6-1-2016

On the second-order asymptotics for entanglement-assisted communication

Nilanjana Datta
University of Cambridge

Marco Tomamichel
Centre for Quantum Technologies

Mark M. Wilde
Louisiana State University

Recommended Citation
Datta, N., Tomamichel, M., & Wilde, M. (2016). On the second-order asymptotics for entanglement-assisted communication. *Quantum Information Processing, 15* (6), 2569-2591. https://doi.org/10.1007/s11128-016-1272-5

On the Second-Order Asymptotics for Entanglement-Assisted Communication

Nilanjana Datta* Marco Tomamichel† Mark M. Wilde‡

August 8, 2018

Abstract

The entanglement-assisted classical capacity of a quantum channel is known to provide the formal quantum generalization of Shannon's classical channel capacity theorem, in the sense that it admits a single-letter characterization in terms of the quantum mutual information and does not increase in the presence of a noiseless quantum feedback channel from receiver to sender. In this work, we investigate second-order asymptotics of the entanglement-assisted classical communication task. That is, we consider how quickly the rates of entanglement-assisted codes converge to the entanglement-assisted classical capacity of a channel as a function of the number of channel uses and the error tolerance. We define a quantum generalization of the mutual information variance of a channel in the entanglement-assisted setting. For covariant channels, we show that this quantity is equal to the channel dispersion, and thus completely characterize the convergence towards the entanglement-assisted classical capacity when the number of channel uses increases. Our results also apply to entanglement-assisted quantum communication, due to the equivalence between entanglement-assisted classical and quantum communication established by the teleportation and super-dense coding protocols.

1 Introduction

Let us consider the transmission of classical information through a memoryless quantum channel. If the sender and receiver initially share entangled states which they may use in their communication protocol, then the information transmission is said to be *entanglement-assisted*. The entanglement-assisted classical capacity $C_{\mathrm{ea}}(\mathcal{N})$ of a quantum channel $\mathcal{N}$ is defined to be the maximum rate at which a sender and receiver can communicate classical information with vanishing error probability by using the channel $\mathcal{N}$ as many times as they wish and by using an arbitrary amount of shared entanglement of an arbitrary form. For a noiseless quantum channel, the entanglement-assisted classical capacity is twice as large as its unassisted one, an enhancement realized by the super-dense coding protocol [BW92]. This is in stark contrast to the setting of classical channels where additional shared randomness or entanglement does not increase the capacity.

*Statistical Laboratory, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0WB, UK

†Centre for Quantum Technologies, National University of Singapore, Singapore 117543, Singapore, and School of Physics, The University of Sydney, Sydney, Australia

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

Similarly, for a noisy quantum channel, the presence of entanglement as an auxiliary resource can also lead to an enhancement of its classical capacity [BSST99, BSST02]. Somewhat surprisingly, entanglement assistance is advantageous even for some entanglement-breaking channels [HSR03], such as depolarizing channels with sufficiently high error probability. Bennett, Shor, Smolin and Thapliyal [BSST02] proved that the entanglement-assisted classical capacity $C_{\rm ea}(\mathcal{N})$ of a quantum channel $\mathcal{N}$ is given by a remarkably simple, single-letter formula in terms of the quantum mutual information (defined in the following section). This is in contrast to the unassisted classical capacity of a quantum channel [Hol02b, SW97], for which the best known general expression involves a regularization of the Holevo formula over infinitely many instances of the channel [Has09]. The regularization renders the explicit evaluation of the capacity for a general quantum channel intractable. The formula for the entanglement-assisted capacity is formally analogous to Shannon's well-known formula [Sha48] for the capacity of a discrete memoryless classical channel, which is given in terms of the mutual information between the channel's input and output. The entanglement-assisted capacity does not increase under the presence of a noiseless quantum feedback channel from receiver to sender [Bow04], much like the capacity of a classical channel does not increase in the presence of a noiseless classical feedback link [Sha56].

The formula for $C_{\rm ea}(\mathcal{N})$ derived in [BSST99], however, is only relevant if the channel is available for as many uses as the sender and receiver wish, with there being no correlations in the noise acting on its successive inputs.[1] To see this, let us consider the practical scenario in which a memoryless channel is used a finite number n times, and let $\mathcal{N}^n \equiv \mathcal{N}^{\otimes n}$. Let $\log M^*_{\rm ea}(\mathcal{N}^n, \varepsilon)$ denote the maximum number of bits of information that can be transmitted through n uses of the channel via an entanglement-assisted communication protocol, such that the average probability of failure is no larger than $\varepsilon \in (0,1)$. Then [BSST99] and the strong converse [BDH+14, BCR11] imply that

$$\lim_{n\to\infty} \frac{1}{n} \log M^*_{\rm ea}(\mathcal{N}^n, \varepsilon) = C_{\rm ea}(\mathcal{N}). \tag{1.1}$$

The strong converse from [GW15] implies that

$$\log M^*_{\rm ea}(\mathcal{N}^n, \varepsilon) = nC_{\rm ea}(\mathcal{N}) + O(\sqrt{n}), \tag{1.2}$$

for all $\varepsilon \in (0,1)$. The results of [CMW14] imply that this same expansion holds even when noiseless quantum feedback communication is allowed from receiver to sender.

We are interested in investigating the behavior of $M^*_{\rm ea}(\mathcal{N}^n, \varepsilon)$ for large but finite n as a function of ε. In this paper, we derive a lower bound on $\log M^*_{\rm ea}(\mathcal{N}^n, \varepsilon)$, for any fixed value of $\varepsilon \in (0,1)$ and n large enough, of the following form:

$$\log M^*_{\rm ea}(\mathcal{N}^n, \varepsilon) \geq nC_{\rm ea}(\mathcal{N}) + \sqrt{n}\, b + O(\log n). \tag{1.3}$$

The coefficient b that we identify in this paper constitutes a second-order coding rate. The second-order coding rate obtained here depends on the channel as well as on the allowed error threshold ε, and we obtain an explicit expression for it in Theorem 3. In addition, we conjecture that in fact $\log M^*_{\rm ea}(\mathcal{N}^n, \varepsilon) = nC_{\rm ea}(\mathcal{N}) + \sqrt{n}\, b + o(\sqrt{n})$ for all quantum channels. We show that this conjecture is true for the class of *covariant channels* [Hol02b].

Our lower bound on $\log M^*_{\rm ea}(\mathcal{N}^n, \varepsilon)$ resembles the asymptotic expansion for the maximum number of bits of information which can be transmitted through n uses of a generic discrete, memoryless

[1] In other words, the channel is assumed to be *memoryless.*

classical channel $\mathcal{W}$, with an average probability of error no larger than ε, denoted $\log M^*(\mathcal{W}^n, \varepsilon)$. The latter was first derived by Strassen in 1962 [Str62] and refined by Hayashi [Hay09] as well as Polyanskiy, Poor and Verdú [PPV10]. It is given by

$$\log M^*(\mathcal{W}^n, \varepsilon) = nC(\mathcal{W}) + \sqrt{nV_\varepsilon(\mathcal{W})}\Phi^{-1}(\varepsilon) + o(\sqrt{n}), \tag{1.4}$$

where $\mathcal{W}^n$ denotes n uses of the channel, $C(\mathcal{W})$ is its capacity given by Shannon's formula [Sha48], Φ^{-1} is the inverse of the cumulative distribution function of a standard normal random variable, and $V_\varepsilon(\mathcal{W})$ is a property of the channel (which depends on ε) called its ε*-dispersion* [PPV10]. The right hand side of (1.4) is called the *Gaussian approximation* of $\log M^*(\mathcal{W}^n, \varepsilon)$. This result has recently been generalized to classical coding for quantum channels [TT15] and it was shown that a formula reminiscent of (1.4) also holds for the classical capacity of quantum channels with product inputs. In fact, the Gaussian approximation is a common feature of the second-order asymptotics for optimal rates of many other quantum information processing tasks such as data compression, communication, entanglement manipulation and randomness extraction (see, e.g., [TH13, KH13, TT15, DL15] and references therein).

Even though we focus our presentation throughout on entanglement-assisted classical communication, we would like to point out that all of the results established in this paper apply to entanglement-assisted quantum communication as well. This is because the protocols of teleportation [BBC+93] and super-dense coding [BW92] establish an equivalence between entanglement-assisted classical and quantum communication. This equivalence was noted in early work on entanglement-assisted communication [BSST99]. That this equivalence applies at the level of individual codes is a consequence of the development, e.g., in Appendix B of [LM15], and as a result, the equivalence applies to second-order asymptotics as well. This point has also been noted in [TBR15].

Finally, we note that a one-shot lower bound on $M^*_{\text{ea}}(\mathcal{N}, \varepsilon)$ has already been derived in [DH13]. Moreover, in [MW14] a one-shot upper bound was obtained. Even though these bounds converge in first order to the formula for the capacity obtained by Bennett *et al.* [BSST02], neither of these works deals with characterizing second-order asymptotics.

This paper is organized as follows. Section 2 introduces the necessary notation and definitions. Section 3 presents our main theorem and our conjecture. The proof of the theorem is given in Section 4. In Section 4, we also provide a proof of our conjecture for the case of covariant channels. We end with a discussion of open questions in Section 5, summarizing the problems encountered when trying to prove the converse for general channels.

2 Notations and Definitions

Let $\mathcal{B}(\mathcal{H})$ denote the algebra of linear operators acting on a finite-dimensional Hilbert space $\mathcal{H}$. Let $\mathcal{P}(\mathcal{H}) \subset \mathcal{B}(\mathcal{H})$ be the set of positive semi-definite operators, and let $\mathcal{D}(\mathcal{H}) \subset \mathcal{P}(\mathcal{H})$ denote the set of *quantum states* (density matrices), $\mathcal{D}(\mathcal{H}) := \{\rho \in \mathcal{P}(\mathcal{H}) : \operatorname{Tr}\rho = 1\}$. We denote the dimension of a Hilbert space $\mathcal{H}_A$ by $|A|$ and write $\mathcal{H}_A \simeq \mathcal{H}_{A'}$ when $\mathcal{H}_A$ and $\mathcal{H}_{A'}$ are isomorphic, i.e., if $|A| = |A'|$. A quantum state ψ is called pure if it is rank one; in this case, we associate with it an element $|\psi\rangle \in \mathcal{H}$ such that $\psi = |\psi\rangle\langle\psi|$. The set of *pure quantum states* is denoted $\mathcal{D}_*(\mathcal{H})$.

