Covert Communication over Classical-Quantum Channels

Azadeh Sheikholeslami*, Boulat A. Bash†, Don Towsley†, Dennis Goeckel*, Saikat Guha†

*Electrical and Computer Engineering Department, University of Massachusetts, Amherst, Massachusetts 01003–9292
†Quantum Information Processing Group, Raytheon BBN Technologies, Cambridge, Massachusetts 02138
‡College of Information and Computer Sciences, University of Massachusetts, Amherst, Massachusetts 01003–9264

Abstract—Recently, the fundamental limits of covert, i.e., reliable-yet-undetectable, communication have been established for general memoryless channels and for lossy-noisy bosonic (quantum) channels with a quantum-limited adversary. The key import of these results was the square-root law (SRL) for covert communication, which states that $O(\sqrt{n})$ covert bits, but no more, can be reliably transmitted over $n$ channel uses with $O(\sqrt{n})$ bits of secret pre-shared between communicating parties. Here we prove the achievability of the SRL for a general memoryless classical-quantum channel, showing that SRL covert communication is achievable over any quantum communication channel with a product-state transmission strategy. We leave open the converse, which, if proven, would show that even using entangled transmissions and entangling measurements, the SRL for covert communication cannot be surpassed over an arbitrary quantum channel.

I. INTRODUCTION

Security is important for many types of communication, ranging from electronic commerce to diplomatic missives. Preventing the extraction of information from a message by an unauthorized party has been extensively studied by the cryptography and information theory communities. However, the standard secure communication tools do not address the situations when not only the content of the signal must be protected, but also the detection of the occurrence of the communication must be prevented. This motivates an exploration of the information-theoretic limits of covert communication, i.e., communicatimg with low probability of detection/interception (LPD/LPI). Communication channels can be divided into three classes over which: (a) covert communication is impossible, (b) constant-rate covert communication is possible, and (c) covert communication is governed by the square root law (SRL). The third class is non-trivial: the authors of [2] examined covert communication over the additive white Gaussian noise (AWGN) channels from the transmitter to the intended recipient and the adversary, and found that $O(\sqrt{n})$ covert bits (but no more) can be reliably transmitted over $n$ channel uses [2]. More recently, the authors in [3] and [4] extended the SRL to arbitrary discrete memoryless channels (DMCs), and characterized the constant in front of $\sqrt{n}$ in terms of the channels’ transition probabilities.

A classical channel stems from a ‘quantum’ channel, i.e., the physical electromagnetic propagation medium, with a choice of the quantum states of the transmitted signal and the receiver measurement, whose quantum description is with the positive operator-valued measure (POVM) operators. For example, a lossy optical (quantum) channel, when paired with laser-light (coherent state) modulation and a heterodyne detection receiver, induces an AWGN channel. Similarly, a lossy optical channel when paired with laser-light signaling and an ideal photon counting receiver induces a continuous-input discrete-output Poisson channel. The classical communication capacity (the Holevo capacity) of the quantum channel itself—without any restrictive assumptions on the transmitted signals and the receiver measurement—is generally greater than the capacities of the classical channels induced by pairing the quantum channel with specific conventional transmitters and receivers [5]. This is because using transmit states that are entangled over multiple channel uses and/or employing joint (entangling, or inseparable) measurements over blocks of multiple channel uses at the output can increase the capacity, even if the underlying quantum channel acts independently and memorylessly on each channel use.

