10-1-2017

Energy-Constrained Quantum Communication and Digital Dynamical Decoupling

Haoyu Qi
Louisiana State University and Agricultural and Mechanical College

ENERGY-CONSTRAINED QUANTUM COMMUNICATION AND DIGITAL DYNAMICAL DECOUPLING

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Physics & Astronomy

by
Haoyu Qi
B.Sc., Beijing Institute of Technology, 2012
December 2017

To my parents, who shaped not only my physique, but also my character subtlety, which I realized after many years;

To my high school teacher Mrs. Zhao, my Ph.D. advisors Professor Dowling and Professor Wilde, who brought me into and helped me survive this 10-year journey of Physics;

To my best friend Chen, who is always there for me. And I want you to know that I will always be there for you as well;

Finally to my beloved 'little friend', who I love more than sharks love blood; who keeps reminding me that 'poetry, beauty, romance, love, these are what we stay alive for'. It is the future with her presence that makes me typing right now.

Acknowledgments

I want to first thank professor Ilya Vekhter who brought me into LSU in the first place. I still remember the time when he invited me to Chimes and explained to me étouffée and hushpuppy, and how we think American football is boring in the same way.

I am greatly grateful to my advisor, Professor Jonathan P. Dowling. I still feel it is unbelievable, even several years has passed, that I could get a RA position by simply exchanging emails with Jon. It was right after I joined the QST group that I somehow learned how to conduct research in an independent way. I still remember clearly the first time when I completed a paper on my own — from generating idea, drafting the paper, to finally posting it on arXiv. It was a true turning point of my PhD. Probably with hindsight, it is Jon's special way of management, which offered me a pressure-free environment, that eventually led to my transformation.

I am grateful to have had the opportunity to work with Professor Lorenza Viola. I still remember her hospitality when I visited Dartmouth college. It is during there I realized that happiness could simply be having a cup of clam chowder during a freezing day. I learned a lot from the way how she devoted herself to research every single day, and how strict and high-standard she could be when it comes to her or her students' research.

I am deeply grateful to my co-supervisor, Professor Mark M. Wilde, who introduced me to the field of quantum Shannon theory. It is the first time that I was immersed in a field with such richness and mathematical sophistication. Mark is always willing to share his encyclopedic knowledge, even his own research experience. When communicating with his students, he is always polite and makes us feel treated as equals. I am also thankful for his support and help so that I had the opportunities to interact with many other top researchers in this field. I learned so much, not only from his expertise in com-

munication theory, but also from his enthusiasm toward research and his self-disciplinary routine.

I am thankful to Masahiro Takeoka, who took care of my visit in Tokyo and taught me CV QKD. His research is always guided by physical intuition and starts with simplified problems, from which I learned a lot. I will never forget the evening he played 'Memory' from the movie 'Departure' for me. I am also thankful to Masahiko Matsuo whom I want to drink more sake with. I always have the feeling that we share something in common. Many thanks to Mineko Asano who looked after and treated me like an older sister. Tokyo is a place where miracles can really happen. I miss you all and I will be back one day to see you guys.

I am thankful to Professor Jens Eisert, who offered me the great opportunity to apply for the Humboldt fellowship with him. Even when I procrastinated and fell behind the time schedule, he was still supportive and encouraged me. I also appreciate his support on my travel in Berlin. The Berlin Cathedral, the breath-taking sounds of violin and the people lying on the grass in front of it, constitutes the most beautiful and touching scene I have ever been within. I am thankful to Kauishik P. Seshadreesan for his accompany in Erlangen, and his wife Varsha who took a lot of time to prepare a great Indian dinner with samosa and curry.

I am grateful to Manish Gupta, who let me try my first samosa and who can talk with me without any holdback. I wish him the best for his life in India. I am grateful to Kunal Sharma, who was there in some of my hardest time, who is a great partner in gym and research, and who taught me how to play racquetball although it looks more and more dangerous to me. I want to thank Zhichao Xue and Chenglong You, who are my best Chinese friends and lent me their hands so many times.

I thank folks from QST group and the physics department: Professor Hwang Lee, Zhihao Xiao, Qingle Wang, Xiaoping Ma, Siddhartha Das, Sushovit Adhikari, Noah Davis,

Kevin Valson Jacob, Eneet Kaur, Nick Lanning, Nicolas Studer. Thanks you guys for all the wonderful memories.

Finally, I am thankful to Arnell Jackson who helped me with all the tedious paperworks since my first day in LSU. Each time when I see those cabinets full of carefully organized folders and the stickers in her office, I see her contribution to this department. I also want to thank Carol Duran and Claire Bullock who helped with my travel reimbursement many times, with great patience and kindness.

Table of Contents

Abstract

This is a two-part thesis glued together by an everlasting theme in Quantum Information Science - to save the quantum state, or the information stored in it, from unavoidably environment-induced noise. The first part of this thesis studies the ultimate rate of reliably transmitting information, stored in quantum systems, through a noisy evolution. Specifically, we consider communication over optical links, upon which future inter-city quantum communication networks will be built. We show how to treat the infinite-dimensional bosonic system rigorously and establish the theory of energy-constrained private and quantum communication over quantum channels. Our result represents important progress in the field of energy-constrained quantum communication theory. As an example of communication over optical channels, we solve the triple trade-off capacity and broadcast capacity of quantum-limited amplifier channels. Our result not only includes two single-letter capacities, which are rare in quantum communication theory, but it is also the only known application of a recently proved minimum output-entropy conjecture.

The second part of my thesis includes two of my works on dynamical decoupling (DD). DD is an open-loop technique to keep a qubit alive during decoherence, which is important for the actual implementation of quantum memory or a quantum computer. Instead of treating quantum evolution as a completely positive trace preserving map like in communication theory, we consider time-dependent evolution of a specific quantum system in quantum control theory. With more than decade of development of the theory of DD, people started to focus on pulse sequences with low sequencing complexity (called digital pulse sequences), which are required for large-scale implementation of quantum computation devices. We propose two unifying frameworks to systematically generate

these engineering-friendly pulse sequences. Surprisingly, we prove that these two frameworks are actually two sides of the same coin, and thus our work greatly deepens our understanding of the underlying structure and the decoupling performance of digital pulse sequences.

Chapter 1
Introduction and Preliminaries

The foundation of Quantum Mechanics was established during the first half of the 20th century by numerous geniuses [1]. However, quantum mechanics is so profound, rich and also so peculiar that even today new directions keep emerging with name in the form of 'quantum $\times$'. A common theme among many of nowadays important fields is the reconciliation of classical subject with quantum mechanics. For example, quantum computation [2] is the hybrid of computer science and quantum mechanics, quantum Shannon theory [3] is the reconciliation of classical information theory and quantum mechanics; quantum control theory [4] enhances classical control theory to harness quantum systems. However, these reconciliation are never merely '$1 + 1$'-like additions: each of these fields finds novel phenomena and consequentially a huge potential of applications. Efficient quantum algorithm of large number factorization [5], quantum key distribution [6], quantum teleportation [7], represent a few. All of these phenomena are possible only due to the unique properties of quantum mechanics. Therefore, to understand the two topics in this thesis, we first review the postulates of quantum mechanics. The richness derived from these simple and elegant postulates is remarkable. Next we briefly review the history, basic concepts and tools of quantum Shannon theory. Actually the viewpoint of quantum channels deepens our understanding of quantum mechanics. Since the first part of this thesis will focus on energy constrained quantum communication, with bosonic system as its most important application, at last we review the basics of Gaussian formalism, where quantum Shannon theory merges with quantum optics.

1.1 QUANTUM MECHANICS

We start with stating the four postulates of quantum mechanics. We summarize some of its important and counterintuitive features. Finally we briefly review the mathematical formalism of qubits.

Postulates

Every theory of physics includes two parts: kinematics and dynamics. Kinematics makes assumptions on how we mathematically describe the existence of matters and how their states evolve with time. The four postulates stated below comprise the kinematics of quantum mechanics. I directly take them from the classic textbook [2] since they are widely accepted.

Postulate 1.1 *Associated to any isolated physical system is Hilbert space known as the state space of the system. The system is completely described by its state vector, which is a unit vector in the system's state space.*

We usually label a quantum system by a capitalized Latin letter. For example, a friend we will repeatedly encounter in Shannon theory is Alice, who usually holds a quantum system A. The Hilbert space of Alice is $\mathcal{H}_A$. A state vector of this Hilbert space is written as a ket $|\psi\rangle_A$, with a subscript indicating the system we refer to. As said nicely in [2] the first postulate sets up the arena in which quantum mechanics takes place. Just like in Newtonian mechanics, the state of a particle is fully specified by its coordinates (maybe along with its momentums) at given time, (x, y, z, t). In quantum mechanics, it is the state vector at given time, $|\psi(t)\rangle$, encloses all the information of that quantum system.

Postulate 1.2 *The evolution of a closed quantum system is described by a unitary transformation. That is, the state $|\psi(t_1)\rangle$ of the system at time t_1 is related to the state $|\psi(t_2)\rangle$ of the system at time*

t_2 by a unitary operator $U(t_2, t_1)$ which depends only on the times t_1 and t_2,

$$|\psi(t_2)\rangle = U(t_2, t_1)|\psi(t_1)\rangle \ . \tag{1.1}$$

Postulate 1.2 plays similar role as Newton's second law. Just as we don't know how to decide the force just by looking at $F = ma$, the second postulate does not tell us what is the actual unitary. However, it does tell us that the evolution of quantum state is governed by a unitary operator, just as $F = ma$ tells us the movement of particles are governed by a second-order differential equation. As to how to decide the actual unitary, it is the task of dynamics.

Any unitary operator can be written in the following form

$$U(t) = e^{-iHt} \ , \tag{1.2}$$

where H is a Hermitian operator called the Hamiltonian of the system. Substitute this form into Eq. (1.1), take the derivative with respect to time, and the result is consistent with the famous Schrödinger Equation.

Postulate 1.3 *Quantum measurements are described by a collection $\{M_m\}$ of measurement operators, acting on the state space of the system being measured. The index m refers to the measurement outcomes that may occur in the experiment. If the state of the quantum system is $|\psi\rangle$ before the measurement, then the probability that result m being read out is given by*

$$p(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle \ , \tag{1.3}$$

and the state of the system after the measurement is

$$\frac{M_m|\psi\rangle}{\sqrt{\langle\psi|M_m^\dagger M_m|\psi\rangle}}\,. \tag{1.4}$$

The measurement operator satisfied the completeness relation,

$$\sum_m M_m^\dagger M_m = I, \tag{1.5}$$

which is imposed by the fact that probabilities sum to one.

The Postulate 1.3 is probably the most intriguing one and is one way in which quantum mechanics truly departs from classical mechanics. The state of a quantum system changes discretely, or collapses, into one of the measurement outcomes. In other words, whenever we measure the system, the measurement itself disturbs the quantum state irreversibly. Many interesting philosophical debates over this so-called 'observer effect' were there since the beginning of quantum mechanics. The most well-known sentence is perhaps this one: 'does the moon exist when we are not looking at it?' It is actually this effect of quantum mechanics that gives the biggest trouble to quantum physicists and engineers. The environment interacting with the qubits in our quantum computer acts just like a observer who keeps measuring the qubits and thus destroys their coherence. On the other hand, if it is treated in clever way, we can use it to detect malicious eavesdropper and thus protect the transmitted information from being stolen.

Postulate 1.4 *The state space of a composite physical system is the tensor product of the state spaces of the component physical systems. Moreover, if we have systems numbered* 1 *through* n*, and system number* i *is prepared in the state* $|\psi_i\rangle$*, then the joint state of the total system is* $|\psi_1\rangle \otimes |\psi_2\rangle \otimes \ldots \otimes |\psi_n\rangle$.

This property is usually not considered as one of the fundamental postulates. However,

the way of describing composite quantum systems is very different from those of classical systems. When you consider two classical particles, the state space is the direct product, instead of tensor product, of the two individual spaces. This assumption has a far reaching impact on many-body theory. On one hand, this tensor product structure gives the notion of 'locality' which is essential for describing many-body systems [8]. On the other hand, tensor product structure leads to exponentially large dimension of Hilbert space which is inefficient for computation. However, the ground states of many-body system usually only occupy a small corner of the full Hibert space. Rough estimation tells us that it takes time longer than the life of Universe for the state of a many-body system to travel through the full Hilbert space. These interesting observations lead to the powerful tool of *tensor network* [9].

Qubit

The reasons we consider qubit systems are threefold. First of all, qubit systems are the simplest and most important quantum systems in quantum information science. They can model various physical systems widely used in real experiments. Second, we would like to use qubit system as an example to illustrate the four postulates above. Third, it gives us a good opportunity to introduce some of the most common quantities and notations which we will use later.

A qubit lives in a two-dimensional Hilbert space $\mathcal{C}^2$. The two orthogonal basis states are usually denoted as $|0\rangle$ and $|1\rangle$. A pure qubit state is a superposition of the basis states,

$$|\psi\rangle = a|0\rangle + b|1\rangle \ . \tag{1.6}$$

Here a and b are in general complex numbers and are called the *amplitudes*. They have to satisfy $|a|^2 + |b|^2 = 1$ since $|\psi\rangle$ is a unit vector. The two basis states can be the ground state and first excited state of a ion trap, an NV center, or a superconducting circuit. Qubit can

also be actual spin systems like electron or nuclei, or it can be the two polarizations of a photon.

The most basic unitary operators acting on qubit systems are the Pauli matrices X, Y, Z given by

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.7}$$

The Pauli X and Pauli Z matrices are also referred as the the *bit flip* and *phase flip* gates, respectively. Usually we take the basis states $|0\rangle$ and $|1\rangle$ to be the eigenstates of the Pauli Z operator. Other basic quantum gates are the CNOT gate and Hardmard gate, for more details the reader should go to textbook [2].

The simplest measurement on a qubit system is formed from the projections onto the eigenbasis of Pauli Z operator: $M_0 = |0\rangle\langle 0|$, $M_1 = |1\rangle\langle 1|$. The probability of getting 0 and 1 (or down and up) is given by $|a|^2$ and $|b|^2$. The state will collapse to either $|0\rangle$ or $|1\rangle$ respectively.

If we have two qubit system A and B, the basis of the composite system are all possible tensor products: $|00\rangle_{AB}, |01\rangle_{AB}, |10\rangle_{AB}, |11\rangle_{AB}$ and any pure bipartite state is a superposition of these basis states.

A quantum system with d basis states is called a *qudit*. There are corresponding generalized Pauli operators for qudit systems. For details see Section 3.7 in [3]. In Sec. 1.3 we will consider quantum states in an infinite dimensional Hilbert space.

1.2 QUANTUM SHANNON THEORY

The section is written based on [3].

Bounded linear operators and quantum states

An operator A is bounded if $\|A\|_1 \equiv \mathrm{Tr}\{|A|\} \leq \infty$. Let $\mathcal{B}(H)$ denote the algebra of bounded linear operators acting on Hilbert space H. Let $\mathcal{SP}(\mathcal{H}) = \{X \in \mathcal{B}(H) : X \geq 0\}$ be the subset of postitive semi-definite operators, and Let $\mathcal{P}(\mathcal{H}) = \{X \in \mathcal{B}(H) : X > 0\}$ be the subset of all positive operators. From the spectral theorem, it follows that every Hermitian operator has a *spectral decomposition*,

$$A = \sum_i a_i |i\rangle\langle i| \,. \tag{1.8}$$

A function of a Hermitian operator is then defined by its action on each eigenvalues:

$$f(A) = \sum_{i:a_i \neq 0} f(a_i) |i\rangle\langle i| \,. \tag{1.9}$$

In Sec. 1.1 we only consider pure quantum states. However, imagine we prepare different pure quantum states $|\psi_i\rangle$ with different probability p_i, then we have a mixture of pure quantum states:

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \,. \tag{1.10}$$

If a linear operator is positive, Hermitian and has unit trace, like the state above, we call it a density operator. We define $\mathcal{D}(\mathcal{H}) = \{\rho \in \mathcal{B}(\mathcal{H}) : \rho \geq 0, \rho^\dagger = \rho, \mathrm{Tr}\{\rho\} = 1\}$ to be the set of density operators. Each density operator has a non-unique decomposition in the form of Eq.(1.10); one of them is the spectral decomposition. The density operator captures both classical uncertainty and quantum uncertainty, thus serving as the most general representation of a quantum state.

