Entanglement-Assisted Covert Communication via Qubit Depolarizing Channels

Elyakim Israel Zlotnick

Entanglement-Assisted Covert Communication via Qubit Depolarizing Channels

Research Thesis

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Electrical Engineering

Elyakim Israel Zlotnick

Submitted to the Senate
of the Technion — Israel Institute of Technology
Nisan 5784 Haifa April 2024

This research was done in the The Andrew and Erna Viterbi Faculty of Electrical and Computer Engineering, under the supervision of Prof. Uzi Pereg.

The results in this thesis have been published as articles by the author and research collaborators in conferences and journals during the course of the author’s research period:

E. Zlotnick, B. Bash, and U. Pereg. Entanglement-assisted covert communication via the qubit depolarizing channel. *Submitted to IEEE Trans. Info. Theory. Preprint available on arXiv:2305.05477*, 2023.

E. Zlotnick, B. Bash, and U. Pereg. Entanglement-assisted covert communication via the qubit depolarizing channel. In *2023 IEEE Int’l Symp. Inf. Theory (ISIT’2020)*, 2023.

Acknowledgements

I want to sincerely thank Prof. Uzi Pereg for his invaluable guidance and support throughout my thesis. His expertise and encouragement were crucial in shaping my research. I am also grateful to Boulat Bash from the University of Arizona for his helpful advice and support. Lastly, heartfelt thanks to my wife and family for their patience, support, and love throughout this journey.

The author of this thesis states that the research, including the collection, processing and presentation of data, addressing and comparing to previous research, etc., was done entirely in an honest way, as expected from scientific research that is conducted according to the ethical standards of the academic world. Also, reporting the research and its results in this thesis was done in an honest and complete manner, according to the same standards.

The generous financial help of Mrs. Ilene Shaw from the United States, and the Helen Diller Quantum Center at the Technion, is gratefully acknowledged.

Contents

List of Figures

Abstract

The standard security approach in communications aims to prevent a malicious eavesdropper from retrieving the information that is transmitted to the legitimate receiver. However, privacy and safety concerns may require an even stronger security criterion. In covert communication, not only the information is kept secret, but the transmission itself must be concealed from detection by an adversary. Despite the severity of such limitations, it is possible to communicate $O(\sqrt{n})$ bits of information in a block of n transmissions via a noisy channel, in all but trivial scenarios. That is, the sender (Alice) can use an error-correction code to map $O(\sqrt{n})$ information bits into codewords of length n, such that the legitimate receiver (Bob) can decode the information reliably, while the adversary (Willie) cannot detect the transmission. Previously, it was shown that pre-shared entanglement can improve the scaling to $O(\sqrt{n}\log(n))$ information bits in the continuous-variable setting, for Gaussian bosonic channels (Gagatsos et al., 2020). In some information-theoretic frameworks, coding scales are larger in continuous-variable models, compared to the discrete-variable setting. Therefore, until now, it was not clear whether the logarithmic factor can be achieved in discrete-variable covert communication.

Here, we study covert communication via the qubit depolarizing channel with entanglement assistance, in different scenarios. In the canonical representation of the qubit depolarizing channel, Bob receives a single qubit, while two qubits dissipate to the environment. We consider three scenarios. If the adversary has full access to the environment (Scenario 1), then we show that covert communication is impossible. On the other hand, if the adversary receives the first qubit, i.e., “half” the environment (Scenario 2), then covert communication turns out to be trivial. The most interesting case is when the adversary receives the other half (Scenario 3). In this case, the number of information bits that can be transmitted, reliably and covertly, scales as $O(\sqrt{n}\log(n))$ when entanglement assistance is available to Alice and Bob, as opposed to $O(\sqrt{n})$ information bits without assistance. We further developed an explicit expression for the lower bound on this result, which highlights the performance improvement enabled by entanglement resources. Thereby, we establish that the logarithmic performance boost is not reserved to continuous-variable systems.

Finally, we interpret these results in terms of energy-constrained communication. Covert communication can be seen as communication under an energy constraint, where

the total transmission energy must not exceed a specified value, thus reducing the detection probability of transmissions.

Chapter 1

Introduction and Background

Privacy and confidentiality are critical in communication systems [1]. The traditional security approaches (s.t., encryption [2], information-theoretic secrecy [3], and quantum key distribution [4, 5, 6]) ensure that an eavesdropper is unable to recover any transmitted information. However, privacy and safety concerns may further require *covertness* [7, 8, 9, 10, 11, 12]. Covertness is a stronger requirement than traditional security: not only is the transmitted information kept secret, but also the transmission itself is concealed from detection by an adversary (a warden) [13, 14]. Despite the severity of limitations imposed by covertness, it is possible to communicate $O(\sqrt{n})$ bits of information both reliably and covertly over n classical channel uses [15, 16, 17]. This property is referred to as the "square root law" (SRL). The SRL has also been observed in covert communication over finite-dimensional classical-quantum channels [18, 19, 20], as well as continuous-variable bosonic channels [21, 22, 23, 24]. Covert sensing is also governed by an SRL [25, 26].

Quantum information theory demonstrates that quantum channels exhibit distinct advantages over classical channels, establishing their superiority in various communication scenarios. Especially, the utilization of pre-shared entanglement resources has been established as a means to enhance performance and increase throughput. [27, 28, 29, 30, 31]. In the scope of covert communication, Gagatsos et al. [23] showed that entanglement assistance allows transmission of $O(\sqrt{n}\log n)$ reliable and covert bits over n uses of continuous-variable bosonic channel, surpassing the SRL scaling. This scaling is also referred to as the square-root-log law.

1.1 Preliminaries

Quantum information theory provides a general probabilistic framework that captures the performance behavior for communication system of a quantum nature. As the quantum theory reduces to the classical description in the classical limit, quantum information theory can be viewed as a generalization of the classical Shannon theory.

1.1.1 Quantum Systems

We use standard notation in quantum information processing, as, e.g., in [32, Ch. 2.2.1]. We label quantum systems by A, B, C, $\ldots$, and classical systems by X, Y, Z, $\ldots$.

The Hilbert space for a system A is denoted by $\mathcal{H}_A$. The space of linear operators $\mathcal{H} \to \mathcal{H}$ is denoted by $\mathcal{L}(\mathcal{H})$. A quantum state can be represented by a density operator, *i.e.*, a unit-trace positive semidefinite operator. The space of density operators is denoted by $\mathscr{S}(\mathcal{H})$. A positive operator-valued measure (POVM) $\{D_m\}_{m=1}^{M}$ is a set of positive semidefinite linear operators in $\mathcal{L}(\mathcal{H})$ such that $\sum_{m=1}^{M} D_m = \mathbb{1}$, where $\mathbb{1}$ is the identity operator on $\mathcal{H}$.

An operator $V : \mathcal{H}_A \to \mathcal{H}_b$ is called an isometry if $V^\dagger V = \mathbb{1}$. If A and B are of the same dimension, that is, $\dim(\mathcal{H}_A) = \dim(\mathcal{H}_B)$, then V is a unitary, i.e., $V^\dagger V = VV^\dagger = \mathbb{1}$

1.1.2 Quantum Channels

A quantum channel is defined as a completely-positive trace-preserving (CPTP) linear map

$$\mathcal{N}_{A\to B} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)\,. \tag{1.1}$$

Every quantum channel has a Stinespring representation: There exists an operator $V : \mathcal{H}_A \to \mathcal{H}_B \otimes \mathcal{H}_E$ such that

$$\mathcal{N}_{A\to B}(\rho) = \mathrm{Tr}_E(V\rho V^\dagger) \tag{1.2}$$

for $\rho \in \mathcal{L}(\mathcal{H}_A)$, where the operator V is an *isometry*, i.e.,

$$V^\dagger V = \mathbb{1} \tag{1.3}$$

The evolution of a quantum state is described by a quantum channel. The Stinespring representation can be understood as follows: when Alice (subsystem A) transmits information, part of the system reaches Bob (subsystem B), while its complementary is leaked to the environment (subsystem E). By eliminating subsystem E (mathematically, by tracing out), we are left with the system that Bob receives.

In addition, every quantum channel can be presented as a Kraus map, i.e., an operator-sum form:

$$\mathcal{N}_{A\to B}(\rho) = \sum_i K_i \rho K_i^\dagger \tag{1.4}$$

with $\sum_i K_i^\dagger K_i = \mathbb{1}$.

The quantum channel generalizes the classical channel. To illustrate this, consider an input distribution $P_X(x)$ and a classical channel $P_{Y|X}(y|x)$. Assuming an input state $\rho = \sum_x P_X(x) |x\rangle\langle x|$, where $\{|x\rangle\}$ forms an orthonormal basis, and applying a quantum channel with the Kraus operators $\left\{K_{x,y} = \sqrt{P_{Y|X}(y|x)}\, |y\rangle\langle x| \right\}$, the resulting output state is:

$$\mathcal{N}_{A\to B}(\rho) = \sum_y P_Y(y)\, |y\rangle\langle y| \tag{1.5}$$

where P_Y is the output distribution,

$$P_Y(y) = \sum_x P_{Y|X}(y|x) P_X(x) \tag{1.6}$$

(see [33, Sec. 4.6.4] for further details).

1.1.3 Scaling

For a given function $g(n)$, we denote by $O(g(n))$ the set of functions $f(n)$ for which there exist positive constants c and n_0 such that $0 \leq f(n) \leq cg(n)$ for all $n \geq n_0$, we write $f(n) = O(g(n))$ to indicate that a function $f(n)$ belongs to the set $O(g(n))$ [34]. Equivalently,

$$f(n) = O(g(n)) \text{ if } \limsup_{n\to\infty} \left| \frac{f(n)}{g(n)} \right| < \infty\,. \tag{1.7}$$

Similarly, for continuous-variable functions, F and G on $[0, \infty)$, we write

$$F(x) = \mathcal{O}(G(x)) \text{ if } \limsup_{x\to 0} \left| \frac{F(x)}{G(x)} \right| < \infty\,. \tag{1.8}$$

1.1.4 Relative Entropy and Information Measures

Given a pair of quantum states $\rho, \sigma \in \mathscr{S}(\mathcal{H})$, the quantum relative entropy is defined as

$$D(\rho||\sigma) = \mathrm{tr}[\rho(\log(\rho) - \log(\sigma)]\,, \tag{1.9}$$

if $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$; and $D(\rho||\sigma) = +\infty$, otherwise. The second and fourth moments of relative entropy are defined as

$$V(\rho||\sigma) = \mathrm{tr}\left[\rho|(\log(\rho) - \log(\sigma) - D(\rho||\sigma)|^2\right]\,, \tag{1.10}$$

$$Q(\rho||\sigma) = \mathrm{tr}\left[\rho|(\log(\rho) - \log(\sigma) - D(\rho||\sigma)|^4\right]\,, \tag{1.11}$$

if $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$; and $D(\rho||\sigma) = +\infty$, otherwise.

The relative entropy and its moments have the following interpretation. In the quantum theory, the expectation value of a measurable operator A on a quantum system in the state ρ can be expressed as,

$$\langle A \rangle = \mathrm{tr}(A\rho)\,. \tag{1.12}$$

Thereby, we can view relative entropy as the expectation value of the measurable $A \equiv \log(\rho) - \log(\sigma)$ on a system in the state ρ. In a similar manner, the second moment can be interpreted as the expectation value of the measurable $[A - \langle A \rangle]^2$ and likewise the fourth moment is $\Big\langle [A - \langle A \rangle]^4 \Big\rangle$.

In addition to the relative entropy and its moments, we define the information measure $\eta(\rho||\sigma)$. Consider a spectral decomposition,

$$\sigma = \sum_i \lambda_i P_i\,, \tag{1.13}$$

where λ_i are the eigenvalues, and P_i are projection operators (projectors) on the corresponding eigenspaces. Then, define [26]:

$$\eta(\rho||\sigma) = \\ \sum_{i\neq j} \frac{\log(\lambda_i) - \log(\lambda_j)}{\lambda_i - \lambda_j} \mathrm{Tr}[(\rho-\sigma)P_i(\rho-\sigma)P_j] + \sum_i \frac{1}{\lambda_i} \mathrm{Tr}[(\rho-\sigma)P_i(\rho-\sigma)P_i]\,. \tag{1.14}$$

Intuitively, the information measure above can be interpreted as the second derivative of the relative entropy between two density matrices. To see this, we recall that the Taylor expansion of a differentiable function around $\alpha = 0$ is,

$$f(\alpha) = f(0) + f'(0)\alpha + \frac{1}{2} f''(0)\alpha^2 + \dots \tag{1.15}$$

Now, according to [26, Eq. 82], for small $\alpha > 0$,

$$D(\alpha\rho_1 + (1-\alpha)\rho_0||\rho_0) = \frac{1}{2}\eta(\rho_1||\rho_0)\alpha^2 + \mathcal{O}(\alpha^3)\,. \tag{1.16}$$

Therefore, $\eta(\rho_1||\rho_0)$ can be viewed as the second derivative of $D(\alpha\rho_1 + (1-\alpha)\rho_0||\rho_0)$ at $\alpha = 0$. Alternatively, it can be understood as the derivative at the point ρ_0 in the direction of ρ_1. Furthermore, the terms $\frac{1}{\lambda_i}$ and $\frac{\log(\lambda_i)-\log(\lambda_j)}{\lambda_i-\lambda_j}$ are reminiscent of the derivative of the $\log(\cdot)$ function, which appears in the entropic formulas. We note that if ρ and σ commute, then $\eta(\rho||\sigma)$ reduces to the chi-square divergence, i.e., $\eta(\rho||\sigma) = \chi^2(\rho||\sigma) \equiv \mathrm{Tr}(\rho^2\sigma^{-1}) - 1$. Therefore, $\eta(\rho||\sigma)$ can also be viewed as a variation of the chi-square divegence.

1.1.5 Entropy and Mutual Information

The von Neumann entropy for a density operator ρ is defined as

$$H(\rho) \equiv -\operatorname{tr}[\rho \log \rho] \,. \tag{1.17}$$

Given a bipartite state ρ_{AB}, the quantum mutual information is defined as

$$I(A;B)_\rho \equiv H(\rho_A) + H(\rho_B) - H(\rho_{AB}) \,. \tag{1.18}$$

Furthermore, the conditional quantum entropy is defined by

$$H(A|B)_\rho = H(\rho_{AB}) - H(\rho_B) \tag{1.19}$$

and similarly, the conditional mutual information is$I(A;B|C)_\rho = H(A|C)_\rho + H(B|C)_\rho - H(A,B|C)_\rho$.

The Holevo information of a channel $\mathcal{N}_{A\to B}$ is defined as,

$$\chi(\mathcal{N}) \equiv \max_\rho I(X;B)_\rho \,, \tag{1.20}$$

where the maximization is carried out over classical-quantum states of the form,

$$\rho_{XB} = \sum_x P_X(x) \, |x\rangle\!\langle x| \, \mathcal{N}_{A\to B}(\psi_A^x) \,, \tag{1.21}$$

and $\{|x\rangle\}$ is an an orthonormal basis.

The hypothesis testing relative entropy $D_H^\varepsilon(\rho||\sigma)$ is defined for $\varepsilon \in [0,1]$ as [35]:

$$D_H^\varepsilon(\rho||\sigma) = -\log \inf_\Lambda \{\operatorname{Tr}\{\Lambda\sigma\} | \operatorname{Tr}\{\Lambda\rho\} \geq 1-\varepsilon, 0 \leq \Lambda \leq \mathbb{1}\} \,. \tag{1.22}$$

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

$$D_H^\varepsilon(\rho^{\otimes n}||\sigma^{\otimes n}) = nD(\rho||\sigma) + \sqrt{nV(\rho||\sigma)}\Phi^{-1}(\varepsilon) + O(\log n) \tag{1.23}$$

where,

$$\Phi^{-1}(\varepsilon) = \sup\{\varepsilon \in [0,1] | \Phi(\varepsilon) \leq \varepsilon\} \,, \qquad \Phi(\varepsilon) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\varepsilon} e^{\frac{x^2}{2}} dx \,. \tag{1.24}$$

Information measures play a crucial role in characterizing communication performance. For instance, the entropy appears in the characterization of information compression limits, and the mutual information in that of the channel capacity, i.e., the ultimate coding rate for reliable transmission over a noisy channel.

1.2 Unassisted Communication

1.2.1 Unassisted Code

The code for quantum communication without entanglement resources and without security requirements is defined as follows.

Definition 1.2.1. An $(\mathsf{M}(n,R),n)$ unassisted code consists of

- a message set $\{1,\dots,\mathsf{M}(n,R)\}$, where $\mathsf{M}(n,R)$ is an integer,
- an encoding map,

$$\mathcal{F}:\{1,\dots,\mathsf{M}(n,R)\}\to\mathscr{S}(\mathcal{H}_A^{\otimes n})\,, \tag{1.25}$$

  for $m\in[1:\mathsf{M}(n,R)]$,

- a decoding POVM,

$$\mathcal{D}_{B^n}=\{D_m\}_{m=1}^{\mathsf{M}(n,R)} \tag{1.26}$$

We denote the code by $(\mathcal{F},\mathcal{D})$.