For a bipartite operator $\omega_{AB} \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$, let $\omega_A := \operatorname{Tr}_B(\omega_{AB})$ denote its restriction to the subsystem A, where Tr_B denotes the partial trace over B. Let I_A denote the identity operator on $\mathcal{H}_A$, and let $\pi_A := I_A/|A|$ be the completely mixed state in $\mathcal{D}(\mathcal{H}_A)$.

A *positive operator-valued measure* (POVM) is a set $\{\Lambda_A^x\}_{x\in\mathcal{X}} \subset \mathcal{P}(\mathcal{H}_A)$ such that $\sum_{x\in\mathcal{X}} \Lambda_A^x = I_A$, where $\mathcal{X}$ denotes any index set. We use the convention that $\mathcal{E}_{A\to B}$ refers to a *completely positive trace-preserving* (CPTP) map $\mathcal{E}_{A\to B} : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$. We call such maps *quantum channels* in the following. The identity map on $\mathcal{B}(\mathcal{H}_A)$ is denoted id_A.

We employ the cumulative distribution function for a standard normal random variable:

$$\Phi(a) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} \mathrm{dx}\, \exp\left(-\frac{\mathrm{x}^2}{2}\right). \tag{2.1}$$

and its inverse $\Phi^{-1}(\varepsilon) := \sup\{a \in \mathbb{R} \,|\, \Phi(a) \leq \varepsilon\}$.

2.1 Entanglement-Assisted Codes

We consider entanglement-assisted classical (EAC) communication through a noisy quantum channel, given by a CPTP map $\mathcal{N}_{A\to B}$. The sender (Alice) and the receiver (Bob) initially share an arbitrary pure state $|\varphi_{A'B'}\rangle$, where without loss of generality we assume that $\mathcal{H}_{A'} \simeq \mathcal{H}_{B'}$, the system A' being with Alice and the system B' with Bob. The goal is to transmit classical messages from Alice to Bob, labelled by the elements of an index set $\mathcal{M}$, through $\mathcal{N}_{A\to B}$.

Without loss of generality, any EAC communication protocol can be assumed to have the following form: Alice encodes her classical messages into states of the system A' in her possession. Let the encoding CPTP map corresponding to message $m \in \mathcal{M}$ be denoted by $\mathcal{E}^m_{A'\to A}$. Alice then sends the system A through $\mathcal{N}_{A\to B}$ to Bob. Subsequently, Bob performs a POVM $\{\Lambda^m_{BB'}\}_{m\in\mathcal{M}}$ on the system BB' in his possession. This yields a classical register $\widehat{M}$ which contains his inference $\hat{m} \in \mathcal{M}$ of the message sent by Alice.

The above defines an *EAC code* for the quantum channel $\mathcal{N}_{A\to B}$, which consists of a quadruple

$$\mathcal{C} = \Big\{\mathcal{M},\, |\varphi_{A'B'}\rangle,\, \{\mathcal{E}^m_{A'\to A}\}_{m\in\mathcal{M}},\, \{\Lambda^m_{BB'}\}_{m\in\mathcal{M}}\Big\}. \tag{2.2}$$

The size of a code is denoted as $|\mathcal{C}| = |\mathcal{M}|$. The average probability of error for $\mathcal{C}$ on $\mathcal{N}_{A\to B}$ is

$$p_{\mathrm{err}}(\mathcal{N}_{A\to B}, \mathcal{C}) := \Pr[M \neq \widehat{M}] = 1 - \frac{1}{|\mathcal{M}|} \sum_m \mathrm{Tr}\left(\Lambda^m_{BB'}\, \mathcal{N}_{A\to B} \otimes \mathrm{id}_{B'}\big(\mathcal{E}^m_{A'\to A} \otimes \mathrm{id}_{B'}(\varphi_{A'B'})\big)\right). \tag{2.3}$$

The following quantity describes the maximum size of an EAC code for transmitting classical information through a single use of $\mathcal{N}_{A\to B}$ with average probability of error at most ε.

Definition 1. Let $\varepsilon \in (0,1)$ and $\mathcal{N} \equiv \mathcal{N}_{A\to B}$ be a quantum channel. We define

$$M^*_{\mathrm{ea}}(\mathcal{N}, \varepsilon) := \max\big\{m \in \mathbb{N} \,\big|\, \exists\, \mathcal{C} : |\mathcal{C}| = m \wedge p_{\mathrm{err}}(\mathcal{N}, \mathcal{C}) \leq \varepsilon\big\}, \tag{2.4}$$

where $\mathcal{C}$ is a code as prescribed in (2.2).

We are particularly interested in the quantity $M^*_{\mathrm{ea}}(\mathcal{N}^n, \varepsilon)$, where $n \in \mathbb{N}$ and $\mathcal{N}^n \equiv \mathcal{N}^n_{A^n\to B^n} = \mathcal{N}_{A_1\to B_1} \otimes \ldots \otimes \mathcal{N}_{A_n\to B_n}$ is the n-fold memoryless repetition of $\mathcal{N}$.

2.2 Information Quantities

For a pair of positive semi-definite operators ρ and σ with $\operatorname{supp}\rho \subseteq \operatorname{supp}\sigma$, the *quantum relative entropy* and the *relative entropy variance* [Li14, TH13] are respectively defined as follows:[2]

$$D(\rho\|\sigma) := \operatorname{Tr}\left[\rho\left(\log\rho - \log\sigma\right)\right], \qquad \text{and} \tag{2.5}$$

$$V(\rho\|\sigma) := \operatorname{Tr}\left[\rho\left(\log\rho - \log\sigma - D(\rho\|\sigma)\right)^2\right]. \tag{2.6}$$

For a bipartite state ρ_{AB}, let us define the *mutual information* $I(A:B)_\rho := D(\rho_{AB}\|\rho_A \otimes \rho_B)$. Similarly, we define the *mutual information variance* $V(A:B)_\rho := V(\rho_{AB}\|\rho_A \otimes \rho_B)$.

The EAC capacity of a quantum channel $\mathcal{N}$ is defined as

$$C_{\text{ea}}(\mathcal{N}) := \lim_{\varepsilon\to 0}\limsup_{n\to\infty}\frac{1}{n}\log M^*_{\text{ea}}(\mathcal{N}^n,\varepsilon). \tag{2.7}$$

Bennett, Shor, Smolin and Thapliyal [BSST02] established that the EAC capacity for a quantum channel $\mathcal{N} \equiv \mathcal{N}_{A\to B}$ satisfies

$$C_{\text{ea}}(\mathcal{N}) = \max_{\psi_{AA'}} I(A':B)_\omega, \qquad \text{where} \quad \omega_{A'B} = \mathcal{N}_{A\to B}\otimes \operatorname{id}_{A'}(\psi_{AA'}), \tag{2.8}$$

and the maximum is taken over all $\psi_{AA'} \in \mathcal{D}_*(\mathcal{H}_A \otimes \mathcal{H}_{A'})$ with $\mathcal{H}_{A'} \simeq \mathcal{H}_A$. Its proof was later simplified by Holevo [Hol02a], and an alternative proof was given in [HDW08].

In analogy with [TT15], the following definitions will be used to characterize our lower bounds on the second-order asymptotic behavior of $M^*_{\text{ea}}(\mathcal{N}^n,\varepsilon)$.

Definition 2. Let $\mathcal{N} \equiv \mathcal{N}_{A\to B}$ be a quantum channel. The set of *capacity achieving resource states* on $\mathcal{N}$ is defined as

$$\Pi(\mathcal{N}) := \arg\max_{\psi_{AA'}} I(A':B)_\omega \subseteq \mathcal{D}_*(\mathcal{H}_A \otimes \mathcal{H}_{A'}), \tag{2.9}$$

where $\omega_{A'B}$ is given in (2.8). The *minimal mutual information variance* and the *maximal mutual information variance* of $\mathcal{N}$ are respectively defined as

$$V_{\text{ea,min}}(\mathcal{N}) := \min_{\psi_{AA'}} V(A':B)_\omega \qquad \text{and} \qquad V_{\text{ea,max}}(\mathcal{N}) := \max_{\psi_{AA'}} V(A':B)_\omega, \tag{2.10}$$

where the optimizations are over $\psi_{AA'} \in \Pi(\mathcal{N})$ and $\omega_{A'B}$ is given in (2.8).

3 Results

Our main result is stated in the following theorem, which provides a second-order lower bound on the maximum number of bits of classical message which can be transmitted through n independent uses of a noisy channel via an entanglement-assisted protocol, for any given allowed error threshold.

Theorem 3. *Let $\varepsilon \in (0,1)$ and let $\mathcal{N} \equiv \mathcal{N}_{A\to B}$ be a quantum channel. Then,*

$$\log M^*_{\text{ea}}(\mathcal{N}^n,\varepsilon) \geq \begin{cases} nC_{\text{ea}}(\mathcal{N}) + \sqrt{nV_{\text{ea,min}}(\mathcal{N})}\ \Phi^{-1}(\varepsilon) + K(n;\mathcal{N},\varepsilon) & \text{if } \varepsilon < \frac{1}{2} \\ nC_{\text{ea}}(\mathcal{N}) + \sqrt{nV_{\text{ea,max}}(\mathcal{N})}\ \Phi^{-1}(\varepsilon) + K(n;\mathcal{N},\varepsilon) & \text{else} \end{cases} \tag{3.1}$$

where $K(n;\mathcal{N},\varepsilon) = O(\log n)$.

[2] All logarithms in this paper are taken to base two.

The proof of Theorem 3 is split into two parts, Proposition 11 in Section 4.2 and Proposition 14 in Section 4.3. We first derive a one-shot lower bound on $\log M^*_{\mathrm{ea}}(\mathcal{N},\varepsilon)$ using a coding scheme that is a one-shot version of the coding scheme given in [HDW08] and reviewed in [Wil13, Sec. 20.4]. Our one-shot bound is expressed in terms of an entropic quantity called the hypothesis testing relative entropy [WR12], which has its roots in early work on the quantum Stein's lemma [HP91] (see Section 4.1 for a definition). The relation between classical coding over a quantum channels and binary quantum hypothesis testing was first pointed out by Hayashi and Nagaoka [HN03].

An asymptotic expansion for this quantity for product states was derived independently by Tomamichel and Hayashi [TH13] and Li [Li14], and we make use of this to obtain our lower bound on $\log M^*_{\mathrm{ea}}(\mathcal{N}^n,\varepsilon)$ in the second step.