POVM

In Postulate 1.3, a measurement is given by a set of operator $\{M_m\}$ satisfying the completeness relation. This formalism can be derived by considering the system of interest interacting unitarily with a probe system and we measure the probe by projectors (See Section 4.2 in [3]). However, in quantum Shannon theory we usually consider a more general form of quantum measurement, since we sometimes do not care about the post-measurement state.

Definition 1.5 (POVM) *A positive operator-valued measure (POVM) is a set of non-negative operators $\{\Lambda_j\}$ that satisfy the completeness relation $\sum_j \Lambda_j = I$. The probability of obtaining outcome j is*

$$\mathrm{Tr}\{\rho\Lambda_j\}\ . \tag{1.11}$$

As we will see later in the error analysis of communication protocols, we usually first construct some POVM and then calculate the error probability.

Quantum channels

The concept of a quantum channel is at the heart of quantum Shannon theory. We begin with the definition of a legitimate quantum channel.

Definition 1.6 *Let $\mathcal{N}_{A\to B}$ be a map from $\mathcal{B}(\mathcal{H}_A)$ to $\mathcal{B}(\mathcal{H}_B)$. It is a quantum channel if the followings are true:*

1. *Linearity. $\mathcal{N}(\alpha X_A + \beta Y_A) = \alpha\mathcal{N}(X_A) + \beta\mathcal{N}(Y_A)$.*
2. *Completely positive. For a reference system R of arbitrary size, $id_R \otimes N(X_{RA}) \in \mathcal{SP}(\mathcal{H}_R \otimes \mathcal{H}_B)$ for all $X_{RA} \in \mathcal{SP}(\mathcal{H}_R \otimes \mathcal{H}_A)$.*
3. *Trace preserving. $\mathrm{Tr}\{X_A\} = \mathrm{Tr}\{\mathcal{N}(X_A)\}$.*

The first condition comes from the linearity of quantum mechanics. The second and third conditions guarantee the output is a valid quantum state.

An equivalent and extremely useful representation of a quantum channel is the Kraus representation.

Theorem 1.7 *A map $\mathcal{N} : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ is a quantum channel if and only if it has a Choi-Kraus representation,*

$$\mathcal{N}(X_A) = \sum_{l=0}^{d-1} V_l X_A V_l^\dagger, \tag{1.12}$$

where $X_A \in \mathcal{D}(\mathcal{H}_A)$, $V_l : \mathcal{H}_A \to \mathcal{H}_B$ is linear for all l, satisfying

$$\sum_{l=0}^{d-1} V_l^\dagger V_l = I_A \tag{1.13}$$

and $d \leq \dim(\mathcal{H}_A)\dim(H_B)$.

Proof. See the proof of Theorem 4.4.1 in [3]. ■

An important perspective of quantum channel is that the quantum noise of a quantum system arises from its unitary interaction with another system (the environment), which we do not have access to. Actually it can be shown that for any quantum channel $\mathcal{N}_{A\to B}$ there is some unitary on a larger Hilbert space such that the following is true,

$$\mathcal{N}(\rho_A) = \mathrm{Tr}_E\{U_{AE}(\rho_A \otimes |0\rangle\langle 0|_E)U_{AE}^\dagger\} \,. \tag{1.14}$$

The definition of a quantum channel is so general that every process in quantum mechanics can be understood as a quantum channel. For instance, the preparation of a quantum state ρ_A on system A can be seen as a quantum channel from trivial Hilbert space $\mathcal{C}$ to $\mathcal{H}_A$. Unitary evolution, or more general an isometry, is the simplest quantum channel, in

the sense that there is only one Kraus operator. Quantum measurement can be considered as a quantum channel from a quantum system to a classical system (or the so-called quantum-classical channel).

The purified theory

As we have seen from last subsection, noisy evolution can be considered as a noiseless unitary evolution in a larger Hilbert space with some inaccessible subsystem. The idea of purification is fundamental and so powerful that it leads to many important techniques in quantum Shannon theory. We can consider the purification of both a quantum state and a quantum channel.

Definition 1.8 *A purification of a density operator $\rho_A \in \mathcal{D}(\mathcal{H}_A)$ is a pure bipartite state $|\psi\rangle_{RA}$ on a reference system R and the original system A, such that*

$$\rho_A = \mathrm{Tr}_R\{|\psi\rangle\langle\psi|_{RA}\} \, . \tag{1.15}$$

The physical interpretation of purification of a quantum state is that we can think of the intrinsic noisiness of a quantum system is due to its entanglement with some reference system to which we do not have access.

An important and useful fact is that all purification of a quantum state are equivalent to each other, up to an isometry. Consider two purification of state ρ_A, $|\psi\rangle_{R_1A}$ and $|\psi\rangle_{R_2A}$, such that $\dim(\mathcal{H}_{R_1}) \leq \dim(\mathcal{H}_{R_2})$. Then there exists an isometry $V_{R_1\to R_2}$ such that

$$|\psi\rangle_{R_2A} = (V_{R_1\to R_2} \otimes I_A)|\psi\rangle_{R_1A} \, . \tag{1.16}$$

Definition 1.9 *Consider a quantum channel $\mathcal{N}_{A\to B}$ and some enviroment system E. An isometric extension or Stinespring dilation $U : \mathcal{H}_A \to \mathcal{H}_B \otimes \mathcal{H}_E$ of the channel $\mathcal{N}$ is an isometry such*

that

$$\mathrm{Tr}_E\{UXU^\dagger\} = \mathcal{N}_{A\to B}(X_A)\ . \tag{1.17}$$

And the following is true for U:

$$U^\dagger U = I_A,\ \ UU^\dagger = \Pi_{BE}\ , \tag{1.18}$$

where Π_{BE} is a projector onto $\mathcal{H}_B \otimes \mathcal{H}_E$.

We usually use the following notation to represent the super-operator of the isomeric extension:

$$\mathcal{U}^{\mathcal{N}}_{A\to BE}(X_A) = UX_AU^\dagger\ . \tag{1.19}$$

The idea of isometric extension leads to the very import concept of *complementary channel*.

Definition 1.10 *Let $\mathcal{N}_{A\to B}$ be a quantum channel and $\mathcal{U}^{\mathcal{N}}_{A\to BE}$ be an isometry. The corresponding complementary channel $\mathcal{N}^c_{A\to E}$ of $\mathcal{N}$ is defined as follows*

$$\mathcal{N}^c(X_A) = \mathrm{Tr}_B\{\mathcal{U}^{\mathcal{N}}_{A\to BE}(X_A)\}\ , \tag{1.20}$$

for any $X_A \in \mathcal{B}(\mathcal{H}_A)$.

The complementary channel describes the evolution from the viewpoint of the environment system. In cryptographic settings, it is assumed that the malicious party Eve controls the environment. So the complementary channel is important when analyzing the security of the protocol. Notice that just like the purifications of a quantum state are

non-unique, so are the isometric extensions and the complementary channels of a quantum channel. They are connected by isometries mapping between different environment systems.

Quantum fidelity and trace distance

Given two quantum states, it is important to measure the distance or the closeness between them. As we will see later, the reliable decoding condition or the security condition of a communication protocol is defined in terms of distance measures. These measures are also necessary in converse proofs (upper bounds on capacity).

The α-norm of an operator is defined as

$$\|X\|_\alpha = \left[\mathrm{Tr}\{ \sqrt[\alpha]{X^\dagger X}\} \right]^{1/\alpha} . \tag{1.21}$$

Thus the trace norm is a special case when $\alpha = 1$ and $\| \cdot \|_2$ is the HIlbert-Schmit norm.

The *fidelity* of two quantum states $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ is defined as [10]

$$F(\rho, \sigma) \equiv \left\| \sqrt{\rho}\sqrt{\sigma} \right\|_1^2 . \tag{1.22}$$

Uhlmann's theorem is the statement that the fidelity has the following alternate expression as a probability overlap [10]:

$$F(\rho, \sigma) = \sup_U |\langle \phi^\rho | U \otimes I_{\mathcal{H}} | \phi^\sigma \rangle|^2 , \tag{1.23}$$

where $|\phi^\rho\rangle \in \mathcal{H}' \otimes \mathcal{H}$ and $|\phi^\sigma\rangle \in \mathcal{H}'' \otimes \mathcal{H}$ are fixed purifications of ρ and σ, respectively, and the optimization is with respect to all partial isometries $U : \mathcal{H}'' \to \mathcal{H}'$. The fidelity is non-decreasing with respect to a quantum channel $\mathcal{N} : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$, in the sense that for all $\rho, \sigma \in \mathcal{D}(\mathcal{H}_A)$:

$$F(\mathcal{N}(\rho), \mathcal{N}(\sigma)) \geq F(\rho, \sigma). \tag{1.24}$$

Application of Uhlmann's theorem together with the monotonicity property in (1.24) implies that for a given extension ρ_{AB} of ρ_A, there exists an extension σ_{AB} of σ_A such that

$$F(\rho_{AB}, \sigma_{AB}) = F(\rho_A, \sigma_A). \tag{1.25}$$

The *trace distance* between states ρ and σ is defined as $\|\rho - \sigma\|_1$. One can normalize the trace distance by multiplying it by $1/2$ so that the resulting quantity lies in the interval $[0, 1]$. The trace distance obeys a direct-sum property: for an orthonormal basis $\{|x\rangle\}_x$ for an auxiliary Hilbert space $\mathcal{H}_X$, probability distributions $p(x)$ and $q(x)$, and sets $\{\rho^x\}_x$ and $\{\sigma^x\}_x$ of states in $\mathcal{D}(\mathcal{H}_B)$, which realize classical–quantum states

$$\rho_{XB} \equiv \sum_x p(x)|x\rangle\langle x|_X \otimes \rho_B^x, \tag{1.26}$$

$$\sigma_{XB} \equiv \sum_x q(x)|x\rangle\langle x|_X \otimes \sigma_B^x, \tag{1.27}$$

the following holds

$$\|\rho_{XB} - \sigma_{XB}\|_1 = \sum_x \|p(x)\rho_B^x - q(x)\sigma_B^x\|_1 . \tag{1.28}$$

The trace distance is monotone non-increasing with respect to a quantum channel $\mathcal{N} : \mathcal{T}(\mathcal{H}_A) \to \mathcal{T}(\mathcal{H}_B)$, in the sense that for all $\rho, \sigma \in \mathcal{D}(\mathcal{H}_A)$:

$$\|\mathcal{N}(\rho) - \mathcal{N}(\sigma)\|_1 \leq \|\rho - \sigma\|_1 . \tag{1.29}$$

This kind of data-processing inequality is important in converse proofs.

What is the relation between fidelity and trace distance? The following equality holds

for any two pure states $\phi, \psi \in \mathcal{D}(\mathcal{H})$:

$$\frac{1}{2}\left\|\phi-\psi\right\|_1 = \sqrt{1-F(\phi,\psi)}. \tag{1.30}$$

For any two arbitrary states $\rho, \sigma \in \mathcal{D}(\mathcal{H})$, the following inequalities hold

$$1-\sqrt{F(\rho,\sigma)} \leq \frac{1}{2}\left\|\rho-\sigma\right\|_1 \leq \sqrt{1-F(\rho,\sigma)}. \tag{1.31}$$

These inequalities are called Fuchs-van-de-Graaf inequalities, as they were established in [11] for finite-dimensional states.

We also state another two useful lemmas in converse proofs. Fano's inequality is important in the converse proofs of classical protocols. It states that for random variables X and Y with alphabets $\mathcal{X}$ and $\mathcal{Y}$, respectively, the following inequality holds

$$H(X|Y) \leq \varepsilon \log_2(|\mathcal{X}|-1) + h_2(\varepsilon), \tag{1.32}$$

where

$$\varepsilon \equiv \Pr\{X \neq Y\}, \tag{1.33}$$

$$h_2(\varepsilon) \equiv -\varepsilon \log_2 \varepsilon - (1-\varepsilon)\log_2(1-\varepsilon). \tag{1.34}$$

Observe that $\lim_{\varepsilon\to 0} h_2(\varepsilon) = 0$.

In quantum Shannon theory the Alicki–Fannes–Winter (AFW) inequality is usually what we need in the converse proofs. Let $\rho_{AB}, \sigma_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ with $\dim(\mathcal{H}_A) < \infty$, $\varepsilon \in [0,1]$, and suppose that $\left\|\rho_{AB}-\sigma_{AB}\right\|_1/2 \leq \varepsilon$. The Alicki–Fannes–Winter inequality [12, 13] is follows:

$$|H(A|B)_\rho - H(A|B)_\sigma| \leq 2\varepsilon \log_2 \dim(\mathcal{H}_A) + g(\varepsilon), \tag{1.35}$$

where

$$g(\varepsilon) \equiv (\varepsilon + 1) \log_2 (\varepsilon + 1) - \varepsilon \log_2 \varepsilon. \tag{1.36}$$

Observe that $\lim_{\varepsilon \to 0} g(\varepsilon) = 0$. If the states are classical on the first system, as in (1.26)–(1.27), and $\dim(\mathcal{H}_X) < \infty$ and $\|\rho_{XB} - \sigma_{XB}\|_1 / 2 \leq \varepsilon$, then the inequality can be strengthened to [14, Theorem 11.10.3]

$$|H(X|B)_\rho - H(X|B)_\sigma| \leq \varepsilon \log_2 \dim(\mathcal{H}_X) + g(\varepsilon). \tag{1.37}$$

Quantum entropies and information

The quantum entropy of a state $\rho \in \mathcal{D}(\mathcal{H})$ is defined as

$$H(\rho) \equiv \mathrm{Tr}\{\eta(\rho)\}, \tag{1.38}$$

where $\eta(x) = -x \log_2 x$ if $x > 0$ and $\eta(0) = 0$. It is a non-negative, concave, lower semi-continuous function on $\mathcal{D}(\mathcal{H})$ [15]. It is also not necessarily finite (see, e.g., [16]). When ρ_A is assigned to a system A, we write $H(A)_\rho \equiv H(\rho_A)$.

The quantum relative entropy $D(\rho\|\sigma)$ of $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ is defined as [17]

$$D(\rho\|\sigma) \equiv \sum_i \langle i|\rho \log_2 \rho - \rho \log_2 \sigma|i\rangle, \tag{1.39}$$

where $\{|i\rangle\}_{i=1}^{\infty}$ is an orthonormal basis of eigenvectors of the state ρ, if $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$ and $D(\rho\|\sigma) = \infty$ otherwise. The quantum relative entropy $D(\rho\|\sigma)$ is non-negative for $\rho, \sigma \in \mathcal{D}(\mathcal{H})$, and is monotone with respect to a quantum channel $\mathcal{N} : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ [18]:

$$D(\rho\|\sigma) \geq D(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)). \tag{1.40}$$

Below we summarize several important information quantities which are the optimal

rates for information tasks we will consider in this thesis.

The quantum mutual information $I(A;B)_\rho$ of a bipartite state $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is defined as [17]

$$I(A;B)_\rho = D(\rho_{AB} \| \rho_A \otimes \rho_B), \tag{1.41}$$

and obeys the bound [17]

$$I(A;B)_\rho \leq 2 \min\{H(A)_\rho, H(B)_\rho\}. \tag{1.42}$$

The coherent information $I(A\rangle B)_\rho$ of ρ_{AB} is defined as [19, 20]

$$I(A\rangle B)_\rho \equiv I(A;B)_\rho - H(A)_\rho. \tag{1.43}$$

For infinite dimensional system we need to further impose the condition $H(A)_\rho < \infty$. This expression reduces to

$$I(A\rangle B)_\rho = H(B)_\rho - H(AB)_\rho \tag{1.44}$$

if $H(B)_\rho < \infty$ [19, 20].