We note that the dependence on n is supressed from the encoding and decoding maps in order to simplify notation.

The error probability is defined as the probability of incorrectly decoding the received message, given by

$$\begin{aligned}P_{\mathrm{e}}^{(n)}(\mathcal{F},\mathcal{D})&=\sum_{m=1}^{\mathsf{M}}\frac{1}{\mathsf{M}}\Pr(\widehat{m}\neq m|m)\\&=\frac{1}{\mathsf{M}}\sum_{m=1}^{\mathsf{M}}\left(1-\mathrm{tr}\left[(\mathcal{D}\circ\mathcal{N}_{A\to B}^{\otimes n}\circ\mathcal{F}(m)\right]\right)\,,\end{aligned} \tag{1.27}$$

with $\mathsf{M}\equiv\mathsf{M}(n,R)$, where $\Pr(m\neq\widehat{m})$ is the probability of decoding error for each message m.

An $(\mathsf{M}(n,R),n,\varepsilon)$-code for unassisted communication satisfies

$$P_{\mathrm{e}}^{(n)}(\mathcal{F},\mathcal{D})\leq\varepsilon\,. \tag{1.28}$$

The coding rate is characterized as

$$L_{\text{no-EA}}\equiv\frac{\log\left(\mathsf{M}(n,R)\right)}{n}\,, \tag{1.29}$$

i.e., the number of bits per channel use. An achievable coding rate is defined as follow:

Definition 1.2.2. A rate $L_{\text{no-EA}}>0$ is achievable if for every $\varepsilon>0$, and sufficiently large n, there exists a $(2^{L_{\text{no-EA}}\cdot n},n,\varepsilon)$ code.

The channel capacity, as defined below, is a central concept in information theory.

Definition 1.2.3. The channel capacity, denoted as $C_{\text{no-EA}}(\mathcal{N}_{A\to B}, \mathsf{M})$, is defined as the supremum of achievable rates.

Remark 1. In communication without covertness, the number of codewords is exponential in the code length n, i.e., $M(n, R) = 2^{nR}$, where R is the communication rate in units of information bits per transmission. In non-standard settings, including covert communication, the code size $\mathsf{M}(n, R)$ is not necessarily exponential. In particular, in some models we have $\mathsf{M}(n, R) = 2^{\sqrt{n}R}$ and $\mathsf{M}(n, R) = 2^{\sqrt{n}\log(n)R}$.

Remark 2. We note that one may also consider the capacity with respect to a different scaling function. In particular, here, the capacity without covertness satisfies

$$C_{\text{no-EA}}(\mathcal{N}_{A\to B}, \mathsf{M}) = +\infty \tag{1.30}$$

in sub-exponential scale $\mathsf{M}(n, R) = o(2^{nR})$, such as $\mathsf{M}(n, R) = 2^{\sqrt{n}R}$, and

$$C_{\text{no-EA}}(\mathcal{N}_{A\to B}, \mathsf{M}) = 0 \tag{1.31}$$

in super-exponential scale $\mathsf{M}(n, R) = \omega(2^{nR})$, such as $\mathsf{M}(n, R) = 2^{n^2 R}$). In general, the optimal scale of the capacity is the one for which $0 < C_{\text{no-EA}}(\mathcal{N}_{A\to B}, \mathsf{M}) < \infty$.

1.2.2 Unassisted Capacity

Recall the definition of the Holevo information in (1.20). The capacity for the transmission of classical information via a channel $\mathcal{N}_{A\to B}$ is characterized by the regularized form of the Holevo information, i.e. (refer to [36, 37]),

$$C_{\text{no-EA}}(\mathcal{N}_{A\to B}, \mathsf{M}) = \lim_{k\to\infty} \frac{1}{k}\chi(\mathcal{N}^{\otimes k}) \tag{1.32}$$

for the exponential scaling function,

$$\mathsf{M}(n, R) = 2^{nR}. \tag{1.33}$$

Remark 3. We note that a single-letter capacity formula, i.e., a formula of the form $C_{\text{no-EA}}(\mathcal{N}, \mathsf{M}) = f(\mathcal{N})$, is an open problem for the communication setting of a quantum channel without covertness and without entanglement assistance [33, Ch. 20.2], [38, Ch. 8.3]. In the special case of a channel with a classical input X, the characterization reduces to $C_{\text{no-EA}}(\mathcal{N}_{X\to B}, \mathsf{M}) = \chi(\mathcal{N})$. In particular, this yields Shannon's celebrated channel coding theorem [39], for classical channels.

1.3 Entanglement Assisted Communication

In practical communication systems, there are often periods when both parties are inactive. These intervals can be utilized to generate shared entanglement resources, which can subsequently enhance communication performance and throughput once transmission continues [27, 28, 29, 30, 31]. In this context, we will discuss a key result within quantum information theory—the entanglement-assisted classical capacity. In this scenario, Alice and Bob aim is to exchange classical information via a quantum channel while utilizing shared entanglement resources. The common assumption in this context is that there is an unlimited supply of shared entanglement resources.

1.3.1 Entanglement-assisted Code

The definition of a code for communication over a quantum channel with entanglement assistance is given below.

Definition 1.3.1. An $(\mathsf{M}(n,R),n)$ entanglement assisted code consists of

- a message set $\{1,\dots,\mathsf{M}(n,R)\}$, where $\mathsf{M}(n,R)$ is an integer,
- a pure entangled state $\Psi_{T_AT_B}$,
- an encoding map,
$$\mathcal{F}^{(m)}_{T_A\to A^n}:\mathscr{S}(\mathcal{H}_{T_A})\to\mathscr{S}(\mathcal{H}_A^{\otimes n}), \tag{1.34}$$
for $m\in[1:\mathsf{M}(n,R)]$,
- a decoding POVM,
$$\mathcal{D}_{B^nT_B}=\{D_m\}_{m=1}^{\mathsf{M}(n,R)} \tag{1.35}$$

We denote the code by $(\Psi,\mathcal{F},\mathcal{D})$.

The error probability is defined as the probability of incorrectly decoding the received message, given by

$$\begin{aligned}P_{\mathrm{e}}^{(n)}(\Psi,\mathcal{F},\mathcal{D})&=\sum_{m=1}^{\mathsf{M}}\frac{1}{\mathsf{M}}\Pr(\widehat{m}\neq m|m) &(1.36)\\ &=\frac{1}{\mathsf{M}}\sum_{m=1}^{\mathsf{M}}\Big(1-\mathrm{tr}\Big[(\mathcal{D}\circ\mathcal{N}_{A\to B}^{\otimes n}\circ\mathcal{F}(m)\Big]\Big)\,, &(1.37)\end{aligned}$$

with $\mathsf{M}\equiv\mathsf{M}(n,R)$ where $\Pr(m\neq\widehat{m})$ is the probability of decoding error for each message m.

An $(\mathsf{M}(n,R),n,\varepsilon)$-code for entanglement assisted communication satisfies

$$P_{\mathrm{e}}^{(n)}(\Psi,\mathcal{F},\mathcal{D}) \leq \varepsilon\,. \tag{1.38}$$

The coding rate for entanglement assisted communication (without a requirement of covertness) is characterized as

$$L_{\mathrm{EA}} \equiv \frac{\log\left(\mathsf{M}(n,R)\right)}{n}\,, \tag{1.39}$$

i.e., the number of bits per channel use. An achievable coding rate is defined as follow:

Definition 1.3.2. A rate $L_{\mathrm{EA}} > 0$ is achievable if for every $\varepsilon > 0$, and sufficiently large n, there exists a $(2^{L_{\mathrm{EA}}\cdot n},n,\varepsilon)$ code.

The channel capacity, is defined below.

Definition 1.3.3. The channel capacity, denoted as $C_{\mathrm{EA}}(\mathcal{N}_{A\to B},\mathsf{M})$, is defined as the supremum of the entanglement assisted achievable rates.

Typically, a codebook can be constructed such that the number of information bits, $\log(\mathsf{M})$, is linear in n, namely, $\log(\mathsf{M}) = O(n)$. However, in the framework of covert communication, $\log(\mathsf{M}(n,R))$ is restricted to $O(\sqrt{n})$, as discussed in the next chapter.

The entanglement-assisted communication setting is depicted in Figure 1.1. Suppose that Alice and Bob share the entangled state $\Psi_{T_AT_B}$, in systems T_A and T_B, respectively. Alice wishes to send one of $\mathsf{M}(n,R)$ equally-likely messages. To encode a message m, she applies the encoding map $\mathcal{F}^{(m)}_{T_A\to A^n}$ to her share T_A of the entanglement resource. This results in a quantum state

$$\rho^{(m)}_{A^nT_B} = (\mathcal{F}^{(m)}_{T_A^n\to A^n}\otimes \mathbb{1}_{T_B})(\Psi_{T_AT_B})\,. \tag{1.40}$$

Alice transmits the part A^n of $\rho^{(m)}_{A^nT_B}$ through n uses of the channel $\mathcal{N}_{A\to B}$. The joint output state is

$$\rho^{(m)}_{B^nT_B} = \left(\mathcal{N}^{\otimes n}_{A\to B}\otimes \mathrm{id}_{T_B}\right)\left(\rho^{(m)}_{A^nT_B}\right)\,. \tag{1.41}$$

Bob decodes the message from the reduced output state $\rho^{(m)}_{B^nT_B} = \mathrm{Tr}\left[\rho^{(m)}_{B^nT_B}\right]$ by applying the POVM $\mathcal{D}_{B^nT_B}$.

Remark 4. In our achievability analysis, we will identify the entanglement resource $\Psi_{T_AT_B}$ with the product state $\psi^{\otimes n}_{A_1A}$, as in (3.11). That is, we use entanglement resources such that Alice and Bob's entangled systems, T_A and T_B, consist of n copies of A and A_1, respectively.

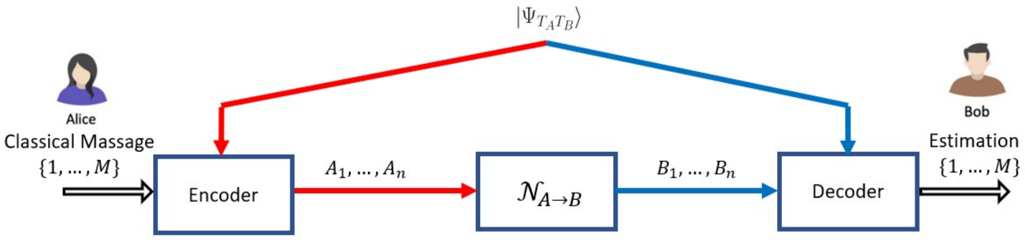

Figure 1.1: Entanglement assisted communication diagram.

1.3.2 Entanglement Assisted Capacity

Let $C_{\text{EA}}(\mathcal{N}_{A\to B}, \mathsf{M})$ be the entanglement assisted classical capacity [33, Ch. 21.3] show that

$$C_{EA}(\mathcal{N}_{A\to B}, \mathsf{M}) = \max_{|\phi_{A_1A}\rangle} I(A_1;B)_\rho\,, \tag{1.42}$$

for the exponential scaling function,

$$\mathsf{M}(n,R) = 2^{nR}\,. \tag{1.43}$$

where the optimization is of pure states $|\psi_{AA_1}\rangle$, with A_1 serving as a reference system, and $\rho_{A_1B} \equiv (\mathbb{1}_{A_1} \otimes \mathcal{N}_{A\to B})(|\phi_{A_1A}\rangle\langle\phi_{A_1A}|)$.

The entanglement-assisted capacity is bounded from above by

$$C_{EA}(\mathcal{N}_{A\to B}, \mathsf{M}) \leq 2 \cdot \log(\dim(\mathcal{H}_A))\,. \tag{1.44}$$

This bound is saturated for a noiseless communication channel. In particular, for a noiseless *qubit* channel, the entanglement-assisted capacity is $C_{EA}() = 2$. Below, we illustrate a well-known protocol that achieves this rate.

1.3.3 Super-Dense Coding

In this section, we describe a well-known protocol, known as *super-dense coding*, for entanglement-assisted communication of 2 classical information bits per transmission via a noiseless qubit channel. That is, superdense coding is capacity achieving for a noiseless qubit channel with entanglement assistance.

Let us denote the four orthogonal Bell states as follows:

$$\left|\Phi^{(i,j)}\right\rangle \equiv (Z^i \otimes X^j)\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \tag{1.45}$$

where Z and X represent the Pauli operators, and $i, j \in 0, 1$.

Suppose that Alice would like to send 2 classical bits to Bob, and Alice and Bob share the Bell state $\left|\Phi^{(0,0)}\right\rangle \equiv \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ (EPR state), and they have a noisless qubit channel.

Now, consider the scenario where Alice intends to transmit 2 classical bits (x_0 and x_1) to Bob, and they share the Bell state $\left|\Phi^{(0,0)}\right\rangle \equiv \frac{1}{\sqrt{2}}(|00\rangle+|11\rangle)$, commonly known as the EPR state. Additionally, they have access to a noiseless qubit channel, as depicted in Fig. 1.2.

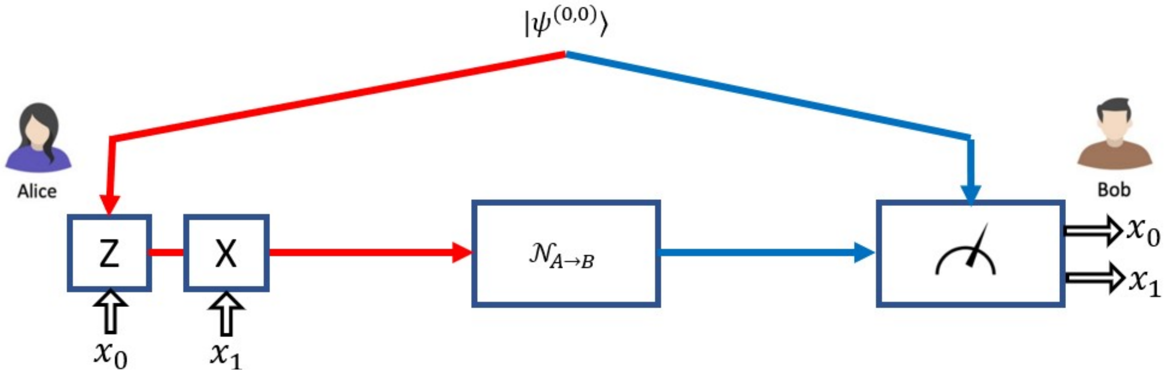

Figure 1.2: Super-dense coding diagram.

For the encoding step, Alice applies the operator $Z^{x_0} \cdot X^{x_1}$ to her qubit. Consequently, the shared entangled resources transform into the Bell state $\Phi^{(x_0,x_1)}$ (as described in [33, 6.1.3]). Subsequently, Alice sends her qubit to Bob, who performs a measurement in the Bell states basis on both Alice's and his qubits to decode x_0 and x_1.

Therefore, through a single utilization of the quantum channel, Alice is able to transmit 2 classical bits to Bob, achieving a data rate that matches the entanglement-assisted capacity of the qubit channel.

1.3.4 Second Order Result

The lemma bellow provides an achievability result for the transmission over a memoryless quantum channel, using entanglement resources, and can be considered as second order result of the achievable number of information bits.

Lemma 1.3.4 (see [23, Lemma 1] [35, 40])**.** *Consider a memoryless quantum channel $\mathcal{N}_{A\to B}$. For every pure entangled state $|\psi_{A_1A}\rangle \in \mathcal{H}_{A_1} \otimes \mathcal{H}_A$, arbitrarily small $\varepsilon > 0$, and sufficiently large n, there exists a coding scheme that employs pre-shared entanglement resources to transmit $\log(\mathsf{M})$ bits over n uses of $\mathcal{N}_{A\to B}$ with decoding error probability ε such that:*

$$\log(\mathsf{M}) \geq nD(\psi_{A_1B}||\psi_{A_1} \otimes \psi_B) + \sqrt{nV(\psi_{A_1B}||\psi_{A_1} \otimes \psi_B)}\Phi^{-1}(\varepsilon) - C_n \tag{1.46}$$

with

$$\psi_{A_1B} = (\mathrm{id}_{A_1} \otimes \mathcal{N}_{A\to B})(\psi_{A_1A}) \tag{1.47}$$

and

$$C_n = \frac{\beta_{\text{B-E}}}{\sqrt{2\pi}} \frac{[Q(\psi_{A_1B}||\psi_{A_1} \otimes \psi_B)]^{\frac{3}{4}}}{V(\psi_{A_1B}||\psi_{A_1} \otimes \psi_B)} + \frac{V(\psi_{A_1B}||\psi_{A_1} \otimes \psi_B)}{\sqrt{2\pi}} + \log(4\varepsilon n) \tag{1.48}$$

where $D(\cdot||\cdot)$ is the quantum relative entropy, $V(\cdot||\cdot)$, $Q(\cdot||\cdot)$ are the second and fourth moments in (1.10)-(1.11), $\beta_{B\text{-}E}$ is the Berry-Esseen constant satisfying $0.40973 \leq \beta_{B\text{-}E} \leq 0.4784$, and $\Phi^{-1}(x)$ is the inverse-Gaussian distribution function.