Remark 4. In particular, Theorem 3 establishes that

$$\liminf_{n\to\infty} \frac{1}{\sqrt{n}}\Big(\log M^*_{\mathrm{ea}}(\mathcal{N}^n,\varepsilon) - nC_{\mathrm{ea}}(\mathcal{N})\Big) \geq \begin{cases} \sqrt{V_{\mathrm{ea,min}}(\mathcal{N})}\ \Phi^{-1}(\varepsilon) & \text{if } \varepsilon < \frac{1}{2} \\ \sqrt{V_{\mathrm{ea,max}}(\mathcal{N})}\ \Phi^{-1}(\varepsilon) & \text{else} \end{cases}. \tag{3.2}$$

In analogy with [PPV10, Eq. (221)] and [TT15, Rm. 4], we define the EAC ε-channel dispersion, for $\varepsilon \in (0,1)\setminus\{\frac{1}{2}\}$ as

$$V_{\mathrm{ea},\varepsilon}(\mathcal{N}) := \limsup_{n\to\infty} \frac{1}{n}\left(\frac{\log M^*_{\mathrm{ea}}(\mathcal{N}^n,\varepsilon) - nC_{\mathrm{ea}}(\mathcal{N})}{\Phi^{-1}(\varepsilon)}\right)^2. \tag{3.3}$$

Theorem 3 shows that $V_{\mathrm{ea},\varepsilon}(\mathcal{N}) \leq V_{\mathrm{ea,min}}(\mathcal{N})$ if $\varepsilon < \frac{1}{2}$ and $V_{\mathrm{ea},\varepsilon}(\mathcal{N}) \geq V_{\mathrm{ea,max}}(\mathcal{N})$ if $\varepsilon > \frac{1}{2}$.

This leads us to the following conjecture:

Conjecture 5. *We conjecture that* (3.1) *is an equality with* $K(n;\mathcal{N},\varepsilon) = o(\sqrt{n})$*. In particular, we conjecture that the EAC* ε*-channel dispersion satisfies*

$$V_{\mathrm{ea},\varepsilon}(\mathcal{N}) = \begin{cases} V_{\mathrm{ea,min}}(\mathcal{N}) & \textit{if } \varepsilon < \frac{1}{2} \\ V_{\mathrm{ea,max}}(\mathcal{N}) & \textit{else} \end{cases} \tag{3.4}$$

and, thus, $\log M^*_{\mathrm{ea}}(\mathcal{N}^n,\varepsilon) = nC_{\mathrm{ea}}(\mathcal{N}) + \sqrt{nV_{\mathrm{ea},\varepsilon}(\mathcal{N})}\ \Phi^{-1}(\varepsilon) + o(\sqrt{n})$.

We show that Conjecture 5 is true for the class of covariant quantum channels. This follows essentially from an analysis by Matthews and Wehner [MW14] which we recapitulate in Section 4.4 and the asymptotic expansion of the hypothesis testing relative entropy.

4 Proofs

4.1 Technical Preliminaries

For given orthonormal bases $\{|i_A\rangle\}_{i=1}^d$ and $\{|i_B\rangle\}_{i=1}^d$ in isomorphic Hilbert spaces $\mathcal{H}_A \simeq \mathcal{H}_B \simeq \mathcal{H}$ of dimension d, we define a maximally entangled state of Schmidt rank d to be

$$|\Phi_{AB}\rangle := \frac{1}{\sqrt{d}}\sum_{i=1}^{d} |i_A\rangle \otimes |i_B\rangle. \tag{4.1}$$

Note that if $d = 1$ then $|\Phi_{AB}\rangle$ is a product state. We often make use of the following identity ("transpose trick"): for any operator M,

$$(M_A \otimes I_B)|\Phi_{AB}\rangle = (I_A \otimes M_B^T)|\Phi_{AB}\rangle, \tag{4.2}$$

where $M_B^T := \sum_{i,j=1}^d |i\rangle_B\langle j|_A M_A |i\rangle_A\langle j|_B$ denotes the transpose.

4.1.1 Distance Measures

The *trace distance* between two states ρ and σ is given by

$$\tfrac{1}{2}\|\rho - \sigma\|_1 = \max_{0\leq Q\leq I} \operatorname{Tr}\big(Q(\rho-\sigma)\big) = \operatorname{Tr}\big[\{\rho \geq \sigma\}(\rho - \sigma)\big] \tag{4.3}$$

where $\{\rho \geq \sigma\}$ denotes the projector onto the subspace where the operator $\rho - \sigma$ is positive semi-definite. The fidelity of ρ and σ is defined as

$$F(\rho, \sigma) := \left\|\sqrt{\rho}\sqrt{\sigma}\right\|_1 . \tag{4.4}$$

For a pair of pure states ϕ and ψ, the trace distance and fidelity satisfy the following relation:

$$\tfrac{1}{2}\|\phi - \psi\|_1 = \sqrt{1 - F^2(\phi, \psi)}. \tag{4.5}$$

4.1.2 Relative Entropies for One-Shot Analysis

We will phrase our one-shot bounds in terms of the following relative entropy.

Definition 6. Let $\varepsilon \in (0, 1)$, $\rho \in \mathcal{D}(\mathcal{H})$ and $\sigma \in \mathcal{P}(\mathcal{H})$. Then, the *hypothesis testing relative entropy* [WR12] is defined as

$$D_H^\varepsilon(\rho\|\sigma) := -\log \beta_\varepsilon(\rho\|\sigma), \tag{4.6}$$

where

$$\beta_\varepsilon(\rho\|\sigma) := \min\big\{ \operatorname{Tr}(Q\sigma) : 0 \leq Q \leq I \wedge \operatorname{Tr}(Q\rho) \geq 1 - \varepsilon\big\}. \tag{4.7}$$

Note that when $\sigma \in \mathcal{D}(\mathcal{H})$, $\beta_\varepsilon(\rho\|\sigma)$ has an interpretation as the smallest type-II error of a hypothesis test between ρ and σ, when the type-I error is at most ε. The following lemma lists some properties of $D_H^\varepsilon(\rho\|\sigma)$.

Lemma 7. *Let $\varepsilon \in (0, 1)$. The hypothesis testing relative entropy has the following properties:*

1. *For any $\rho \in \mathcal{D}(\mathcal{H})$, $\sigma' \geq \sigma \geq 0$ we have $D_H^\varepsilon(\rho\|\sigma) \geq D_H^\varepsilon(\rho\|\sigma')$.*

2. *For any $\rho \in \mathcal{D}(\mathcal{H})$, $\sigma \geq 0$, $\alpha > 0$, we have $D_H^\varepsilon(\rho\|\alpha\sigma) = D_H^\varepsilon(\rho\|\sigma) - \log \alpha$.*

3. *For any classical-quantum state*

$$\rho_{XB} = \sum_{x\in\mathcal{X}} p(x)|x\rangle\langle x| \otimes \rho_B^x \in \mathcal{D}(\mathcal{H}_X \otimes \mathcal{H}_B), \tag{4.8}$$

for any $\sigma_X = \sum_{x\in\mathcal{X}} q(x)|x\rangle\langle x|$ (with $\{p(x)\}_{x\in\mathcal{X}}$ and $\{q(x)\}_{x\in\mathcal{X}}$ probability distributions on $\mathcal{X}$), and for any $\sigma_B \in \mathcal{D}(\mathcal{H}_B)$, we have

$$D_H^\varepsilon(\rho_{XB}\|\sigma_X \otimes \sigma_B) \geq \min_{x\in\mathcal{X}} D_H^\varepsilon(\rho_B^x\|\sigma_B), \tag{4.9}$$

4. *For any $\delta \in (0, 1-\varepsilon)$, $\rho, \rho' \in \mathcal{D}(\mathcal{H})$ with $\frac{1}{2}\|\rho - \rho'\|_1 \leq \delta$, and $\sigma \in \mathcal{P}(\mathcal{H})$, we have $D_H^{\varepsilon}(\rho'\|\sigma) \leq D_H^{\varepsilon+\delta}(\rho\|\sigma)$.*

Properties 1–3 can be verified by close inspection and we omit their proofs.

Proof of Property 4. Consider Q to be the operator achieving the minimum in the definition of $\beta_\varepsilon(\rho'\|\sigma)$, i.e.

$$D_H^{\varepsilon}(\rho'\|\sigma) = -\log \operatorname{Tr}(Q\sigma) \qquad \text{and} \qquad \operatorname{Tr}(Q\rho') \geq 1 - \varepsilon. \tag{4.10}$$

From (4.3), we have $\operatorname{Tr}[Q(\rho' - \rho)] \leq \frac{1}{2}\|\rho - \rho'\|_1 \leq \delta$. Hence, $\operatorname{Tr}(Q\rho) \geq \operatorname{Tr}(Q\rho') - \frac{1}{2}\|\rho - \rho'\|_1 \geq 1 - \varepsilon - \delta$, and

$$D_H^{\varepsilon}(\rho'\|\sigma) \leq \max_{\substack{0 \leq Q' \leq I \\ \operatorname{Tr}(Q'\rho) \geq 1-\varepsilon-\delta}} \left[-\log \operatorname{Tr}(Q'\sigma)\right] = D_H^{\varepsilon+\delta}(\rho\|\sigma), \tag{4.11}$$

which concludes the proof. □

The following result, established independently in [TH13, Eq. (34)] and [Li14], plays a central role in our analysis.

Lemma 8 ([TH13, Li14])**.** *Let $\varepsilon \in (0,1)$ and let $\rho, \sigma \in \mathcal{D}(\mathcal{H})$. Then,*

$$D_H^{\varepsilon}(\rho^{\otimes n}\|\sigma^{\otimes n}) = nD(\rho\|\sigma) + \sqrt{nV(\rho\|\sigma)}\Phi^{-1}(\varepsilon) + O(\log n). \tag{4.12}$$

Two other generalized relative entropies which are relevant for our analysis are the *collision relative entropy* and the *information-spectrum relative entropy* [TH13, Def. 8]. For any pair of positive semi-definite operators ρ and σ satisfying the condition $\operatorname{supp}\rho \subseteq \operatorname{supp}\sigma$, they are respectively defined as follows:

$$D_2(\rho\|\sigma) := \log\left(\operatorname{Tr}\left(\sigma^{-1/4}\rho\sigma^{-1/4}\right)^2\right), \tag{4.13}$$

and, for any $\varepsilon \in (0,1)$,

$$D_s^{\varepsilon}(\rho\|\sigma) := \sup\left\{R \mid \operatorname{Tr}\left(\rho\{\rho \leq 2^R\sigma\}\right) \leq \varepsilon\right\}, \tag{4.14}$$

where we write $A \geq B$ if $A - B$ is positive semidefinite. The following result, proved in [BG14, Thm. 4], relates these quantities.

