http://arxiv.org/abs/1801.00821>)
12. Anshu A, Arad I, Vidick T. 2016 Simple proof of the detectability lemma and spectral gap amplification. *Phys. Rev. B* **93**, 205142. (doi:10.1103/PhysRevB.93.205142)
13. Kastoryano MJ, Lucia A. 2018 Divide and conquer method for proving gaps of frustration free Hamiltonians. *J. Stat. Mech: Theory Exp.* **2018**, 033105. (doi:10.1088/1742-5468/aaa793)
14. Serafini A. 2017 *Quantum continuous variables*. Boca Raton, FL: CRC Press.
15. Hayashi M, Nagaoka H. 2003 General formulas for capacity of classical-quantum channels. *IEEE Trans. Inform. Theory* **49**, 1753–1768. (doi:10.1109/TIT.2003.813556)
16. Tomamichel M, Hayashi M. 2013 A hierarchy of information quantities for finite block length analysis of quantum tasks. *IEEE Trans. Inform. Theory* **59**, 7693–7710. (doi:10.1109/TIT.2013.2276628)
17. Tomamichel M, Tan VYF. 2015 Second-order asymptotics for the classical capacity of image-additive quantum channels. *Commun. Math. Phys.* **338**, 103–137. (doi:10.1007/s00220-015-2382-0)
18. Datta N, Leditzky F. 2015 Second-order asymptotics for source coding, dense coding, and pure-state entanglement conversions. *IEEE Trans. Inform. Theory* **61**, 582–608. (doi:10.1109/TIT.2014.2366994)
19. Datta N, Tomamichel M, Wilde MM. 2016 On the second-order asymptotics for entanglement-assisted communication. *Quant. Inform. Process.* **15**, 2569–2591. (doi:10.1007/s11128-016-1272-5)
20. Wilde MM, Renes JM, Guha S. 2016 Second-order coding rates for pure-loss bosonic channels. *Quant. Inform. Process.* **15**, 1289–1308. (doi:10.1007/s11128-015-0997-x)

21. Tomamichel M, Berta M, Renes JM. 2016 Quantum coding with finite resources. *Nat. Commun.* **7**, 11419. (doi:10.1038/ncomms11419)
22. Leditzky F. 2016 Relative entropies and their use in quantum information theory. PhD thesis, University of Cambridge, Girton College. (<http://arxiv.org/abs/1611.08802>)
23. Wilde MM, Tomamichel M, Berta M. 2017 Converse bounds for private communication over quantum channels. *IEEE Trans. Inform. Theory* **63**, 1792–1817. (doi:10.1109/TIT.2017.2648825)
24. Wilde MM. 2017 Position-based coding and convex splitting for private communication over quantum channels. *Quant. Inform. Process.* **16**, 264. (doi:10.1007/s11128-017-1718-4)
25. Aaronson S. 2006 QMA/qpoly \subseteq PSPACE/poly: de-Merlinizing quantum protocols. In *Twenty-First Annual IEEE Conf. on Computational Complexity*, pp. 261–273, Prague, Czech Republic, July 2006. IEEE. (<http://arxiv.org/abs/quant-ph/0510230>)
26. Giovannetti V, Lloyd S, Maccone L. 2012 Achieving the Holevo bound via sequential measurements. *Phys. Rev. A* **85**, 012302. (doi:10.1103/PhysRevA.85.012302)
27. Sen P. 2011 Achieving the Han-Kobayashi inner bound for the quantum interference channel by sequential decoding. (<http://arxiv.org/abs/1109.0802>)
28. Wilde MM. 2013 Sequential decoding of a general classical-quantum channel. *Proc. R. Soc. A* **469**, 20130259. (doi:10.1098/rspa.2013.0259)
29. Anshu A, Jain R, Warsi NA. 2017 One shot entanglement assisted classical and quantum communication over noisy quantum channels: A hypothesis testing and convex split approach. (<http://arxiv.org/abs/1702.01940>)
30. Holevo AS. 2003 Entanglement-assisted capacity of constrained channels. In *Proc. of SPIE, First Int. Symp. on Quantum Informatics*, vol. 5128, pp. 62–69. (<http://arxiv.org/abs/quant-ph/0211170>)
31. Holevo AS. 2004 Entanglement-assisted capacities of constrained quantum channels. *Theory Probab. Appl.* **48**, 243–255. (doi:10.1137/S0040585X97980415)