For a large class of practical quantum channels, which can be modeled as lossy, additive-thermal-noise bosonic channels, entangled inputs are known not to help attain any capacity advantage [6], i.e., transmitting individually-modulated laser-light pulses of complex-amplitude $\alpha$ on each channel use (i.e., a product-state input), with $\alpha$ drawn i.i.d. from a complex Gaussian distribution, is optimal. On the other hand, using collective measurements (over many channel uses) at the receiver does increase the capacity of such Gaussian bosonic channels—not only over what is achievable using any standard optical receiver, but also over what is achievable with an arbitrary measurement allowed by quantum mechanics that acts on single channel uses at a time. The SRL governs covert communication over Gaussian bosonic channels [7], which motivates its generalization to the class of classical-quantum (cq) channels which do not transmit entangled inputs, i.e., where the transmitter Alice maps a classical random variable $x \in \mathcal{X}$ to a transmitted quantum state...
The classical-quantum channel we consider is the map $x \mapsto \sigma_x^B \in \mathcal{D}(\mathcal{H})$, where $x \in \mathcal{X}$ is Alice’s classical input, $\mathcal{X}$ being the input alphabet, and $\mathcal{D}(\mathcal{H})$ is the set of density operators on a $d$-dimensional Hilbert space $\mathcal{H}$. The classical-quantum channel from Alice to Bob is the map $x \mapsto \sigma_x^B \in \mathcal{D}(\mathcal{H}_B)$, where $\sigma_x^B = \text{Tr}_B(\rho_x^W)$ is the state that Bob receives, and the classical-quantum channel from Alice to Willie is the map $x \mapsto \rho_x^W \in \mathcal{D}(\mathcal{H}_W)$, where $\rho_x^W = \text{Tr}_B(\rho_x^W)$ is the state that Willie receives, and $\text{Tr}_C(\cdot)$ is the partial trace over system $C$. For simplicity, we consider binary inputs, i.e., $\mathcal{X} = \{0, 1\}$. The symbol 0 is called the innocent symbol, which is the input of the channel when no communication occurs, and the symbol 1 is called the non-innocent symbol. For simplicity of notation, we will drop the system-label superscripts, i.e., we denote $\rho^W$ by $\rho$, $\sigma^B$ by $\sigma$, and $\rho^W$ by $\rho$. We consider communication over a memoryless cq channel. Hence, the output state corresponding to the input sequence $x = (x_1, \ldots, x_n) \in \mathcal{X}^n$, $x_i \in \{0, 1\}$, at Bob is given by,
\[\sigma^n(x) = \sigma_{x_1} \otimes \cdots \otimes \sigma_{x_n} \in \mathcal{D}(\mathcal{H}_B^{\otimes n}),\]
and at Willie is given by,
\[\rho^n(x) = \rho_{x_1} \otimes \cdots \otimes \rho_{x_n} \in \mathcal{D}(\mathcal{H}_W^{\otimes n}).\]

**A. Reliability metric**

Consider a random codebook such that a code of block-length $n$ consists of an encoding map $\{1, \ldots, M\} \rightarrow \mathcal{X}^n$, where $M$ is the size of codebook, and, a decoding POVM $\Lambda = \{\Lambda_m\}_{m=1}^M$ that Bob performs on his system such that $\sum_m \Lambda_m \leq I$, where $I$ is the identity channel. Bob can distinguish between the channel outputs even if Alice transmits states that are entangled. This cq channel is completely specified by the map $x \mapsto \sigma_x^B \in \mathcal{D}(\mathcal{H}_B)$, and its capacity is given by the HSW theorem,
\[C = \max_{p(x)} \left\{ \frac{1}{n} \sum_x p(x) \text{Tr} \left( \sigma_x^B \log \sigma_x^B \right) \right\} - \sum_x p(x) \text{Tr} (\sigma_x^B \log \sigma_x^B),\]
where $\text{Tr}$ is the von Neumann entropy of the state $\sigma_x$ [8], [9].

Here we consider the information-theoretic limits of covert communication over a cq channel $x \mapsto \sigma_x^B$, from Alice to Bob and Willie. We develop explicit conditions that differentiate classes of cq channels over which, covert communication is impossible, constant-rate covert communication is possible, and covert communication is governed by the SRL. We limit our analysis to a binary-input cq channel $\mathcal{X} \xrightarrow{\sigma^B_x} \mathcal{Y}$, where one classical symbol is input to the channel with no noise in the case of a bosonic channel. We term this the ‘innocent’ input. This is not a restrictive assumption for the proof of achievability. We leave open the converse for case (c), which would show that for an arbitrary non-trivial TPCP map $\mathcal{N} : \mathcal{X} \mapsto \mathcal{Y}$ from Alice to Bob and Willie (i.e., Alice-to-Willie channels with non-zero classical capacity), no more O(\sqrt{n}) bits can be sent both reliably and covertly over the channel—even if Alice transmits states that are entangled over multiple channel uses, and Bob uses arbitrary collective measurements over multiple channel-use blocks at the receiver.