The derivation is based on a *position-based* coding scheme, where Alice associates each message with n entangled systems ψ_{A_1A}, and sends it over n channel uses. Bob receives the enatngled state ψ_{A_1B}, and uses *sequential decoding* for each message consecutively [40, 23]. Its analysis in [35, Sec. 5] prove that,

$$\log(\mathsf{M}) \geq D_H^{\varepsilon-\zeta}\left((\psi_{A_1B})^{\otimes n} \,||\, (\psi_{A_1})^{\otimes n} \otimes (\psi_B)^{\otimes n}\right) - \log\left(4\varepsilon/\zeta^2\right), \tag{1.49}$$

for $\zeta \in (0, \varepsilon)$, where $D_H^{\varepsilon-\zeta}(\rho||\sigma)$ is the hypothesis testing relative entropy (1.22). the derivation of the lemma is completed by using the expansion of the hypothesis testing relative entropy (1.23), and setting $\zeta = 1/\sqrt{n}$.

1.4 Energy Constraints

Energy constraints in communication and information theory refer to limitations on the amount of energy that can be expended or transmitted in a communication system. A state ρ satisfies the energy constraint E, with the Hamiltonian $\hat{\mathsf{H}}$, if

$$\mathrm{tr}(\hat{\mathsf{H}}\rho) \leq E\,. \tag{1.50}$$

In information theory, the concept of channel capacity under energy constraints involves determining the maximum rate at which information can be reliably transmitted over a communication channel, with a limitation on the energy in the system. We denote the capacity of the channel $\mathcal{N}$, with the energy constraint E as $C(\mathcal{N}, E, \mathsf{M})$.

Consider communication over a finite-dimensional channel under an energy constraint, E . Then, the capacities with and without entanglement assistance, are given by [41]

$$C_{\text{no-EA}}(\mathcal{N}, E, \mathsf{M}) = \max_{\rho_{XA}:\mathrm{tr}(\hat{\mathsf{H}}\rho_A)\leq E} I(X;B)_\rho \tag{1.51}$$

$$C_{\text{EA}}(\mathcal{N}, E, \mathsf{M}) = \max_{|\psi_{A_1A}\rangle:\mathrm{tr}(\hat{\mathsf{H}}\psi_A)\leq E} I(A_1;B)_\omega \tag{1.52}$$

with the observable (Hamiltonian) $\hat{\mathsf{H}} = |1\rangle\langle 1|$, where

$$\rho_{XA} = \sum_{x \in \mathcal{X}} p_X(x)\, |x\rangle\langle x| \otimes \phi_A^{(x)}\,, \tag{1.53}$$

$$\rho_{XB} = (\mathrm{id}_X \otimes \mathcal{N}_{A\to B})(\rho_{XA})\,, \tag{1.54}$$

and

$$\omega_{A_1 B} = (\mathrm{id}_X \otimes \mathcal{N}_{A\to B})(\psi_{A_1 A}) \tag{1.55}$$

The maximization in (1.51) is over all the input classical-quantum states ρ_{XA} such that the reduced average state $\rho_A \equiv \mathrm{tr}_X(\rho_{XA})$ satisfies the energy constraint $\mathrm{tr}(\hat{\mathsf{H}}\rho_A) \leq E$. Similarly, the maximization in (1.52) is over all the entangled input states $|\psi_{A_1 A}\rangle$ with a reduced state ψ_A such that $\mathrm{tr}(\hat{\mathsf{H}}\psi_A) \leq E$. The scaling function for both assisted and unassisted cases is $\mathsf{M}(n, R) = 2^{nR}$.

Remark 5. When considering unassisted communication, the capacity under energy constraints can be interpreted as the Holevo information (1.20) of the channel, where the maximization is restricted to states that comply with the specified energy constraint.

1.5 Depolarizing Channel

The depolarizing channel is a natural model for noise in quantum systems [27, 42, 43]. The qubit depolarizing channel with parameter q transmits the input qubit perfectly with probability $1 - q$, and outputs a completely mixed state with probability q. Consider a qubit depolarizing channel from Alice to Bob expressed as:

$$\begin{aligned} \mathcal{N}_{A\to B}(\rho_A) &= (1-q)\rho_A + q\frac{\mathbb{1}}{2} \\ &= \left(1 - \frac{3q}{4}\right)\rho_A + \frac{q}{4}\left(X\rho_A X + Y\rho_A Y + Z\rho_A Z\right), \end{aligned} \tag{1.56}$$

with $0 < q < 1$, dimensions $\dim(\mathcal{H}_A) = \dim(\mathcal{H}_B) = 2$, where X, Y, and Z are the Pauli operators, and the second equality follows from the Pauli twirl identity [33, Ch. 4.7.4].

1.5.1 Unassisted Capacity

The unassisted capacity of the qubit depolarizing channel with a parameter $0 < q < 1$ is given by (see [33, Ch. 20.4.4]),

$$C_{\text{no-EA}}(\mathcal{N}_{A\to B}, \mathsf{M}) = 1 - h_2(\frac{q}{2})\,, \tag{1.57}$$

where the binary entropy $h_2(\cdot)$ is defined as

$$h_2(q) \equiv -q\log(q) - (1-q)\log(1-q)\,. \tag{1.58}$$

Achievability is shown by employing the symmetric ensemble:

$$\left\{p_X = \left(\frac{1}{2}, \frac{1}{2}\right), \phi^{(0)} = |0\rangle\langle 0| \ , \ \phi^{(1)} = |1\rangle\langle 1|\right\} \tag{1.59}$$

1.5.2 Entanglement Assisted Classical Capacity

The enatnglement assisted classical capacity of the qubit depolarizing channel with a parameter $0 < q < 1$ is given by,

$$C_{\mathrm{EA}}(\mathcal{N}_{A\to B}, \mathsf{M}) = 2 - H\Big(1 - \frac{3q}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\Big) , \tag{1.60}$$

where $H(\mathbf{P}) = -\sum_i p_i \log(p_i)$ represents the classical Shannon entropy corresponding to the probability vector $\mathbf{P}$.

Achievability is shown by employing the EPR state (see Bell states in (1.45)):

$$\left|\Phi^{(0,0)}\right\rangle \equiv \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) , \tag{1.61}$$

which results in the output state,

$$\begin{aligned} \omega_{A_1 B} &\equiv (\mathrm{id} \otimes \mathcal{N}_{A\to B})(\Phi^{(0,0)}) \\ &= (1 - \frac{3q}{4})\Phi^{(0,0)} + \frac{1}{4}\Phi^{(1,0)} + \frac{1}{4}\Phi^{(0,1)} + \frac{1}{4}\Phi^{(1,1)} . \end{aligned} \tag{1.62}$$

with $\Phi^{(i,j)} \equiv \left|\Phi^{(i,j)}\right\rangle\!\left\langle\Phi^{(i,j)}\right|$. The reduced states are then $\omega_{A_1} = \omega_B = \frac{\mathbb{1}}{2}$. Thus,

$$\begin{aligned} I(A_1; B)_\omega &= H(A_1) + H(B) - H(A_1 B) \\ &= 1 + 1 - H\Big(1 - \frac{3q}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\Big) \\ &= 2 - H\Big(1 - \frac{3q}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\Big) \end{aligned} \tag{1.63}$$

1.6 Bosonic Channel

The bosonic channel represents a fundamental component in the realm of continuous-variable communication. Within this framework, channels are characterized by an infinite-dimensional Hilbert space, facilitating the transmission of quantum information encoded in continuous variables. Here we focus on the single-mode bosonic channel, where the system comprises a single mode of a quantum harmonic oscillator.

The qunatum harmonic oscillator is often used to describe a photon beam. The Hamiltonian

$$\hat{H}_{\mathrm{HO}} = \hat{a}^\dagger \hat{a} , \tag{1.64}$$

is associated with the number of photons, where $\hat{a}$ is the annihilation operator which holds the canonical commutation relation:

$$[\hat{a}, \hat{a}^\dagger] = \mathbb{1} \tag{1.65}$$

(see [44, Ch. II]).

A quantum Gaussian state is a state proportional to the exponential of a quadratic polynomial in $\hat{a}$ and $\hat{a}^\dagger$. A thermal Gaussian state $\rho_{\bar{n}}$, is a Gaussian state with a polynomial that proportional to the Hamiltonian $\hat{H}_{\mathrm{HO}}$:

$$\rho_{\bar{n}} = \frac{1}{(\bar{n}-1)} \left(\frac{\bar{n}}{\bar{n}+1} \right)^{\hat{H}_{\mathrm{HO}}}, \tag{1.66}$$

where $\bar{n}$ is the mean photon numbr.

The single-mode bosonic channel is characterized by a beam splitter with a transmissivity parameter η and annihilation operators $\hat{a}$, $\hat{b}$, $\hat{e}_{\bar{n}}$, $\hat{w}$, as in Fig. 1.3. Here, $\hat{a}$ represents the mode of the transmission sent by Alice, $\hat{b}$ represents the output received by Bob, $\hat{e}_{\bar{n}}$ signifies noise originating from the environment, manifested as a thermal Gaussian mode with a mean photon number $\bar{n}$, and $\hat{w}$ denotes the output leaked to the environment.

The beam splitter is represented by the unitary operator,

$$\hat{U}_\eta = \exp\left(\left(\hat{a}^\dagger \hat{e}_{\bar{n}} - \hat{e}_{\bar{n}}^\dagger \hat{a} \right) \arccos(\sqrt{\eta}) \right), \tag{1.67}$$

which establishes the relationship between the inputs and outputs,

$$\hat{b} = \sqrt{\eta}\hat{a} + \sqrt{1-\eta}\hat{e}_{\bar{n}} \tag{1.68}$$

$$\hat{w} = -\sqrt{1-\eta}\hat{a} + \sqrt{\eta}\hat{e}_{\bar{n}} \tag{1.69}$$

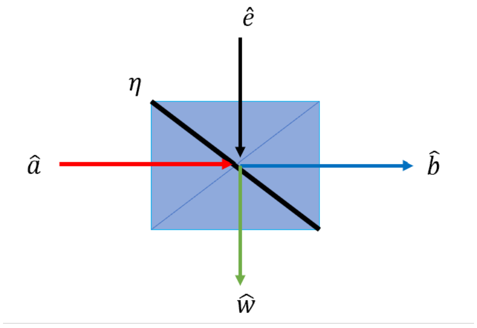

Figure 1.3: Beam splitter diagram.

The entropy of a thermal Gaussian state $\rho_{\bar{n}}$ is expressed as:

$$g(\bar{n}) = (1+\bar{n})\log(1+\bar{n}) - \bar{n}\log(\bar{n}),. \tag{1.70}$$

The unassisted capacity of a bosonic channel with a transmissivity parameter η is given by:

$$C_{\text{no-EA}}(\mathcal{N}_{Bos}, \mathsf{M}) = g(\eta\bar{N}_A + (1-\eta)\bar{N}_E) - g((1-\eta)\bar{N}_E), , \tag{1.71}$$

where $\bar{N}_A$ and $\bar{N}_E$ represent the mean photon number of Alice and the environment, respectively.

Chapter 2

Covert Communication

2.1 Reliability and Covertness

In covert communication, Alice decides whether to transmit information to Bob or not, and Willie, the warden, has the task of discerning whether Alice and Bob communicate through the channel (see Figure 2.1). To this end, he performs a binary measurement $\{\Delta_{\mathsf{H0}}, \Delta_{\mathsf{H1}}\}$, where the outcome $\mathsf{H1}$ represents the hypothesis that Alice sent information, while $\mathsf{H0}$ indicates the contrary hypothesis.

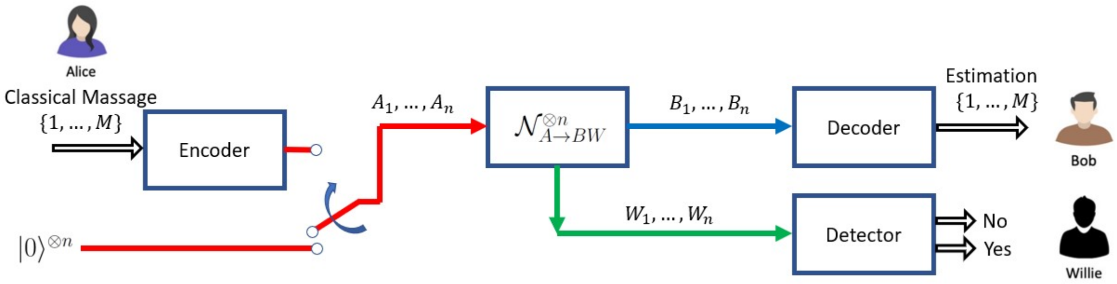

Figure 2.1: Covert communication diagram (without assistance).

Willie fails by either accusing Alice of transmitting when she is not (false alarm), or missing Alice's transmission (missed detection). Denoting the probabilities of these errors by $P_{\text{FA}} = P(\text{choose } \mathsf{H1}|\mathsf{H0} \text{ is true})$ and $P_{\text{MD}} = P(\text{choose } \mathsf{H0}|\mathsf{H1} \text{ is true})$, respectively, and assuming equally likely hypotheses, Willie's average probability of error is $E^{(n)} = \frac{P_{\text{FA}}+P_{\text{MD}}}{2}$. A random choice yields an ineffective detector with $E^{(n)} = \frac{1}{2}$. The goal of covert communication is to design a sequence of codes such that Willie's detector is forced to be arbitrarily close to ineffective. Let us denote the average state that Willie receives by $\overline{\rho}_{W^n}$, and Willie's state corresponding to innocent input by $\omega_0^{\otimes n}$ (both for n channel uses).

As a result of hypothesis tests [45] and with the adaptation to the quantum case

[18, 19], Willie's optimal tests satisfies

$$P_{\mathrm{FA}} + P_{\mathrm{MD}} \geq 1 - \sqrt{D(\overline{\rho}_{W^n} || \omega_0^{\otimes n})} \, . \tag{2.1}$$

Therefore, a sufficient condition to render *any* detector ineffective for Willie is

$$D(\overline{\rho}_{W^n} || \omega_0^{\otimes n}) \approx 0 \, . \tag{2.2}$$

Covert communication is characterized by two key elements:

1. *Covertness*, measured by $D(\overline{\rho}_{W^n} || \omega_0^{\otimes n})$, which can be viewed as the deviation of Willie's output from the ”no communication” state, $\omega_0^{\otimes n}$. Thereby, a smaller value indicates higher covertness.

2. *Reliability*, represented by the decoding error probability of Bob.

A communication is covert and reliable if both *covertness* and *reliability* approach 0 as the number of channel uses n tends to infinity.

2.2 Unassisted Covert Communication

2.2.1 Unassisted Covert Code

In the context of covert communication, the code mappings, $\mathcal{F}$ and $\mathcal{D}$, are defined in the same manner as for standard communication without covertness, as described in Definition 1.2.1. The error probability $P_e^{(n)}(\mathcal{F}, \mathcal{D})$ is determined as described in (1.27).

We denote the average state that Willie receives by

$$\overline{\rho}_{W^n} = \frac{1}{M} \sum_{m=1}^{M} \rho_{W^n}^{(m)} \tag{2.3}$$

where $\rho_{W^n}^{(m)}$ is the reduced state of the joint output $\rho_{B^n W^n}^{(m)}$. Then, *covertness* is measured through the relative entropy, $D(\overline{\rho}_{W^n} || \omega_0^{\otimes n})$, where $\omega_0 \equiv \mathcal{N}_{A \to W}(|0\rangle\langle 0|)$ is the output corresponding to the innocent input.

Formally, an $(M, n, \varepsilon, \delta)$-code for unassisted covert communication satisfies

$$P_e^{(n)}(\mathcal{F}, \mathcal{D}) \leq \varepsilon \tag{2.4}$$

and

$$D(\overline{\rho}_{W^n} || \omega_0^{\otimes n}) \leq \delta \, . \tag{2.5}$$

2.2.2 Unassisted Covert Capacity

In traditional communication problems, the coding rate is defined as $R = \frac{\mathsf{log(M)}}{n}$, i.e., the number of bits per channel use. In covert communication, however, the best achievable rate is zero, since the number of information bits is sublinear in n. [18, 19] prove that unassisted communication allows reliable transmission of $\log(M) = O(\sqrt{n})$ covert bits. Hence, the covert coding rate is characterized as in [19]:

$$L_{\text{no-EA}} = \frac{\log(\mathsf{M})}{\sqrt{\delta n}}. \tag{2.6}$$

where δ is the covertness level in (2.5).