32. Holevo AS. 2011 *Probabilistic and statistical aspects of quantum theory*. Pisa, Italy: Scuola Normale Superiore.
33. Buscemi F, Datta N. 2010 The quantum capacity of channels with arbitrarily correlated noise. *IEEE Trans. Inform. Theory* **56**, 1447–1460. (doi:10.1109/TIT.2009.2039166)
34. Wang L, Renner R. 2012 One-shot classical-quantum capacity and hypothesis testing. *Phys. Rev. Lett.* **108**, 200501. (doi:10.1103/PhysRevLett.108.200501)
35. Lindblad G. 1973 Entropy, information and quantum measurements. *Commun. Math. Phys.* **33**, 305–322. (doi:10.1007/BF01646743)
36. Li K. 2014 Second order asymptotics for quantum hypothesis testing. *Ann. Stat.* **42**, 171–189. (doi:10.1214/13-AOS1185)
37. Kaur E, Wilde MM. 2017 Upper bounds on secret key agreement over lossy thermal bosonic channels. *Phys. Rev. A* **96**, 062318. (doi:10.1103/PhysRevA.96.062318)
38. Datta N, Pautrat Y, Rouzé C. 2016 Second-order asymptotics for quantum hypothesis testing in settings beyond i.i.d.–quantum lattice systems and more. *J. Math. Phys.* **57**, 062207. (doi:10.1063/1.4953582)
39. Holevo AS. 1998 The capacity of the quantum channel with general signal states. *IEEE Trans. Inform. Theory* **44**, 269–273. (doi:10.1109/18.651037)
40. Schumacher B, Westmoreland MD. 1997 Sending classical information via noisy quantum channels. *Phys. Rev. A* **56**, 131–138. (doi:10.1103/PhysRevA.56.131)
41. Bennett CH, Shor PW, Smolin JA, Thapliyal AV. 1999 Entanglement-assisted classical capacity of noisy quantum channels. *Phys. Rev. Lett.* **83**, 3081–3084. (doi:10.1103/PhysRevLett.83.3081)
42. Bennett CH, Shor PW, Smolin JA, Thapliyal AV. 2002 Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem. *IEEE Trans. Inform. Theory* **48**, 2637–2655. (doi:10.1109/TIT.2002.802612)
43. Holevo AS. 2002 On entanglement assisted classical capacity. *J. Math. Phys.* **43**, 4326–4333. (doi:10.1063/1.1495877)
44. Qi H, Wang Q, Wilde MM. 2018 Applications of position-based coding to classical communication over quantum channels. *J. Phys. A* **51**, 444002. (doi:10.1088/1751-8121/aae290)
45. Verdu S. 1990 On channel capacity per unit cost. *IEEE Trans. Inform. Theory* **36**, 1019–1030. (doi:10.1109/18.57201)
46. Cariolaro G, Erseghe T. 2003 *Pulse position modulation*. New York, NY: John Wiley & Sons Inc.
47. Holevo AS. 2012 *Quantum systems, channels, information: a mathematical introduction*, vol. 16. Berlin, Germany: Walter de Gruyter.

48. Holevo AS, Shirokov ME. 2013 On the entanglement-assisted classical capacity of infinite-dimensional quantum channels. *Probl. Inform. Transm.* **49**, 15–31. (doi:10.1134/S003294601301002X)
49. Holevo AS, Werner RF. 2001 Evaluating capacities of bosonic Gaussian channels. *Phys. Rev. A* **63**, 032312. (doi:10.1103/PhysRevA.63.032312)
50. Giovannetti V, Lloyd S, Maccone L, Shor PW. 2003 Entanglement assisted capacity of the broadband lossy channel. *Phys. Rev. Lett.* **91**, 047901. (doi:10.1103/PhysRevLett.91.047901)
51. Holevo AS, Giovannetti V. 2012 Quantum channels and their entropic characteristics. *Rep. Prog. Phys.* **75**, 046001. (doi:10.1088/0034-4885/75/4/046001)
52. Wilde MM, Tomamichel M, Lloyd S, Berta M. 2017 Gaussian hypothesis testing and quantum illumination. *Phys. Rev. Lett.* **119**, 120501. (doi:10.1103/PhysRevLett.119.120501)
53. Korolev VYu, Shevtsova IG. 2010 On the upper bound for the absolute constant in the Berry-Esseen inequality. *Theory Probab. Appl.* **54**, 638–658. (doi:10.1137/S0040585X97984449)
54. Shevtsova I. 2011 On the absolute constants in the Berry-Esseen type inequalities for identically distributed summands. (<http://arxiv.org/abs/1111.6554>)
55. Bauschke HH. 2001 Projection algorithms: results and open problems. *Stud. Comput. Math.* **8**, 11–22. (doi:10.1016/S1570-579X(01)80003-3)