The mutual information of a quantum channel $\mathcal{N} : \mathcal{T}(\mathcal{H}_A) \to \mathcal{T}(\mathcal{H}_B)$ with respect to a state $\rho \in \mathcal{D}(\mathcal{H}_A)$ is defined as [19]

$$I(\rho, \mathcal{N}) \equiv I(R;B)_\omega, \tag{1.45}$$

where $\omega_{RB} \equiv (\mathrm{id}_R \otimes \mathcal{N}_{A\to B})(\psi^\rho_{RA})$ and $\psi^\rho_{RA} \in \mathcal{D}(\mathcal{H}_R \otimes \mathcal{H}_A)$ is a purification of ρ, with $\mathcal{H}_R \simeq \mathcal{H}_A$. The coherent information of a quantum channel $\mathcal{N} : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ with

respect to a state $\rho \in \mathcal{D}(\mathcal{H}_A)$ is defined as [19]

$$I_c(\rho, \mathcal{N}) \equiv I(R\rangle B)_\omega, \tag{1.46}$$

with ω_{RB} defined as above. These quantities obey a data processing inequality, which is that for a quantum channel $\mathcal{M} : \mathcal{B}(\mathcal{H}_B) \to \mathcal{B}(\mathcal{H}_C)$ and ρ and $\mathcal{N}$ as before, the following holds [19]

$$I(\rho, \mathcal{N}) \geq I(\rho, \mathcal{M} \circ \mathcal{N}), \tag{1.47}$$

$$I_c(\rho, \mathcal{N}) \geq I_c(\rho, \mathcal{M} \circ \mathcal{N}). \tag{1.48}$$

The conditional quantum mutual information (CQMI) of a finite-dimensional tripartite state ρ_{ABC} is defined as

$$I(A; B|C)_\rho \equiv H(AC)_\rho + H(BC)_\rho - H(ABC)_\rho - H(C)_\rho. \tag{1.49}$$

In the general case, it is defined as [21, 22]

$$I(A; B|C)_\rho \equiv \sup_{P_A} \left\{ I(A; BC)_{Q\rho Q} - I(A; C)_{Q\rho Q} : Q = P_A \otimes I_{BC} \right\}, \tag{1.50}$$

where the supremum is with respect to all finite-rank projections $P_A \in \mathcal{B}(\mathcal{H}_A)$ and we take the convention that $I(A; BC)_{Q\rho Q} = \lambda I(A; BC)_{Q\rho Q/\lambda}$ where $\lambda = \mathrm{Tr}\{Q\rho_{ABC}Q\}$. The above definition guarantees that many properties of CQMI in finite dimensions carry over to the general case [21, 22]. In particular, the following chain rule holds for a four-party state $\rho_{ABCD} \in \mathcal{D}(\mathcal{H}_{ABCD})$:

$$I(A; BC|D)_\rho = I(A; C|D)_\rho + I(A; B|CD)_\rho. \tag{1.51}$$

1.3 GAUSSIAN BOSONIC SYSTEMS

The counterpart of the qudit system is infinite-dimensional quantum system, in which quantum information come in a continuous form. The most common model for continuous quantum information is the quantized harmonic oscillator, which can be described by continuous variables such as position and momentum (the *quadratures*). Since it describes the propagation of electromagnetic field, continuous variable quantum systems are particularly relevant for quantum communication and quantum metrology. Continuous variables have also been considered in the field of quantum computation. One of the well-known continuous variable quantum computation model is the *measurement-based one-way quantum computation*, where the difficulty of realizing high-fidelity quantum gates is transferred to the preparation of highly entangled cluster states.

The most practically relevant and mathematically simple framework to describe the bosonic system is the Gaussian formalism. In the Gaussian framework, we only consider the Gaussian states, which are fully characterized by their mean and its second-order momentum; and Gaussian quantum channels, which are the transformations that preserve the Gaussianity. Gaussian states include the most common optical states encountered in the experiment, like coherent states, squeezed states and two-mode squeeze vacuum. The Gaussian channels, especially the phase-insensitive ones (defined later), model natural physical processes such as photon loss, photon amplification, thermalizing noise, or random kicks in phase space.

In this subsection we briefly review Gaussian states and channels (see [23, 24] for more comprehensive reviews, but note that here we mostly follow the conventions of [23]). Let

$$\hat{R} \equiv [\hat{q}_1, \ldots, \hat{q}_m, \hat{p}_1, \ldots, \hat{p}_m] \equiv [\hat{x}_1, \ldots, \hat{x}_{2m}] \tag{1.52}$$

denote a row vector of position- and momentum-quadrature operators, satisfying the

canonical commutation relations:

$$\left[\hat{R}_j, \hat{R}_k\right] = i\Omega_{j,k}, \quad \text{where} \quad \Omega \equiv \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \otimes I_m, \tag{1.53}$$

and I_m denotes the $m \times m$ identity matrix. We take the annihilation operator for the jth mode as $\hat{a}_j = (\hat{q}_j + i\hat{p}_j)/\sqrt{2}$. For an m-mode quantum oscillator, the Hamiltonian, or the Gibbs observable defined later, is given by

$$\hat{E}_m \equiv \sum_{j=1}^{m} \omega_j \hat{a}_j^\dagger \hat{a}_j, \tag{1.54}$$

where $\omega_j > 0$ is the frequency of the jth mode and $\hat{a}_j$ is the photon annihilation operator for the jth mode, so that $\hat{a}_j^\dagger \hat{a}_j$ is the photon number operator for the jth mode.

For $z \in \mathbb{R}^{2m}$, we define the unitary displacement operator $D(z) = D^\dagger(-z) \equiv \exp(i\hat{R}z)$ [1]. Displacement operators satisfy the following relation:

$$D(z)D(z') = D(z+z')^\dagger \exp(-\frac{1}{2} z^T i\Omega z'). \tag{1.55}$$

Every state $\rho \in \mathcal{D}(\mathcal{H})$ has a corresponding Wigner characteristic function, defined as

$$\chi_\rho(z) \equiv \mathrm{Tr}\{D(z)\rho\}, \tag{1.56}$$

and from which we can obtain the state ρ as

$$\rho = \int \frac{d^{2m}z}{(2\pi)^m}\, \chi_\rho(z)\, D^\dagger(\xi). \tag{1.57}$$

[1] Another common convention appears in the literature, like in [24], is to define $\hat{R} \equiv (\hat{q}_1, \ldots, \hat{q}_m, \hat{p}_1, \ldots, \hat{p}_m)^T$ and $\Omega = \oplus_{k=1}^m \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Then the displacement operator is defined as $D(\xi) = \exp(\xi i\Omega\hat{R})$. To translate back and forth between this convention and the one we adopt in this work, use the correspondence $z = -i\Omega\xi$.

A quantum state ρ is Gaussian if its Wigner characteristic function has a Gaussian form as

$$\chi_\rho(\xi) = \exp\left(-\frac{1}{4} z^T V^\rho z + i\left[\mu^\rho\right]^T z\right), \tag{1.58}$$

where μ^ρ is the $2m \times 1$ mean vector of ρ, whose entries are defined by $\mu_j^\rho \equiv \left\langle \hat{R}_j \right\rangle_\rho$ and V^ρ is the $2m \times 2m$ covariance matrix of ρ, whose entries are defined as

$$V_{j,k}^\rho \equiv \langle \{\hat{R}_j - \mu_j^\rho, \hat{R}_k - \mu_k^\rho\} \rangle. \tag{1.59}$$

The following condition holds for a valid covariance matrix: $V \geq i\Omega$, which is a manifestation of the uncertainty principle.

A thermal Gaussian state θ_β of m modes with respect to $\hat{E}_m$ from (1.54) and having inverse temperature $\beta > 0$ thus has the following form:

$$\theta_\beta = e^{-\beta \hat{E}_m} / \operatorname{Tr}\{e^{-\beta \hat{E}_m}\} \tag{1.60}$$

and has a mean vector equal to zero and a diagonal $2m \times 2m$ covariance matrix. One can calculate that the photon number in this state is equal to

$$\sum_j \frac{1}{e^{\beta \omega_j} - 1}. \tag{1.61}$$

It is also well known that thermal states can be written as a Gaussian mixture of displacement operators acting on the vacuum state:

$$\theta_\beta = \int d^{2m}\xi \; p(\xi) \; D(\xi) \left[|0\rangle\langle 0|\right]^{\otimes m} D^\dagger(\xi), \tag{1.62}$$

where $p(\xi)$ is a zero-mean, circularly symmetric Gaussian distribution. From this, it also follows that randomly displacing a thermal state in such a way leads to another thermal

state of higher temperature:

$$\theta_\beta = \int d^{2m}\xi \; q(\xi) \; D(\xi)\theta_{\beta'} D^\dagger(\xi), \tag{1.63}$$

where $\beta' \geq \beta$ and $q(\xi)$ is a particular circularly symmetric Gaussian distribution.

A $2m \times 2m$ matrix S is symplectic if it preserves the symplectic form: $S\Omega S^T = \Omega$. According to Williamson's theorem [25], there is a diagonalization of the covariance matrix V^ρ of the form,

$$V^\rho = S^\rho \left(D^\rho \oplus D^\rho\right) \left(S^\rho\right)^T, \tag{1.64}$$

where S^ρ is a symplectic matrix and $D^\rho \equiv \operatorname{diag}(\nu_1, \ldots, \nu_m)$ is a diagonal matrix of symplectic eigenvalues such that $\nu_i \geq 1$ for all $i \in \{1, \ldots, m\}$. Computing this decomposition is equivalent to diagonalizing the matrix $iV^\rho\Omega$ [26, Appendix A].

The entropy $H(\rho)$ of a quantum Gaussian state ρ is a direct function of the symplectic eigenvalues of its covariance matrix V^ρ [23]:

$$H(\rho) = \sum_{j=1}^{m} g((\nu_j - 1)/2) \equiv g(V^\rho), \tag{1.65}$$

where $g(\cdot)$ is defined in (1.36) and we have indicated a shorthand for this entropy as $g(V^\rho)$.

A Gaussian quantum channel $\mathcal{N}_{X,Y}$ from m modes to m modes has the following effect on a displacement operator $D(z)$ [23]:

$$D(z) \longmapsto D(Xz) \exp\left(-\frac{1}{4} z^T Y z + i d^T z\right), \tag{1.66}$$

where X is a real $2m \times 2m$ matrix, Y is a real $2m \times 2m$ positive semi-definite matrix, and $d \in \mathbb{R}^{2m}$, such that they satisfy

$$Y - i\Omega + iX^T\Omega X \geq 0. \tag{1.67}$$

The effect of the channel on the mean vector μ^ρ and the covariance matrix V^ρ is thus as follows:

$$\mu^\rho \longmapsto X^T \mu^\rho + d, \tag{1.68}$$

$$V^\rho \longmapsto X^T V^\rho X + Y. \tag{1.69}$$

All Gaussian channels are covariant with respect to displacement operators. That is, the following relation holds

$$\mathcal{N}_{X,Y}(D(z)\rho D^\dagger(z)) = D(X^T z)\mathcal{N}_{X,Y}(\rho)D^\dagger(X^T z). \tag{1.70}$$

Just as every quantum channel can be implemented as a unitary transformation on a larger space followed by a partial trace, so can Gaussian channels be implemented as a Gaussian unitary on a larger space with some extra modes prepared in the vacuum state, followed by a partial trace [27]. Given a Gaussian channel $\mathcal{N}_{X,Y}$ with Z such that $Y = ZZ^T$ we can find two other matrices X_E and Z_E such that there is a symplectic matrix

$$S = \begin{bmatrix} X^T & Z \\ X_E^T & Z_E \end{bmatrix}, \tag{1.71}$$

which corresponds to a Gaussian unitary transformation on a larger space. The complementary channel $\hat{\mathcal{N}}_{X_E,Y_E}$ from input to the environment then effects the following transformation on mean vectors and covariance matrices:

$$\mu^\rho \longmapsto X_E^T \mu^\rho, \tag{1.72}$$

$$V^\rho \longmapsto X_E^T V^\rho X_E + Y_E, \tag{1.73}$$

where $Y_E \equiv Z_E Z_E^T$.

A quantum Gaussian channel for which $X = X' \oplus X'$, $Y = Y' \oplus Y'$, and $d = d' \oplus d'$ is known as a *phase-insensitive Gaussian channel*, because it does not have a bias to either quadrature when applying noise to the input state. Phase-insensitive Gaussian channels can be further classified into *gauge-covariant* and *gauge-contravariant*, which are best characterized by how they transform the characteristic function:

$$\begin{cases} \chi_{\rho_{out}}(z) = \chi_{\rho_{in}}(\sqrt{\tau}z)\exp(-y|z|^2/2)\ , \text{for } \tau > 0\ , \\ \chi_{\rho_{out}}(z) = \chi_{\rho_{in}}(-\sqrt{|\tau|}z^*)\exp(-y|z|^2/2)\ , \text{for } \tau < 0\ . \end{cases} \tag{1.74}$$

In both cases, $y \geq |\tau - 1|$ is required for them to be valid Gaussian channels. The gauge-contravariant channel is also called the *phase conjugation channel*.

1.4 Outline and contributions

The rest of this thesis is organized as follows

- Chapter 2 — Energy-constrained private and quantum capacities of quantum channels. In this chapter, we rigorously consider quantum information in infinite-dimensional Hilbert space and define the concept of energy observable. We then proceed to define several energy-constrained quantum and private communication tasks. We clarify the relationship between these tasks by developing several code conversions between them. The most technical part of this chapter is to prove that the regularized energy-constrained coherent information is an achievable rate for all the tasks we defined previously. By establishing a converse bound on the private capacity, we prove the capacity — the single-letter energy-constrained coherent information — of all the tasks for degraded quantum channels. Finally, we apply our theory to phase-insensitive Gaussian channels, and thus clarify some of the folklore in this field which has been around for many years now. This is joint work with Mark Wilde and can be found in [28].

- Chapter 3 — Capacities of quantum-limited amplifier channels. In this chapter, we consider a specific example of communication over infinite-dimensional quantum system. We solve both triple trade-off capacity and broadcast capacity over quantum-limited amplifier channels. We also show that the trade-off capacity outperforms those achieved by time-sharing strategy, and the broadcast capacity outperforms those achieved by homodyne or heterodyne detections. Our main contributions are two converse proofs, which crucially rely on the recent progress on the minimum output-entropy conjecture (MOE). We present a brief review on MOE in Appendix A, in which we discuss the history of MOEs and their relations. We also consider a specific example to illustrate why MOEs are crucial to capacity proofs. Finally we introduce the entropy photon-number inequality, which encompasses all MOEs as special cases, and some recent effort to prove it. This chapter and the Appendix A are based on joint work with Mark Wilde and Saikat Guha, and can be found in [29, 30].

- Chapter 4 — Review of dynamical decoupling. In this chapter we motivate the importance of DD for the realization of fault-tolerant quantum computation. We first review the history of DD and point out a recent change of focus toward digital pulse sequences, which is of the subject of next two chapters. We then review the basic framework of DD control theory, including the toggling frame, Magnus expansion and the definition of cancellation order.

- Chapter 5 — Concatenated-projection dynamical decoupling (CPDD). In this chapter we propose a unifying framework for digital pulse sequences. We first define the projection sequences, which are the basic constituents in our framework. Projection sequence along each direction reduces the error Hamiltonian in the perpendicular directions. By concatenating different projection sequences with different order, an

arbitrary CPDD sequence can be constructed. We find that almost all previously known digital DD schemes are special cases of our CPDD framework. We can also systematically construct high-performance pulse sequences which were found previously by simulations with large overhead. And more importantly, our framework can help to understand why some pulse sequences are superior than others. This work is joint with Jonathan Dowling, which can be found in [31].

- Chapter 6 — General Walsh dynamical decoupling. In this chapter we generalize the theory of Walsh DD, which was proposed previously to protect qubits from pure-dephasing noise, to settings with general decoherence. Surprisingly we prove that GWDD is equivalent to CPDD, although they are constructed by using very different language. By leveraging this equivalence, we derive an analytical formula for the cancellation order of an arbitrary GWDD sequence and explicit procedure to calculate an upper bound on the decoupling error. Finally we find optimal a GWDD sequences which outperform the best known DD schemes in a certain regime of parameters. This work is joint with Jonathan Dowling and Lorenza Viola, which can be found in [32].

Chapter 2
Energy-constrained Private and Quantum Capacities of Quantum Channels

2.1 Introduction

We first explain the practical importance of energy-constrained quantum communication and briefly review developments along this direction. We then introduce and present our work on energy-constrained private and quantum communication [28].