Lemma 9 ([BG14])**.** *Let $\varepsilon, \lambda \in (0,1)$ and $\rho, \sigma \in \mathcal{D}(\mathcal{H})$. Then,*

$$2^{D_2(\rho\|\lambda\rho+(1-\lambda)\sigma)} \geq (1-\varepsilon)\left[\lambda + (1-\lambda)2^{-D_s^{\varepsilon}(\rho\|\sigma)}\right]^{-1}. \tag{4.15}$$

Finally, the following lemma provides a useful relation between the hypothesis testing relative entropy and the information spectrum relative entropy [TH13, Lm. 12].

Lemma 10 ([TH13])**.** *Let $\varepsilon \in (0,1)$, $\delta \in (0, 1-\varepsilon)$, $\rho \in \mathcal{D}(\mathcal{H})$, and $\sigma \in \mathcal{P}(\mathcal{H})$. Then, $D_H^{\varepsilon}(\rho\|\sigma) \geq D_s^{\varepsilon}(\rho\|\sigma) \geq D_H^{\varepsilon+\delta}(\rho\|\sigma) + \log\delta$.*

4.2 One-Shot Achievability

Our protocol is modeled after [BSST02]. Consider a quantum channel $\mathcal{N}_{A\to B}$ and introduce an auxiliary Hilbert space $\mathcal{H}_{A'} \simeq \mathcal{H}_A$. Let

$$\mathcal{H}_A \otimes \mathcal{H}_{A'} = \bigoplus_t \mathcal{H}_A^t \otimes \mathcal{H}_{A'}^t, \qquad \mathcal{H}_A^t \simeq \mathcal{H}_{A'}^t \tag{4.16}$$

be a decomposition of $\mathcal{H}_A \otimes \mathcal{H}_{A'}$, and set $d_t = |\mathcal{H}_A^t|$. We assume that $|\vartheta_{AA'}\rangle$ can be written as a superposition of maximally entangled states:

$$|\vartheta_{AA'}\rangle = \sum_t \sqrt{p(t)}\, |\Phi^t\rangle, \tag{4.17}$$

where $|\Phi^t\rangle$ denotes a maximally entangled state of Schmidt rank d_t in $\mathcal{H}_A^t \otimes \mathcal{H}_{A'}^t$ and $p(t)$ is some probability distribution so that $\sum_t p(t) = 1$.

Proposition 11. *Let $\varepsilon \in (0,1)$, $\delta \in (0, \frac{\varepsilon}{2})$ and let $\mathcal{N} \equiv \mathcal{N}_{A\to B}$ be a quantum channel. Then for any $\vartheta_{AA'}$ of the form* (4.17)*, we have*

$$\log M^*_{\mathrm{ea}}(\mathcal{N}, \varepsilon) \geq D_H^{\varepsilon-2\delta}(\mathcal{N}_{A\to B}(\vartheta_{AA'}) \,\|\, \mathcal{N}_{A\to B}(\kappa_{AA'})) - f(\varepsilon, \delta), \tag{4.18}$$

where $f(\varepsilon,\delta) := \log \frac{1-\varepsilon}{\delta^2}$, $\kappa_{AA'} := \sum_t p(t)\, \pi_A^t \otimes \pi_{A'}^t$, and π_A^t is the maximally mixed state on $\mathcal{H}_A^t$.

Remark 12. Note that the hypothesis testing relative entropy on the right hand side is not reminiscent of a mutual information type quantity since the second argument is not a product state.

Proof. Consider the set

$$\mathcal{S} := \left\{ ((x_t, z_t, b_t))_t \,\middle|\, x_t, z_t \in \{0, 1, \cdots, d_t - 1\}, b_t \in \{0, 1\} \right\}, \tag{4.19}$$

where the index t labels the Hilbert spaces of the decomposition in (4.16). For any $s \in \mathcal{S}$, consider the following unitary operator in $\mathcal{B}(\mathcal{H}_A)$:

$$U_A(s) := \bigoplus_t (-1)^{b_t} X(x_t) Z(z_t), \tag{4.20}$$

where $X(x_t)$ and $Z(z_t)$ are the Heisenberg-Weyl operators defined in Appendix A.

For any $M \in \mathbb{N}$, we now construct a random code as follows. Let $\mathcal{M} = \{1, 2, \ldots, M\}$. We set $A' \equiv A$ (i.e., we use the labels interchangeably), and $\mathcal{H}_{B'} \simeq \mathcal{H}_A$. We consider the resource state $\varphi_{AB'} = \mathrm{id}_{A'\to B'}(\vartheta_{AA'})$. For each message $m \in \mathcal{M}$, choose a *codeword*, s_m, uniformly at random from the set $\mathcal{S}$. The encoding operation, $\{\mathcal{E}_A^m\}_{m\in\mathcal{M}}$, is then given by the (random) unitary $U(s_m)$ as prescribed above. In particular,

$$\mathcal{E}_A^m \otimes \mathrm{id}_{B'}(\varphi_{AB'}) = \phi_{AB'}^{s_m}, \quad \text{where} \quad |\phi_{AB'}^{s_m}\rangle := (U_A(s_m) \otimes I_{B'})\, |\varphi_{AB'}\rangle. \tag{4.21}$$

We denote the corresponding channel output state by $\rho_{BB'}^{s_m} := \mathcal{N}_{A\to B}\left(\phi_{AB'}^{s_m}\right)$ and use "pretty good" measurements for decoding. These are given by the POVM $\{\Lambda_{BB'}^m\}_{m\in\mathcal{M}}$, where

$$\Lambda_{BB'}^m := \left(\sum_{m'\in\mathcal{M}} \rho_{BB'}^{s_{m'}} \right)^{-\frac{1}{2}} \rho_{BB'}^{s_m} \left(\sum_{m'\in\mathcal{M}} \rho_{BB'}^{s_{m'}} \right)^{-\frac{1}{2}}. \tag{4.22}$$

Let us now analyze the code $\mathcal{C} = \{\mathcal{M}, \varphi_{AB'}, \{\mathcal{E}^m_{A'}\}_{m\in\mathcal{M}}, \{\Lambda^m_{BB'}\}_{m\in\mathcal{M}}\}$ given by (4.21) and (4.22), where we recall that s_m is a random variable. For this purpose, consider the random state

$$\sigma_{MSBB'} := \frac{1}{M}\sum_{m\in\mathcal{M}} |m\rangle\langle m|_M \otimes |s_m\rangle\langle s_m|_S \otimes \rho^{s_m}_{BB'}. \tag{4.23}$$

Then, following Beigi and Gohari [BG14, Thm. 5], we find that the average probability of successfully inferring the sent message can be expressed as

$$p_{\text{succ}}(\mathcal{C},\mathcal{N}) := 1 - p_{\text{err}}(\mathcal{C},\mathcal{N}) = \frac{1}{M}\sum_{m\in\mathcal{M}} \operatorname{Tr}(\Lambda^m_{BB'}\rho^{s_m}_{BB'}) \tag{4.24}$$

$$= \frac{1}{M} 2^{D_2(\sigma_{MSBB'}\|\sigma_{MS}\otimes\sigma_{BB'})}. \tag{4.25}$$

Moreover employing both the data-processing inequality and joint convexity of the collision relative entropy as in [BG14], we establish the following lower bound on the expected value of p_{succ} with respect to the randomly chosen codewords:

$$\mathbb{E}\left(p_{\text{succ}}(\mathcal{C},\mathcal{N})\right) \geq \frac{1}{M} 2^{D_2(\mathbb{E}(\sigma_{SBB'})\|\mathbb{E}(\sigma_S\otimes\sigma_{BB}))}. \tag{4.26}$$

Note that

$$\mathbb{E}(\sigma_S\otimes\sigma_{BB'}) = \mathbb{E}\left(\frac{1}{M^2}\sum_{m\in\mathcal{M}} |s_m\rangle\langle s_m| \otimes \rho^{s_m}_{BB'}\right) + \mathbb{E}\left(\frac{1}{M^2}\sum_{\substack{m,m'\in\mathcal{M}\\ m'\neq m}} |s_m\rangle\langle s_m| \otimes \rho^{s_{m'}}_{BB'}\right) \tag{4.27}$$

$$= \frac{1}{M}\rho_{SBB'} + \left(1-\frac{1}{M}\right)\rho_S\otimes\rho_{BB'}, \tag{4.28}$$

where

$$\rho_{SBB'} := \mathbb{E}(\sigma_{SBB'}) = \mathcal{N}_{A\to B}\left(\frac{1}{|\mathcal{S}|}\sum_{s\in\mathcal{S}} |s\rangle\langle s|\otimes U_A(s)\varphi_{AB'}U^\dagger_A(s)\right) \tag{4.29}$$

$$= \frac{1}{|\mathcal{S}|}\sum_{s\in\mathcal{S}} |s\rangle\langle s|\otimes\mathcal{N}_{A\to B}\left(U_A(s)\varphi_{AB'}U^\dagger_A(s)\right), \tag{4.30}$$

and ρ_S and $\rho_{BB'}$ are the corresponding reduced states on the systems S and BB', respectively. In particular, defining $V(x_t,z_t) := X(x_t)Z(z_t)$, using the decomposition (4.17) of the state $|\varphi_{AB'}\rangle$ and the definition (4.20) of the unitary operators $U_A(s)$, we find that

$$\rho_{BB'} = \mathcal{N}_{A\to B}\left(\frac{1}{|\mathcal{S}|}\sum_{s\in\mathcal{S}} U_A(s)\left(\sum_{t,t'}\sqrt{p(t)p(t')}|\Phi^t\rangle\langle\Phi^{t'}|\right)U^\dagger_A(s)\right) \tag{4.31}$$

$$= \mathcal{N}_{A\to B}\left(\sum_t p(t)\frac{1}{d_t^2}\sum_{x_t,z_t=0}^{d_t-1} V(x_t,z_t)|\Phi^t\rangle\langle\Phi^t|V^\dagger(x_t,z_t)\right) \tag{4.32}$$

$$+\,\mathcal{N}_{A\to B}\left(\sum_{\substack{t,t'\\ t'\neq t}}\sqrt{p(t)p(t')}\,\frac{1}{4}\sum_{b_t,b_{t'}\in\{0,1\}}(-1)^{b_t+b_{t'}}\frac{1}{d_t^2 d_{t'}^2}\sum_{x_t,z_t=0}^{d_t-1}\sum_{x_{t'},z_{t'}=0}^{d_{t'}-1} V(x_t,z_t)|\Phi^t\rangle\langle\Phi^{t'}|V^\dagger(x_{t'},z_{t'})\right) \tag{4.33}$$

can be written as the sum of a diagonal ($t = t'$) and an off-diagonal ($t \neq t'$) term. It can be verified (see, e.g., [Wil13, pp. 504–505]) that the off-diagonal term vanishes and in fact

$$\rho_{BB'} = \sum_t p(t)\mathcal{N}_{A\to B}(\pi_A^t) \otimes \pi_{B'}^t, \tag{4.34}$$

where $\pi_A^t = \mathrm{Tr}_{B'}(\Phi^t)$ and $\pi_{B'}^t = \mathrm{Tr}_A(\Phi^t)$ are completely mixed states. The above identity follows from the fact that applying a Heisenberg-Weyl operator uniformly at random completely randomizes a quantum state, yielding a completely mixed state.