Definition 2.2.1. A covert rate $L_{\text{no-EA}} > 0$ is achievable without assistance if for every $\varepsilon, \delta > 0$, and sufficiently large n, there exists a $(2^{L_{\text{no-EA}} \cdot \sqrt{\delta n}}, n, \varepsilon, \delta)$ code.

Definition 2.2.2. The unassisted covert capacity is defined as the supremum of achievable covert rates. We denote this capacity by $C_{\text{cov-no-EA}}(\mathcal{N})$, where the subscript stands for covert communication without assistance.

2.2.3 Square-Root-Law (SRL)

As mentioned earlier, in the context of unassisted communication, it is possible to transmit approximately $O(\sqrt{n})$ covert bits over n channel uses. This is commonly referred to as the square-root-law (SRL).

Proving the achievability of the SRLs involves the following principles. In the finite-dimensional case, both classical and quantum [16, 17, 18, 19, 20], a symbol (say, 0) in the input alphabet is designated as "innocent." Random coding is employed such that a non-innocent symbol is transmitted with probability $\sim {}^{1}/_{\sqrt{n}}$ to ensure covertness.

We can offer an intuitive explanation as follows (see [46, Ch. 1.1.2]). Let us use a random codebook, where each codeword is generated using an i.i.d. distribution. For simplicity, we will focus the example of a qubit channel, where each codeword is chosen with a Bernoulli(α) distribution. This codeword can be represented by the input state,

$$\varphi_\alpha \equiv (1-\alpha)\left|0\right\rangle\!\left\langle 0\right| + \alpha \left|1\right\rangle\!\left\langle 1\right| . \tag{2.7}$$

This leads to the outcome where Willie's average state $\overline{\rho}_{W^n}$ is equivalent to $\omega_\alpha^{\otimes n}$. Now, if α remains independent of n, the condition for covert communication becomes

$$\begin{aligned} D(\overline{\rho}_{W^n}||\omega_0^{\otimes n}) &= D(\omega_\alpha^{\otimes n}||\omega_0^{\otimes n}) \\ &= nD(\omega_\alpha||\omega_0) \end{aligned} \tag{2.8}$$

which does not tend towards zero as n approaches infinity. Hence, we require α to be dependent on n.

Let $\gamma_n = o(1) \cap \omega\left(\frac{\log n}{n^{1/6}}\right)$, that is, as $n \to \infty, \gamma_n \to 0$ and $\frac{n^{1/6}}{\log n} \cdot \gamma_n \to +\infty$. Bash *et. al.* [18, 19] show that by choosing $\alpha = \alpha_n$, where

$$\alpha_n \equiv \frac{\gamma_n}{\sqrt{n}} \ , \tag{2.9}$$

we get the upper bound

$$n\alpha_n^2\zeta \geq D(\overline{\rho}_{W^n}||\omega_0^{\otimes n})\,, \tag{2.10}$$

where $\zeta \geq 0$ is a constant. In this case, taking n to infinity ensure covertness.

Ref. [19] shows that in the case of unassisted covert communication over finite-dimension channels, the covert capacity is given by,

$$C_{\text{cov-no-EA}}(\mathcal{N}) = \frac{D(\rho_1||\rho_0)}{\sqrt{\frac{1}{2}\eta(\omega_1||\omega_0)}}\,, \tag{2.11}$$

where $\rho_x = \mathcal{N}_{\mathcal{A}\to\mathcal{B}}(|x\rangle\langle x|)$, $\omega_x = \mathcal{N}_{\mathcal{A}\to\mathcal{W}}(|x\rangle\langle x|)$, and $\eta(\cdot||\cdot)$ is defined in Equation (1.14).

For the case of continuous-variable covert communication, innocent symbol corresponding to zero transmitted power occurs naturally in communication over classical additive white Gaussian noise (AWGN) [15, 16, 17] and classical-quantum bosonic [21, 22, 23, 24] channels. Maintaining average transmitted power $O(1/\sqrt{n})$ correspondingly measured in Watts and in the emitted photon number ensures covertness.

2.3 Entanglement Assisted Covert Communication

2.3.1 Entanglement Assisted Covert Code

In the context of entanglement assisted covert communication, the code $(\Psi, \mathcal{F}, \mathcal{D})$ is defined in the same manner as for entanglement assisted communication without covertness, as described in Definition 1.3.1. The error probability $P_e^{(n)}(\Psi, \mathcal{F}, \mathcal{D})$ is determined as described in (1.27).

We denote the average state that Willie receives by

$$\overline{\rho}_{W^n} = \frac{1}{M}\sum_{m=1}^{M}\rho_{W^n}^{(m)} \tag{2.12}$$

where $\rho_{W^n}^{(m)}$ is the reduced state of the joint output $\rho_{B^nW^nT_B}^{(m)}$. Then, *covertness* is measured through the relative entroy, $D(\overline{\rho}_{W^n}||\omega_0^{\otimes n})$, where $\omega_0 \equiv \mathcal{N}_{A\to W}(|0\rangle\langle 0|)$ is the output corresponding to innocent input. Formally, an $(M, n, \varepsilon, \delta)$-code for entanglement-assisted covert communication satisfies

$$P_e^{(n)}(\Psi, \mathcal{F}, \mathcal{D}) \leq \varepsilon \tag{2.13}$$

and

$$D(\overline{\rho}_{W^n}||\omega_0^{\otimes n}) \leq \delta\,. \tag{2.14}$$

2.4 Entanglement Assisted Covert Capacity

As previously discussed, the number of information bits in covert communication, is sublinear in n, while specifically, without assistance it is $O(\sqrt{n})$. Here we prove that entanglement assistance allows reliable transmission of $\log(M) = O(\sqrt{n}\log n)$ covert bits. Hence, the covert coding rate is characterized as in [23]:

$$L_{\text{EA}} = \frac{\log(\mathsf{M})}{\sqrt{\delta n}\log n}. \tag{2.15}$$

where δ is the covertness level in (2.14).

Definition 2.4.1. A covert rate $L_{\text{EA}} > 0$ is achievable with entanglement assistance if for every $\varepsilon, \delta > 0$, and sufficiently large n, there exists a $(2^{L\cdot\sqrt{\delta n}\log n}, n, \varepsilon, \delta)$ code.

Remark 6. Achievable rates correspond to error and covertness levels that tend to zero in the limit of $n \to \infty$. That is, one may rewrite Definition 2.4.1 as follows [18]. A rate L_{EA} is asymptotically achievable if there exists a sequence of codes such that

$$\frac{\log(\mathsf{M})}{\log n\sqrt{nD(\overline{\rho}_{W^n}||\omega_0^{\otimes n})}} \geq L_{\text{EA}} - \zeta_n \quad \forall n \geq n_0 \tag{2.16}$$

for some $n_0 > 0$ and sequence ζ_n that tends to zero as $n \to \infty$, while the error probability satisfies

$$\lim_{n\to\infty} P_e^{(n)}(\Psi, \mathcal{F}, \mathcal{D}) = 0\,, \tag{2.17}$$

and the covertness,

$$\lim_{n\to\infty} D(\overline{\rho}_{W^n}||\omega_0^{\otimes n}) = 0\,. \tag{2.18}$$

Definition 2.4.2. The entanglement-assisted covert capacity is defined as the supremum of achievable covert rates. We denote this capacity by $C_{\text{cov-EA}}(\mathcal{N})$, where the subscript stands for covert communication with entanglement assistance.

Consider the following state, with $\alpha \in [0, 1]$:

$$\varphi_\alpha \equiv (1-\alpha)\,|0\rangle\!\langle 0| + \alpha\,|1\rangle\!\langle 1|\ . \tag{2.19}$$

Let $\gamma_n = o(1) \cap \omega\left(\frac{\log n}{n^{1/6}}\right)$, that is, as $n \to \infty$,$\gamma_n \to 0$ and $\frac{n^{1/6}}{\log n} \cdot \gamma_n \to +\infty$. Choosing $\alpha = \alpha_n$ where

$$\alpha_n \equiv \frac{\gamma_n}{\sqrt{n}} \tag{2.20}$$

ensures covertness [18, 19]. That is, if the average state of the input system A^n is given by $\rho_{A^n} = (\varphi_{\alpha_n})^{\otimes n}$, then the covertness requirement (2.14) is satisfied for large n.

2.5 Entanglement Assisted Covert Capacity of Bosonic Channels

Gagatsos et al. [23] showed that entanglement assistance allows transmission of $O(\sqrt{n}\log n)$ reliable and covert bits over n uses of continuous-variable bosonic channel, surpassing the SRL scaling. This scaling is also referred to as the square-root-log law. As in the unassisted setting, the transmission is limited to $O({}^1\!/\!\sqrt{n})$ mean photon number. However, so far it has remained open whether such a performance boost can be achieved in communication over finite-dimensional quantum channels.

The entanglement-assisted covert capacity of the bosonic channel $\mathcal{N}_{\text{Bos}}$ according to [23] is,

$$C_{\text{cov-EA}}(\mathcal{N}_{\text{Bos}}) = \frac{\sqrt{2\eta\bar{N}_E(1+\eta\bar{N}_E)}}{1-\eta} \cdot \frac{\eta}{2(1+(1-\eta)\bar{N}_E)}\,, \tag{2.21}$$

where $\bar{N}_E$ is the mean photon number of the environment noise.

The result of [23] is based on Lemma 1.3.4, which provides an achievability result for the transmission over a memoryless quantum channel, regardless of covertness.

Chapter 3

Entanglement-Assisted Covert Communication

3.1 Depolarizing Channel With a Warden

Consider a covert communication quantum channel $\mathcal{N}_{A\to BW}$, which maps a quantum input state ρ_A to a joint output state ρ_{BW}. The systems A, B, and W are associated with the transmitter, the legitimate receiver, and an adversarial warden, referred to as Alice, Bob, and Willie. The marginal channels $\mathcal{N}_{A\to B}$ and $\mathcal{N}_{A\to W}$, from Alice to Bob, and from Alice to Willie, respectively, satisfy $\mathcal{N}_{A\to B}(\rho_A) = \mathrm{Tr}_W\left(\mathcal{N}_{A\to BW}(\rho_A)\right)$ and $\mathcal{N}_{A\to W}(\rho_A) = \mathrm{Tr}_B\left(\mathcal{N}_{A\to BW}(\rho_A)\right)$ for $\rho_A \in \mathscr{S}(\mathcal{H}_A)$. Our channel is memoryless: for ρ_{A^n} occupying input systems $A^n = (A_1, \ldots, A_n)$, the joint output state is $\mathcal{N}^{\otimes n}_{A\to BW}(\rho_{A^n})$.

Here, we investigate covert communication over a depolarizing channel $\mathcal{V}_{A\to BE_1E_2}$ given by the Stinespring dilation:

$$\mathcal{V}_{A\to BE_1E_2}(\rho_A) = V\rho_A V^\dagger, \tag{3.1}$$

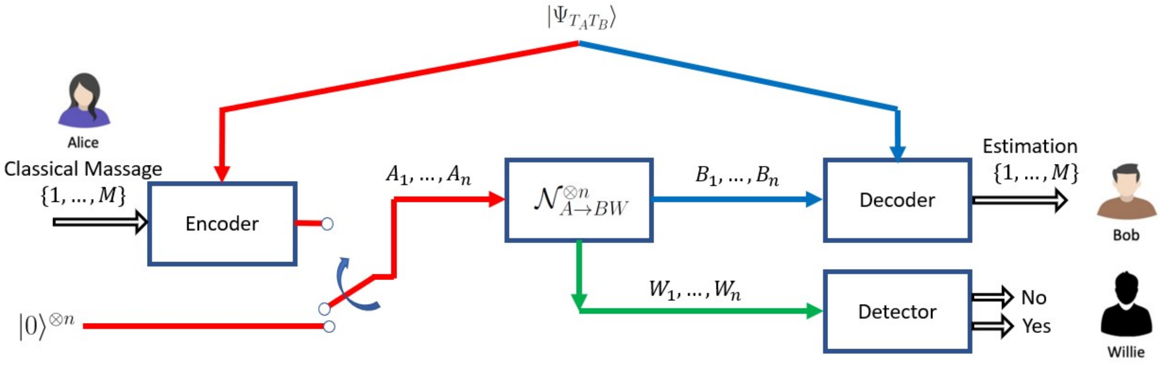

Figure 3.1: Entanglement-assisted coding for covert communication over a quantum channel $\mathcal{N}_{A\to BW}$.

where $V : \mathcal{H}_A \to \mathcal{H}_B \otimes \mathcal{H}_{E_1} \otimes \mathcal{H}_{E_2}$ is an isometry defined by

$$V \equiv \sqrt{1-\frac{3q}{4}}\mathbb{1} \otimes |00\rangle + \sqrt{\frac{q}{4}}X \otimes |01\rangle + \sqrt{\frac{q}{4}}Y \otimes |11\rangle + \sqrt{\frac{q}{4}}Z \otimes |10\rangle \ . \tag{3.2}$$

Remark 7. The canonical Stinespring dilation for the qubit depolarizing channel is defined by $\widetilde{\mathcal{V}}_{A\to BE}(\rho) = \widetilde{V}\rho\widetilde{V}^\dagger$, where $\widetilde{V} \equiv \sqrt{1-\frac{3q}{4}}\mathbb{1} \otimes |0\rangle + \sqrt{\frac{q}{4}}X \otimes |1\rangle + \sqrt{\frac{q}{4}}Y \otimes |2\rangle + \sqrt{\frac{q}{4}}Z \otimes |3\rangle$ (see [43, Eq. (13)]). As we identify $E \equiv (E_1, E_2)$, our definition in (3.2) is equivalent to this canonical description. Note, however, that any other Stinespring representation is equivalent to (3.2) up to an isometry on the environment E [47, Sec. III-B].

We consider three cases:

- Scenario 1: Willie receives (E_1, E_2)
- Scenario 2: Willie receives E_2
- Scenario 3: Willie receives E_1

Remark 8. In any depolarizing channel model, Scenario 1 represents the worst-case scenario where Willie is given access to Bob's entire environment, $E = (E_1, E_2)$. This is the maximum amount of information that Willie can acquire in the quantum setting. It is important to note that the no-cloning theorem applies in the quantum setting and prohibits a channel model where Willie is given a copy of Bob's output state, whereas in the classical setting, Willie could have a copy of Bob's output. Hence, the quantum channel from Alice to Willie is *not* a depolarizing channel.

Remark 9. In the boundary case of $q = 0$, Bob receives the qubit state as is, while Willie obtains no information, in agreement with the no-cloning theorem. Conversely, if $q = 1$, Willie receives the qubit state, and Bob gets only noise. As one may expect, covert communication is trivial in the former case, and impossible in the latter.

Remark 10. We assume without loss of generality that the innocent state is represented by $|0\rangle$. However, it is important to note that this choice is arbitrary. Since the depolarizing channel is symmetric with respect to the input state, our findings can easily be extended to any product state $|\psi_{\text{idle}}\rangle^{\otimes n}$ that corresponds to an idle transmission system.

We present our results on Entanglement-Assisted Covert Communication. We address the three scenarios presented above.

3.2 Willie Receives (E_1, E_2)

We begin with the case where Willie receives the entire environment, i.e., both E_1 and E_2 (see Fig. 3.2). This can be viewed as the worst-case scenario (see Remark 8).

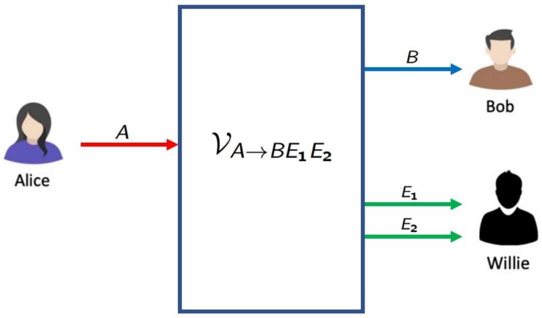

Figure 3.2: Scenario 1: Willie receives (E_1, E_2).

Theorem 3.1. *Covert communication is impossible in Scenario 1. Hence, if* $W = (E_1, E_2)$*, then* $C_{\textit{cov-EA}}(\mathcal{N}, \mathsf{M}) = 0$*.*

Proof of Theorem 3.1. Let ω_0 and ω_1 denote Willie's output states corresponding to the inputs $|0\rangle$ and $|1\rangle$, respectively. That is $\omega_x \equiv \mathcal{N}_{A\to W}(|x\rangle\langle x|)$ for $x \in \{0, 1\}$.