**II. System Model and Metric**

The classical-quantum channel we consider is the map $x \mapsto \sigma_x^B \in \mathcal{D}(\mathcal{H})$, where $x \in \mathcal{X}$ is Alice’s classical input, $\mathcal{X}$ being the input alphabet, and $\mathcal{D}(\mathcal{H})$ is the set of density operators on a $d$-dimensional Hilbert space $\mathcal{H}$. The classical-quantum channel from Alice to Bob is the map $x \mapsto \sigma_x^B \in \mathcal{D}(\mathcal{H}_B)$, where $\sigma_x^B = \text{Tr}_B \{\tau_x^W\}$ is the state that Bob receives, and the classical-quantum channel from Alice to Willie is the map $x \mapsto \rho_x^W \in \mathcal{D}(\mathcal{H}_W)$, where $\rho_x^W = \text{Tr}_B \{\rho_x^W\}$ is the state that Willie receives, and $\text{Tr}_C(\cdot)$ is the partial trace over system $C$. For simplicity, we consider binary inputs, i.e., $\mathcal{X} = \{0, 1\}$. The symbol 0 is called the innocent symbol, which is the input of the channel when no communication occurs, and the symbol 1 is called the non-innocent symbol. For simplicity of notation, we will drop the system-label superscripts, i.e., we denote $\rho^W$ by $\rho$, $\sigma^B$ by $\sigma$, and $\rho^W$ by $\rho$. We consider communication over a memoryless channel, hence, the output state corresponding to the input sequence $x = (x_1, \ldots, x_n) \in \mathcal{X}^n$, $x_i \in \{0, 1\}$, at Bob is given by,
\[\sigma^n(x) = \sigma_{x_1} \otimes \cdots \otimes \sigma_{x_n} \in \mathcal{D}(\mathcal{H}_B^{\otimes n}),\]
and at Willie is given by,
\[\rho^n(x) = \rho_{x_1} \otimes \cdots \otimes \rho_{x_n} \in \mathcal{D}(\mathcal{H}_W^{\otimes n}).\]
III. MAIN RESULTS

Depending on certain conditions on the cq channel between Alice and Willie as we specify below, the following three different scenarios are possible.

A. No covert communication

Consider the case that support $\rho_1$ is not contained in the support of $\rho_0$, i.e. $\text{sup}(\rho_1) \not\subseteq \text{sup}(\rho_0)$. The following theorem shows that in this situation Alice cannot communicate reliably and covertly to Bob.

**Theorem 2.** When $\text{sup}(\rho_1) \not\subseteq \text{sup}(\rho_0)$, reliable covert communication between Alice and Bob is impossible.

**Proof.** See [13, Appendix B].

B. Constant rate covert communication

Consider the case when Willie’s output is fixed, i.e., $\rho_1 = \rho_0$. In this case, $D(\rho_1 \| \rho_0) = 0$, and thus $D(\rho^x \| \rho_0^x) = 0$. Theorem 2 gives that Alice cannot communicate reliably and covertly to Bob. In other words, what Willie sees is irrelevant to what Alice transmits. Hence, by the HSW theorem, the Holevo capacity of the Alice-to-Bob channel can be achieved covertly [12].

C. Square-root law covert communication

Consider the case that $\rho_1 \neq \rho_0$, and the support of $\rho_1$ is contained in the support of $\rho_0$, i.e. $\text{sup}(\rho_1) \subseteq \text{sup}(\rho_0)$. In the remainder of this paper, we will determine the number of bits that can be sent reliably and covertly over such a classical-quantum channel from Alice to Bob. The following theorem establishes the achievability of $O(1/\sqrt{n})$ covert information bits over $n$ uses of a classical-quantum channel that satisfies the conditions described above.