Hence, for any $0 < \delta < \varepsilon$, we have

$$\mathbb{E}\left(p_{\mathrm{succ}}(\mathcal{C}, \mathcal{N})\right) \geq \frac{1}{M} 2^{D_2\left(\rho_{SBB'} \| \frac{1}{M}\rho_{SBB'} + (1-\frac{1}{M})(\rho_S \otimes \rho_{BB'}\right)} \tag{4.35}$$

$$\geq \frac{1-(\varepsilon-\delta)}{1+(M-1)\, 2^{-D_s^{\varepsilon-\delta}(\rho_{SBB'} \| \rho_S \otimes \rho_{BB'})}}, \tag{4.36}$$

where the last line follows from Lemma 9. Thus, provided that

$$M \leq \frac{\delta}{1-\varepsilon} 2^{D_s^{\varepsilon-\delta}(\rho_{SBB'} \| \rho_S \otimes \rho_{BB'})} + 1 \tag{4.37}$$

the random code satisfies $\mathbb{E}\left(p_{\mathrm{succ}}(\mathcal{C}, \mathcal{N})\right) \geq 1 - \varepsilon$. In particular, there exists a (deterministic) code which satisfies $p_{\mathrm{succ}}(\mathcal{C}, \mathcal{N}) \geq 1 - \varepsilon$. Hence, we conclude that

$$\log M^*_{\mathrm{ea}}(\mathcal{N}, \varepsilon) \geq D_s^{\varepsilon-\delta}(\rho_{SBB'} \| \rho_S \otimes \rho_{BB'}) + \log \frac{\delta}{1-\varepsilon} \tag{4.38}$$

$$\geq D_H^{\varepsilon-2\delta}(\rho_{SBB'} \| \rho_S \otimes \rho_{BB'}) - f(\varepsilon, \delta), \tag{4.39}$$

where we require that $\varepsilon > 2\delta$ and use

$$f(\varepsilon, \delta) = \log \frac{1-\varepsilon}{\delta^2}. \tag{4.40}$$

The inequality in (4.39) follows from Lemma 10. Further, since $\rho_{SBB'}$ is a classical-quantum state as seen in (4.30), by item 3 of Lemma 7 we have

$$D_H^{\varepsilon-2\delta}(\rho_{SBB'} \| \rho_S \otimes \rho_{BB'}) \geq \min_{s\in\mathcal{S}} D_H^{\varepsilon-2\delta}(\rho^s_{BB'} \| \rho_{BB'}), \tag{4.41}$$

where

$$\rho^s_{BB'} = \mathcal{N}_{A\to B}\left(U_A(s)\varphi_{AB'}U_A^\dagger(s)\right). \tag{4.42}$$

Using the decomposition (4.17) of the state $|\varphi_{AB'}\rangle$ and the transpose trick (4.2) we can write

$$\rho^s_{BB'} = U^T_{B'}(s)\mathcal{N}_{A\to B}\left(\varphi_{AB'}\right)U^{T\dagger}_{B'}(s). \tag{4.43}$$

Further, from (4.34) it follows that

$$U^T_{B'}(s)\rho_{BB'}U^{T\dagger}_{B'}(s) = \rho_{BB'}. \tag{4.44}$$

Hence, (4.43), (4.44), (4.34), and the invariance of the hypothesis testing relative entropy under the same unitary on both states imply that

$$D_H^{\varepsilon-2\delta}(\rho^s_{BB'} \| \rho_{BB'}) = D_H^{\varepsilon-2\delta}\left(\mathcal{N}_{A\to B}\left(\varphi_{AB'}\right) \middle\| \sum_t p(t)\left(\mathcal{N}_{A\to B}(\pi_A^t)\right) \otimes \pi_{B'}^t\right), \tag{4.45}$$

From (4.39) and (4.45) we obtain the statement of the proposition. □

Remark 13. Alternatively, one may also employ the one-shot achievability result of Hayashi and Nagaoka [HN03] (in the form of [WR12]), which leads to the following bound on the one-shot ε-error entanglement-assisted capacity. Let $\varepsilon \in (0,1)$. Then, for any $\delta \in (0,\varepsilon)$ and for any $|\vartheta_{AA'}\rangle$ with decomposition (4.17), we have

$$\log M^*_{\rm ea}(\mathcal{N},\varepsilon) \geq D_H^{\varepsilon-\delta}\left(\mathcal{N}_{A\to B}\left(\vartheta_{AA'}\right) \,\middle\|\, \sum_t p(t)\left(\mathcal{N}_{A\to B}(\pi^t_A)\right)\otimes \pi^t_{A'}\right) - \log\frac{4\varepsilon}{\delta^2}. \tag{4.46}$$

The proof of this lower bound uses the same coding scheme as given above while employing the error analysis and decoder given in [HN03].

4.3 Second-Order Analysis for Achievability

Theorem 3 is a direct corollary of the following result, for an appropriate choice of $\psi_{AA'}$.

Proposition 14. *Let $\varepsilon \in (0,1)$, $\mathcal{N} \equiv \mathcal{N}_{A\to B}$ be a quantum channel, and $\psi_{AA'} \in \mathcal{D}_*(\mathcal{H}_A \otimes \mathcal{H}'_A)$, where $\mathcal{H}_{A'} \simeq \mathcal{H}_A$. Then, we have*

$$\log M^*_{\rm ea}(\mathcal{N}^n,\varepsilon) \geq nI(A':B)_\omega + \sqrt{nV(A':B)_\omega}\,\Phi^{-1}(\varepsilon) + K(n;\mathcal{N},\varepsilon,\psi_{AA'}), \tag{4.47}$$

where $\omega_{A'B} = \mathcal{N}_{A\to B}\otimes {\rm id}_{A'}(\psi_{AA'})$ and $K(n;\mathcal{N},\varepsilon,\psi_{AA'}) = O(\log n)$.

Proof. We intend to apply Proposition 11 to the channel $\mathcal{N}^n := \mathcal{N}^{\otimes n}$ for a fixed n. For this purpose, let us first construct an appropriate resource state $\vartheta_{A^nA'^n}$. We write

$$|\psi_{AA'}\rangle = \sum_{x\in\mathcal{X}} \sqrt{q(x)}\,|x\rangle_A \otimes |x\rangle_{A'} \tag{4.48}$$

in its Schmidt decomposition, where $\mathcal{X} = \{1,2,\cdots,d\}$ with $d = |\mathcal{H}_A| = |\mathcal{H}_{A'}|$ and define $\rho_{A'} = {\rm Tr}_A(\psi_{AA'})$. For a sequence $x^n = (x_1,x_2,\ldots,x_n) \in \mathcal{X}^n$, we write $|x^n\rangle_{A^n} = |x_1\rangle_{A_1}\otimes|x_2\rangle_{A_2}\otimes\cdots\otimes|x_n\rangle_{A_n}$. We denote type classes (for sequences of length n) by $\mathcal{T}^t$, i.e. $\mathcal{T}^t = \{x^n \in \mathcal{X}^n : P_{x^n} = t\}$ where P_{x^n} denotes the empirical distribution of the sequence $x^n \in \mathcal{X}^n$. The set of empirical distributions is denoted $\mathcal{P}_n$. (We refer to Appendix B for a short overview of the method of types and relevant results.) We consider the decomposition

$$(\mathcal{H}_A \otimes \mathcal{H}_{A'})^{\otimes n} = \bigoplus_{t\in\mathcal{P}_n} \mathcal{H}^t_{A^n}\otimes\mathcal{H}^t_{A'^n}, \tag{4.49}$$

where $\mathcal{H}^t_{A^n} = {\rm span}\big\{|x^n\rangle_{A^n}\,\big|\, x^n \in \mathcal{T}^t\big\}$ as in (4.16). Notably, since $\psi^{\otimes n}_{AA'}$ is a tensor-power state, we can write

$$|\psi_{AA'}\rangle^{\otimes n} = \sum_{t\in\mathcal{P}_n}\sqrt{p'(t)}\,|\Phi^t\rangle, \tag{4.50}$$

where $|\Phi^t\rangle \in \mathcal{H}^t_{A^n}\otimes\mathcal{H}^t_{A'^n}$ denotes a maximally entangled state of Schmidt rank $d_t = |\mathcal{T}^t|$, and

$$p'(t) := \sum_{x^n\in\mathcal{T}^t} q^n(x^n), \quad \text{where} \quad q^n(x^n) = \prod_{i=1}^n q(x_i)\,. \tag{4.51}$$

Now, fix a small $\mu > 0$ and consider a restriction of $|\psi_{AA'}\rangle^{\otimes n}$ to types μ-close to q. More precisely, we consider the set $\mathcal{P}_n^{q,\mu} := \{t \in \mathcal{P}_n \,|\, D(t\|q) \le \mu\}$ and define

$$|\vartheta_{A^n A'^n}\rangle := \sum_{t\in\mathcal{P}_n^{q,\mu}} \sqrt{p(t)}\,|\Phi^t\rangle, \quad \text{where} \quad p(t) = \frac{p'(t)}{\alpha}, \quad \text{and} \tag{4.52}$$

$$\alpha := \sum_{t\in\mathcal{P}_n^{q,\mu}} p'(t) \;=\; \sum_{\substack{x^n\in\mathcal{X}^n \\ D(P_{x^n}\|q)\le\mu}} q^n(x^n) \ge 1 - 2^{-n\left(\mu - |\mathcal{X}|\frac{\log(n+1)}{n}\right)}, \tag{4.53}$$

where the last inequality follows from (B.6) in Appendix B. Note that

$$\frac{1}{2}\left\|\vartheta_{A'^n B'^n} - \psi_{A'B'}^{\otimes n}\right\|_1 = \sqrt{1 - F^2\left(\vartheta_{A'^n B'^n}, \psi_{A'B'}^{\otimes n}\right)} = \sqrt{1-\alpha} \tag{4.54}$$

$$\le 2^{-\frac{n}{2}\left(\mu - |\mathcal{X}|\frac{\log(n+1)}{n}\right)} =: g(n,\mu), \tag{4.55}$$

where the last inequality follows from (4.53).