For the given scenario where Willie receives the entire environment, it is possible to demonstrate that (see [43]),

$$\omega_0 = \begin{pmatrix} 1-\frac{3q}{4} & 0 & 0 & \sqrt{\frac{q}{4}(1-\frac{3q}{4})} \\ 0 & \frac{q}{4} & -i\frac{q}{4} & 0 \\ 0 & i\frac{q}{4} & \frac{q}{4} & 0 \\ \sqrt{\frac{q}{4}(1-\frac{3q}{4})} & 0 & 0 & \frac{q}{4} \end{pmatrix}, \tag{3.3}$$

and

$$\omega_1 = \begin{pmatrix} 1-\frac{3q}{4} & 0 & 0 & -\sqrt{\frac{q}{4}(1-\frac{3q}{4})} \\ 0 & \frac{q}{4} & i\frac{q}{4} & 0 \\ 0 & -i\frac{q}{4} & \frac{q}{4} & 0 \\ -\sqrt{\frac{q}{4}(1-\frac{3q}{4})} & 0 & 0 & \frac{q}{4} \end{pmatrix}. \tag{3.4}$$

The null spaces of ω_0 and ω_1 contain vectors,

$$|e_0\rangle \equiv \begin{pmatrix} 0 \\ i \\ 1 \\ 0 \end{pmatrix}, \tag{3.5}$$

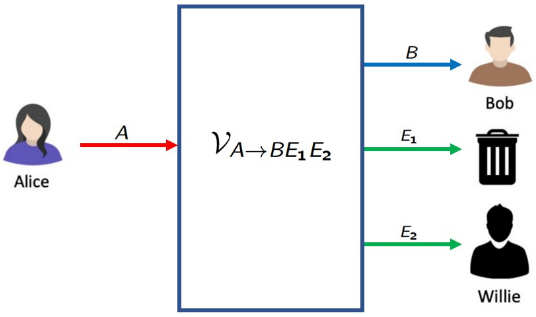

Figure 3.3: Scenario 2: Willie receives E_2.

and

$$|e_1\rangle \equiv \begin{pmatrix} 0 \\ -i \\ 1 \\ 0 \end{pmatrix}, \tag{3.6}$$

respectively. Since $\langle e_0|e_1\rangle = 0$, it follows that $\text{supp}(\omega_1) \not\subseteq \text{supp}(\omega_0)$.

Therefore, Willie can perform a measurement to detect a non-zero transmission with certainty. ■

Essentially, in Scenario 1, Willie's entanglement with the transmitted qubit is strong enough for him to detect any encoding operation.

3.3 Willie Receives E_2

Next, we consider another extreme setting (see Fig. 3.3).

Theorem 3.2. *Covert communication is trivial in Scenario 2. That is, if $W = E_2$, then Alice can communicate information as without the covertness requirement, and send $O(n)$ bits.*

Proof of Theorem 3.2. Suppose Alice transmits the general state $\rho = (1-a)|0\rangle\langle 0| + a|1\rangle\langle 1| + b|0\rangle\langle 1| + b^*|1\rangle\langle 0|$. Then, Willie receives the state,

$$\mathcal{N}_{A\to W}(\rho) = \left(1-\frac{q}{2}\right)|0\rangle\langle 0| + \frac{q}{2}|1\rangle\langle 1| \\ + 2\,\text{Re}\{b\}\left(\left(\sqrt{\left(1-\frac{3q}{4}\right)\frac{q}{4}} + i\frac{q}{4}\right)|0\rangle\langle 1| + \left(\sqrt{\left(1-\frac{3q}{4}\right)\frac{q}{4}} - i\frac{q}{4}\right)|1\rangle\langle 0|\right). \tag{3.7}$$

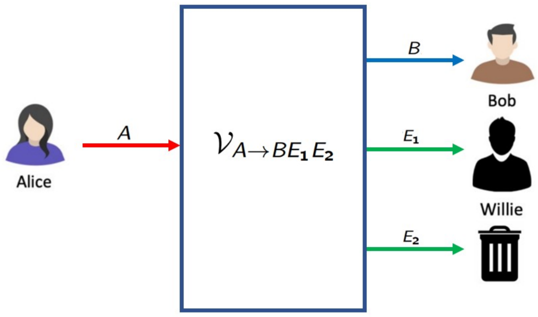

Figure 3.4: Scenario 3: Willie receives E_1.

Substituting $\rho = |0\rangle\langle 0|$ and $\rho = |1\rangle\langle 1|$ into (3.7), respectively, yields:

$$\begin{aligned}\omega_0 &= \mathcal{N}_{A\to W}(|0\rangle\langle 0|)\\ &= \left(1-\frac{q}{2}\right)|0\rangle\langle 0| + \frac{q}{2}|1\rangle\langle 1| \,, \end{aligned} \tag{3.8}$$

and

$$\begin{aligned}\omega_1 &= \mathcal{N}_{A\to W}(|1\rangle\langle 1|)\\ &= \left(1-\frac{q}{2}\right)|0\rangle\langle 0| + \frac{q}{2}|1\rangle\langle 1| \,. \end{aligned} \tag{3.9}$$

Therefor, if $W = E_2$, then Willie receives $\omega_0 = \omega_1 = (1-\frac{q}{2})\,|0\rangle\langle 0| + \frac{q}{2}\,|1\rangle\langle 1|$. In this scenario, even without entanglement assistance, Alice can transmit classical codewords as in the standard non-covert model, while Willie cannot discern between zero and non-zero inputs. ■

3.4 Willie Receives E_1

We proceed to our main result on the entanglement-assisted covert capacity $C_{\text{cov-EA}}$ of the depolarizing channel. From this point on, we focus on Scenario 3, where Willie receives the first qubit of the environment (see Section 3.1).

Theorem 3.3. *Consider a qubit depolarizing channel $\mathcal{N}_{A\to BW}$ as specified in Section 3.1 above, where $W = E_1$. The entanglement-assisted covert capacity is bounded as*

$$C_{cov\text{-}EA}(\mathcal{N}, \mathsf{M}) \geq \frac{4\sqrt{2}}{3}\frac{(1-q)^2}{(2-q)\sqrt{\eta(\omega_1||\omega_0)}} \tag{3.10}$$

where $\omega_0 \equiv \mathcal{N}_{A\to W}(|0\rangle\langle 0|)$ and $\omega_1 \equiv \mathcal{N}_{A\to W}(|1\rangle\langle 1|)$.

Note that $\eta(\omega_1||\omega_0)$ is defined in (1.14). Our lower bound is depicted in Figure 3.5. As can be seen in the figure, our lower bound has the expected behavior for the covert capacity in the boundary points (see Remark 9). For $q = 0$, we have $C_{\text{cov-EA}}(\mathcal{N}) = +\infty$ in the $\sqrt{n}\log n$ scale, because the warden only receives noise and Alice can transmit a linear number of information bits (effectively, there is no warden). Whereas, for $q = 1$, the covert and non-covert capacities are zero.

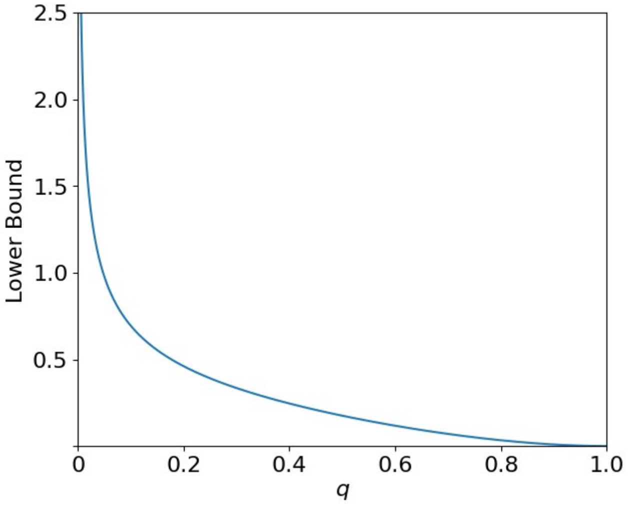

Figure 3.5: The lower bound on the entanglement-assisted covert capacity of Scenario 3 in Theorem 3.3, as a function of the noise parameter q.

Following the definitions in Section 2.4, a bound of the form $C_{\text{cov-EA}} \geq L_0$ implies that it is possible to transmit $L_0\sqrt{\delta n}\log n$ information bits reliably and covertly (see Definitions 2.4.1 and 2.4.2). Recall that without entanglement assistance, covert communication requirements limit the message to $O(\sqrt{n})$ information bits [18, 19]. Thereby, we have established that entanglement assistance increases the message scale in covert communication, from $O(\sqrt{n})$ to $O(\sqrt{n}\log n)$ information bits. A similar result has been shown for continuous-variable bosonic channels by Gagatsos et al. [23]. To the best of our knowledge, our result in Theorem 3.3, on the depolarizing channel, is the first demonstration of such a property for a finite-dimensional channel.

Remark 11. In covert communication with entanglement assistance, the scale of $\sqrt{n}\log n$ has already been observed in a continuous-variable model, i.e., the bosonic channel [23]. However, until now, it has remained unclear whether this performance boost can also be achieved in finite dimensions. As we explain below, the answer is far from obvious.

In some communication settings, the coding scale is larger for continuous-variable channels. For example, in deterministic identification, the code size is super-exponential and scales as $2^{n\log nR}$ for Gaussian channels [48] and Poisson channels [49]. On the other hand, deterministic identification is limited to an exponential scale for finite-dimensional channels [50].

Nevertheless, we show here that in covert communication over a qubit depolarizing channel, entanglement assistance can increase the number of information bits from $O(\sqrt{n})$ to $O(\sqrt{n}\log n)$, as in the bosonic case. In other words, the $\log n$ performance boost is not reserved to continuous variable systems.

3.5 Proof of Main Theorem

3.5.1 Proof Idea

We begin with the proof idea. Consider Scenario 3 presented in Section 3.1. First, we identify an entangled state that meets the above condition for covertness. As opposed to previous work, we do *not* encode a random bit sequence with $\sim {}^{1}\!/\!_{\sqrt{n}}$ frequency (or probability) of 1's. Instead, we employ "weakly" entangled states of the form

$$|\psi_{A_1A}\rangle = \sqrt{1-\alpha}\,|00\rangle + \sqrt{\alpha}\,|11\rangle\,, \tag{3.11}$$

such that the squared amplitude of this quantum superposition of states describing innocent and non-innocent symbols is $\alpha = O\left({}^{1}\!/\!_{\sqrt{n}}\right)$. In order to gurantee covertness, the probability amplitude must be such that the state of the transmission is very close to that of a sequence of innocent states $|0\rangle^{\otimes n}$.

Furthermore, we modify the approach in [23] to analyze the order of the number of covert information bits using Lemma 1.3.4

3.5.2 Analysis

In this section, we give the proof for Theorem 3.3. We present the main stages of the proof, while the technical details are deferred to the appendix. We begin with the following lemma.

Lemma 3.5.1. *Let $\gamma_n = n^{\nu-\frac{1}{6}}$, where $0 < \nu < \frac{1}{6}$ is arbitrary and does not depend on n. Then, there exists an entanglement-assisted covert coding scheme for qubit depolarizing channel with blocklength n, size M, and average error probability ε that satisfies*

$$\log(\mathsf{M}) \geq 2\left(\frac{2}{3}-\nu\right)\frac{(1-q)^2}{2-q}\gamma_n\sqrt{n}\log n + O(\sqrt{n}\gamma_n)\,. \tag{3.12}$$

Proof. To prove the lemma, we need to show that, for arbitrarily small $\varepsilon, \delta > 0$ and large n, there exists an $(\mathsf{M}, n, \varepsilon, \delta)$ code for the depolarizing channel with entanglement assistance, with a code size M as in (3.12). To this end, we apply Lemma 1.3.4 with $|\psi_{A_1A}\rangle$ as in (3.11), with a parameter $\alpha = \alpha_n$ as in (2.9). Note that setting $\gamma_n = n^{\nu-\frac{1}{6}}$ as in the lemma statement yields

$$\alpha_n = \frac{\gamma_n}{\sqrt{n}} = n^{\nu-\frac{2}{3}}. \tag{3.13}$$

Intuitively, as the value of α_n is small, the input state that Alice sends through the channel is close to the innocent state, i.e., $\psi_A \approx |0\rangle\langle 0|$. Given the joint state $\psi_{A_1A} \equiv |\psi_{A_1A}\rangle\langle\psi_{A_1A}|$, the channel input A is in the reduced state $\psi_A \equiv \mathrm{Tr}_{A_1}\left[|\psi_{A_1A}\rangle\langle\psi_{A_1A}|\right] =$

φ_{α_n}, with φ_{α_n} as in (2.19). That is, the reduced input state fits the achievability proof for the covert capacity without entanglement assistance in [18, 19], i.e., without entanglement assistance. Based on the analysis therein, this input state meets the covertness requirement. As the covertness requirement does not involve the entanglement resources, it follows that covertness holds here as well, i.e., $D(\overline{\rho}_{W^n}||\omega_0^{\otimes n})$ tends to zero as $n \to \infty$.

Having established both reliability and covertness, it remains to estimate the code size. To this end, consider the joint state ψ_{A_1B} of the output system B and the reference system A_1, as in (1.47). In order to estimate each term on the right-hand side of (1.46), we first derive expressions for the operator logarithms, $\log(\psi_{A_1B})$ and $\log(\psi_{A_1} \otimes \psi_B)$, and then we approximate the relative entropy $D(\psi_{A_1B}||\psi_{A_1} \otimes \psi_B)$, and its second and fourth moments $V(\psi_{A_1B}||\psi_{A_1} \otimes \psi_B)$ and $Q(\psi_{A_1B}||\psi_{A_1} \otimes \psi_B)$.

The full technical details are given in the appendices. In Appendix A, we analyze the spectral decompositions, and then use the Taylor expansions near $\alpha = 0$. Throughout the derivation, we maintain the exact value of the dominant terms and reduce the approximation error to its order class, following the asymptotic notation in Section 1.1.3. In Appendix B, we estimate the quantum relative entropy and its moments, and show that

$$\begin{aligned} D(\psi_{A_1B}||\psi_{A_1} \otimes \psi_B) &= -2\frac{(1-q)^2}{2-q}\alpha_n \log(\alpha_n) + O(\alpha_n)\,, \\ V(\psi_{A_1B}||\psi_{A_1} \otimes \psi_B) &= O(\alpha_n \log^2(\alpha_n))\,, \\ Q(\psi_{A_1B}||\psi_{A_1} \otimes \psi_B) &= O(\alpha_n \log^2(\alpha_n))\,, \end{aligned} \tag{3.14}$$

for $\alpha = \alpha_n$ as chosen above (see (3.13)).

The proof is concluded by placing the approximations above into (1.46) , as detailed bellow.

We observe that when choosing $\alpha = \alpha_n = \frac{\gamma_n}{\sqrt{n}}$ with $\gamma_n = n^{\nu - \frac{1}{6}}$, we obtain the following approximation for the first term on the right-hand side of (1.46),

$$\begin{aligned} nD(\psi_{A_1B}||\psi_{A_1} \otimes \psi_B) &= -2\frac{(1-q)^2}{2-q}\sqrt{n}\gamma_n \log\left(\frac{\gamma_n}{\sqrt{n}}\right) + O(\sqrt{n}\gamma_n) \\ &= -2\frac{(1-q)^2}{2-q} n^{\nu+\frac{1}{3}} \log\left(n^{\nu-\frac{2}{3}}\right) + O(n^{\nu+\frac{1}{3}}) \\ &= 2\left(\frac{2}{3} - \nu\right)\frac{(1-q)^2}{2-q} n^{\nu+\frac{1}{3}} \log n + O(n^{\nu+\frac{1}{3}})\,. \end{aligned} \tag{3.15}$$

In a similar manner, we approximate the second term by

$$\begin{aligned}\sqrt{nV(\psi_{A_1B}||\psi_{A_1}\otimes\psi_B)} &= O\left(\sqrt{\sqrt{n}\gamma_n\log^2(n^{\nu-\frac{2}{3}})}\right)\\ &= O\left(\sqrt{n^{\nu+\frac{1}{3}}\log n}\right)\\ &= O\left(n^{\frac{\nu}{2}+\frac{1}{6}}\log n\right)\,. \qquad (3.16)\end{aligned}$$

Finally, the last term (1.46) is C_n, as defined in (1.48). To show that this term vanishes, we write

$$\begin{aligned}\frac{[Q(\psi_{A_1B}||\psi_{A_1}\otimes\psi_B)]^{\frac{3}{4}}}{V(\psi_{A_1B}||\psi_{A_1}\otimes\psi_B)} &= O\left(\frac{\left(n^{\frac{\nu}{2}+\frac{1}{6}}\log n\right)^{\frac{3}{4}}}{n^{\frac{\nu}{2}+\frac{1}{6}}\log n}\right)\\ &= O\left(n^{-\frac{\nu}{8}-\frac{1}{6}}\log^{-\frac{1}{4}}(n)\right) \qquad (3.17)\end{aligned}$$

which tends to zero as $n\to\infty$. Hence,

$$\log(M)\geq 2\left(\frac{2}{3}-\nu\right)\frac{(1-q)^2}{2-q}n^{\nu+\frac{1}{3}}\log n+O(n^{\nu+\frac{1}{3}}) \qquad (3.18)$$

for every $0<\nu<\frac{1}{6}$.

We are now ready for the proof of Theorem 3.3.