Energy-constrained quantum Shannon theory

Bosonic or optical quantum channels are ubiquitous in today's telecom networks. Therefore they are the most important application arena in quantum Shannon theory. However, bosonic systems, which are modeled as quantum harmonic oscillators, live infinite dimensional Hilbert space. As a consequence, many established results in finite dimensional quantum Shannon theory break down (infinities appear) if there is no constraint on the input energy. On the other hand, although there are indeed finite results in the infinite input-energy limit, for example the quantum capacity of bosonic Gaussian channels [33, 34], they have limited applicability to realistic scenarios. Although the final energy-constrained capacity formula looks just like the formula in the finite dimensional case plus a energy constraint on the optimization, rigorous proofs of these results are highly non-trivial and usually involve more advanced tools than those used in finite dimension theory. Therefore, from both the practical and technical points of view, studying energy-constrained quantum Shannon theory is important.

It is Holevo who first established the energy-constrained classical capacity over quantum channels [35]. Later Holevo and Shirokov adapted the Bennett-Shor-Smolin-Thaplyal

theorem [36] on the entanglement-assisted classical communication over quantum channels to the energy-constrained scenario [37]. Wilde and I established the energy constrained private and quantum capacity of quantum channels in [28], which we will present in detail later. Another line of development is the continuity bound on information quantities with energy constraint. Winter modified the method of Alicki and Fannes [12] to derive a tight continuity bound for the conditional entropy of states with bounded energy [13]. Later Shirokov generalized Winter's result to any *locally almost affine* functions [38]. People also considered the continuity bounds for capacities of quantum channels [39, 40]. As an important application of these continuity bounds, recently authors of Ref. [41] established several upper bounds on the quantum capacity of thermal channels which are very near to known lower bounds.

Energy-constrained private and quantum capacities of quantum channels

The quantum capacity is essential for understanding how fast we will be able to perform distributed quantum computations between remote locations, and the private capacity is connected to the ability to generate secret key between remote locations, as in quantum key distribution (see, e.g., [42] for a review). In general, there are connections between private capacity and quantum capacity [43] (see also [44]), but the results of [45, 46, 47, 48] demonstrated that these concepts and the capacities can be very different. In fact, the most striking examples are channels for which their quantum capacity is equal to zero but their private capacity is strictly greater than zero [47, 48].

We have also seen advances related to quantum capacity of bosonic channels. Important statements, discussions, and critical steps concerning quantum capacity of single-mode quantum-limited attenuator and amplifier channels were reported in [33, 34]. In particular, these papers stated a formula for the quantum capacity of these channels, whenever infinite energy is available at the transmitter. These formulas have been supported with a proof in Theorem 8 of [49] and [50]. However, in practice, no transmitter

could ever use infinite energy to transmit quantum information, and so the results from [33, 34] have limited applicability to realistic scenarios. Given that the notion of quantum capacity itself is already somewhat removed from practice, as argued in [51], it seems that supplanting a sender and receiver with infinite energy in addition to perfect quantum computers and an infinite number of channel uses only serves to push this notion much farther away from practice. One of the main aims of our work is to continue the effort of bringing this notion closer to practice, by developing a general theory of energy-constrained quantum and private communication.

In light of the above discussion, we are thus motivated to understand both quantum and private communication over quantum channels with realistic energy constraints. Refs. [52, 53] were some of the earlier works to discuss quantum and private communication with energy constraints, in addition to other kinds of communication tasks. The more recent efforts in [54, 49, 29] have considered energy-constrained communication in more general trade-off scenarios, but as special cases, they also furnished proofs for energy-constrained quantum and private capacities of quantum-limited attenuator and amplifier channels (see [49, Theorem 8] and [29]). In more detail, let $Q(\mathcal{N}, N_S)$ and $P(\mathcal{N}, N_S)$ denote the respective quantum and private capacities of a quantum channel $\mathcal{N}$, such that the mean input photon number for each channel use cannot exceed $N_S \in [0, \infty)$. Ref. [49, Theorem 8] established that the quantum capacity of a pure-loss channel $\mathcal{L}_\eta$ with transmissivity parameter $\eta \in [0, 1]$ is equal to

$$Q(\mathcal{L}_\eta, N_S) = \max\left[g(\eta N_S) - g((1-\eta)N_S), 0\right], \tag{2.1}$$

where $g(x)$ is the entropy of a thermal state with mean photon number x, defined as

$$g(x) \equiv (x+1)\log_2(x+1) - x\log_2 x. \tag{2.2}$$

We (see (2.220)) establishes the private capacity formula for $\mathcal{L}_\eta$:

$$P(\mathcal{L}_\eta, N_S) = \max\left[g(\eta N_S) - g((1-\eta)N_S), 0\right]. \tag{2.3}$$

A special case of the results of [29] established that the quantum and private capacities of a quantum-limited amplifier channel $\mathcal{A}_\kappa$ with gain parameter $\kappa \in [1, \infty)$ are equal to

$$\begin{aligned} Q(\mathcal{A}_\kappa, N_S) &= P(\mathcal{A}_\kappa, N_S) && (2.4)\\ &= g(\kappa N_S + \kappa - 1) - g([\kappa - 1][N_S + 1]). && (2.5) \end{aligned}$$

Taking the limit as $N_S \to \infty$, these formulas respectively converge to

$$\max\left[\log_2(\eta/\left[1-\eta\right]), 0\right], \tag{2.6}$$

$$\log_2\left(\kappa/\left[\kappa - 1\right]\right), \tag{2.7}$$

which were stated in [33, 34] in the context of quantum capacity. Figure 2.1 plots the gap between the unconstrained and constrained quantum capacity formulas in (2.6) and (2.1), respectively. We would like to suggest that our contribution on this topic is timely. At the least, we think it should be a useful resource for the community of researchers working on related topics to have such a formalism and associated results written down explicitly, even though a skeptic might argue that they have been part of the folklore of quantum information theory for many years now. To support our viewpoint, we note that there have been several papers released in the past few years which suggest that energy-constrained quantum and private capacities have not been sufficiently clarified in the existing literature. For example, in [55], one of the main results contributed was a non-tight upper bound on the private capacity of a pure-loss bosonic channel, in spite of the fact that (2.3) was already part of the folklore of quantum information theory. In [56], it is stated that the

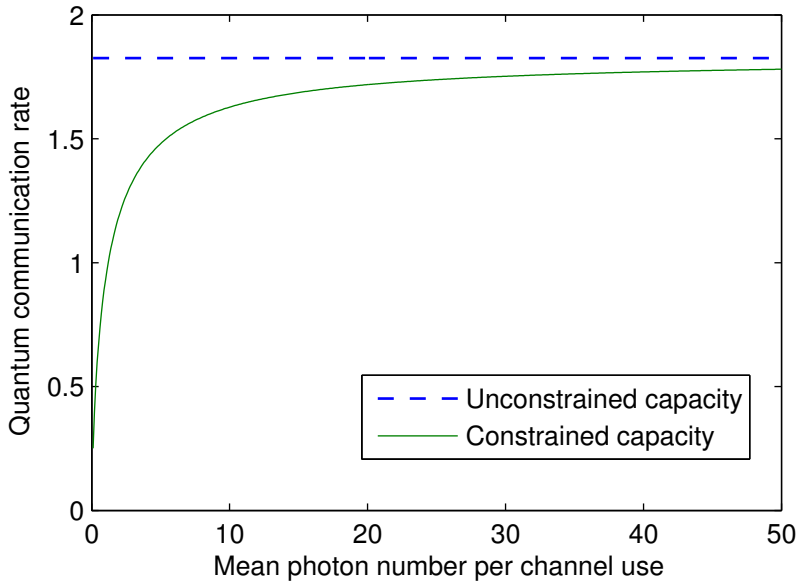

Figure 2.1: Comparison between the unconstrained (dashed line) and constrained (solid line) quantum and private capacities of the pure-loss channel for $\eta = 0.78$. For lower photon numbers, there is a large gap between these capacities.

"entropy photon-number inequality turns out to be crucial in the determining the classical capacity regions of the quantum bosonic broadcast and wiretap channels," in spite of the fact that no such argument is needed to establish the quantum or private capacity of the pure-loss channel. Similarly, it is stated in [57] that the entropy photon-number inequality "conjecture is of particular significance in quantum information theory since if it were true then it would allow one to evaluate classical capacities of various bosonic channels, e.g., the bosonic broadcast channel and the wiretap channel." Thus, it seems timely and legitimate to confirm that no such entropy photon-number inequality or minimum output-entropy conjecture is necessary in order to establish the results regarding quantum or private capacity of the pure-loss channel—the existing literature (specifically, Theorem 8 in Ref. [49] and now the previously folklore (2.220)) has established these capacities. The same is the case for the quantum-limited amplifier channel due to the results of [29]. The entropy photon-number inequality indeed implies formulas for quantum and

private capacities of the quantum-limited attenuator and amplifier channels, but it appears to be much stronger than what is actually necessary to accomplish this goal. The different proof of these formulas that we give in the present paper (see Section 2.8) is based on the monotonicity of quantum relative entropy, concavity of coherent information of degradable channels with respect to the input density operator, and covariance of Gaussian channels with respect to displacement operators.

The rest of this chapter is organized as follows. In Section 2.2, we define the concept of energy observable which is crucial for quantum information in infinite-dimensional spaces. We also give some results we will use later. We then begin our development in Section 2.3 by defining several energy-constrained communication tasks, including quantum communication with a uniform energy constraint, entanglement transmission with an average energy constraint, private communication with a uniform energy constraint, and secret key transmission with an average energy constraint. In Section 2.4, we develop several code conversions between these various communication tasks, which allow us to conclude non-trivial relations between the capacities corresponding to them. Section 2.6 proves that the regularized, energy-constrained coherent information is an achievable rate for all of the tasks, whenever the energy observable satisfies the Gibbs condition of having a well defined thermal state for all temperatures (Definition 2.9) and the channel satisfies a finite output-entropy condition (Condition 2.10). For degradable channels satisfying these conditions, we find in Section 2.7 that the single-letter energy-constrained coherent information is equal to all of the capacities. We finally apply our results to quantum Gaussian channels in Section 2.8 and recover several results already established in the literature on Gaussian quantum information. In some cases, we establish new results, like the formula for private capacity in (2.3).

As discussed earlier, to meaningfully study quantum communication over infinite dimension system, it is necessary to impose an energy constraint. Therefore, to make the concept of 'energy' rigorous, we start by defining an Hamiltonian-like operator, called energy observable (see [58, Definition 11.3]):

Definition 2.1 (Energy observable) *Let G be a positive semi-definite operator, i.e., $G \in \mathcal{SP}(\mathcal{H}_A)$. Throughout, we refer to G as an energy observable. In more detail, we define G as follows: let $\{|e_j\rangle\}_j$ be an orthonormal basis for a Hilbert space $\mathcal{H}$, and let $\{g_j\}_j$ be a sequence of non-negative real numbers bounded from below. Then the following formula*

$$G|\psi\rangle = \sum_{j=1}^{\infty} g_j |e_j\rangle\langle e_j|\psi\rangle \tag{2.8}$$

defines a self-adjoint operator G on the dense domain $\{|\psi\rangle : \sum_{j=1}^{\infty} g_j^2 \left|\langle e_j|\psi\rangle\right|^2 < \infty\}$, for which $|e_j\rangle$ is an eigenvector with corresponding eigenvalue g_j.

For a state $\rho \in \mathcal{D}(\mathcal{H}_A)$, we follow the convention [59] and define its energy as

$$\mathrm{Tr}\{G\rho\} \equiv \sup_n \mathrm{Tr}\{\Pi_n G \Pi_n \rho\}, \tag{2.9}$$

where Π_n denotes the spectral projection of G corresponding to the interval $[0, n]$. Since we always consider many channel uses, we need to define some sort of 'average' energy of the state of n input systems.

Definition 2.2 *The nth extension $\overline{G}_n$ of an energy observable G is defined as*

$$\overline{G}_n \equiv \frac{1}{n}[G \otimes I \otimes \cdots \otimes I + \cdots + I \otimes \cdots \otimes I \otimes G]. \tag{2.10}$$

We require the following proposition for some of the developments in this chapter:

Proposition 2.3 *Let $\mathcal{N}$ be a degradable quantum channel and $\hat{\mathcal{N}}$ a complementary channel for it. Let ρ_0 and ρ_1 be states and let $\rho_\lambda = \lambda\rho_0 + (1-\lambda)\rho_1$ for $\lambda \in [0,1]$. Suppose that the entropies $H(\rho_\lambda)$, $H(\mathcal{N}(\rho_\lambda))$ and $H(\hat{\mathcal{N}}(\rho_\lambda))$ are finite for all $\lambda \in [0,1]$. Then the coherent information of $\mathcal{N}$ is concave with respect to these inputs, in the sense that*

$$\lambda I_c(\rho_0, \mathcal{N}) + (1-\lambda) I_c(\rho_1, \mathcal{N}) \leq I_c(\rho_\lambda, \mathcal{N}). \tag{2.11}$$

Proof. See the proof of Proposition 1 in [28]. ■

2.3 Energy-constrained quantum and private capacities

In the subsections that follow, let $\mathcal{N} : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ denote a quantum channel, and let G be an energy observable. Let $n \in \mathbb{N}$ denote the number of channel uses, $M \in \mathbb{N}$ the size of a code, $P \in [0, \infty)$ an energy parameter, and $\varepsilon \in [0,1]$ an error parameter. In what follows, we discuss four different notions of capacity: quantum communication with a uniform energy constraint, entanglement transmission with an average energy constraint, private communication with a uniform energy constraint, and secret key transmission with an average energy constraint.

Quantum communication with a uniform energy constraint

An $(n, M, G, P, \varepsilon)$ code for quantum communication with uniform energy constraint consists of an encoding channel $\mathcal{E}^n : \mathcal{T}(\mathcal{H}_S) \to \mathcal{T}(\mathcal{H}_A^{\otimes n})$ and a decoding channel $\mathcal{D}^n : \mathcal{T}(\mathcal{H}_B^{\otimes n}) \to \mathcal{T}(\mathcal{H}_S)$, where $M = \dim(\mathcal{H}_S)$. The energy constraint is uniform, in the sense that the following bound is required to hold for all states resulting from the output of the encoding channel $\mathcal{E}^n$:

$$\operatorname{Tr}\left\{\overline{G}_n \mathcal{E}^n(\rho_S)\right\} \leq P, \tag{2.12}$$

where $\rho_S \in \mathcal{D}(\mathcal{H}_S)$. Note that

$$\operatorname{Tr}\left\{\overline{G}_n \mathcal{E}^n(\rho_S)\right\} = \operatorname{Tr}\left\{G\overline{\rho}_n\right\}, \tag{2.13}$$

where

$$\overline{\rho}_n \equiv \frac{1}{n}\sum_{i=1}^{n} \operatorname{Tr}_{A^n \setminus A_i}\{\mathcal{E}^n(\rho_S)\}. \tag{2.14}$$

due to the i.i.d. nature of the observable $\overline{G}_n$. Furthermore, the encoding and decoding channels are good for quantum communication, in the sense that for all pure states $\phi_{RS} \in \mathcal{D}(\mathcal{H}_R \otimes \mathcal{H}_S)$, where $\mathcal{H}_R$ is isomorphic to $\mathcal{H}_S$, the following entanglement fidelity criterion holds

$$F(\phi_{RS}, (\mathrm{id}_R \otimes [\mathcal{D}^n \circ \mathcal{N}^{\otimes n} \circ \mathcal{E}^n])(\phi_{RS})) \geq 1 - \varepsilon. \tag{2.15}$$

A rate R is achievable for quantum communication over $\mathcal{N}$ subject to the uniform energy constraint P if for all $\varepsilon \in (0,1)$, $\delta > 0$, and sufficiently large n, there exists an $(n, 2^{n[R-\delta]}, G, P, \varepsilon)$ quantum communication code with uniform energy constraint. The quantum capacity $Q(\mathcal{N}, G, P)$ of $\mathcal{N}$ with uniform energy constraint is equal to the supremum of all achievable rates.