Next, recall that $\mathcal{N}^n \equiv (\mathcal{N}_{A\to B})^{\otimes n}$. Then by Proposition 11, for fixed $\varepsilon > 0$ and $0 < 2\delta < \varepsilon$ and $\vartheta_{A^n A'^n}$ given in (4.52), we establish that

$$\log M^*_{\text{ea}}(\mathcal{N}^n, \varepsilon) \ge D_H^{\varepsilon-2\delta}\left(\mathcal{N}^n\left(\vartheta_{A^n A'^n}\right) \,\middle\|\, \sum_{t\in\mathcal{P}_n^{q,\mu}} p(t)\left(\mathcal{N}^n(\pi^t_{A^n})\right)\otimes \pi^t_{A'^n}\right) - f(\varepsilon,\delta), \tag{4.56}$$

where $f(\varepsilon,\delta)$ is given by (4.40), and $\pi^t_{A^n}$ and $\pi^t_{A'^n}$ are completely mixed states. In particular, for any t with $D(t\|q) \le \mu$, we have

$$\pi^t_{A'^n} = \frac{1}{d_t}\sum_{x^n\in\mathcal{T}^t} |x^n\rangle\langle x^n| \le (n+1)^{|\mathcal{X}|}2^{n\mu} \sum_{x^n\in T_t} q^n(x^n)|x^n\rangle\langle x^n| \tag{4.57}$$

$$\le \underbrace{(n+1)^{|\mathcal{X}|}2^{n\mu}}_{=:\,\gamma_{n,\mu}} \sum_{x^n\in\mathcal{X}^n} q^n(x^n)|x^n\rangle\langle x^n| = \gamma_{n,\mu}\,\rho_{A'}^{\otimes n}\,. \tag{4.58}$$

The first inequality in (4.58) follows from (B.5) in Appendix B, which is a consequence of the fact that $D(t\|q) \le \mu$.

Next, we use (4.56) and (4.58) to obtain

$$\log M^*_{\text{ea}}(\mathcal{N}^n, \varepsilon) \tag{4.59}$$

$$\ge D_H^{\varepsilon-2\delta}\left(\mathcal{N}^n\left(\vartheta_{A^n A'^n}\right) \,\middle\|\, \sum_{t\in\mathcal{P}_n^{q,\mu}} p(t)\left(\mathcal{N}^n(\pi^t_{A^n})\right)\otimes \gamma_{n,\mu}\,\rho_{A'}^{\otimes n}\right) - f(\varepsilon,\delta) \tag{4.60}$$

$$= D_H^{\varepsilon-2\delta}\left(\mathcal{N}^n\left(\vartheta_{A^n A'^n}\right) \,\middle\|\, \sum_{t\in\mathcal{P}_n^{q,\mu}} p(t)\left(\mathcal{N}^n(\pi^t_{A^n})\right)\otimes \rho_{A'}^{\otimes n}\right) - f(\varepsilon,\delta) - \log\gamma_{n,\mu} \tag{4.61}$$

$$\ge D_H^{\varepsilon-2\delta-g(n,\mu)}\left(\left(\mathcal{N}(\psi_{AA'})\right)^{\otimes n} \,\middle\|\, \sum_{t\in\mathcal{P}_n^{q,\mu}} p(t)\,\mathcal{N}^n(\pi^t_{A^n})\otimes \rho_{A'}^{\otimes n}\right) - f(\varepsilon,\delta) - \log\gamma_{n,\mu}, \tag{4.62}$$

$$\ge D_H^{\varepsilon-2\delta-g(n,\mu)}\left(\left(\mathcal{N}(\psi_{AA'})\right)^{\otimes n} \,\middle\|\, \sum_{t\in\mathcal{P}_n} p(t)\,\mathcal{N}^n(\pi^t_{A^n})\otimes \rho_{A'}^{\otimes n}\right) - f(\varepsilon,\delta) - \log\gamma_{n,\mu}, \tag{4.63}$$

$$= D_H^{\varepsilon-2\delta-g(n,\mu)}\left(\left(\mathcal{N}(\psi_{AA'})\right)^{\otimes n} \,\middle\|\, \left(\mathcal{N}(\rho_A)\right)^{\otimes n}\otimes \rho_{A'}^{\otimes n}\right) - f(\varepsilon,\delta) - \log\gamma_{n,\mu}. \tag{4.64}$$

The first and second lines follow from items 1 and 2 of Lemma 7, respectively. The third line follows from item 4 of Lemma 7. The fourth line also follows from item 2 of Lemma 7, since

$$\sum_{t\in\mathcal{P}_n} p(t)\mathcal{N}^n(\pi^t_{A^n}) \otimes \rho_{A'}^{\otimes n} \geq \sum_{t\in\mathcal{P}_n^{q,\mu}} p(t)\mathcal{N}^n(\pi^t_{A^n}) \otimes \rho_{A'}^{\otimes n}. \tag{4.65}$$

The last line follows from the linearity of $\mathcal{N}^n$ and the fact that

$$\sum_{t\in\mathcal{P}_n} p(t)\pi^t_{A^n} = \mathrm{Tr}_{A'^n}(\psi_{AA'}^{\otimes n}) = \rho_A^{\otimes n}. \tag{4.66}$$

Let us choose $\delta = 1/\sqrt{n}$ and $\mu = ((|\mathcal{X}|+1)\log(n+1))/n$. Then

$$g(n,\mu) = \frac{1}{\sqrt{(n+1)}} \leq \frac{1}{\sqrt{n}}, \quad \text{and} \quad \varepsilon - 2\delta - g(n,\mu) \geq \varepsilon - 3/\sqrt{n}. \tag{4.67}$$

Since $D_H^{\varepsilon}(\rho\|\sigma) \geq D_H^{\varepsilon'}(\rho\|\sigma)$ for $\varepsilon > \varepsilon'$, we obtain the following bound from (4.64)

$$\log M^*_{\mathrm{ea}}(\mathcal{N}^n, \varepsilon) \geq D_H^{\varepsilon - 3/\sqrt{n}}\Big((\mathcal{N}(\psi_{AA'}))^{\otimes n} \Big\| (\mathcal{N}(\rho_A))^{\otimes n} \otimes \rho_{A'}^{\otimes n}\Big) - f(\varepsilon,\delta) - \log\gamma_{n,\mu}, \tag{4.68}$$

thus arriving at an expression involving the hypothesis testing relative entropy for product states. Substituting the above choices for δ and μ in the expressions (4.40) for $f(\varepsilon,\delta)$ and in $\gamma_{n,\mu}$, we find that

$$f(\varepsilon,\delta) + \log\gamma_{n,\mu} = O(\log n). \tag{4.69}$$

Crucially, Lemma 8 applied to (4.68) now implies that

$$\begin{aligned} \log M^*_{\mathrm{ea}}(\mathcal{N}^n, \varepsilon) &\geq nD\big(\mathcal{N}(\psi_{AA'}) \,\big\|\, \mathcal{N}(\rho_A) \otimes \rho_{A'}\big) \\ &\qquad + \sqrt{nV\big(\mathcal{N}(\psi_{AA'}) \,\big\|\, \mathcal{N}(\rho_A) \otimes \rho_{A'}\big)}\, \Phi^{-1}(\varepsilon - 3/\sqrt{n}) + K'(n;\mathcal{N},\varepsilon,\psi_{AA'}) \qquad (4.70) \\ &= nI(A':B)_\omega + \sqrt{nV(A':B)_\omega}\, \Phi^{-1}(\varepsilon - 3/\sqrt{n}) + K'(n;\mathcal{N},\varepsilon,\psi_{AA'}), \qquad (4.71) \end{aligned}$$

where $K'(n;\mathcal{N},\varepsilon,\psi_{AA'}) = O(\log n)$ due to (4.69). To conclude the proof, note that Φ^{-1} is continuously differentiable around $\varepsilon > 0$, and thus $\Phi^{-1}(\varepsilon - 3/\sqrt{n}) = \Phi^{-1}(\varepsilon) + O(1/\sqrt{n})$. □

4.4 Second-Order Converse for Covariant Quantum Channels

In this section, we observe that the Gaussian approximation is valid for the entanglement-assisted capacity of covariant quantum channels (i.e., Conjecture 5 is true for this class of channels). Holevo first defined the class of covariant quantum channels [Hol02b], and it is now known that many channels fall within this class, including depolarizing channels, transpose depolarizing channels [WH02, FHMV04], Pauli channels, cloning channels [Bra11], etc. Note that the following argument up to (4.81) has already essentially been proven in Section III-E of Matthews and Wehner [MW14]. However, we give a brief exposition in this section for completeness. We leave open the question of determining whether the Gaussian approximation is valid for the entanglement-assisted capacity of general discrete memoryless quantum channels.

Let $\mathcal{N}_{A\to B}$ be a quantum channel mapping density operators acting on an input Hilbert space $\mathcal{H}_A$ to those acting on an output Hilbert space $\mathcal{H}_B$. Let G be a compact group, and for every $g \in G$, let

$g \to U_A(g)$ and $g \to V_B(g)$ be continuous projective unitary representations of G in $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. Then the channel $\mathcal{N}_{A\to B}$ is said to be covariant with respect to these representations if the following relation holds for all $g \in G$ and input density operators ρ:

$$\mathcal{N}_{A\to B}\Big(U_A(g)\rho U_A^\dagger(g)\Big) = V_B(g)\mathcal{N}_{A\to B}(\rho)V_B^\dagger(g). \tag{4.72}$$

We restrict our attention in this section to covariant channels for which the representation acting on the input space is irreducible.