Proof of Theorem 3.3. First, we observe that in this scenario, $\text{supp}(\omega_1)\subseteq\text{supp}(\omega_0)$ and, in addition, $\omega_0\neq\omega_1$ (see derivation in Appendix 3.4), therefore, covert communication is possible, and not trivial. Then, even if Willie's output state is ω_1, there is still ambiguity whether the input is innocent or not.

By Lemma 3.5.1, we have established achievability for the following covert rate:

$$L_n=\frac{2\left(\frac{2}{3}-\nu\right)\frac{(1-q)^2}{2-q}\gamma_n+O\left(\frac{\gamma_n}{\log n}\right)}{\sqrt{D(\overline{\rho}_{W^n}||\omega_0^{\otimes n})}}\,. \qquad (3.19)$$

As mentioned above, in the proof of Lemma 3.5.1, covertness is guaranteed due to the fact that the reduced input state in our coding scheme with entanglement assistance is the same as the average input state in the previous coding scheme without assistance [18, 19]. Furthermore, the following property extends as well: there exists $\zeta>0$ such that,

$$|D(\overline{\rho}_{W^n}||\omega_0^{\otimes n})-nD(\omega_{\alpha_n}||\omega_0)|\leq e^{-\zeta\gamma_n^{\frac{3}{2}}n^{\frac{1}{4}}}\,, \qquad (3.20)$$

where $\overline{\rho}_{W^n}$ is the actual state of Willie's system as defined in (2.12), the state $\omega_0=\mathcal{N}_{A\to W}(|0\rangle\langle 0|)$ is the Willie's output corresponding to the innocent input, and $\omega_{\alpha_n}\equiv\mathcal{N}_{A\to W}(\varphi_{\alpha_n})$, with φ_{α_n} as in (2.19) (see achievability proof in [18], [19, Theorem 1]).

Based on a result that was recently developed for covert sensing using entangled states [26, Lemma 5],

$$D(\omega_{\alpha_n}||\omega_0) = \frac{\alpha_n^2}{2}\eta(\omega_1||\omega_0) + O(\alpha_n^3) \tag{3.21}$$

for sufficiently small α_n. Thus, by (3.20) and (3.21),

$$D(\overline{\rho}_{W^n}||\omega_0^{\otimes n}) \leq \frac{\gamma_n^2}{2}\eta(\omega_1||\omega_0) + e^{-\zeta\gamma_n^{\frac{3}{2}}n^{\frac{1}{4}}} + O\left(\frac{\gamma_n^3}{\sqrt{n}}\right). \tag{3.22}$$

By applying this bound to the denominator in (3.19), we have:

$$L_n \geq \frac{2\left(\frac{2}{3}-\nu\right)\frac{(1-q)^2}{2-q}\gamma_n + O\left(\frac{\gamma_n}{\log n}\right)}{\sqrt{\frac{\gamma_n^2}{2}\eta(\omega_1||\omega_0) + e^{-\zeta\gamma_n^{\frac{3}{2}}n^{\frac{1}{4}}} + O\left(\frac{\gamma_n^3}{\sqrt{n}}\right)}}. \tag{3.23}$$

Hence, in the limit of $n \to \infty$, we achieve

$$L \geq \frac{2\left(\frac{2}{3}-\nu\right)\frac{(1-q)^2}{2-q}}{\sqrt{\frac{1}{2}\eta(\omega_1||\omega_0)}} \tag{3.24}$$

for arbitrarily small $\nu > 0$, which completes the proof. ■

3.6 Interpretations

3.6.1 Energy Constraint Interpretation

We provide interpretation for the logarithmic advantage. In the bosonic case, the ratio between the entanglement-assisted capacity and the unassisted capacity, follows a logarithmic trend of $\log(1/E)$, where E is the limit on the transmission mean photon number [41, 51]. Yet, to ensure covertness, the mean photon number must be restricted to $E_n = O(\frac{1}{\sqrt{n}})$. Consequently, an $O(\log n)$ factor arises [52]. Based on our derivation, a similar phenomenon is observed for the qubit depolarizing channel.

Indeed, consider communication over a finite-dimensional channel under an energy constraint, E, without the covertness constraint [41, Sec. 2]. Then, the capacities with and without entanglement assistance, are given by [41] (see Section 1.4),

$$C_0(\mathcal{N}, E) = \max_{\{p_x(x), \phi_A^{(x)}\}:\text{tr}(\mathsf{F}\rho_A)\leq E} I(X;B)_\rho \tag{3.25}$$

$$C_{\text{EA}}(\mathcal{N}, E) = \max_{|\psi_{A_1 A}\rangle:\text{tr}(\mathsf{F}\psi_A)\leq E} I(A_1;B)_\omega \tag{3.26}$$

Let $0 < E \leq \frac{1}{2}$. The capacity without entanglement assistance is given by

$$C_0(\mathcal{N}) = h_2\left(E * \frac{q}{2}\right) - h_2(\frac{q}{2}) \tag{3.27}$$

where '$*$' denotes the binary convolution operation: $\alpha * \beta = (1-\alpha)\beta + \alpha(1-\beta)$.

The direct part follows by choosing the ensemble $\{(1-E, E), |0\rangle, |1\rangle\}$, which results in the average input state $\widetilde{\rho}_A \equiv (1-E)|0\rangle\langle 0| + E|1\rangle\langle 1|$. To see the converse part, consider a general state $\rho_A \equiv (1-p)|0\rangle\langle 0| + p|1\rangle\langle 1| + b|0\rangle\langle 1| + b^*|1\rangle\langle 0|$ which satisfies the maximization in (1.51), and let us define $\widetilde{\rho}_A \equiv (1-E)|0\rangle\langle 0| + E|1\rangle\langle 1|$. For the unassisted case, we have,

$$\begin{aligned} C_0(\mathcal{N}) &= H(B)_\rho - H(B|X)_\rho \\ &\leq H(B)_\rho - H^{\min}(\mathcal{N}) \end{aligned} \tag{3.28}$$

where $H^{\min}(\mathcal{N})$ is the minimum output entropy, $H^{\min}(\mathcal{N}) \equiv \min_{\rho_A} H(\mathcal{N}(\rho_A))$. By [33, Sec. 20.4.4], $H^{\min}(\mathcal{N}) = h_2\left(\frac{q}{2}\right)$. Based on the derivation in Appendix C, the first entropy term is maximal for the input state $\widetilde{\rho}_A \equiv (1-E)|0\rangle\langle 0| + E|1\rangle\langle 1|$. Hence, $C_0(\mathcal{N}, E) \leq h_2(E * \frac{q}{2}) - h_2(\frac{q}{2})$. Together, the direct and converse parts imply (3.27).

Given entanglement assistance, we can restrict our attention to input states of the form $|\psi_{A_1A}\rangle = \sqrt{1-a}|00\rangle + \sqrt{a}|11\rangle$, since the depolarizing channel is unitarily covariant (see [33, Section 24.8]). In particular, the maximum is attained by the entangled state $|\psi_{A_1A}\rangle = \sqrt{1-E}|00\rangle + \sqrt{E}|11\rangle$, which is associated with an energy value $\mathrm{tr}(\mathsf{F}\psi_A) = E$. This results in the energy-constrained entanglement-assisted capacity formula,

$$C_{\mathrm{EA}}(\mathcal{N}, E) = h_2(E) + h_2(E * \frac{q}{2}) - H(\psi_{A_1B}) \tag{3.29}$$

where $\psi_{A_1B} \equiv (\mathrm{id} \otimes \mathcal{N})(|\psi_{A_1A}\rangle\langle\psi_{A_1A}|)$. Now, based on the derivations in Lemma 3.5.1, for $E \to 0$, we have

$$\begin{aligned} \frac{C_{\mathrm{EA}}(\mathcal{N}, E)}{C_0(\mathcal{N}, E)} &= \frac{h_2(E) + h_2(E * \frac{q}{2}) - H(\psi_{A_1B})}{h_2(E * \frac{q}{2}) - h_2(\frac{q}{2})} \\ &\sim \frac{-E\log(E)}{E} \\ &= -\log(E), \end{aligned} \tag{3.30}$$

by taking $\alpha = E$. To satisfy the covert constraint, we effectively impose an energy constraint $E_n \sim \frac{1}{\sqrt{n}}$, which results in the following ratio between the entanglement-assisted and unassisted covert capacities,

$$\frac{C_{\text{EA-cov}}(\mathcal{N})}{C_{\text{0-cov}}(\mathcal{N})} \sim \log n\,. \tag{3.31}$$

3.6.2 Decoder Performance Interpretation

The square root law for the unassisted cases (both classical and quantum) were derived for the non-trivial scenario, in which Bob cannot determine with certainty if Alice sends a non-innocent symbol. However, if Bob has this capability, i.e., $\text{supp}(\mathcal{N}_{A\to B}(|1\rangle\langle 1|)) \nsubseteq \text{supp}(\mathcal{N}_{A\to B}(|0\rangle\langle 0|))$, then the scaling law becomes $O(\sqrt{n}\log n)$, even for a classical channel [18, 19, 16]. The depolarizing channel is fair in this sense. Yet, entanglement assistance has a similar effect as granting Bob the capability of identifying a non-innocent transmission with certainty.

The quantum erasure channel is another fundamental model in Quantum Shannon Theory, [53] where for an input state ρ, Bob receives the original state with probability $1-q$, or an erasure state $|e\rangle\langle e|$, which is orthogonal to the qubit space, with probability q. For this channel, Bob can determine that Alice sent $|1\rangle$ with certainty, as $\text{supp}(\mathcal{N}_{A\to B}(|1\rangle\langle 1|)) \nsubseteq \text{supp}(\mathcal{N}_{A\to B}(|0\rangle\langle 0|))$. Thereby, the scaling law becomes $O(\sqrt{n}\log n)$ information bits, even without entanglement resources.

Chapter 4

Conclusion and Future Direction

We have studied covert communication through the qubit depolarizing channel, where Alice and Bob share entanglement resources and wish to communicate, while an adversarial warden, Willie, is trying to detect their communication. We addressed three scenarios. In the first scenario, Willie can determine with certainty whether Alice has transmitted a non-innocent state, making covert communication impossible. In the second, Willie cannot distinguish between the $|0\rangle$ and $|1\rangle$ inputs, making covert communication effortless. The outcomes of our study mainly pertain to the third scenario, wherein covert communication is both feasible and non-trivial. Our results show that it is possible to transmit $O(\sqrt{n}\log n)$ bits reliably and covertly. This result surpasses the maximum scaling of $O(\sqrt{n})$ reliable and covert bits in both the classical and quantum cases without entanglement assistance.

A future interesting direction is to consider a more general model, where the covert communication channel is formed by a concatenation of the depolarizing channel $\mathcal{V}_{A\to BE}$ with a general channel $\mathcal{P}_{E\to W}$ to Willie, namely, $\mathcal{N}_{A\to BW} = (\mathrm{id}_B \otimes \mathcal{P}_{E\to W}) \circ \mathcal{V}_{A\to BE}$. Furthermore, an important direction for future research involves deriving lower bounds of entanglement assisted covert capacity for a general channel. Additionally, Here we investigate the utilization of quantum resources for transmitting classical information. Future research could delve into covert communication techniques for transmission of quantum information, including covert entanglement distribution.

An additional potential future research direction involves exploring the upper bound of the entanglement-assisted covert capacity of the qubit depolarizing channel, as our study currently addresses only the lower bound, which may not be optimal. Yet, the coding technique we utilize achieves the optimal covert capacity observed in scenarios without entanglement assistance, as demonstrated in [18, 19].

Our results can be viewed as a step forward towards understanding covert communication via general quantum channels in the presence of pre-shared entanglement resources. Following the past literature, the preliminary results on entanglement-assisted communication via the depolarizing and erasure channels [27] have led to a complete characterization for a general quantum channel [28]. We can only hope to see the same progress in the study of covert communication.

Appendix A

Matrix Logarithms Estimation

A.1 Approximation Tools

We provide the approximation tools that will be used throughout the derivation, using the "big $\mathcal{O}$-notation" in Section 1.1.3.

- Useful Taylor expansions (at $x = 0$):

$$\sqrt{a + b \cdot x + c \cdot x^2} = \sqrt{a} + \frac{b}{2\sqrt{a}} x + \mathcal{O}(x^2)\,, \tag{A.1}$$

$$\log\Big(a + b \cdot x + c \cdot x^2\Big) = \frac{\ln(a)}{\ln(2)} + \frac{b}{a\ln(2)} x - \frac{b^2 - 2ac}{a^2 \ln(4)} x^2 + \mathcal{O}(x^3)\,, \tag{A.2}$$

$$\sqrt{x(1-x)} = \sqrt{x} + \mathcal{O}(x^{\frac{3}{2}})\,, \tag{A.3}$$

$$\frac{x}{\sqrt{x(1-x)}} = \sqrt{x} + \mathcal{O}(x^{\frac{3}{2}})\,, \tag{A.4}$$

$$\frac{1}{\sqrt{x(1-x)}} = \frac{1}{\sqrt{x}} + \mathcal{O}(\sqrt{x})\,, \tag{A.5}$$

$$\frac{1}{\sqrt{c \cdot x + 1}} = 1 + \mathcal{O}(x)\,, \tag{A.6}$$

$$\frac{1}{\sqrt{\frac{c}{x} + 1}} = \frac{1}{\sqrt{c}} \sqrt{x} + \mathcal{O}(x^{\frac{3}{2}})\,. \tag{A.7}$$

- The spectral decomposition of an Hermitian operator,

$$P = a\,|00\rangle\langle 00| + b\,|01\rangle\langle 01| + c\,|10\rangle\langle 10| + d\,|11\rangle\langle 11| + s(|00\rangle\langle 11| + |11\rangle\langle 00|) \tag{A.8}$$

consists of the eigenvalues

$$\begin{aligned}
\lambda_1 &= \frac{1}{2}\left(a + d + \sqrt{(a+d)^2 - 4(ad - s^2)}\right),\\
\lambda_4 &= \frac{1}{2}\left(a + d - \sqrt{(a+d)^2 - 4(ad - s^2)}\right),\\
\lambda_2 &= b,\\
\lambda_3 &= c\,. \qquad\qquad (A.9)
\end{aligned}$$

and the associated eigenvectors,

$$\begin{aligned}
|\lambda_1\rangle &= C_1\left(\widetilde{\lambda}_1\,|00\rangle + |11\rangle\right),\\
|\lambda_4\rangle &= C_4\left(\widetilde{\lambda}_4\,|00\rangle + |11\rangle\right),\\
|\lambda_2\rangle &= |01\rangle\,,\\
|\lambda_3\rangle &= |10\rangle\ . \qquad\qquad (A.10)
\end{aligned}$$

where

$$\begin{aligned}
\widetilde{\lambda}_1 &\equiv -\frac{a - \lambda_1}{s}\,, && (A.11)\\
\widetilde{\lambda}_4 &\equiv -\frac{a - \lambda_4}{s}\,, && (A.12)\\
C_1 &\equiv \frac{1}{\sqrt{\widetilde{\lambda}_1^2 + 1}}\,, && (A.13)\\
C_4 &\equiv \frac{1}{\sqrt{\widetilde{\lambda}_4^2 + 1}}\,. && (A.14)
\end{aligned}$$

A.2 Output density operators

The joint state ψ_{A_1B} of the reference system and Bob's output is obtained by applying the depolarizing channel:

$$\begin{aligned}
\psi_{A_1B} &= (\mathbb{1}_{A_1} \otimes \mathcal{N}_{A\to B})(\psi_{A_1A})\\
&= \left(1 - \frac{3}{4}q\right)\psi_{A_1A} + \frac{q}{4}[(\mathbb{1}_{A_1} \otimes X)\psi_{A_1A}(\mathbb{1}_{A_1} \otimes X) + (\mathbb{1}_{A_1} \otimes Y)\psi_{A_1A}(\mathbb{1}_{A_1} \otimes Y)\\
&\quad + (\mathbb{1}_{A_1} \otimes Z)\psi_{A_1A}(\mathbb{1}_{A_1} \otimes Z)]\,. \qquad\qquad (A.15)
\end{aligned}$$

Algebraic manipulations yield

$$\begin{aligned}
\psi_{A_1B} = &\left(1 - \frac{q}{2}\right)(1-\alpha)\,|00\rangle\!\langle 00| + \left(1 - \frac{q}{2}\right)\alpha\,|11\rangle\!\langle 11| + \frac{q}{2}(1-\alpha)\,|01\rangle\!\langle 01| + \frac{q}{2}\alpha\,|10\rangle\!\langle 10|\\
&+ (1-q)\sqrt{\alpha}\sqrt{1-\alpha}(|00\rangle\!\langle 11| + |11\rangle\!\langle 00|)\,. \qquad\qquad (A.16)
\end{aligned}$$