**Theorem 3.** For any stationary memoryless classical-quantum channel with $\text{sup}(\rho_1) \subseteq \text{sup}(\rho_0)$, there exists a coding scheme such that for $n$ sufficiently large and $\omega_n = o(1) \cap \omega \left(\frac{1}{\sqrt{n}}\right)$,

\[
\log M = (1 - \epsilon)\omega_n \sqrt{n} D(\sigma_1 \| \sigma_0),
\]

\[
\log K = \omega_n \sqrt{n} \left[(1 + \epsilon)D(\rho_1 \| \rho_0) - (1 - \epsilon)D(\sigma_1 \| \sigma_0)\right]^+, \]

and,

\[
P^B_e \leq e^{-\chi_1 \omega_n \sqrt{n}},
\]

\[
|D(\rho^x \| \rho_0^x) - D(\rho_1^x \| \rho_0^x)| \leq e^{-\chi_2 \omega_n \sqrt{n}},
\]

\[
D(\rho_1^x \| \rho_0) \leq \chi_3 \omega_n^2,
\]

where $\epsilon \in (0, 1)$, $\chi_1 > 0$, $\chi_2 > 0$, and $\chi_3 > 0$ are constants, and $|x|^+ = \max\{x, 0\}$.

Before we proceed to the proof of Theorem 3, we state some important definitions and lemmas in Section IV.

IV. PREREQUISITES

A. Prior Probability Distribution

We consider the following distribution on $X = \{0, 1\}$:

\[
p(x) = \begin{cases} 
\alpha_n & \text{if } x = 1, \\
1 - \alpha_n & \text{if } x = 0,
\end{cases}
\]

where 1 is the non- innocent symbol, 0 is the innocent symbol, and $\alpha_n$ is the probability of transmitting 1. The output of the classical-quantum channel corresponding to this input distribution is denoted by,

\[
\tau_{\alpha_n} = \sum_{x \in X} p(x) \tau_x = (1 - \alpha_n)\tau_0 + \alpha_n\tau_1.
\]

Hence, the state corresponding to this input distribution that Bob receives is $\sigma_{\alpha_n} = Tr_T \{\tau_{\alpha_n}\}$, and that Willie receives is $\rho_{\alpha_n} = Tr_B \{\tau_{\alpha_n}\}$, respectively. From linearity of the trace, $\sigma_{\alpha_n} = \sum_{x \in X} p(x)\sigma_x = (1 - \alpha_n)\sigma_0 + \alpha_n\sigma_1$.

and,

\[
\rho_{\alpha_n} = \sum_{x \in X} p(x)\rho_x = (1 - \alpha_n)\rho_0 + \alpha_n\rho_1.
\]

B. Characterization of $\alpha_n$

In this section we show that for a specific choice of $\alpha_n$, the quantum relative entropy between the state induced by $p(x)$ over $n$ channel-uses, $\rho_{\alpha_n}^\otimes n$, and the state induced by the innocent symbol over $n$ channel uses, $\rho_0^\otimes n$, vanishes as $n$ tends to infinity. This is the generalization of a similar concept introduced in [4], to classical-quantum systems.

First we recall a lemma from [1].

**Lemma 4.** For any states $S$ and $T$ and any number $c \geq 0$,

\[
D(S \| T) \leq e^{-1} \text{Tr} \left\{ S^{1+c}T^{-c} - S \right\},
\]

where $\chi_3 > 0$ is a constant.

**Proof.** See [13, Appendix C].

V. PROOF OF THEOREM 3

This section is dedicated to the proof of Theorem 3. The proof has two parts. First the reliability of the coding scheme, and then its covertness, are established.

A. Reliability Analysis

In this section our goal is to prove the reliability part of Theorem 3. First we recall a lemma (Lemma 2 in [14]) which we will use in the analysis of the error probability.

**Lemma 6.** For operators $0 < S < I$ and $T > 0$, we have,

\[
I - \sqrt{S + T^{-1}} S \sqrt{S + T^{-1}} \leq (1 + c)(I - S) + (2 + c + e^{-1}) T,
\]

where $c$ is an arbitrary strictly positive real number.

Next, we prove a lemma that will be used in proving both the reliability and covertness. First, consider a self-adjoint operator $A$ and its spectral decomposition $A = \sum \lambda_i |a_i\rangle \langle a_i|$, where $\{\lambda_i\}$ are eigenvalues, and $|a_i\rangle \langle a_i|$ are the associated eigenspaces. Then, the non-negative spectral projection on $A$ is defined as in [14],

\[
\{A \geq 0\} = \sum_{\epsilon \lambda_i \geq 0} |a_i\rangle \langle a_i|,
\]

where $\epsilon$ is the probability of transmitting 0. The output of the classical-quantum channel corresponding to this input distribution is denoted by,

\[
\tau_{\alpha_n} = \sum_{x \in X} p(x) \tau_x = (1 - \alpha_n)\tau_0 + \alpha_n\tau_1.
\]