Entanglement transmission with an average energy constraint

An $(n, M, G, P, \varepsilon)$ code for entanglement transmission with average energy constraint is defined very similarly as above, except that the requirements are less stringent. The energy constraint holds on average, in the sense that it need only hold for the maximally mixed state π_S input to the encoding channel $\mathcal{E}^n$:

$$\operatorname{Tr}\left\{\overline{G}_n \mathcal{E}^n(\pi_S)\right\} \leq P. \tag{2.16}$$

Furthermore, we only demand that the particular maximally entangled state $\Phi_{RS} \in \mathcal{D}(\mathcal{H}_R \otimes \mathcal{H}_S)$, defined as

$$\Phi_{RS} \equiv \frac{1}{M} \sum_{m,m'=1}^{M} |m\rangle\langle m'|_R \otimes |m\rangle\langle m'|_S, \tag{2.17}$$

is preserved with good fidelity:

$$F(\Phi_{RS}, (\mathrm{id}_R \otimes [\mathcal{D}^n \circ \mathcal{N}^{\otimes n} \circ \mathcal{E}^n])(\Phi_{RS})) \geq 1 - \varepsilon. \tag{2.18}$$

A rate R is achievable for entanglement transmission over $\mathcal{N}$ subject to the average energy constraint P if for all $\varepsilon \in (0, 1)$, $\delta > 0$, and sufficiently large n, there exists an $(n, 2^{n[R-\delta]}, G, P, \varepsilon)$ entanglement transmission code with average energy constraint. The entanglement transmission capacity $E(\mathcal{N}, G, P)$ of $\mathcal{N}$ with average energy constraint is equal to the supremum of all achievable rates.

From definitions, it immediately follows that quantum capacity with uniform energy constraint can never exceed entanglement transmission capacity with average energy constraint:

$$Q(\mathcal{N}, G, P) \leq E(\mathcal{N}, G, P). \tag{2.19}$$

In Section 2.5, we establish the opposite inequality.

Private communication with a uniform energy constraint

An $(n, M, G, P, \varepsilon)$ code for private communication consists of a set $\{\rho^m_{A^n}\}_{m=1}^M$ of quantum states, each in $\mathcal{D}(\mathcal{H}_A^{\otimes n})$, and a POVM $\{\Lambda^m_{B^n}\}_{m=1}^M$ such that

$$\mathrm{Tr}\left\{\overline{G}_n \rho^m_{A^n}\right\} \leq P, \tag{2.20}$$

$$\mathrm{Tr}\{\Lambda^m_{B^n} \mathcal{N}^{\otimes n}(\rho^m_{A^n})\} \geq 1 - \varepsilon, \tag{2.21}$$

$$\frac{1}{2} \left\| \hat{\mathcal{N}}^{\otimes n}(\rho^m_{A^n}) - \omega_{E^n} \right\|_1 \leq \varepsilon, \tag{2.22}$$

for all $m \in \{1, \ldots, M\}$, with ω_{E^n} some fixed state in $\mathcal{D}(\mathcal{H}_E^{\otimes n})$. In the above, $\hat{\mathcal{N}}$ is a channel complementary to $\mathcal{N}$. Observe that

$$\mathrm{Tr}\left\{\overline{G}_n \rho_{A^n}^m\right\} = \mathrm{Tr}\left\{G \overline{\rho}_A^m\right\}, \tag{2.23}$$

where

$$\overline{\rho}_A^m \equiv \frac{1}{n} \sum_{i=1}^{n} \mathrm{Tr}_{A^n \setminus A_i}\{\rho_{A^n}^m\}. \tag{2.24}$$

A rate R is achievable for private communication over $\mathcal{N}$ subject to uniform energy constraint P if for all $\varepsilon \in (0,1)$, $\delta > 0$, and sufficiently large n, there exists an $(n, 2^{n[R-\delta]}, G, P, \varepsilon)$ private communication code. The private capacity $P(\mathcal{N}, G, P)$ of $\mathcal{N}$ with uniform energy constraint is equal to the supremum of all achievable rates.

Secret key transmission with an average energy constraint

An $(n, M, G, P, \varepsilon)$ code for secret key transmission with average energy constraint is defined very similarly as above, except that the requirements are less stringent. The energy constraint holds on average, in the sense that it need only hold for the average input state:

$$\frac{1}{M} \sum_{m=1}^{M} \mathrm{Tr}\left\{\overline{G}_n \rho_{A^n}^m\right\} \leq P. \tag{2.25}$$

Furthermore, we only demand that the conditions in (2.21)–(2.22) hold on average:

$$\frac{1}{M} \sum_{m=1}^{M} \mathrm{Tr}\{\Lambda_{B^n}^m \mathcal{N}^{\otimes n}(\rho_{A^n}^m)\} \geq 1 - \varepsilon, \tag{2.26}$$

$$\frac{1}{M} \sum_{m=1}^{M} \frac{1}{2} \left\| \hat{\mathcal{N}}^{\otimes n}(\rho_{A^n}^m) - \omega_{E^n} \right\|_1 \leq \varepsilon, \tag{2.27}$$

with ω_{E^n} some fixed state in $\mathcal{D}(\mathcal{H}_E^{\otimes n})$.

A rate R is achievable for secret key transmission over $\mathcal{N}$ subject to the average en-

ergy constraint P if for all $\varepsilon \in (0,1)$, $\delta > 0$, and sufficiently large n, there exists an $(n, 2^{n[R-\delta]}, G, P, \varepsilon)$ secret key transmission code with average energy constraint. The secret key transmission capacity $K(\mathcal{N}, G, P)$ of $\mathcal{N}$ with average energy constraint is equal to the supremum of all achievable rates.

From definitions, it immediately follows that private capacity with uniform energy constraint can never exceed secret key transmission capacity with average energy constraint

$$P(\mathcal{N}, G, P) \leq K(\mathcal{N}, G, P). \tag{2.28}$$

In Section 2.5, we establish the opposite inequality.

2.4 Code conversions

In this section, we establish several code conversions, which allow for converting one type of code into another type of code along with some loss in the code parameters. In particular, in the forthcoming subsections, we show how to convert

1. an entanglement transmission code with an average energy constraint to a quantum communication code with a uniform energy constraint,

2. a quantum communication code with a uniform energy constraint to a private communication code with a uniform energy constraint,

3. and a secret key transmission code with an average energy constraint to a private communication code with a uniform energy constraint.

These code conversions then allow us to establish several non-trivial relations between the corresponding capacities, which we do in Section 2.5.

Entanglement transmission with an average energy constraint to quantum communication with a uniform energy constraint

In this subsection, we show how an entanglement transmission code with an average energy constraint implies the existence of a quantum communication code with a uniform energy constraint, such that there is a loss in performance in the resulting code with respect to several code parameters.

A result like this was first established in [60] and reviewed in [61, 62, 63], under the assumption that there is no energy constraint. Here we follow the proof approach available in [62, 63], but we make several modifications in order to deal with going from an average energy constraint to a uniform energy constraint.

Theorem 2.4 *For all $\delta \in (1/M, 1/2)$, the existence of an $(n, M, G, P, \varepsilon)$ entanglement transmission code with average energy constraint implies the existence of an $(n, \lfloor \delta M \rfloor, G, P/(1-2\delta), 2\sqrt{\varepsilon/[\delta - 1/M]})$ quantum communication code with uniform energy constraint.*

Proof. Suppose that an $(n, M, G, P, \varepsilon)$ entanglement transmission code with average energy constraint exists. This implies that the conditions in (2.16) and (2.18) hold. Let $\mathcal{C}^n : \mathcal{T}(\mathcal{H}_S) \to \mathcal{T}(\mathcal{H}_S)$ denote the finite-dimensional channel consisting of the encoding, communication channel, and decoding:

$$\mathcal{C}^n \equiv \mathcal{D}^n \circ \mathcal{N}^{\otimes n} \circ \mathcal{E}^n. \tag{2.29}$$

We proceed with the following algorithm:

1. Set $k = M$, $\mathcal{H}_M = \mathcal{H}_S$, and $\delta \in (1/M, 1/2)$. Suppose for now that δM is a positive integer.

2. Set $|\phi_k\rangle \in \mathcal{H}_k$ to be a state vector such that the input-output fidelity is minimized:

$$|\phi_k\rangle \equiv \arg \min_{|\phi\rangle \in \mathcal{H}_k} \langle \phi | \mathcal{C}^n(|\phi\rangle\langle\phi|)|\phi\rangle, \tag{2.30}$$

and set the fidelity F_k and energy E_k of $|\phi_k\rangle$ as follows:

$$F_k \equiv \min_{|\phi\rangle \in \mathcal{H}_k} \langle \phi | \mathcal{C}^n(|\phi\rangle\langle\phi|)|\phi\rangle \tag{2.31}$$

$$= \langle \phi_k | \mathcal{C}^n(|\phi_k\rangle\langle\phi_k|)|\phi_k\rangle, \tag{2.32}$$

$$E_k \equiv \mathrm{Tr}\{\overline{G}_n \mathcal{E}^n(|\phi_k\rangle\langle\phi_k|)\}. \tag{2.33}$$

3. Set

$$\mathcal{H}_{k-1} \equiv \mathrm{span}\{|\psi\rangle \in \mathcal{H}_k : |\langle\psi|\phi_k\rangle| = 0\}. \tag{2.34}$$

That is, $\mathcal{H}_{k-1}$ is set to the orthogonal complement of $|\phi_k\rangle$ in $\mathcal{H}_k$, so that $\mathcal{H}_k = \mathcal{H}_{k-1} \oplus \mathrm{span}\{|\phi_k\rangle\}$. Set $k := k - 1$.

4. Repeat steps 2-3 until $k = (1-\delta)\, M$ after step 3.

5. Let $|\phi_k\rangle \in \mathcal{H}_k$ be a state vector such that the input energy is maximized:

$$|\phi_k\rangle \equiv \arg \max_{|\phi\rangle \in \mathcal{H}_k} \mathrm{Tr}\{\overline{G}_n \mathcal{E}^n(|\phi\rangle\langle\phi|)\}, \tag{2.35}$$

and set the fidelity F_k and energy E_k of $|\phi_k\rangle$ as follows:

$$F_k \equiv \langle \phi_k | \mathcal{C}^n(|\phi_k\rangle\langle\phi_k|)|\phi_k\rangle \tag{2.36}$$

$$E_k \equiv \max_{|\phi\rangle \in \mathcal{H}_k} \mathrm{Tr}\{\overline{G}_n \mathcal{E}^n(|\phi\rangle\langle\phi|)\} \tag{2.37}$$

$$= \mathrm{Tr}\{\overline{G}_n \mathcal{E}^n(|\phi_k\rangle\langle\phi_k|)\}. \tag{2.38}$$

6. Set

$$\mathcal{H}_{k-1} \equiv \text{span}\{|\psi\rangle \in \mathcal{H}_k : |\langle\psi|\phi_k\rangle| = 0\}. \tag{2.39}$$

Set $k := k - 1$.

7. Repeat steps 5-6 until $k = 0$ after step 6.

The idea behind this algorithm is to successively remove minimum fidelity states from $\mathcal{H}_S$ until $k = (1-\delta)\,M$. By the structure of the algorithm and some analysis given below, we are then guaranteed for this k and lower that

$$1 - \min_{|\phi\rangle \in \mathcal{H}_k} \langle\phi|\mathcal{C}^n(|\phi\rangle\langle\phi|)|\phi\rangle \leq \varepsilon/\delta. \tag{2.40}$$

That is, the subspace $\mathcal{H}_k$ is good for quantum communication with fidelity at least $1 - \varepsilon/\delta$. After this k, we then successively remove maximum energy states from $\mathcal{H}_k$ until the algorithm terminates. Furthermore, the algorithm implies that

$$F_M \leq F_{M-1} \leq \cdots \leq F_{(1-\delta)M+1}, \tag{2.41}$$

$$E_{(1-\delta)M} \geq E_{(1-\delta)M-1} \geq \cdots \geq E_1, \tag{2.42}$$

$$\mathcal{H}_M \supseteq \mathcal{H}_{M-1} \supseteq \cdots \supseteq \mathcal{H}_1. \tag{2.43}$$

Also, $\{|\phi_k\rangle\}_{k=1}^{l}$ is an orthonormal basis for $\mathcal{H}_l$, where $l \in \{1, \ldots, M\}$.

We now analyze the result of this algorithm by employing Markov's inequality and some other tools. From the condition in (2.18) that the original code is good for entanglement transmission, we have that

$$F(\Phi_{RS}, (\text{id}_R \otimes \mathcal{C}^n)(\Phi_{RS})) \geq 1 - \varepsilon. \tag{2.44}$$

Since $\{|\phi_k\rangle\}_{k=1}^{M}$ is an orthonormal basis for $\mathcal{H}_M$, we can write

$$|\Phi\rangle_{RS} = \frac{1}{\sqrt{M}} \sum_{k=1}^{M} |\phi_k^*\rangle_R \otimes |\phi_k\rangle_S, \tag{2.45}$$

where $*$ denotes complex conjugate with respect to the basis in (2.17), and the reduced state can be written as $\Phi_S = \frac{1}{M} \sum_{k=1}^{M} |\phi_k\rangle\langle\phi_k|_S$. A consequence of [14, Exercise 9.5.1] is that

$$\begin{aligned} F(\Phi_{RS}, (\mathrm{id}_R \otimes \mathcal{C}^n)(\Phi_{RS})) &\leq \frac{1}{M} \sum_k \langle\phi_k|\mathcal{C}^n(|\phi_k\rangle\langle\phi_k|)|\phi_k\rangle \\ &= \frac{1}{M} \sum_k F_k. \end{aligned} \tag{2.46}$$

So this means that

$$\frac{1}{M} \sum_k F_k \geq 1 - \varepsilon \quad \Leftrightarrow \quad \frac{1}{M} \sum_k (1 - F_k) \leq \varepsilon. \tag{2.47}$$

Now taking K as a uniform random variable with realizations $k \in \{1, \ldots, M\}$ and applying Markov's inequality, we find that

$$\Pr_K\{1 - F_K \geq \varepsilon/\delta\} \leq \frac{\mathbb{E}_K\{1 - F_K\}}{\varepsilon/\delta} \leq \frac{\varepsilon}{\varepsilon/\delta} = \delta. \tag{2.48}$$

So this implies that $(1 - \delta)\, M$ of the F_k values are such that $F_k \geq 1 - \varepsilon/\delta$. Since they are ordered as given in (2.41), we can conclude that $\mathcal{H}_{(1-\delta)M}$ is a subspace good for quantum communication in the following sense:

$$\min_{|\phi\rangle \in \mathcal{H}_{(1-\delta)M}} \langle\phi|\mathcal{C}^n(|\phi\rangle\langle\phi|)|\phi\rangle \geq 1 - \varepsilon/\delta. \tag{2.49}$$

Now consider from the average energy constraint in (2.16) that

$$P \geq \operatorname{Tr}\left\{\overline{G}_n \mathcal{E}^n(\pi_S)\right\} \tag{2.50}$$

$$= \frac{1}{M}\sum_{k=1}^{M} \operatorname{Tr}\left\{\overline{G}_n \mathcal{E}^n(|\phi_k\rangle\langle\phi_k|_S)\right\} \tag{2.51}$$

$$= \frac{1}{M}\sum_{k=1}^{M} E_k \tag{2.52}$$

$$\geq \frac{1-\delta}{(1-\delta)\,M}\sum_{k=1}^{(1-\delta)M} E_k, \tag{2.53}$$

which we can rewrite as

$$\frac{1}{(1-\delta)\,M}\sum_{k=1}^{(1-\delta)M} E_k \leq P/\left(1-\delta\right). \tag{2.54}$$

Taking K' as a uniform random variable with realizations $k \in \{1,\ldots,(1-\delta)\,M\}$ and applying Markov's inequality, we find that

$$\Pr_{K'}\left\{E_{K'} \geq P/\left(1-2\delta\right)\right\} \leq \frac{P/\left(1-\delta\right)}{P/\left(1-2\delta\right)} \tag{2.55}$$

$$= \frac{1-2\delta}{1-\delta}. \tag{2.56}$$

Rewriting this, we find that

$$\Pr_{K'}\left\{E_{K'} \leq P/\left(1-2\delta\right)\right\} \geq 1 - \frac{1-2\delta}{1-\delta} \tag{2.57}$$

$$= \frac{\delta}{1-\delta}. \tag{2.58}$$

Thus, a fraction $\delta/\left(1-\delta\right)$ of the remaining $(1-\delta)\,M$ state vectors $|\phi_k\rangle$ are such that $E_k \leq P/\left(1-2\delta\right)$. Since they are ordered as in (2.42), this means that $\{|\phi_{\delta M}\rangle,\ldots,|\phi_1\rangle\}$ have this property.