In [MW14, Thm. 14], Matthews and Wehner establish the following upper bound on the one-shot entanglement-assisted capacity of a channel $\mathcal{N} \equiv \mathcal{N}_{A\to B}$.

$$\log M^*_{\text{ea}}(\mathcal{N}, \varepsilon) \leq \max_{\rho_A} \min_{\sigma_B} D^\varepsilon_H\big(\mathcal{N}(\phi^\rho_{AA'})\big\|\, \rho_{A'} \otimes \sigma_B\big), \tag{4.73}$$

where $\phi^\rho_{AA'}$ is a purification of ρ_A and $\rho_{A'}$ is the reduction of $\phi^\rho_{AA'}$ to A'. They also prove [MW14, Thm. 19] that the quantity $\beta_\varepsilon\left(\mathcal{N}_{A\to B}(\phi^\rho_{AA'})\|\rho_{A'}\otimes\sigma_B\right)$ defined through (4.7) is convex in the input density operator ρ_A for any σ_B, from which it follows that the quantity

$$\beta_\varepsilon(\mathcal{N}_{A\to B}, \rho_A) := \max_{\sigma_B} \beta_\varepsilon\big(\mathcal{N}_{A\to B}(\phi^\rho_{AA'})\big\|\, \rho_{A'} \otimes \sigma_B\big) \tag{4.74}$$

is convex in ρ_A because it is the pointwise maximum of a set of convex functions.

We would now like to apply these results to the entanglement-assisted capacity of any discrete memoryless covariant channel $\mathcal{N}_{A^n\to B^n} \equiv \mathcal{N}^{\otimes n}$. By definition, such channels have the following covariance:

$$\begin{aligned}\mathcal{N}_{A^n\to B^n}&\left([U_{A_1}(g_1)\otimes\cdots\otimes U_{A_n}(g_n)]\,\rho_{A^n}\,[U_{A_1}(g_1)\otimes\cdots\otimes U_{A_n}(g_n)]^\dagger\right)\\ &= [V_{B_1}(g_1)\otimes\cdots\otimes V_{B_n}(g_n)]\,\mathcal{N}_{A^n\to B^n}(\rho_{A^n})\,[V_{B_1}(g_1)\otimes\cdots\otimes V_{B_n}(g_n)]^\dagger. \end{aligned}\tag{4.75}$$

Let T_{A^n} be a shorthand for a sequence of local unitaries of the form $U_{A_1}(g_1)\otimes\cdots\otimes U_{A_n}(g_n)$. Let $\mathbb{E}$ denote the expectation over all such unitaries T_{A^n}, with the measure being the product Haar measure $\mu(g_1)\times\cdots\times\mu(g_n)$. Then following [MW14, Sec. III-E], we can conclude the following chain of inequalities:

$$\beta_\varepsilon(\mathcal{N}_{A^n\to B^n}, \rho_{A^n}) = \mathbb{E}\left\{\beta_\varepsilon\left(\mathcal{N}_{A^n\to B^n}, T_{A^n}\rho_{A^n}T^\dagger_{A^n}\right)\right\} \tag{4.76}$$

$$\geq \beta_\varepsilon\left(\mathcal{N}_{A^n\to B^n}, \mathbb{E}\left\{T_{A^n}\rho_{A^n}T^\dagger_{A^n}\right\}\right) \tag{4.77}$$

$$= \beta_\varepsilon\left(\mathcal{N}_{A^n\to B^n}, \pi_{A_1}\otimes\cdots\otimes\pi_{A_n}\right), \tag{4.78}$$

where π is the maximally mixed state. The first equality is a result of [MW14, Prop. 29] (this follows directly from the assumption of channel covariance with respect to the operations T_{A^n}). The sole inequality exploits convexity as mentioned above. The last equality follows because the state $\mathbb{E}\big\{T_{A^n}\rho_{A^n}T^\dagger_{A^n}\big\}$ commutes with all local unitaries $U_{A_1}(g_1)\otimes\cdots\otimes U_{A_n}(g_n)$. As a consequence of Schur's lemma and the irreducibility of the representation on the input space, the only state which possesses such invariances is the tensor-power maximally mixed state. Note that we require irreducibility of the representation on only the input space in order for this argument to hold. So, by using the definition of D_H^ε, we can then conclude that

$$\log M^*_{\mathrm{ea}}(\mathcal{N}_{A^n\to B^n},\varepsilon) \le \max_{\rho_{A^n}} \min_{\sigma_{B^n}} D_H^\varepsilon(\mathcal{N}_{A^n\to B^n}(\phi^\rho_{A^n A'^n}) \| \rho_{A'^n}\otimes\sigma_{B^n}) \tag{4.79}$$

$$\le \min_{\sigma_{B^n}} D_H^\varepsilon((\mathcal{N}_{A\to B}(\Phi_{AA'}))^{\otimes n} \| \pi_{A'}^{\otimes n}\otimes\sigma_{B^n}) \tag{4.80}$$

$$\le D_H^\varepsilon((\mathcal{N}_{A\to B}(\Phi_{AA'}))^{\otimes n} \| \pi_A^{\otimes n}\otimes[\mathcal{N}_{A\to B}(\pi_A)]^{\otimes n}) \tag{4.81}$$

$$= nI(A':B)_\omega + \sqrt{nV(A':B)_\omega}\,\Phi^{-1}(\varepsilon) + O(\log n), \tag{4.82}$$

where the information quantities in the final line are with respect to the state $\omega_{A'B} := \mathcal{N}_{A\to B}(\Phi_{AA'})$. The final equality uses the asymptotic expansion in Lemma 8.

5 Discussion

We have established the direct part of the Gaussian approximation in Theorem 3 and conjectured that the converse also holds in Conjecture 5. We again note that all of our results apply to entanglement-assisted quantum communication as well, due to the teleportation [BBC+93] and super-dense coding [BW92] protocols and the results of [LM15]. In the following we will discuss some of the approaches taken and difficulties encountered when trying to prove the converse for general channels.

Arimoto Converse: Converse proofs using Arimoto's approach [Ari73] and quantum generalizations of the Rényi divergence [MLDS+13, WWY14] as in [GW15] can be used to establish that the probability of successful decoding goes to zero exponentially fast for codes with $\frac{1}{n}\log|M| > C_{\mathrm{ea}}$. However, they only yield trivial results when $\frac{1}{n}\log|M| = C_{\mathrm{ea}} \pm O(1/\sqrt{n})$, as is the case in the Gaussian approximation.

De Finetti Theorems: Following Matthews and Wehner [MW14], we find the following converse bound for n uses of the channel employing the arguments presented in Section 4.4 and (4.73).

$$\log M^*_{\mathrm{ea}}(\mathcal{N}^n,\varepsilon) \le \max_{\rho_{A^n}} \min_{\sigma_{B^n}} D_H^\varepsilon\big(\mathcal{N}(\phi^\rho_{A^n A'^n}) \big\| \rho_{A'^n}\otimes\sigma_{B^n}\big), \tag{5.1}$$

where ρ_{A^n} and σ_B^n are invariant under permutations of the n systems, and $\phi^\rho_{A^n A'^n}$ is chosen to have this property as well. One may now try to approximate the state $\phi_{A^n A'^n}$ by a convex combination of product states using the de Finetti theorem or the exponential de Finetti theorem [Ren07]. However, the problem is that the number of systems that need to be sacrificed is at least of the order $\sqrt{n}$, and thus affects the second-order term significantly.

Relation to Channel Simulation: EAC coding is closely related to the classical communication cost in entanglement-assisted channel simulation [BDH+14] and [BCR11]. In the latter paper, some bounds on the classical communication cost of entanglement-assisted channel simulation for a finite number of channels n are given. However, these bounds turn out to be unsuitable for our purposes since the error is scaled by a factor polynomial in n as a result of applying the post-selection technique [CKR09]. It is not clear how the proof in [BCR11] can be adapted to yield a statement for fixed error.

We believe that establishing Conjecture 5 thus requires new techniques and that this constitutes an interesting open problem.

Acknowledgements. We are especially grateful to Milan Mosonyi for insightful discussions and for his help in establishing the proof of Propositions 11 and 14. We acknowledge discussions with Mario Berta, Ke Li, Will Matthews, and Andreas Winter, and we thank the Isaac Newton Institute (Cambridge) for its hospitality while part of this work was completed. MT is funded by the Ministry of Education (MOE) and National Research Foundation Singapore, as well as MOE Tier 3 Grant "Random numbers from quantum processes" (MOE2012-T3-1-009). MMW acknowledges startup funds from the Department of Physics and Astronomy at LSU, support from the NSF through Award No. CCF-1350397, and support from the DARPA Quiness Program through US Army Research Office award W31P4Q-12-1-0019.

A Heisenberg-Weyl Operators

For any $x, z \in \{0, 1, \cdots, d\}$ the Heisenberg-Weyl Operators $X(x)$ and $Z(z)$ are defined through their actions on the vectors of the qudit computational basis $\{|j\rangle\}_{j\in\{0,1,\cdots,d-1\}}$ as follows:

$$X(x)|j\rangle = |j \oplus x\rangle, \tag{A.1}$$

$$Z(z)|j\rangle = e^{2\pi i z j/d}|j\rangle, \tag{A.2}$$

where $j \oplus x = (j + x) \bmod d$. Also note that if $d = 1$, then both $X(x)$ and $Z(z)$ are equal to the identity operator.

B The Method of Types

In our proofs we employ the notion of *types* [Csi98], and hence we briefly recall certain relevant definitions and properties here.

Let $\mathcal{X}$ denote a discrete alphabet and fix $n \in \mathbb{N}$. The *type* (or empirical probability distribution) P_{x^n} of a sequence $x^n \in \mathcal{X}^n$ is the empirical frequency of occurrences of each letter of $\mathcal{X}$, i.e., $P_{x^n}(a) := \frac{1}{n}\sum_{i=1}^n \delta_{x_i,a}$ for all $a \in \mathcal{X}$. Let $\mathcal{P}_n$ denote the set of all types. The number of types, $|\mathcal{P}_n|$, satisfies the bound [CT91, Thm. 11.1.1]

$$|\mathcal{P}_n| \le (n+1)^{|\mathcal{X}|}. \tag{B.1}$$

For any type $t \in \mathcal{P}_n$, the *type class* $\mathcal{T}^t$ of t is the set of sequences of type t, i.e.

$$\mathcal{T}^t := \{x^n \in \mathcal{X}^n \,:\, P_{x^n} = t\}. \tag{B.2}$$

The number of types in a type class $\mathcal{T}^t$ satisfies the following lower bound [Csi98, Lm. II.2]:

$$|\mathcal{T}^t| \ge \frac{2^{nH(t)}}{(n+1)^{|\mathcal{X}|}}, \tag{B.3}$$

where $H(t) := -\sum_{a\in\mathcal{X}} t(a)\log t(a)$, is the Shannon entropy of the type.