Hence, the state corresponding to this input distribution that Bob receives is $\sigma_{\alpha_n} = Tr_T \{\tau_{\alpha_n}\}$, and that Willie receives is $\rho_{\alpha_n} = Tr_B \{\tau_{\alpha_n}\}$, respectively. From linearity of the trace, $\sigma_{\alpha_n} = \sum_{x \in X} p(x)\sigma_x = (1 - \alpha_n)\sigma_0 + \alpha_n\sigma_1$.
which is the projection to the eigenspace corresponding to non-negative eigenvalues of $A$. The projections \( \{ A > 0 \} \), \( \{ A \leq 0 \} \), and \( \{ A < 0 \} \) are defined similarly.

**Lemma 7.** For any Hermitian matrix $A$ and positive-definite matrix $B$,
\[
\text{Tr} \{ BA \{ A < 0 \} \} \leq 0, \tag{13}
\]
and,
\[
\text{Tr} \{ BA \{ A > 0 \} \} \geq 0. \tag{14}
\]

**Proof.** See [13, Appendix D]. □

Consider the encoding map \( \{ 1, \ldots, M \} \rightarrow \mathbf{x} \in \mathcal{X}^n \) and the square-root measurement decoding POVM for $n$ channel uses,
\[
\Lambda^n_m = \left( \sum_{k=1}^{M} \Pi_k \right)^{-1/2} \Pi_m \left( \sum_{k=1}^{M} \Pi_k \right)^{-1/2}, \tag{15}
\]
where we define the projector $\Pi_m$ as,
\[
\Pi_m = \{ \hat{\sigma}^n(m) - e^a\sigma^0_0 > 0 \}. \tag{16}
\]
Here \( \hat{\sigma}^n(m) = \mathcal{E}_{\sigma^0_0}(\sigma^n(m)) \) is the pinching of $\sigma^n(m)$ as defined in [13, Appendix A], and $a > 0$ is a real number to be determined later.

The average probability of decoding error at Bob over the random codebook is characterized in the next lemma.

**Lemma 8.** For any $a > 0$ and $c > 0$,
\[
\mathbb{E}[P^B_c] \leq (1 + c) \sum_{\mathbf{x}} p(\mathbf{x}) \text{Tr} \{ \sigma^n(\mathbf{x}) \{ \hat{\sigma}^n(\mathbf{x}) - e^a\sigma^0_0 \leq 0 \} \} + (2 + c + e^{-1})M e^{-a} \exp \left( \omega_n^2 \text{Tr} \{ \sigma^{-1}_0 \} \right). \tag{17}
\]

**Proof.** See [13, Appendix E]. □

Now we evaluate the first term of the right-hand side of (17). In [15] it is shown that for any tensor product states $\sigma^n$ and $T^n$ and any number $p > 0$ and $0 \leq \nu \leq 1$,
\[
\text{Tr} \{ S^n \{ S^n - pT^n \leq 0 \} \} \leq (n + 1)^{p \nu} \text{Tr} \left\{ S^{\nu/2} (T^{\nu/2})^{-q} (T^{\nu/2})^{-q} \right\}, \tag{18}
\]
where \( \hat{S}^n = \mathcal{E}_T(S^n) \). Applying this to states $\sigma^n(\mathbf{x})$ and $T^n = \sigma^0_0$ and setting $p = e^a$ yields,
\[
\sum_{\mathbf{x}} p(\mathbf{x}) \text{Tr} \{ \sigma^n(\mathbf{x}) \{ \hat{\sigma}^n(\mathbf{x}) - e^a\sigma^0_0 \leq 0 \} \} \leq (n + 1)^{e^a} \sum_{\mathbf{x}} p(\mathbf{x})(n + 1)^{e^a} \exp \left( aq + \log \text{Tr} \left\{ \sigma(\mathbf{x})(\sigma_0^0)^{q/2}(\sigma(\mathbf{x}))^{-q}(\sigma_0^0)^{-q/2} \right\} \right)
\]
\[
= (n + 1)^{e^a} \sum_{\mathbf{x}} p(\mathbf{x}) \exp \left( aq + \sum_{i=1}^{n} \log \text{Tr} \left\{ \sigma(x_i)\sigma_0^{q/2}(\sigma(x_i))^{-q}\sigma_0^{q/2} \right\} \right), \tag{19}
\]
where the equality follows from the memoryless property of the channel. Let us define the function
\[
\varphi(\sigma(x), q) = -\log \text{Tr} \left\{ (\sigma(x)\sigma_0^{q/2})(\sigma(x))^{-q}\sigma_0^{q/2} \right\}.
\]
Since $\varphi(\sigma_0, q) = 0$, only terms with $x_i = 1$ contribute to the sum in (19). Define the random variable $L$ indicating the number of non-innocent symbols in $\mathbf{x}$. We define the set similar to [4],
\[
C^n_\mu = \{ l \in \mathbb{N} : l - \mu_\omega_n \sqrt{n} < \omega_n \sqrt{n} \}, \tag{20}
\]