We can then conclude that the subspace $\mathcal{H}_{\delta M}$ is such that

$$\dim(\mathcal{H}_{\delta M}) = \delta M, \tag{2.59}$$

$$\min_{|\phi\rangle \in \mathcal{H}_{\delta M}} \langle \phi | \mathcal{C}^n(|\phi\rangle\langle\phi|) | \phi \rangle \geq 1 - \varepsilon/\delta, \tag{2.60}$$

$$\max_{|\phi\rangle \in \mathcal{H}_{\delta M}} \mathrm{Tr}\{\overline{G}_n \mathcal{E}^n(|\phi\rangle\langle\phi|)\} \leq P/\left(1 - 2\delta\right). \tag{2.61}$$

Now applying Proposition 2.17 to (2.60), we can conclude that the minimum entanglement fidelity obeys the following bound:

$$\min_{|\psi\rangle \in \mathcal{H}'_{\delta M} \otimes \mathcal{H}_{\delta M}} \langle \psi | (\mathrm{id}_{\mathcal{H}'_{\delta M}} \otimes \mathcal{C}^n)(|\psi\rangle\langle\psi|) | \psi \rangle \geq 1 - 2\sqrt{\varepsilon/\delta}. \tag{2.62}$$

To finish off the proof, suppose that δM is not an integer. Then there exists a $\delta' < \delta$ such that $\delta' M = \lfloor \delta M \rfloor$ is a positive integer. By the above reasoning, there exists a code with parameters as given in (2.59)–(2.62), except with δ replaced by δ'. Then the code dimension is equal to $\lfloor \delta M \rfloor$. Using that $\delta' M = \lfloor \delta M \rfloor > \delta M - 1$, we find that $\delta' > \delta - 1/M$, which implies that $1 - 2\sqrt{\varepsilon/\delta'} > 1 - 2\sqrt{\varepsilon/[\delta - 1/M]}$. We also have that $P/\left(1 - 2\delta'\right) < P/\left(1 - 2\delta\right)$. This concludes the proof. ■

Quantum communication with a uniform energy constraint implies private communication with a uniform energy constraint

This subsection establishes that a quantum communication code with uniform energy constraint can always be converted to one for private communication with uniform energy constraint, such that there is negligible loss with respect to code parameters.

Theorem 2.5 *The existence of an $(n, M, G, P, \varepsilon)$ quantum communication code with uniform energy constraint implies the existence of an $(n, \lfloor M/2 \rfloor, G, P, 2\sqrt{\varepsilon})$ code for private communication with uniform energy constraint.*

Proof. Starting from an $(n, M, G, P, \varepsilon)$ quantum communication code with uniform energy constraint, we can use it to transmit a maximally entangled state

$$\Phi_{RS} \equiv \frac{1}{M} \sum_{m,m'=1}^{M} |m\rangle\langle m'|_R \otimes |m\rangle\langle m'|_S \tag{2.63}$$

of Schmidt rank M faithfully, by applying (2.15):

$$F(\Phi_{RS}, (\mathrm{id}_R \otimes \mathcal{D}^n \circ \mathcal{N}^{\otimes n} \circ \mathcal{E}^n)(\Phi_{RS})) \geq 1 - \varepsilon. \tag{2.64}$$

Consider that the state

$$\sigma_{RSE^n} \equiv (\mathrm{id}_R \otimes \mathcal{D}^n \circ [\mathcal{U}^{\mathcal{N}}]^{\otimes n} \circ \mathcal{E}^n)(\Phi_{RS}) \tag{2.65}$$

extends the state output from the actual protocol. By Uhlmann's theorem (see (1.25)), there exists an extension of Φ_{RS} such that the fidelity between this extension and the state σ_{RSE^n} is equal to the fidelity in (2.64). However, the maximally entangled state Φ_{RS} is "unextendible" in the sense that the only possible extension is a tensor-product state $\Phi_{RS} \otimes \omega_{E^n}$ for some state ω_{E^n}. So, putting these statements together, we find that

$$F(\Phi_{RS} \otimes \omega_{E^n}, (\mathrm{id}_R \otimes \mathcal{D}^n \circ [\mathcal{U}^{\mathcal{N}}]^{\otimes n} \circ \mathcal{E}^n)(\Phi_{RS})) \geq 1 - \varepsilon. \tag{2.66}$$

Furthermore, measuring the R and S systems locally in the Schmidt basis of Φ_{RS} only increases the fidelity, so that

$$F(\overline{\Phi}_{RS} \otimes \omega_{E^n}, (\mathrm{id}_R \otimes \overline{\mathcal{D}}^n \circ [\mathcal{U}^{\mathcal{N}}]^{\otimes n} \circ \mathcal{E}^n)(\overline{\Phi}_{RS})) \geq 1 - \varepsilon, \tag{2.67}$$

where $\overline{\mathcal{D}}^n$ denotes the concatenation of the original decoder $\mathcal{D}^n$ followed by the local

measurement:

$$\overline{\mathcal{D}}^n(\cdot) \equiv \sum_m |m\rangle\langle m|\mathcal{D}^n(\cdot)|m\rangle\langle m| \tag{2.68}$$

$$= \sum_m \mathrm{Tr}\{\mathcal{D}^{n\dagger}[|m\rangle\langle m|](\cdot)\}|m\rangle\langle m|. \tag{2.69}$$

Observe that $\{\mathcal{D}^{n\dagger}[|m\rangle\langle m|]\}_m$ is a valid POVM. Employing the inequalities in (1.31), we can conclude that

$$\frac{1}{2}\left\|\overline{\Phi}_{RS} \otimes \omega_{E^n} - (\mathrm{id}_R \otimes \overline{\mathcal{D}}^n \circ [\mathcal{U}^{\mathcal{N}}]^{\otimes n} \circ \mathcal{E}^n)(\overline{\Phi}_{RS})\right\|_1 \leq \sqrt{\varepsilon}. \tag{2.70}$$

Using the direct sum property of the trace distance from (1.28) and defining $\rho^m_{A^n} \equiv \mathcal{E}^n(|m\rangle\langle m|_S)$, we can then rewrite this as

$$\frac{1}{2M}\sum_{m=1}^{M}\left\||m\rangle\langle m|_S \otimes \omega_{E^n} - (\overline{\mathcal{D}}^n \circ [\mathcal{U}^{\mathcal{N}}]^{\otimes n})(\rho^m_{A^n})\right\|_1 \leq \sqrt{\varepsilon}. \tag{2.71}$$

Markov's inequality then guarantees that there exists a subset $\mathcal{M}'$ of $[M]$ of size $\lfloor M/2 \rfloor$ such that the following condition holds for all $m \in \mathcal{M}'$:

$$\frac{1}{2}\left\||m\rangle\langle m|_S \otimes \omega_{E^n} - (\overline{\mathcal{D}}^n \circ [\mathcal{U}^{\mathcal{N}}]^{\otimes n})(\rho^m_{A^n})\right\|_1 \leq 2\sqrt{\varepsilon}. \tag{2.72}$$

We now define the private communication code to consist of codewords $\{\rho^m_{A^n} \equiv \mathcal{E}^n(|m\rangle\langle m|_S)\}_{m\in\mathcal{M}'}$ and the decoding POVM to be

$$\{\Lambda^m_{B^n} \equiv \mathcal{D}^{n\dagger}(|m\rangle\langle m|)\}_{m\in\mathcal{M}'} \cup \left\{\Lambda^0_{B^n} \equiv \mathcal{D}^{n\dagger}\left(\sum_{m\notin\mathcal{M}'}|m\rangle\langle m|\right)\right\}. \tag{2.73}$$

Note that the energy constraint holds for all codewords

$$\operatorname{Tr}\{\overline{G}_n \rho_{A^n}^m\} \leq P, \tag{2.74}$$

due to the assumption that we start from a quantum communication code with uniform energy constraint as given in (2.12). Applying monotonicity of partial trace to (2.72) with respect to system S, we find that the following condition holds for all $m \in \mathcal{M}'$:

$$\frac{1}{2}\left\| \omega_{E^n} - \hat{\mathcal{N}}^{\otimes n}(\rho_{A^n}^m) \right\|_1 \leq 2\sqrt{\varepsilon}, \tag{2.75}$$

which gives the desired security condition in (2.22). Applying monotonicity of partial trace to (2.72) with respect to system E^n gives that

$$\frac{1}{2}\left\| |m\rangle\langle m|_S - (\overline{\mathcal{D}}^n \circ \mathcal{N}^{\otimes n})(\rho_{A^n}^m) \right\|_1 \leq 2\sqrt{\varepsilon}, \tag{2.76}$$

for all $m \in \mathcal{M}'$. Abbreviating $\Gamma_{B^n}^{m'} \equiv \mathcal{D}^{n\dagger}(|m'\rangle\langle m'|)$, consider then that for all $m \in \mathcal{M}'$

$$\begin{aligned}
&\frac{1}{2}\left\| |m\rangle\langle m|_S - (\overline{\mathcal{D}}^n \circ \mathcal{N}^{\otimes n})(\rho_{A^n}^m) \right\|_1 \\
&= \frac{1}{2}\left\| |m\rangle\langle m|_S - \sum_{m'=1}^{M} \operatorname{Tr}\{\Gamma_{B^n}^{m'} \mathcal{N}^{\otimes n}(\rho_{A^n}^m)\} |m'\rangle\langle m'| \right\|_1 \\
&= \frac{1}{2}\left\| p_e |m\rangle\langle m|_S - \sum_{m' \neq m} \operatorname{Tr}\{\Gamma_{B^n}^{m'} \mathcal{N}^{\otimes n}(\rho_{A^n}^m)\} |m'\rangle\langle m'| \right\|_1 \\
&= \frac{1}{2}\left(p_e + \sum_{m' \neq m} \operatorname{Tr}\{\Gamma_{B^n}^{m'} \mathcal{N}^{\otimes n}(\rho_{A^n}^m)\} \right) \\
&= 1 - \operatorname{Tr}\{\Lambda_{B^n}^m \mathcal{N}^{\otimes n}(\rho_{A^n}^m)\},
\end{aligned} \tag{2.77}$$

where $p_e \equiv 1 - \operatorname{Tr}\{\Lambda_{B^n}^m \mathcal{N}^{\otimes n}(\rho_{A^n}^m)\}$. Combining this equality with (2.76) gives the desired

reliable decoding condition in (2.21) for all $m \in \mathcal{M}'$

$$\mathrm{Tr}\{\Lambda^m_{B^n}\mathcal{N}^{\otimes n}(\rho^m_{A^n})\} \geq 1 - 2\sqrt{\varepsilon}. \tag{2.78}$$

Thus, we have shown that from an $(n, M, G, P, \varepsilon)$ quantum communication code with uniform energy constraint, one can realize an $(n, \lfloor M/2 \rfloor, G, P, 2\sqrt{\varepsilon})$ code for private communication with uniform energy constraint. ■

Remark 2.6 *That a quantum communication code can be easily converted to a private communication code is part of the folklore of quantum information theory. Ref. [43] proved that the unconstrained quantum capacity never exceeds the unconstrained private capacity, but we are not aware of an explicit code conversion statement of the form given in Theorem 2.5.*

Secret key transmission with an average energy constraint implies private communication with a uniform energy constraint

We finally establish that a secret key transmission code with average energy constraint can be converted to a private communication code with uniform energy constraint.

Theorem 2.7 *For $\delta \in (1/M, 1/3)$, the existence of an $(n, M, G, P, \varepsilon)$ secret key transmission code with average energy constraint implies the existence of an $(n, \lfloor \delta M \rfloor, G, P/(1-3\delta), \varepsilon/[\delta - 1/M])$ private communication code with uniform energy constraint.*

Proof. To begin with, suppose that δM is an integer. The existence of an $(n, M, G, P, \varepsilon)$ secret key transmission code with average energy constraint implies that the following three conditions hold:

$$\frac{1}{M}\sum_{m=1}^{M} E_m \leq P\,, \quad \frac{1}{M}\sum_{m=1}^{M} T_m \geq 1 - \varepsilon\,, \tag{2.79}$$

$$\frac{1}{M}\sum_{m=1}^{M} D_m \leq \varepsilon\,, \tag{2.80}$$

where

$$E_m \equiv \text{Tr}\{\overline{G}_n \rho_{A^n}^m\} \, , \tag{2.81}$$

$$T_m \equiv \text{Tr}\{\Lambda_{B^n}^m \mathcal{N}^{\otimes n}(\rho_{A^n}^m)\} \, , \tag{2.82}$$

$$D_m \equiv \frac{1}{2} \left\| \hat{\mathcal{N}}^{\otimes n}(\rho_{A^n}^m) - \omega_{E^n} \right\|_1 \, . \tag{2.83}$$

Now taking $\hat{M}$ as a uniform random variable with realizations $m \in \{1, \ldots, M\}$ and applying Markov's inequality, we have for $\delta \in (0, 1/3)$ that

$$\Pr_{\hat{M}}\{1 - T_{\hat{M}} \geq \varepsilon/\delta\} \leq \frac{\mathbb{E}_{\hat{M}}\{1 - T_{\hat{M}}\}}{\varepsilon/\delta} \leq \frac{\varepsilon}{\varepsilon/\delta} \, . \tag{2.84}$$

This implies that $(1-\delta)M$ of the T_m values are such that $T_m \geq 1 - \varepsilon/\delta$. We then rearrange the order of T_m, D_m, and E_m using a label m' such that the first $(1-\delta)M$ of the $T_{m'}$ variables satisfy the condition $T_{m'} \geq 1 - \varepsilon/\delta$. Now from (2.79), we have that

$$\varepsilon \geq \frac{1}{M} \sum_{m'=1}^{M} D_{m'} \geq \frac{1-\delta}{(1-\delta)M} \sum_{m'=1}^{(1-\delta)M} D_{m'} \, , \tag{2.85}$$

which can be rewritten as

$$\frac{1}{(1-\delta)M} \sum_{m=1}^{(1-\delta)M} D_{m'} \leq \frac{\varepsilon}{1-\delta}. \tag{2.86}$$

Now taking $\hat{M}'$ as a uniform random variable with realizations $m' \in \{1, \ldots, (1-\delta)M\}$

and applying Markov's inequality, we find that

$$\Pr_{\hat{M}'}\{D_{\hat{M}'} \geq \varepsilon/\delta\} \leq \frac{\mathbb{E}_{\hat{M}'}\{D_{\hat{M}'}\}}{\varepsilon/\delta} \tag{2.87}$$

$$\leq \frac{\varepsilon/(1-\delta)}{\varepsilon/\delta} \tag{2.88}$$

$$= \frac{\delta}{1-\delta} \, . \tag{2.89}$$

Thus a fraction $1 - [\delta/(1-\delta)] = (1-2\delta)/(1-\delta)$ of the first $(1-\delta)M$ variables $D_{m'}$ satisfy $D_{\hat{M}'} \leq \varepsilon/\delta$. Now rearrange the order of $T_{m'}$, $D_{m'}$, and $E_{m'}$ with label m'' such that the first $(1-2\delta)M$ of them satisfy

$$T_{m''} \geq 1 - \varepsilon/\delta \, , \tag{2.90}$$

$$D_{m''} \leq \varepsilon/\delta \, . \tag{2.91}$$

From (2.79), we get that

$$P \geq \frac{1}{M} \sum_{m''=1}^{M} E_{m''} \geq \frac{1-2\delta}{(1-2\delta)M} \sum_{m''=1}^{(1-2\delta)M} E_{m''} \, , \tag{2.92}$$

which can be rewritten as

$$\frac{1}{(1-2\delta)M} \sum_{m''=1}^{(1-2\delta)M} E_{m''} \leq \frac{P}{1-2\delta} \, . \tag{2.93}$$

Taking $\hat{M}''$ as a uniform random variable with realizations $m'' \in \{1, ..., (1-2\delta)M\}$ and

applying Markov's inequality, we find that

$$\Pr_{\hat{M}''}\left\{E_{\hat{M}''} \geq P/(1-3\delta)\right\} \leq \frac{\mathbb{E}_{\hat{M}''}\{E_{\hat{M}''}\}}{P/(1-\delta)} \tag{2.94}$$

$$\leq \frac{P/(1-2\delta)}{P/(1-3\delta)} \tag{2.95}$$

$$= \frac{1-3\delta}{1-2\delta}\,. \tag{2.96}$$

Thus a fraction $1-(1-3\delta)/(1-2\delta) = \delta/(1-2\delta)$ of the first $(1-2\delta)M$ variables $E_{m''}$ satisfy the condition $E_{\hat{M}''} \leq P/(1-3\delta)$. We can finally relabel $T_{m''}$, $D_{m''}$, and $E_{m''}$ with a label m''' such that the first δM of them satisfy

$$E_{m'''} \leq P/(1-3\delta)\,, \tag{2.97}$$

$$T_{m'''} \geq 1-\varepsilon/\delta\,, \tag{2.98}$$

$$D_{m'''} \leq \varepsilon/\delta\,. \tag{2.99}$$

The corresponding codewords then constitute an $(n, \delta M, G, P/(1-3\delta), \varepsilon/\delta)$ private communication code with uniform energy constraint.