Let q be any probability distribution on $\mathcal{X}$. For any sequence $x^n = (x_1, x_2, \ldots, x_n) \in \mathcal{X}^n$, let $q^n(x^n) = \prod_{i=1}^n q(x_i)$. Then, we have

$$q^n(x^n) = 2^{-n(H(t)+D(t\|q))}, \qquad \text{where} \quad t = P_{x^n} \tag{B.4}$$

is the type of x^n and $D(t\|q) := \sum_{a\in\mathcal{X}} t(a)\log\frac{t(a)}{q(a)}$ is the Kullback-Leibler divergence of the probability distributions t and q. From (B.1), (B.3) and (B.4) it follows that for any sequence $x^n \in \mathcal{X}^n$ of type t,

$$(n+1)^{|\mathcal{X}|}2^{nD(t\|q)}q^n(x^n) = 2^{-nH(t)}(n+1)^{|\mathcal{X}|} \geq \frac{1}{|\mathcal{T}^t|}. \tag{B.5}$$

Finally, for any $\mu > 0$ we have [CT91, Eq. (11.98)]

$$\sum_{\substack{x^n\in\mathcal{X}^n \\ D(P_{x^n}\|q)>\mu}} q^n(x^n) \leq 2^{-n\left(\mu-|\mathcal{X}|\frac{\log(n+1)}{n}\right)}. \tag{B.6}$$

References

[Ari73] Suguru Arimoto. On the converse to the coding theorem for discrete memoryless channels. *IEEE Transactions on Information Theory*, 19(3):357–359, May 1973.

[BBC+93] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Physical Review Letters*, 70:1895–1899, 1993.

[BCR11] Mario Berta, Matthias Christandl, and Renato Renner. The quantum reverse Shannon theorem based on one-shot information theory. *Communications in Mathematical Physics*, 306(3):579–615, August 2011. arXiv:0912.3805.

[BDH+14] Charles H. Bennett, Igor Devetak, Aram W. Harrow, Peter W. Shor, and Andreas Winter. Quantum reverse Shannon theorem. *IEEE Transactions on Information Theory*, 60(5):1–34, May 2014. arXiv:0912.5537.

[BG14] Salman Beigi and Amin Gohari. Quantum achievability proof via collision relative entropy. *IEEE Transactions on Information Theory*, 60(12):7980–7986, December 2014. arXiv:1312.3822.

[Bow04] Garry Bowen. Quantum feedback channels. *IEEE Transactions on Information Theory*, 50(10):2429–2434, October 2004. arXiv:quant-ph/0209076.

[Bra11] Kamil Bradler. An infinite sequence of additive channels: The classical capacity of cloning channels. *IEEE Transactions on Information Theory*, 57(8):5497–5503, August 2011. arXiv:0903.1638.

[BSST99] Charles H. Bennett, Peter W. Shor, John A. Smolin, and Ashish V. Thapliyal. Entanglement-assisted classical capacity of noisy quantum channels. *Physical Review Letters*, 83(15):3081–3084, October 1999. arXiv:quant-ph/9904023.

[BSST02] Charles H. Bennett, Peter W. Shor, John A. Smolin, and Ashish V. Thapliyal. Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem. *IEEE Transactions on Information Theory*, 48:2637–2655, October 2002. arXiv:quant-ph/0106052.

[BW92] Charles H. Bennett and Stephen J. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Physical Review Letters*, 69(20):2881–2884, November 1992.

[CKR09] Matthias Christandl, Robert König, and Renato Renner. Postselection technique for quantum channels with applications to quantum cryptography. *Physical Review Letters*, 102(2):020504, January 2009. arXiv:0809.3019.

[CMW14] Tom Cooney, Milan Mosonyi, and Mark M. Wilde. Strong converse exponents for a quantum channel discrimination problem and quantum-feedback-assisted communication. August 2014. arXiv:1408.3373.

[Csi98] Imre Csiszár. The method of types. *IEEE Transactions on Information Theory*, 44(6):2505–2523, October 1998.

[CT91] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley-Interscience, 1991.

[DH13] Nilanjana Datta and Min-Hsiu Hsieh. One-shot entanglement-assisted quantum and classical communication. *IEEE Transactions on Information Theory*, 59(3):1929–1939, March 2013. arXiv:1105.3321.

[DL15] Nilanjana Datta and Felix Leditzky. Second-order asymptotics for source coding, dense coding and pure-state entanglement conversions. *IEEE Transactions on Information Theory*, 61(1):582–608, January 2015. arXiv:1403.2543.

[FHMV04] M. Fannes, B. Haegeman, M. Mosonyi, and D. Vanpeteghem. Additivity of minimal entropy output for a class of covariant channels. October 2004. arXiv:quant-ph/0410195.

[GW15] Manish K. Gupta and Mark M. Wilde. Multiplicativity of completely bounded p-norms implies a strong converse for entanglement-assisted capacity. *Communications in Mathematical Physics*, 334(2):867–887, March 2015. arXiv:1310.7028.

[Has09] Matthew B. Hastings. Superadditivity of communication capacity using entangled inputs. *Nature Physics*, 5:255–257, April 2009. arXiv:0809.3972.

[Hay09] Masahito Hayashi. Information spectrum approach to second-order coding rate in channel coding. *IEEE Transactions on Information Theory*, 55(11):4947–4966, November 2009. arXiv:0801.2242.

[HDW08] Min-Hsiu Hsieh, Igor Devetak, and Andreas Winter. Entanglement-assisted capacity of quantum multiple-access channels. *IEEE Transactions on Information Theory*, 54(7):3078–3090, July 2008. arXiv:quant-ph/0511228.

[HN03] Masahito Hayashi and Hiroshi Nagaoka. General formulas for capacity of classical-quantum channels. *IEEE Transactions on Information Theory*, 49(7):1753–1768, July 2003. arXiv:quant-ph/0206186.

[Hol02a] Alexander S. Holevo. On entanglement assisted classical capacity. *Journal of Mathematical Physics*, 43(9):4326–4333, September 2002. arXiv:quant-ph/0106075.

[Hol02b] Alexander S. Holevo. Remarks on the classical capacity of quantum channel. December 2002. arXiv:quant-ph/0212025.

[HP91] Fumio Hiai and Dénes Petz. The proper formula for relative entropy and its asymptotics in quantum probability. *Communications in Mathematical Physics*, 143(1):99–114, December 1991.

[HSR03] Michael Horodecki, Peter W. Shor, and Mary Beth Ruskai. Entanglement breaking channels. *Reviews in Mathematical Physics*, 15(6):629–641, August 2003. arXiv:quant-ph/0302031.

[KH13] Wataru Kumagai and Masahito Hayashi. Entanglement concentration is irreversible. *Physical Review Letters*, 111(13):130407, September 2013. arXiv:1305.6250.

[Li14] Ke Li. Second order asymptotics for quantum hypothesis testing. *Annals of Statistics*, 42(1):171–189, February 2014. arXiv:1208.1400.

[LM15] Debbie Leung and William Matthews. On the power of PPT-preserving and non-signalling codes. *IEEE Transactions on Information Theory*, 61(8):4486–4499, August 2015. arXiv:1406.7142.

[MLDS$^+$13] Martin Müller-Lennert, Frederic Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel. On quantum Rényi entropies: A new generalization and some properties. *Journal of Mathematical Physics*, 54(12):122203, December 2013. arXiv:1306.3142.

[MW14] William Matthews and Stephanie Wehner. Finite blocklength converse bounds for quantum channels. *IEEE Transactions on Information Theory*, 60(11):7317–7329, November 2014. arXiv:1210.4722.

[PPV10] Yury Polyanskiy, H. Vincent Poor, and Sergio Verdú. Channel coding rate in the finite blocklength regime. *IEEE Transactions on Information Theory*, 56(5):2307–2359, May 2010.

[Ren07] Renato Renner. Symmetry of large physical systems implies independence of subsystems. *Nature Physics*, 3(9):645–649, July 2007. arXiv:quant-ph/0703069.

[Sha48] Claude E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 1948.

[Sha56] Claude E. Shannon. The zero-error capacity of a noisy channel. *IRE Transactions on Information Theory*, IT-2:8–19, 1956.

[Str62] Volker Strassen. Asymptotische Abschätzungen in Shannons Informationstheorie. In *Trans. Third Prague Conf. Inf. Theory*, pages 689–723, Prague, 1962.

[SW97] Benjamin Schumacher and Michael D. Westmoreland. Sending classical information via noisy quantum channels. *Physical Review A*, 56(1):131–138, July 1997.

[TBR15] Marco Tomamichel, Mario Berta, and Joseph M. Renes. Quantum coding with finite resources. April 2015. arXiv:1504.04617.

[TH13] Marco Tomamichel and Masahito Hayashi. A hierarchy of information quantities for finite block length analysis of quantum tasks. *IEEE Transactions on Information Theory*, 59(11):7693–7710, November 2013. arXiv:1208.1478.

[TT15] Marco Tomamichel and Vincent Y. F. Tan. Second-order asymptotics for the classical capacity of image-additive quantum channels. *Communications in Mathematical Physics*, 338(1):103–137, August 2015. arXiv:1308.6503, DOI: 10.1007/s00220-015-2382-0.

[WH02] Reinhard F. Werner and Alexander S. Holevo. Counterexample to an additivity conjecture for output purity of quantum channels. *Journal of Mathematical Physics*, 43(9):4353–4357, September 2002. arXiv:quant-ph/0203003.

[Wil13] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, June 2013. arXiv:1106.1445.

[WR12] Ligong Wang and Renato Renner. One-shot classical-quantum capacity and hypothesis testing. *Physical Review Letters*, 108(20):200501, May 2012. arXiv:1007.5456.

[WWY14] Mark M. Wilde, Andreas Winter, and Dong Yang. Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy. *Communications in Mathematical Physics*, 331(2):593–622, October 2014. arXiv:1306.1586.