describing the values that the random variable $L$ takes, where $0 \leq \mu < 1$ is a constant. Using a Chernoff bound,
\[
P(L \notin C^n_\mu) \leq 2e^{-n^2\mu_\omega_n^2/2}. \tag{21}
\]

Hence,
\[
\sum_{\mathbf{x}} p(\mathbf{x}) \exp \left( aq - \sum_{i=1}^{n} \varphi(\sigma(x_i), q) \right) = \mathbb{E}_L \sum_{\mathbf{x}} p(\mathbf{x}) \exp \left( aq - \sum_{i=1}^{L} \varphi(\sigma(x_i), q) \right) + P(L \notin C^n_\mu) \leq \exp (aq - (1 - \mu)\omega_n \sqrt{n}\varphi(\sigma_1, q)) + 2e^{-n^2\mu_\omega_n^2/2}. \tag{22}
\]

In [13, Appendix G], it is shown that the derivative of $\varphi(\sigma_1, q)$ with respect to $q$ is uniformly continuous, and,
\[
\frac{\partial}{\partial q} \varphi(\sigma_1, 0) = D(\sigma_1||\sigma_0).
\]

Moreover, we have $\varphi(\sigma_1, 0) = 0$. Now let $\varepsilon > 0$ be an arbitrary constant. Because differentiation of $\varphi(\sigma_1, q)$ is uniformly continuous, there exists $0 < \delta < 1$ s.t.,
\[
\left| \varphi(\sigma_1, q) - \varphi(\sigma_1, 0) \right| = D(\sigma_1||\sigma_0) < \varepsilon \text{ for } 0 < q \leq \delta. \tag{23}
\]

Substituting (22) and (23) into (19), and letting $a = (1 - \nu)(1 - \mu)\omega_n \sqrt{n}D(\sigma_1||\sigma_0)$ where $\nu > 0$ is a constant, and realizing that $q \leq \delta$, yields,
\[
\sum_{\mathbf{x}} p(\mathbf{x}) \text{Tr} \{ \sigma^n(\mathbf{x}) \{ \hat{\sigma}^n(\mathbf{x}) - e^a\sigma^0_0 \leq 0 \} \} \leq (n + 1)^{e^a} \left( e^{-\nu\delta(1 - \mu)\omega_n \sqrt{n} + 2e^{-n^2\mu_\omega_n^2/2}} \right). \tag{24}
\]

Consequently, substituting (24) into (17) yields,
\[
\mathbb{E}[P^B_c] \leq (1 + c)(n + 1)^{e^a} \left( e^{-\nu\delta(1 - \mu)\omega_n \sqrt{n} + 2e^{-n^2\mu_\omega_n^2/2}} \right) + (2 + c + e^{-1})M e^{-\nu\delta(1 - \mu)\omega_n \sqrt{n}D(\sigma_1||\sigma_0)\omega_n^2} \text{Tr} \{ \sigma_0^{-1} \}. \tag{25}
\]

Hence, if
\[
\log M = (1 - \varepsilon)\omega_n \sqrt{n}D(\sigma_1||\sigma_0). \tag{26}
\]

where $1 - \varepsilon = (1 - \gamma)(1 - \mu)(1 - \nu)$, and for sufficiently large $n$ there must exist a constant $\xi > 0$ such that the expected error probability is upper-bounded as,
\[
\mathbb{E}[P^B_c] \leq e^{-\xi\omega_n \sqrt{n}}. \tag{27}
\]
B. Covertness Analysis

The goal is now to show that the average state that Willie receives over $n$ channel uses when communication occurs, $\rho^n = \frac{1}{M^n} \sum_{m=1}^M \sum_{k=1}^K \rho^n(m,k)$, is close to the state he receives when no communication occurs, i.e., $\rho_0^n$ in $\mathcal{H}_1$. In order to show this, we first prove the following lemma.