To finish off the proof, suppose that δM is not an integer. Then there exists a $\delta' < \delta$ such that $\delta' M = \lfloor \delta M \rfloor$ is a positive integer. By the above reasoning, there exists a code with parameters as given in (2.97)–(2.99), except with δ replaced by δ'. Then the code size is equal to $\lfloor \delta M \rfloor$. Using that $\delta' M = \lfloor \delta M \rfloor > \delta M - 1$, we find that $\delta' > \delta - 1/M$, which implies that $1-\varepsilon/\delta' > 1-\varepsilon/[\delta - 1/M]$ and $\varepsilon/\delta' < \varepsilon/\left[\delta - 1/M\right]$. We also have that $P/\left(1-3\delta'\right) < P/\left(1-3\delta\right)$. This concludes the proof. ■

2.5 Implications of code conversions for capacities

In this brief section, we show how the various code conversions from Section 2.4 have implications for the capacities defined in Section 2.3. The main result is the following

theorem:

Theorem 2.8 *Let $\mathcal{N} : \mathcal{T}(\mathcal{H}_A) \to \mathcal{T}(\mathcal{H}_B)$ be a quantum channel, $G \in \mathcal{P}(\mathcal{H}_A)$ an energy observable, and $P \in [0, \infty)$. Then the following relations hold for the capacities defined in Section 2.3:*

$$\begin{aligned} Q(\mathcal{N}, G, P) &= E(\mathcal{N}, G, P) \\ &\leq P(\mathcal{N}, G, P) = K(\mathcal{N}, G, P). \end{aligned} \tag{2.100}$$

Proof. As a consequence of the definitions of these capacities and as remarked in (2.19) and (2.28), we have that

$$Q(\mathcal{N}, G, P) \leq E(\mathcal{N}, G, P), \tag{2.101}$$

$$P(\mathcal{N}, G, P) \leq K(\mathcal{N}, G, P). \tag{2.102}$$

So it suffices to prove the following three inequalities:

$$Q(\mathcal{N}, G, P) \geq E(\mathcal{N}, G, P), \tag{2.103}$$

$$Q(\mathcal{N}, G, P) \leq P(\mathcal{N}, G, P), \tag{2.104}$$

$$P(\mathcal{N}, G, P) \geq K(\mathcal{N}, G, P). \tag{2.105}$$

These follow from Theorems 2.4, 2.5, and 2.7, respectively. Let us establish (2.103). Suppose that R is an achievable rate for entanglement transmission with an average energy constraint. This implies the existence of a sequence of $(n, M_n, G, P, \varepsilon_n)$ codes such that

$$\liminf_{n\to\infty} \frac{1}{n} \log M_n = R, \tag{2.106}$$

$$\lim_{n\to\infty} \varepsilon_n = 0. \tag{2.107}$$

Suppose that the sequence is such that M_n is non-decreasing with n (if it is not the

case, then pick out a subsequence for which it is the case). Fix a constant $\delta \in (0, 1/2)$. Now pick n large enough such that $\delta \geq 1/M_n$. Invoking Theorem 2.4, there exists an $(n, \lfloor \delta M_n \rfloor, G, P/(1-2\delta), 2\sqrt{\varepsilon_n / [\delta - 1/M_n]})$ quantum communication code with uniform energy constraint. From the facts that

$$\liminf_{n\to\infty} \frac{1}{n} \log (\lfloor \delta M_n \rfloor) = \liminf_{n\to\infty} \frac{1}{n} \log M_n \tag{2.108}$$

$$= R, \tag{2.109}$$

$$\limsup_{n\to\infty} 2\sqrt{\varepsilon_n / [\delta - 1/M_n]} = 0, \tag{2.110}$$

we can conclude that R is an achievable rate for quantum communication with uniform energy constraint $P/(1-2\delta)$. However, since we have shown this latter statement to be true for all $\delta \in (0, 1/2)$, we can then conclude that the rate R is achievable with uniform energy constraint $\inf_{\delta \in (0,1/2)} P/(1-2\delta) = P$. So this implies (2.103). We can argue the other inequalities in (2.104) and (2.105) similarly, by applying Theorems 2.5 and 2.7, respectively. ■

2.6 Achievability of regularized, energy-constrained coherent information

The main result of this section is Theorem 2.11, which shows that the regularized energy-constrained coherent information is achievable for energy-constrained quantum communication. In order to do so, we need to restrict the energy observables and channels that we consider. We impose two arguably natural constraints: that the energy observable be a Gibbs observable as given in Definition 2.9 and that the channel have finite output entropy as given in Condition 2.10. Gibbs observables have been considered in several prior works [64, 65, 66, 67, 58, 13] as well as finite output-entropy channels [64, 65, 58].

When defining a Gibbs observable, we follow [58, Lemma 11.8] and [13, Section IV]:

Definition 2.9 (Gibbs observable) *Let G be an energy observable as given in Definition 2.1. Such an operator G is a Gibbs observable if for all $\beta > 0$, the following holds*

$$\mathrm{Tr}\{\exp(-\beta G)\} < \infty. \tag{2.111}$$

The above condition implies that a Gibbs observable G always has a finite value of the partition function $\mathrm{Tr}\{\exp(-\beta G)\}$ for all $\beta > 0$ and thus a well defined thermal state for all $\beta > 0$, given by $e^{-\beta G}/\mathrm{Tr}\{e^{-\beta G}\}$.

Condition 2.10 (Finite output entropy) *Let G be a Gibbs observable and $P \in [0, \infty)$. A quantum channel $\mathcal{N}$ satisfies the finite-output entropy condition with respect to G and P if*

$$\sup_{\rho:\mathrm{Tr}\{G\rho\}\leq P} H(\mathcal{N}(\rho)) < \infty, \tag{2.112}$$

Lemma 2.1 *Let $\mathcal{N}$ denote a quantum channel satisfying Condition 2.10, G a Gibbs observable, and $P \in [0, \infty)$. Then any complementary channel $\hat{\mathcal{N}}$ of $\mathcal{N}$ satisfies the finite-entropy condition*

$$\sup_{\rho:\mathrm{Tr}\{G\rho\}\leq P} H(\hat{\mathcal{N}}(\rho)) < \infty. \tag{2.113}$$

Proof. See the proof of condition 10 in [28]. ■

Now we state an important contribution of this work.

Theorem 2.11 *Let $\mathcal{N} : \mathcal{T}(\mathcal{H}_A) \to \mathcal{T}(\mathcal{H}_B)$ denote a quantum channel satisfying Condition 2.10, G a Gibbs observable, and $P \in [0, \infty)$. Then the energy-constrained entanglement transmission capacity $E(\mathcal{N}, G, P)$ is bounded from below by the regularized energy-constrained coherent information of the channel $\mathcal{N}$:*

$$E(\mathcal{N}, G, P) \geq \lim_{k\to\infty} \frac{1}{k} I_c(\mathcal{N}^{\otimes k}, \overline{G}_k, P),$$

where the energy-constrained coherent information of $\mathcal{N}$ is defined as

$$I_c(\mathcal{N}, G, P) \equiv \sup_{\rho:\mathrm{Tr}\{G\rho\}\leq P} H(\mathcal{N}(\rho)) - H(\hat{\mathcal{N}}(\rho)), \tag{2.114}$$

and $\hat{\mathcal{N}}$ denotes a complementary channel of $\mathcal{N}$.

Proof. The main challenge in proving this theorem is to have codes achieving the coherent information while meeting the average energy constraint. We prove the theorem by combining Klesse's technique for constructing entanglement transmission codes [62, 68] with an adaptation of Holevo's technique of approximation and constructing codes meeting an energy constraint [64, 65]. We follow their arguments very closely and show how to combine the techniques to achieve the desired result.

First, we recall what Klesse accomplished in [62] (see also the companion paper [68]). Let $\mathcal{M} : \mathcal{T}(\mathcal{H}_A) \to \mathcal{T}(\mathcal{H}_B)$ denote a quantum channel satisfying Condition 2.10 for some Gibbs observable and energy constraint, so that the receiver entropy is finite, as well as the environment entropy by Lemma 2.1. This implies that entropy-typical subspaces and sequences corresponding to these entropies are well defined and finite, a fact of which we make use. Let V denote a finite-dimensional linear subspace of $\mathcal{H}_A$. Set $L \equiv \dim(V)$, and let $\mathcal{L}$ denote a channel defined to be the restriction of $\mathcal{M}$ to states with support contained in V. Let $\{K_y\}_y$ be a set of Kraus operators for $\mathcal{M}$ and define the probability $p_Y(y)$ by

$$p_Y(y) \equiv \frac{1}{L}\,\mathrm{Tr}\{\Pi_V K_y^\dagger K_y \Pi_V\}, \tag{2.115}$$

where Π_V is a projection onto V. As discussed in [62], there is unitary freedom in the choice of the Kraus operators, and they can be chosen "diagonal," so that $\mathrm{Tr}\{\Pi_V K_y^\dagger K_x \Pi_V\} = 0$ for $x \neq y$. Let $T_Y^{n,\delta}$ denote the δ-entropy-typical set for p_Y, defined as

$$T_Y^{n,\delta} \equiv \left\{y^n : \left|-\left[\log p_{Y^n}(y^n)\right]/n - H(Y)\right| \leq \delta\right\}, \tag{2.116}$$

for integer $n \geq 1$ and real $\delta > 0$, where $p_{Y^n}(y^n) \equiv p_Y(y_1)p_Y(y_2)\cdots p_Y(y_n)$. Let $K_{y^n} \equiv K_{y_1} \otimes K_{y_2} \otimes \cdots \otimes K_{y_n}$. Now define the (trace-non-increasing) quantum operation $\mathcal{L}^{n,\delta}$ to be a map consisting of only the entropy-typical Kraus operators K_{y^n} such that $y^n \in T_Y^{n,\delta}$. The number of such Kraus operators is no larger than $2^{n[H(Y)+\delta]}$, and one can show that $H(Y) = H(\hat{\mathcal{M}}(\pi_V))$, where $\hat{\mathcal{M}}$ is a channel complementary to $\mathcal{M}$ and $\pi_V \equiv \Pi_V/L$ denotes the maximally mixed state on V [62].

One can then further reduce the quantum operation $\mathcal{L}^{n,\delta}$ to another one $\widetilde{\mathcal{L}}^{n,\delta}$ defined by projecting the output of $\mathcal{L}^{n,\delta}$ to the entropy-typical subspace of the density operator $\mathcal{L}(\pi_V) = \mathcal{M}(\pi_V)$. The entropy-typical subspace of a density operator σ with spectral decomposition $\sigma = \sum_z p_Z(z)|z\rangle\langle z|$ is defined as

$$T_\sigma^{n,\delta} \equiv \text{span}\{|z^n\rangle : |-\left[\log p_{Z^n}(z^n)\right]/n - H(\sigma)| \leq \delta\}, \tag{2.117}$$

for integer $n \geq 1$ and real $\delta > 0$. The resulting quantum operation $\widetilde{\mathcal{L}}^{n,\delta}$ is thus finite-dimensional and has a finite number of Kraus operators. We then have the following bounds argued in [62]:

$$\widetilde{L}^{n,\delta} \leq 2^{n\left[H(\hat{\mathcal{M}}(\pi_V))+\delta\right]}, \tag{2.118}$$

$$\text{Tr}\{\widetilde{\mathcal{L}}^{n,\delta}(\pi_{V^{\otimes n}})\} \geq 1 - \varepsilon_1, \tag{2.119}$$

$$\left\|\widetilde{\mathcal{L}}^{n,\delta}(\pi_{V^{\otimes n}})\right\|_2^2 \leq 2^{-n[H(\mathcal{M}(\pi_V))-3\delta]}, \tag{2.120}$$

$$F_e(C_n, \mathcal{L}^{\otimes n}) \geq F_e(C_n, \widetilde{\mathcal{L}}^{n,\delta}), \tag{2.121}$$

where $\widetilde{L}^{n,\delta}$ denotes the number of Kraus operators for $\widetilde{\mathcal{L}}^{n,\delta}$ and the second inequality inequality holds for all $\varepsilon_1 \in (0,1)$ and sufficiently large n. Note that for this latter estimate, we require the law of large numbers to hold when we only know that the entropy is finite (this can be accomplished using the technique discussed in [69]). In the last line, we have

written the entanglement fidelity of a code C_n (some subspace of $V^{\otimes n}$), which is defined as

$$F_e(C_n, \mathcal{L}^{\otimes n}) \equiv \sup_{\mathcal{R}^n} \langle \Phi_{C_n} | (\mathrm{id} \otimes [\mathcal{R}^n \circ \mathcal{L}^{\otimes n}])(\Phi_{C_n}) | \Phi_{C_n} \rangle, \tag{2.122}$$

where $|\Phi_{C_n}\rangle$ denotes a maximally entangled state built from an orthonormal basis of C_n and the optimization is with respect to recovery channels $\mathcal{R}^n$. Let $K_n \equiv \dim C_n$. From the developments in [62], the following bound holds

$$\mathbb{E}_{U_{K_n}(V^{\otimes n})}\{F_e(U_{K_n} C_n, \widetilde{\mathcal{L}}^{n,\delta})\} \geq \mathrm{Tr}\{\widetilde{\mathcal{L}}^{n,\delta}(\pi_{V^{\otimes n}})\} - \sqrt{K \widetilde{L}^{n,\delta} \left\| \widetilde{\mathcal{L}}^{n,\delta}(\pi_{V^{\otimes n}}) \right\|_2^2}, \tag{2.123}$$

where $\mathbb{E}_{U_{K_n}(V^{\otimes n})}$ denotes the expected entanglement fidelity when we apply a randomly selected unitary U_{K_n} to the codespace C_n, taking it to some different subspace of $V^{\otimes n}$. The unitary U_K is selected according to the unitarily invariant measure on the group $\mathbf{U}(V^{\otimes n})$ of unitaries acting on the subspace $V^{\otimes n}$. Combining with the inequalities in (2.118)–(2.121), we find that

$$\mathbb{E}_{U_{K_n}(V^{\otimes n})}\{F_e(U_{K_n} C_n, \mathcal{L}^{\otimes n})\} \geq 1 - \varepsilon_1 - \left[2^{-n\left[H(\mathcal{M}(\pi_V)) - \hat{\mathcal{M}}(\pi_V)) - R - 4\delta\right]} \right]^{\frac{1}{2}}, \tag{2.124}$$

where the rate R of entanglement transmission is defined as $R \equiv [\log K_n]/n$. Thus, if we choose

$$R = H(\mathcal{M}(\pi_V)) - \hat{\mathcal{M}}(\pi_V)) - 5\delta, \tag{2.125}$$

then we find that

$$\mathbb{E}_{U_{K_n}(V^{\otimes n})}\{F_e(U_{K_n} C_n, \widetilde{\mathcal{L}}^{n,\delta})\} \geq 1 - \varepsilon_1 - 2^{-n\delta/2}, \tag{2.126}$$

and we see that the RHS can be made arbitrarily close to one by taking n large enough.

We can then conclude that there exists a unitary U_{K_n}, such that the codespace defined by $U_{K_n}C_n$ achieves the same entanglement fidelity given above, implying that the rate $H(\mathcal{M}(\pi_V)) - \hat{\mathcal{M}}(\pi_V))$ is achievable for entanglement transmission over $\mathcal{M}$.