The reduced matrices ψ_{A_1} and ψ_B are, thus,

$$\psi_B = \left[\left(1-\frac{q}{2}\right)*\alpha\right]|0\rangle\langle 0| + \left[\frac{q}{2}*\alpha\right]|1\rangle\langle 1|\ , \tag{A.17}$$

$$\psi_{A_1} = (1-\alpha)\,|0\rangle\langle 0| + \alpha\,|1\rangle\langle 1| \tag{A.18}$$

where $\alpha * \beta = (1-\alpha)\beta + \alpha(1-\beta)$. Then,

$$\begin{aligned}\psi_{A_1}\otimes\psi_B = (1-\alpha)&\left[\left(1-\frac{q}{2}\right)*\alpha\right]|00\rangle\langle 00| + (1-\alpha)\left[\frac{q}{2}*\alpha\right]|01\rangle\langle 01|\\ &+\alpha\left[\left(1-\frac{q}{2}\right)*\alpha\right]|10\rangle\langle 10| + \alpha\left[\frac{q}{2}*\alpha\right]|11\rangle\langle 11|\ .\end{aligned} \tag{A.19}$$

The logarithm of $\psi_{A_1}\otimes\psi_B$ can be computed directly as it is diagonal in the standard basis. This is not the case for ψ_{A_1B}. Using (A.9), the spectral decomposition consists of the following eigenvalues:

$$\begin{aligned}\lambda_1 &= \frac{1}{2}\left(1-\frac{q}{2}+\sqrt{\left[1-\frac{q}{2}\right]^2 - 4q\left[1-\frac{3q}{4}\right]\alpha + 4q\left[1-\frac{3q}{4}\right]\alpha^2}\right),\\ \lambda_4 &= \frac{1}{2}\left(1-\frac{q}{2}-\sqrt{\left[1-\frac{q}{2}\right]^2 - 4q\left[1-\frac{3q}{4}\right]\alpha + 4q\left[1-\frac{3q}{4}\right]\alpha^2}\right),\\ \lambda_2 &= \frac{q}{2}(1-\alpha),\\ \lambda_3 &= \frac{q}{2}\alpha\,.\end{aligned} \tag{A.20}$$

Using the Taylor approximation in (A.1), we approximate λ_1 and λ_4 by

$$\begin{aligned}\lambda_1 &= \frac{1}{2}\left(1-\frac{q}{2}+\left(1-\frac{q}{2}\right)-\frac{4q\left(1-\frac{3}{4}q\right)}{2\left(1-\frac{q}{2}\right)}\alpha+\mathcal{O}(\alpha^2)\right)\ ,\\ \lambda_4 &= \frac{1}{2}\left(1-\frac{q}{2}-\left(1-\frac{q}{2}\right)-\frac{4q\left(1-\frac{3}{4}q\right)}{2\left(1-\frac{q}{2}\right)}\alpha+\mathcal{O}(\alpha^2)\right)\ .\end{aligned} \tag{A.21}$$

That is,

$$\lambda_1 = 1-\frac{q}{2}-\frac{q\left(1-\frac{3}{4}q\right)}{\left(1-\frac{q}{2}\right)}\alpha+\mathcal{O}(\alpha^2)\,, \tag{A.22}$$

$$\lambda_4 = \frac{q\left(1-\frac{3}{4}q\right)}{\left(1-\frac{q}{2}\right)}\alpha+\mathcal{O}(\alpha^2)\,. \tag{A.23}$$

The eigenvectors of ψ_{A_1B} are given in (A.10), with $\widetilde{\lambda}_1$ and $\widetilde{\lambda}_4$ satisfying

$$\widetilde{\lambda}_1 = \frac{q-2}{2(q-1)\sqrt{\alpha}} + \mathcal{O}(\sqrt{\alpha})\,, \tag{A.24}$$

$$\widetilde{\lambda}_4 = -\frac{2(q-1)}{(q-2)}\sqrt{\alpha} + \mathcal{O}(\sqrt{\alpha}^{\frac{3}{2}})\,, \tag{A.25}$$

by (A.5). and

$$C_1^2 = 4\frac{(q-1)^2}{(q-2)^2}\alpha + \mathcal{O}(\alpha^2)\,, \tag{A.26}$$

$$C_4^2 = 1 + \mathcal{O}(\alpha)\,, \tag{A.27}$$

by (A.7).

By applying (A.2), we approximate the logarithm of the eigenvalues as follows. For the joint state ψ_{A_1B},

$$\log(\lambda_1) = \log\left(1 - \frac{q}{2}\right) + \mathcal{O}(\alpha)\,, \tag{A.28}$$

$$\log(\lambda_2) = \log\left(\frac{q}{2}\right) + \mathcal{O}(\alpha)\,, \tag{A.29}$$

$$\log(\lambda_3) = \log\left(\frac{q}{2}\right) + \log(\alpha) + \mathcal{O}(\alpha^2)\,, \tag{A.30}$$

$$\log(\lambda_4) = \log(C(q)) + \log(\alpha) + \mathcal{O}(\alpha^2)\,. \tag{A.31}$$

As for the product state $\psi_{A_1} \otimes \psi_B$, we have

$$\log\Bigg((1-\alpha)\left[\left(1-\frac{q}{2}\right) * \alpha\right]\Bigg) = \log\left(1 - \frac{q}{2}\right) + \mathcal{O}(\alpha)\,, \tag{A.32}$$

$$\log\Bigg((1-\alpha)\left[\left(\frac{q}{2}\right) * \alpha\right]\Bigg) = \log\left(\frac{q}{2}\right) + \mathcal{O}(\alpha)\,, \tag{A.33}$$

$$\log\Bigg((\alpha)\left[\left(1-\frac{q}{2}\right) * \alpha\right]\Bigg) = \log\left(1 - \frac{q}{2}\right) + \log(\alpha) + \mathcal{O}(\alpha)\,, \tag{A.34}$$

$$\log\Bigg((\alpha)\left[\left(\frac{q}{2}\right) * \alpha\right]\Bigg) = \log\left(\frac{q}{2}\right) + \log(\alpha) \tag{A.35}$$

Hence, the operator-logarithm for ψ_{A_1B} satisfies

$$\begin{aligned}\log(\psi_{A_1B}) &= \log(\lambda_1)\,|\lambda_1\rangle\langle\lambda_1| + \log(\lambda_2)\,|\lambda_2\rangle\langle\lambda_2| + \log(\lambda_3)\,|\lambda_3\rangle\langle\lambda_3| + \log(\lambda_3 4)\,|\lambda_4\rangle\langle\lambda_4| \\ &= \Bigg[\left(1-\frac{q}{2}\right) + \frac{4(q-1)^2}{(q-2)^2}\alpha\log(\alpha)\ +\mathcal{O}(\sqrt{\alpha})]\ |00\rangle\langle 00| \\ &\quad + \left[\log\left(\frac{q}{2}\right) + \mathcal{O}(\alpha)\right] |01\rangle\langle 01| \\ &\quad + \left(\log\left(\frac{q}{2}\right) + \log(\alpha) + \mathcal{O}(\alpha^2)\right) |10\rangle\langle 10| \\ &\quad + [\log(C(q)) + \log(\alpha) + \mathcal{O}(\alpha\log(\alpha))]\ |11\rangle\langle 11| \\ &\quad + \left[-\frac{2(q-1)}{(q-2)}\sqrt{\alpha}\log(\alpha) + \mathcal{O}(\sqrt{\alpha})\right] |00\rangle\langle 11| \\ &\quad + \left[-\frac{2(q-1)}{(q-2)}\sqrt{\alpha}\log(\alpha) + \mathcal{O}(\sqrt{\alpha})\right] |11\rangle\langle 00|\ , \end{aligned} \tag{A.36}$$

and for $\psi_{A_1} \otimes \psi_B$,

$$\begin{aligned}\log(\psi_{A_1} \otimes \psi_B) &= \left[\log\left(1-\frac{q}{2}\right) + \mathcal{O}(\alpha)\right] |00\rangle\langle 00| \\ &\quad + \left[\log\left(\frac{q}{2}\right) + \mathcal{O}(\alpha)\right] |01\rangle\langle 01| \\ &\quad + \left[\log\left(1-\frac{q}{2}\right) + \log(\alpha) + \mathcal{O}(\alpha)\right] |10\rangle\langle 10| \\ &\quad + \left[\log\left(\frac{q}{2}\right) + \log(\alpha) + \mathcal{O}(\alpha)\right] |11\rangle\langle 11|\ . \end{aligned} \tag{A.37}$$

Appendix B

Relative Entropy and Moments

In this chapter, we develop the approximations for the relative entropy $D(\psi_{A_1B}||\psi_{A_1} \otimes \psi_B)$, and its second and fourth moments, $V(\psi_{A_1B}||\psi_{A_1} \otimes \psi_B)$ and $Q(\psi_{A_1B}||\psi_{A_1} \otimes \psi_B)$.

B.1 Relative Entropy

Consider the relative entropy, $D(\psi_{A_1B}||\psi_{A_1} \otimes \psi_B)$. By subtracting (A.37) from (A.36),

$$\begin{aligned}\log(\psi_{A_1B}) - \log(\psi_{A_1} \otimes \psi_B) = &\left[\frac{4(q-1)^2}{(q-2)^2}\alpha\log(\alpha) + \mathcal{O}(\sqrt{\alpha})\right]|00\rangle\langle 00| \\ &+ [\mathcal{O}(\alpha)]\,|01\rangle\langle 01| \\ &+ \left[\log\left(\frac{q}{2}\right) - \log\left(1-\frac{q}{2}\right) + \mathcal{O}(\alpha)\right]|10\rangle\langle 10| \\ &+ \left[\log(C(q)) - \log\left(\frac{q}{2}\right) + \mathcal{O}(\alpha\log(\alpha))\right]|11\rangle\langle 11| \\ &+ \left[-\frac{2(q-1)}{(q-2)}\sqrt{\alpha}\log(\alpha) + \mathcal{O}(\sqrt{\alpha})\right]|00\rangle\langle 11| \\ &+ \left[-\frac{2(q-1)}{(q-2)}\sqrt{\alpha}\log(\alpha) + \mathcal{O}(\sqrt{\alpha})\right]|11\rangle\langle 00|\ . \end{aligned} \tag{B.1}$$

Multiplying by ψ_{A_1B}, we have

$$\begin{aligned}\psi_{A_1B}[\log(\psi_{A_1B}) - \log(\psi_{A_1} \otimes \psi_B)] = &\Bigg[\left(1-\frac{q}{2}\right)\frac{4(q-1)^2}{(q-2)^2}\alpha\log(\alpha) \\ &- 2(1-q)\frac{2(q-1)}{(q-2)}\alpha\log(\alpha) + \mathcal{O}(\sqrt{\alpha})\Bigg]|00\rangle\langle 00| \\ &+\mathcal{O}(\alpha)\,(|01\rangle\langle 01| + |10\rangle\langle 10| + |11\rangle\langle 11|) \\ &+ [\mathcal{O}(\sqrt{\alpha}\log(\alpha))]\,|00\rangle\langle 11| \\ &+ [\mathcal{O}(\sqrt{\alpha}\log(\alpha))]\,|11\rangle\langle 00|\ . \end{aligned} \tag{B.2}$$

Applying the trace, we approximate the relative entropy:

$$\begin{aligned}D(\psi_{A_1B}||\psi_{A_1}\otimes\psi_B) &= \mathrm{tr}\left[\psi_{A_1B}\log(\psi_{A_1B}) - \psi_{A_1B}\log(\psi_{A_1}\otimes\psi_B)\right]\\ &= \left[\left(1-\frac{q}{2}\right)\frac{4(q-1)^2}{(q-2)^2} - (1-q)\frac{2(q-1)}{(q-2)}\right]\alpha\log(\alpha) + \mathcal{O}(\alpha)\\ &= -2\frac{(1-q)^2}{2-q}\alpha\log(\alpha) + \mathcal{O}(\alpha)\,. \end{aligned} \tag{B.3}$$

B.2 Second Moment

Next, we consider the second moment of the relative entropy. By squaring (B.1), we have:

$$\begin{aligned}|\log(\psi_{A_1B}) - \log(\psi_{A_1}\otimes\psi_B)|^2 = &\left[\frac{4(q-1)^2}{(q-2)^2}\alpha\log^2(\alpha) + \mathcal{O}(\alpha\log(\alpha))\right]|00\rangle\langle 00|\\ &+\left[\mathcal{O}(\alpha^2)\right]|01\rangle\langle 01|\\ &+\left[\left(\log\left(\frac{q}{2}\right) - \log\left(1-\frac{q}{2}\right)\right)^2 + \mathcal{O}(\alpha)\right]|10\rangle\langle 10|\\ &+\Bigg[\left(\log(C(q)) - \log\left(\frac{q}{2}\right)\right)^2 + \frac{4(q-1)^2}{(q-2)^2}\alpha\log^2(\alpha)\\ &+\mathcal{O}(\alpha\log(\alpha))\Bigg]|11\rangle\langle 11|\\ &+\left[\mathcal{O}(\sqrt{\alpha}\log(\alpha))\right](|00\rangle\langle 11| + |11\rangle\langle 00|)\,. \end{aligned} \tag{B.4}$$

As we multiply by ψ_{A_1B},

$$\begin{aligned}\psi_{A_1B}|\log(\psi_{A_1B}) - \log(\psi_{A_1}\otimes\psi_B)|^2 = &\left[\left(1-\frac{q}{2}\right)\frac{4(q-1)^2}{(q-2)^2}\alpha\log^2(\alpha) + \mathcal{O}(\alpha\log(\alpha))\right]|00\rangle\langle 00|\\ &+\left[\mathcal{O}(\alpha^2)\right]|01\rangle\langle 01| + \left[\mathcal{O}(\alpha)\right]|10\rangle\langle 10|\\ &+\left[\mathcal{O}(\alpha\log(\alpha))\right]|11\rangle\langle 11|\\ &+\left[\mathcal{O}(\sqrt{\alpha}\log(\alpha)\right](|00\rangle\langle 11| + |11\rangle\langle 00|)\,. \end{aligned} \tag{B.5}$$

Using (B.3) and applying the trace to the above, we obtain an approximation of the second moment:

$$\begin{aligned}V(\psi_{A_1B}||\psi_{A_1}\otimes\psi_B) &= \mathrm{tr}\left[\psi_{A_1B}\,|\log(\psi_{A_1B}) - \log(\psi_{A_1}\otimes\psi_B)\ \ - D(\psi_{A_1B}||\psi_{A_1}\otimes\psi_B)|^2\right]\\ &= \mathrm{tr}\left[\psi_{A_1B}\,|\log(\psi_{A_1B}) - \log(\psi_{A_1}\otimes\psi_B)|^2\right] - 2D(\psi_{A_1B}||\psi_{A_1}\otimes\psi_B)\\ &\qquad\times\mathrm{tr}\left[\psi_{A_1B}\,|\log(\psi_{A_1B}) - \log(\psi_{A_1}\otimes\psi_B)|\right] + D(\psi_{A_1B}||\psi_{A_1}\otimes\psi_B)^2\\ &= \mathrm{tr}\left[\psi_{A_1B}\,|\log(\psi_{A_1B}) - \log(\psi_{A_1}\otimes\psi_B)|^2\right] - D(\psi_{A_1B}||\psi_{A_1}\otimes\psi_B)^2\\ &= \left(1-\frac{q}{2}\right)\frac{4(q-1)^2}{(q-2)^2}\alpha\log^2(\alpha) + \mathcal{O}(\alpha\log(\alpha)) + \mathcal{O}(\alpha^2\log^2(\alpha))\\ &= \frac{2(q-1)^2}{q-2}\alpha\log^2(\alpha) + \mathcal{O}(\alpha\log(\alpha))\,. \end{aligned} \tag{B.6}$$

B.3 Fourth Moment

Consider

$$\begin{aligned} Q(\psi_{A_1B}||\psi_{A_1}\otimes\psi_B) = \operatorname{tr}\big[\psi_{A_1B}\big|\log(\psi_{A_1B}) - \log(\psi_{A_1}\otimes\psi_B) \\ - D(\psi_{A_1B}||\psi_{A_1}\otimes\psi_B)\big|^4\big]\,. \end{aligned} \tag{B.7}$$

We use the binomial identity: $(X-c)^4 = X^4 - 4cX^3 + 6c^2X^2 - 4c^3X + c^4$, for an Hermitian operator $X \in \mathcal{L}(\mathcal{H})$ and a real number c. Substituting $X = \log(\psi_{A_1B}) - \log(\psi_{A_1}\otimes\psi_B)$, and $c = D(\psi_{A_1B}||\psi_{A_1}\otimes\psi_B)$, we obtain

$$\begin{aligned} Q(\psi_{A_1B}||\psi_{A_1}\otimes\psi_B) = \operatorname{tr}\Big[\psi_{A_1B}\left(\log(\psi_{A_1B}) - \log(\psi_{A_1}\otimes\psi_B)\right)^4\Big] \\ -4D(\psi_{A_1B}||\psi_{A_1}\otimes\psi_B) \times \operatorname{tr}\Big[\psi_{A_1B}\big(\log(\psi_{A_1B}) - \log(\psi_{A_1}\otimes\psi_B)\big)^3\Big] \\ + \mathcal{O}(\alpha^3\log^4(\alpha))\,. \end{aligned} \tag{B.8}$$