Lemma 9. For sufficiently large $n$ there exists a coding scheme with

$$\log M + \log K = (1 + \epsilon)\omega_n \sqrt{n}D(\rho_1 || \rho_0), \quad (28)$$

such that,

$$D(\rho^n || \rho_0^n) \leq e^{-\zeta \omega_n \sqrt{n}}, \quad (29)$$

where $\zeta$ is a constant and $\omega_n = o(1) \cap \Omega \left(\frac{1}{n} \right)$.

Proof. See [13, Appendix F]. □

C. Identification of a Specific Code

We choose $\epsilon$, $\zeta$ and $\xi$, $M$, and $K$ such that both (26) and (28) are satisfied. In [13, Appendix H] we use Markov's inequality, to show that, for constants $\chi_1 > 0$ and $\chi_2 > 0$, and sufficiently large $n$, there exists at least one coding scheme such that:

$$P_e^B \leq e^{-\chi_1 \omega_n \sqrt{n}} \text{ and } D(\rho^n || \rho_0^n) \leq e^{-\chi_2 \omega_n \sqrt{n}}. \quad (30)$$

The quantum relative entropy between $\rho^n$ and $\rho_0^n$ is:

$$D(\rho^n || \rho_0^n) = D(\rho^n || \rho_0^n) + D(\rho_0^n || \rho_0^n) + \text{Tr} \left\{ (\rho^n - \rho_0^n) \left( \log \rho^n_0 - \log \rho_0^n \right) \right\}. \quad (31)$$

To show that the last term in right-hand side of (31) vanishes as $n$ tends to infinity, let the eigenvalues of $A = \rho^n - \rho_0^n$ and $B = \log \rho^n_0 - \log \rho_0^n$ be enumerated in decreasing order as $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_d$ and $\delta_1 \geq \delta_2 \geq \cdots \geq \delta_d$, respectively. Then:

$$\text{Tr} \left\{ (\rho^n - \rho_0^n) \left( \log \rho^n_0 - \log \rho_0^n \right) \right\} \leq \sum_{i=1}^d \gamma_i \delta_i \leq \left( \sum_{i=1}^d \gamma_i \right)^\frac{1}{2} \left( \sum_{i=1}^d \delta_i \right)^\frac{1}{2}, \quad (32)$$

where (a) is von Neumann’s trace inequality [16], and (b) is Cauchy-Schwarz inequality.

$$\sum_{i=1}^d \gamma_i^2 = \text{Tr} \left\{ (\rho^n - \rho_0^n)^2 \right\} \leq \text{Tr} \left\{ (\rho^n - \rho_0^n)^2 \right\} \quad (33)$$

Hence, setting $j = 1$,

$$\sum_{i=1}^d \delta_i^2 \leq \sum_{i=1}^d \left( \log (a_i^n) - \log (b_i^n) \right)^2 \quad (34)$$

Substituting (33) and (34) in (32) yields:

$$\text{Tr} \left\{ (\rho^n - \rho_0^n) \left( \log \rho_0^n - \log \rho_0^n \right) \right\} \leq \sqrt{d} \left( \log \frac{a_1}{b_1} \right)^2 \quad (35)$$

Re-arranging (31), substituting (35) and the result of Lemma 9, and appropriately choosing a constant $\chi_2 > 0$ yields:

$$D(\rho^n || \rho_0^n) = D(\rho^n || \rho_0^n) \leq e^{-\chi_2 \omega_n \sqrt{n}}. \quad (36)$$

Application of Lemma 5 completes the proof of Theorem 3, the achievability of square-root-law covert communication over a cq channel. We leave the proof of the converse for future work.

REFERENCES