Using (B.1) and (B.4),

$$\begin{aligned} (\log(\psi_{A_1B}) - \log(\psi_{A_1}\otimes\psi_B))^4 = \Big[\mathcal{O}(\alpha\log^2(\alpha))\Big]\,|00\rangle\!\langle 00| \\ + \Big[\mathcal{O}(\alpha^4)\Big]\,|01\rangle\!\langle 01| \\ + [\mathcal{O}(1)]\,|10\rangle\!\langle 10| \\ +[\mathcal{O}(1)]\,|11\rangle\!\langle 11| \\ + [\mathcal{O}(\sqrt{\alpha}\log(\alpha)]\,|00\rangle\!\langle 11| \\ + [\mathcal{O}(\sqrt{\alpha}\log(\alpha)]\,|11\rangle\!\langle 00|\,, \end{aligned} \tag{B.9}$$

and

$$\begin{aligned} (\log(\psi_{A_1B}) - \log(\psi_{A_1}\otimes\psi_B))^3 = \Big[\mathcal{O}(\alpha\log^2(\alpha))\Big]\,|00\rangle\!\langle 00| \\ + \Big[\mathcal{O}(\alpha^3)\Big]\,|01\rangle\!\langle 01| \\ +[\mathcal{O}(1)]\,|10\rangle\!\langle 10| + [\mathcal{O}(1)]\,|11\rangle\!\langle 11| \\ + [\mathcal{O}(\sqrt{\alpha}\log(\alpha)]\,|00\rangle\!\langle 11| \\ + [\mathcal{O}(\sqrt{\alpha}\log(\alpha)]\,|11\rangle\!\langle 00|\,. \end{aligned} \tag{B.10}$$

Multiplying by ψ_{A_1B}, we have

$$\begin{aligned} \psi_{A_1B}\left|(\log(\psi_{A_1B}) - \log(\psi_{A_1}\otimes\psi_B))^4\right| = \Big[\mathcal{O}(\alpha\log^2(\alpha))\Big]\,|00\rangle\!\langle 00| \\ + \Big[\mathcal{O}(\alpha^4)\Big]\,|01\rangle\!\langle 01| \\ + [\mathcal{O}(\alpha)]\,(|10\rangle\!\langle 10| + |11\rangle\!\langle 11|) \\ + [\mathcal{O}(\sqrt{\alpha}\log(\alpha)]\,|00\rangle\!\langle 11| \\ + [\mathcal{O}(\sqrt{\alpha}\log(\alpha)]\,|11\rangle\!\langle 00|\,, \end{aligned} \tag{B.11}$$

and

$$\begin{aligned}\psi_{A_1B}\left|(\log(\psi_{A_1B})-\log(\psi_{A_1}\otimes\psi_B))^3\right| = &\left[\mathcal{O}(\alpha\log^2(\alpha))\right]|00\rangle\langle 00| \\ &+\left[\mathcal{O}(\alpha^3)\right]|01\rangle\langle 01| \\ &+\left[\mathcal{O}(\alpha)\right](|10\rangle\langle 10|+|11\rangle\langle 11|) \\ &+\left[\mathcal{O}(\sqrt{\alpha}\log(\alpha)\right]|00\rangle\langle 11|+|11\rangle\langle 00| \\ &+\left[\mathcal{O}(\sqrt{\alpha}\log(\alpha)\right]|11\rangle\langle 00|\ . \end{aligned} \tag{B.12}$$

Finally, by tracing out, we obtain the order of the fourth moment:

$$Q(\psi_{A_1B}||\psi_{A_1}\otimes\psi_B)=\mathcal{O}(\alpha\log^2(\alpha))\,. \tag{B.13}$$

Appendix C

Entropy of a General Qubit in output of depolarizing channel

Consider a general input state,

$$\rho_A \equiv \begin{pmatrix} 1-a & b \\ b^* & a \end{pmatrix} . \tag{C.1}$$

The state at the output of the qubit depolarizing channel with a noise parameter q is,

$$\rho_B \equiv \begin{pmatrix} (1-q)(1-a)+\frac{q}{2} & (1-q)b \\ (1-q)b^* & (1-q)a+\frac{q}{2} \end{pmatrix} . \tag{C.2}$$

The eigenvalues of ρ_B are

$$e_1 = \frac{1}{2}\left(1+\sqrt{1-4\left(((1-q)(1-a)+\frac{q}{2})((1-q)a+\frac{q}{2})\right)+4|b|^2}\right) , \tag{C.3}$$

$$e_2 = \frac{1}{2}\left(1-\sqrt{1-4\left(((1-q)(1-a)+\frac{q}{2})((1-q)a+\frac{q}{2})\right)+4|b|^2}\right) . \tag{C.4}$$

Thus, the entropy of ρ_B is

$$H(\rho_B) = -e_1 \log(e_1) - e_2 \log(e_2) . \tag{C.5}$$

Notice that the eigenvalues do not depend on the phase of the off-diagonal entry, b, hence the entropies of ρ_A and $Z\rho_A Z$ are the same, and then,

$$H(\rho_B) = H(Z\rho_B Z) \tag{C.6}$$

with

$$Z\rho_B Z \equiv \begin{pmatrix} (1-q)(1-a)+\frac{q}{2} & -(1-q)b \\ -(1-q)b^* & (1-q)a+\frac{q}{2} \end{pmatrix}. \tag{C.7}$$

Since the entropy is concave, we have

$$\begin{aligned} H(\rho_B) &= \frac{1}{2}H(\rho_B)+\frac{1}{2}H(Z\rho_B Z) \\ &\leq H\left(\frac{1}{2}\rho_B+\frac{1}{2}Z\rho_B Z\right) \\ &= H\left(\left[(1-q)(1-a)+\frac{q}{2}\right]|0\rangle\langle 0| + \left[(1-q)a+\frac{q}{2}\right]|1\rangle\langle 1|\right) \\ &= H\Big(\mathcal{N}_{A\to B}\left((1-a)\,|0\rangle\langle 0| + a\,|1\rangle\langle 1|\right)\Big) \end{aligned} \tag{C.8}$$

Therefore, the maximal output entropy can be achieved with $b = 0$. i.e., for an input state of the form

$$\rho_A \equiv \begin{pmatrix} 1-a & 0 \\ 0 & a \end{pmatrix}. \tag{C.9}$$

Since the energy constraint requires $a \leq E$,

$$H(\rho_B) \leq \max_{0\leq a\leq E} H\Big(\mathcal{N}_{A\to B}\left((1-a)\,|0\rangle\langle 0| + a\,|1\rangle\langle 1|\right)\Big) \tag{C.10}$$

$$= \max_{0\leq a\leq E} h_2\Big(a * \frac{q}{2}\Big) \tag{C.11}$$

$$= h_2\Big(E * \frac{q}{2}\Big) \tag{C.12}$$

Bibliography

[1] M. Wang, T. Zhu, T. Zhang, J. Zhang, S. Yu, and W. Zhou, “Security and privacy in 6G networks: New areas and new challenges,” *Digital Commun. Netw.*, vol. 6, no. 3, pp. 281–291, 2020.

[2] J. Talbot, D. Welsh, and D. J. A. Welsh, *Complexity and cryptography: An Introduction.* Cambridge University Press, 2006, vol. 13.

[3] M. Bloch and J. Barros, *Physical-layer Security: From Information Theory to Security engineering.* Cambridge University Press, 2011.

[4] C. H. Bennett and G. Brassard, “Quantum cryptography: Public key distribution and coin tossing,” in *Proc. IEEE Int. Conf. Comp.*, 1984.

[5] R. Renner, “Security of quantum key distribution,” *Int. J. Quantum Inf.*, vol. 6, no. 01, pp. 1–127, 2008.

[6] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, “The security of practical quantum key distribution,” *Rev. Mod. Phys.*, vol. 81, no. 3, p. 1301, 2009.

[7] M. Bloch, O. Günlü, A. Yener, F. Oggier, H. V. Poor, L. Sankar, and R. F. Schaefer, “An overview of information-theoretic security and privacy: Metrics, limits and applications,” *IEEE J. Sel. Areas Inf. Theory*, vol. 2, no. 1, pp. 5–22, 2021.

[8] R. F. Schaefer, H. Boche, and H. V. Poor, “Secure communication under channel uncertainty and adversarial attacks,” *Proc. IEEE*, vol. 103, no. 10, pp. 1796–1813, 2015.

[9] L. Bai, J. Xu, and L. Zhou, “Covert communication for spatially sparse mmwave massive mimo channels,” *IEEE Trans. Comm.*, vol. 71, no. 3, pp. 1615–1630, 2023.

[10] J. Bai, J. He, Y. Chen, Y. Shen, and X. Jiang, “On achievable covert communication performance under csi estimation error and feedback delay,” 2024.

[11] W. Chen, H. Ding, S. Wang, F. Gong, and G. Xia, “On the limits of covert backscatter communication over undecodable ambient signals,” *IEEE Trans. Inf. Forensics Security*, vol. 18, pp. 4198–4213, 2023.

[12] S.-Y. Wang, M.-C. Chang, and M. R. Bloch, "Covert joint communication and sensing under variational distance constraint," in *2024 58th Annual Conference on Information Sciences and Systems (CISS)*. IEEE, 2024, pp. 1–6.

[13] B. A. Bash, D. Goeckel, D. Towsley, and S. Guha, "Hiding information in noise: Fundamental limits of covert wireless communication," *IEEE Commun. Mag.*, vol. 53, no. 12, pp. 26–31, 2015.

[14] M. Tahmasbi, A. Savard, and M. R. Bloch, "Covert capacity of non-coherent rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 66, no. 4, pp. 1979–2005, 2020.

[15] B. A. Bash, D. Goeckel, and D. Towsley, "Limits of reliable communication with low probability of detection on AWGN channels," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1921–1930, 2013.

[16] M. R. Bloch, "Covert communication over noisy channels: A resolvability perspective," *IEEE Trans. Inf. Theory*, vol. 62, no. 5, pp. 2334–2354, 2016.

[17] L. Wang, G. W. Wornell, and L. Zheng, "Fundamental limits of communication with low probability of detection," *IEEE Trans. Inf. Theory*, vol. 62, no. 6, pp. 3493–3503, 2016.

[18] A. Sheikholeslami, B. A. Bash, D. Towsley, D. Goeckel, and S. Guha, "Covert communication over classical-quantum channels," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Barcelona, Spain, Jul. 2016.

[19] M. S. Bullock, A. Sheikholeslami, M. Tahmasbi, R. C. Macdonald, S. Guha, and B. A. Bash, "Covert communication over classical-quantum channels," arXiv:1601.06826 [quant-ph], 2023.

[20] M. Tahmasbi and M. R. Bloch, "Framework for covert and secret key expansion over classical-quantum channels," *Phys. Rev. A*, vol. 99, no. 5, p. 052329, 2019.

[21] B. A. Bash, A. H. Gheorghe, M. Patel, J. L. Habif, D. Goeckel, D. Towsley, and S. Guha, "Quantum-secure covert communication on bosonic channels," *Nat. commun.*, vol. 6, no. 1, pp. 1–9, 2015.

[22] M. S. Bullock, C. N. Gagatsos, S. Guha, and B. A. Bash, "Fundamental limits of quantum-secure covert communication over bosonic channels," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 3, pp. 471–482, 2020.

[23] C. N. Gagatsos, M. S. Bullock, and B. A. Bash, "Covert capacity of bosonic channels," *IEEE J. Sel. Areas Inf. Theory*, vol. 1, no. 2, pp. 555–567, 8 2020.

[24] S.-Y. Wang, T. Erdoğan, and M. Bloch, "Towards a characterization of the covert capacity of bosonic channels under trace distance," in *IEEE Int. Symp. Inf. Theory (ISIT)*, 2022, pp. 318–323.

[25] B. A. Bash, C. N. Gagatsos, A. Datta, and S. Guha, "Fundamental limits of quantum-secure covert optical sensing," in *IEEE Int. Symp. Inf. Theory (ISIT)*. IEEE, 2017, pp. 3210–3214.

[26] M. Tahmasbi and M. R. Bloch, "On covert quantum sensing and the benefits of entanglement," *IEEE J. Sel. Areas Inf. Theory*, vol. 2, no. 1, pp. 352–365, 2021.

[27] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, "Entanglement-assisted classical capacity of noisy quantum channels," *Phys. Rev. Lett*, vol. 83, no. 15, p. 3081, 1999.

[28] ——, "Entanglement-assisted capacity of a quantum channel and the reverse shannon theorem," *IEEE Trans. Inf. Theory*, vol. 48, no. 10, pp. 2637–2655, 2002.

[29] S. Hao, H. Shi, W. Li, J. H. Shapiro, Q. Zhuang, and Z. Zhang, "Entanglement-assisted communication surpassing the ultimate classical capacity," *Phys. Rev. Lett.*, vol. 126, no. 25, p. 250501, 2021.

[30] A. Chiuri, S. Giacomini, C. Macchiavello, and P. Mataloni, "Experimental achievement of the entanglement-assisted capacity for the depolarizing channel," *Phys. Rev. A*, vol. 87, no. 2, p. 022333, 2013.

[31] U. Pereg, C. Deppe, and H. Boche, "Quantum channel state masking," *IEEE Trans. Inf. Theory*, vol. 67, no. 4, pp. 2245–2268, 2021.

[32] M. Tomamichel, *Quantum information processing with finite resources: mathematical foundations*. Springer, 2015, vol. 5.

[33] M. M. Wilde, *Quantum information theory*, 2nd ed. Cambridge University Press, 2017.

[34] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to algorithms*. MIT press, 2022.

[35] S. Khabbazi Oskouei, S. Mancini, and M. M. Wilde, "Union bound for quantum information processing," *Proc. Royal Soc. A*, vol. 475, no. 2221, p. 20180612, 2019.

[36] B. Schumacher and M. D. Westmoreland, "Sending classical information via noisy quantum channels," *Phys. Rev. A*, vol. 56, pp. 131–138, Jul 1997.

[37] A. S. Holevo, "The capacity of quantum channel with general signal states," 1996.

[38] ——, *Quantum Systems, Channels, Information*. Berlin, Boston: De Gruyter, 2013.

[39] C. E. Shannon, “A mathematical theory of communication,” *The Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948.

[40] M. M. Wilde, “Position-based coding and convex splitting for private communication over quantum channels,” *Quantum Inf. Proc.*, vol. 16, no. 10, p. 264, 2017.

[41] A. S. Holevo, “Entanglement-assisted capacity of constrained channels,” in *1st Int. Symp. Quantum Info.*, vol. 5128. SPIE, 2003, pp. 62–69.

[42] C. King, “The capacity of the quantum depolarizing channel,” *IEEE Trans. Inf. Theory*, vol. 49, no. 1, pp. 221–229, 2003.

[43] D. Leung and J. Watrous, “On the complementary quantum capacity of the depolarizing channel,” *Quantum*, vol. 1, 10 2015.

[44] G. De Palma, “New lower bounds to the output entropy of multi-mode quantum gaussian channels,” *IEEE Transactions on Information Theory*, vol. 65, no. 9, pp. 5959–5968, 2019.

[45] E. L. Lehmann and J. P. Romano, *Testing statistical hypotheses*, 3rd ed., ser. Springer Texts in Statistics. New York: Springer, 2005.

[46] M. Tahmasbi, “Covert communication: From classical channels to quantum channels,” 2020.

[47] D. Kretschmann, D. Schlingemann, and R. F. Werner, “The information-disturbance tradeoff and the continuity of stinespring’s representation,” *IEEE Trans. Inf. Theory*, vol. 54, no. 4, pp. 1708–1717, 2008.

[48] M. J. Salariseddigh, U. Pereg, H. Boche, and C. Deppe, “Deterministic identification over fading channels,” in *IEEE Inf. Theory Workshop (ITW)*, 2021, pp. 1–5.

[49] M. J. Salariseddigh, U. Pereg, H. Boche, C. Deppe, V. Jamali, and R. Schober, “Deterministic identification for molecular communications over the poisson channel,” *arXiv:2203.02784*, 2022.

[50] R. Ahlswede and G. Dueck, “Identification via channels,” *IEEE Trans. Inf. Theory*, vol. 35, pp. 15–29, 1989.

[51] S. Guha, Q. Zhuang, and B. A. Bash, “Infinite-fold enhancement in communications capacity using pre-shared entanglement,” in *IEEE Int. Symp. Inf. Theory (ISIT)*, 2020, pp. 1835–1839.

[52] H. Shi, Z. Zhang, and Q. Zhuang, “Practical route to entanglement-assisted communication over noisy bosonic channels,” *Phys. Rev. App.*, vol. 13, no. 3, mar 2020.

[53] C. H. Bennett, D. P. DiVincenzo, and J. A. Smolin, “Capacities of quantum erasure channels,” *Phys. Rev. Lett.*, vol. 78, no. 16, p. 3217, 1997.