Now we apply the methods of Holevo [65] and further arguments of Klesse [62] to see how to achieve the rate given in the statement of the theorem for the channel $\mathcal{N}$ while meeting the desired energy constraint. We follow the reasoning in [65] very closely. Consider that G is a non-constant operator. Thus, the image of the convex set of all density operators under the map $\rho \to \mathrm{Tr}\{G\rho\}$ is an interval. Suppose first that P is not equal to the minimum eigenvalue of G. Then there exists a real number P' and a density operator ρ in $\mathcal{D}(\mathcal{H}_A)$ such that

$$\mathrm{Tr}\{G\rho\} \leq P' < P. \tag{2.127}$$

Let $\rho = \sum_{j=1}^{\infty} \lambda_j |j\rangle\langle j|$ be a spectral decomposition of ρ, and define

$$\rho_d \equiv \sum_{j=1}^{d} \tilde{\lambda}_j |j\rangle\langle j|, \text{ where} \tag{2.128}$$

$$\tilde{\lambda}_j \equiv \lambda_j \left(\sum_{j=1}^{d} \lambda_j \right)^{-1}. \tag{2.129}$$

Then $\|\rho - \rho_d\|_1 \to 0$ as $d \to \infty$. Let $g(j) \equiv \langle j|G|j\rangle$, so that

$$\mathrm{Tr}\{G\rho_d\} = \sum_{j=1}^{d} \tilde{\lambda}_j g(j) = P' + \varepsilon_d, \tag{2.130}$$

where $\varepsilon_d \to 0$ as $d \to \infty$. Consider the density operator $\rho_d^{\otimes m}$, and let $\Pi_d^{m,\delta}$ denote its strongly typical projector, defined as the projection onto the strongly typical subspace

$$\mathrm{span}\{|j^m\rangle : \left| N(j|j^m)/m - \tilde{\lambda}_j \right| \leq \delta\}, \tag{2.131}$$

where $|j^m\rangle \equiv |j_1\rangle \otimes \cdots \otimes |j_m\rangle$ and $N(j|j^m)$ denotes the number of appearances of the symbol j in the sequence j^m. Let

$$\pi_d^{m,\delta} \equiv \Pi_d^{m,\delta} / \operatorname{Tr}\{\Pi_d^{m,\delta}\} \tag{2.132}$$

denote the maximally mixed state on the strongly typical subspace. We then find that for positive integers m and n,

$$\begin{aligned}
&\operatorname{Tr}\left\{\overline{G}_{mn}\left(\left[\pi_d^{m,\delta}\right]^{\otimes n} - \rho_d^{\otimes mn}\right)\right\} \\
&= \operatorname{Tr}\left\{\overline{\left(\overline{G}_m\right)}_n\left(\left[\pi_d^{m,\delta}\right]^{\otimes n} - \rho_d^{\otimes mn}\right)\right\} &(2.133)\\
&= \operatorname{Tr}\left\{\overline{G}_m\left(\pi_d^{m,\delta} - \rho_d^{\otimes m}\right)\right\} \leq \delta \max_{j\in[d]} g(j), &(2.134)
\end{aligned}$$

where $[d] \equiv \{1, \ldots, d\}$ and the inequality follows from applying a bound from [70] (also called "typical average lemma" in [71]). Now we can apply the above inequality to find that

$$\begin{aligned}
&\operatorname{Tr}\left\{\overline{G}_{mn}\left[\pi_d^{m,\delta}\right]^{\otimes n}\right\} \\
&\leq \operatorname{Tr}\{\overline{G}_m \rho_d^{\otimes m}\} + \delta \max_{j\in[d]} g(j) &(2.135)\\
&= \operatorname{Tr}\{G\rho_d\} + \delta \max_{j\in[d]} g(j) &(2.136)\\
&= P' + \varepsilon_d + \delta \max_{j\in[d]} g(j). &(2.137)
\end{aligned}$$

For all d large enough, we can then find δ_0 such that the last line above is $\leq P/(1+\delta_1)$ for $\delta, \delta_1 \in (0, \delta_0]$.

The quantum coding scheme we use is that of Klesse [62] discussed previously, now setting $\mathcal{M} = \mathcal{N}^{\otimes m}$ and the subspace V to be the frequency-typical subspace of $\rho_d^{\otimes m}$, so

that $\Pi_V = \Pi_d^{m,\delta}$. Letting π_{C_n} denote the maximally mixed projector onto the codespace $C_n \subset V^{\otimes n}$, we find that [62, Section 5.3]

$$\mathbb{E}_{U_{K_n}(V^{\otimes n})}\{U_{K_n}\pi_{C_n}U_{K_n}^\dagger\} = \pi_{V^{\otimes n}} = \left[\pi_d^{m,\delta}\right]^{\otimes n}. \tag{2.138}$$

So this and the reasoning directly above imply that

$$\mathbb{E}_{U_{K_n}(V^{\otimes n})}\{\mathrm{Tr}\{\overline{G_{mn}}U_{K_n}\pi_{C_n}U_{K_n}^\dagger\}\} \leq P/(1+\delta_1), \tag{2.139}$$

for $\delta, \delta_1 \leq \delta_0$. Furthermore, from (2.126), for arbitrary $\varepsilon \in (0,1)$ and sufficiently large n, we find that

$$\mathbb{E}_{U_{K_n}(V^{\otimes n})}\{1 - F_e(U_{K_n}C_n, \mathcal{N}^{\otimes mn})\} \leq \varepsilon, \tag{2.140}$$

as long as the rate

$$R = [H(\mathcal{N}^{\otimes m}(\pi_d^{m,\delta})) - H(\hat{\mathcal{N}}^{\otimes m}(\pi_d^{m,\delta}))]/m - \delta' \tag{2.141}$$

for $\delta' > 0$. At this point, we would like to argue the existence of a code that has arbitrarily small error and meets the energy constraint. Let E_0 denote the event $1 - F_e(U_{K_n}C_n, \mathcal{N}^{\otimes mn}) \leq \sqrt{\varepsilon}$ and let E_1 denote the event $\mathrm{Tr}\{\overline{G_{mn}}U_{K_n}\pi_{C_n}U_{K_n}^\dagger\} \leq P$. We can apply the union bound

and Markov's inequality to find that

$$\begin{aligned}
&\Pr_{U_{K_n}(V^{\otimes n})}\{\overline{E_0 \cap E_1}\} \\
&= \Pr_{U_{K_n}(V^{\otimes n})}\{E_0^c \cup E_1^c\} && (2.142)\\
&\leq \Pr_{U_{K_n}(V^{\otimes n})}\{1 - F_e(U_{K_n} C_n, \mathcal{N}^{\otimes mn}) \geq \sqrt{\varepsilon}\} \\
&\quad + \Pr_{U_{K_n}(V^{\otimes n})}\left\{\mathrm{Tr}\{\overline{G_{mn}} U_{K_n} \pi_{C_n} U_{K_n}^\dagger\} \geq P\right\} && (2.143)\\
&\leq \frac{1}{\sqrt{\varepsilon}} \mathbb{E}_{U_{K_n}(V^{\otimes n})}\{1 - F_e(U_{K_n} C_n, \mathcal{N}^{\otimes mn})\} \\
&\quad + \frac{1}{P} \mathbb{E}_{U_{K_n}(V^{\otimes n})}\{\mathrm{Tr}\{\overline{G_{mn}} U_{K_n} \pi_{C_n} U_{K_n}^\dagger\}\} && (2.144)\\
&\leq \sqrt{\varepsilon} + 1/(1+\delta_1). && (2.145)
\end{aligned}$$

Since we can choose n large enough to have ε arbitrarily small, there exists such an n such that the last line is strictly less than one. This then implies the existence of a code C_n such that $F_e(C_n, \mathcal{N}^{\otimes mn}) \geq 1 - \sqrt{\varepsilon}$ and $\mathrm{Tr}\{\overline{G_{mn}} \pi_{C_n}\} \leq P$ (i.e., it has arbitrarily good entanglement fidelity and meets the average energy constraint). Furthermore, the rate achievable using this code is equal to $[H(\mathcal{N}^{\otimes m}(\pi_d^{m,\delta})) - H(\hat{\mathcal{N}}^{\otimes m}(\pi_d^{m,\delta}))]/m$. We have shown that this rate is achievable for all $\delta > 0$ and all integer $m \geq 1$. By applying the limiting argument from [70] (see also [72]), we thus have that the following is an achievable rate as well:

$$\lim_{\delta \to 0} \lim_{m \to \infty} \frac{1}{m} [H(\mathcal{N}^{\otimes m}(\pi_d^{m,\delta})) - H(\hat{\mathcal{N}}^{\otimes m}(\pi_d^{m,\delta}))] = H(\mathcal{N}(\rho_d)) - H(\hat{\mathcal{N}}(\rho_d)), \tag{2.146}$$

where $\mathrm{Tr}\{G\rho_d\} \leq P' + \varepsilon_d \leq P$. Given that both $H(\mathcal{N}(\rho_d))$ and $H(\hat{\mathcal{N}}(\rho_d))$ are finite, we can apply (1.43)–(1.46) and rewrite

$$H(\mathcal{N}(\rho_d)) - H(\hat{\mathcal{N}}(\rho_d)) = I_c(\rho_d, \mathcal{N}). \tag{2.147}$$

Finally, we take the limit $d \to \infty$ and find that

$$\liminf_{d\to\infty} I_c(\rho_d, \mathcal{N}) \geq I_c(\rho, \mathcal{N}), \tag{2.148}$$

where we have used the representation

$$I_c(\rho_d, \mathcal{N}) = I(\rho_d, \mathcal{N}) - H(\rho_d), \tag{2.149}$$

applied that the mutual information is lower semicontinuous [19, Proposition 1], the entropy H is continuous for all states σ such that $\mathrm{Tr}\{G\sigma\} < P$ (following from a variation of [58, Lemma 11.8]), and the fact that a purification $|\psi_d^\rho\rangle \equiv \sum_{j=1}^d \tilde{\lambda}_j^{1/2} |j\rangle \otimes |j\rangle$ has the convergence $\| |\psi_d^\rho\rangle\langle\psi_d^\rho| - |\psi^\rho\rangle\langle\psi^\rho| \|_1 \to 0$ as $d \to \infty$. Now since $H(\mathcal{N}(\rho))$ and $H(\hat{\mathcal{N}}(\rho))$ are each finite, we can rewrite

$$I_c(\rho, \mathcal{N}) = H(\mathcal{N}(\rho)) - H(\hat{\mathcal{N}}(\rho)). \tag{2.150}$$

We have thus proven that the rate $H(\mathcal{N}(\rho)) - H(\hat{\mathcal{N}}(\rho))$ is achievable for entanglement transmission with average energy constraint for all ρ satisfying $\mathrm{Tr}\{G\rho\} < P$.

We can extend this argument to operators ρ such that $\mathrm{Tr}\{G\rho\} = P$ by approximating them with operators $\rho_\xi = (1-\xi)\rho + \xi |e\rangle\langle e|$, where $|e\rangle$ is chosen such that $\langle e|G|e\rangle < P$. Suppose now that P is the minimum eigenvalue of G. In this case, the condition $\mathrm{Tr}\{G\rho\} \leq P$ reduces to the support of ρ being contained in the spectral projection of G corresponding to this minimum eigenvalue. The condition in Definition 2.9 implies that the eigenvalues of G have finite multiplicity, and so the support of ρ is a fixed finite-dimensional subspace. Thus we can take $\rho_d = \rho$, and we can repeat the above argument with the equality $\mathrm{Tr}\{G\rho\} = P$ holding at each step.

As a consequence, we can conclude that

$$\sup_{\mathrm{Tr}\{G\rho\}\leq P} H(\mathcal{N}(\rho)) - H(\hat{\mathcal{N}}(\rho)) \tag{2.151}$$

is achievable as well. Finally, we can repeat the whole argument for all $\rho^{(k)} \in \mathcal{D}(\mathcal{H}_A^{\otimes k})$ satisfying $\mathrm{Tr}\{\overline{G}_k \rho^{(k)}\} \leq P$, take the channel as $\mathcal{N}^{\otimes k}$, and conclude that the following rate is achievable:

$$\frac{1}{k} \sup_{\mathrm{Tr}\{\overline{G}_k \rho^{(k)}\}\leq P} H(\mathcal{N}^{\otimes k}(\rho^{(k)})) - H(\hat{\mathcal{N}}^{\otimes k}(\rho^{(k)})). \tag{2.152}$$

Taking the limit as $k \to \infty$ gives the statement of the theorem. ■

2.7 Energy-constrained quantum and private capacity of degradable channels

It is unknown how to compute the quantum and private capacities of general channels, but if they are degradable, the task simplifies considerably. That is, it is known from [73] and [74], respectively, that both the unconstrained quantum and private capacities of a degradable channel $\mathcal{N}$ are given by the following formula:

$$Q(\mathcal{N}) = P(\mathcal{N}) = \sup_{\rho} I_c(\rho, \mathcal{N}). \tag{2.153}$$

Here we prove the following theorem, which holds for the energy-constrained quantum and private capacities of a channel $\mathcal{N}$. Since we already have a lower bound, we only need to prove that the energy-constrained coherent information is also an upper bound on the private capacity.

Theorem 2.12 *Let G be a Gibbs observable and $P \in [0, \infty)$. Let a quantum channel $\mathcal{N}$ be degradable and satisfy Condition 2.10. Then the energy-constrained capacities $Q(\mathcal{N}, G, P)$, $E(\mathcal{N}, G, P)$,*

$P(\mathcal{N}, G, P)$, and $K(\mathcal{N}, G, P)$ are finite, equal, and given by the following formula:

$$\sup_{\rho:\mathrm{Tr}\{G\rho\}\leq P} H(\mathcal{N}(\rho)) - H(\hat{\mathcal{N}}(\rho)), \tag{2.154}$$

where $\hat{\mathcal{N}}$ denotes a complementary channel of $\mathcal{N}$.

Proof. That the quantity in (2.154) is finite follows directly from the assumption in Condition 2.10 and Lemma 2.1. From Theorem 2.8, we have that

$$\begin{aligned} Q(\mathcal{N}, G, P) &= E(\mathcal{N}, G, P) \\ &\leq P(\mathcal{N}, G, P) = K(\mathcal{N}, G, P). \end{aligned} \tag{2.155}$$

Theorem 2.11 implies that the rate in (2.154) is achievable. So this gives that

$$\sup_{\rho:\mathrm{Tr}\{G\rho\}\leq P} H(\mathcal{N}(\rho)) - H(\hat{\mathcal{N}}(\rho)) \leq Q(\mathcal{N}, G, P) = E(\mathcal{N}, G, P). \tag{2.156}$$

To establish the theorem, it thus suffices to prove the following converse inequality

$$K(\mathcal{N}, G, P) \leq \sup_{\rho:\mathrm{Tr}\{G\rho\}\leq P} H(\mathcal{N}(\rho)) - H(\hat{\mathcal{N}}(\rho)). \tag{2.157}$$

To do so, we make use of several ideas from [43, 73, 74, 75]. Consider an $(n, M, G, P, \varepsilon)$ code for secret key transmission with an average energy constraint, as described in Section 2.3.4. Using such a code, we take a uniform distribution over the codewords, and the state resulting from an isometric extension of the channel is as follows:

$$\sigma_{\hat{M}B^nE^n} \equiv \frac{1}{M}\sum_{m=1}^{M} |m\rangle\langle m|_{\hat{M}} \otimes [\mathcal{U}^{\mathcal{N}}]^{\otimes n}(\rho^m_{A^n}). \tag{2.158}$$

Now consider that each codeword in such a code has a spectral decomposition as follows:

$$\rho_{A^n}^m \equiv \sum_{l=1}^{\infty} p_{L|\hat{M}}(l|m)|\psi^{l,m}\rangle\langle\psi^{l,m}|_{A^n}, \tag{2.159}$$

for a probability distribution $p_{L|\hat{M}}$ and some orthonormal basis $\{|\psi^{l,m}\rangle_{A^n}\}_l$ for $\mathcal{H}_{A^n}$. Then the state $\sigma_{\hat{M}B^nE^n}$ has the following extension:

$$\sigma_{L\hat{M}B^nE^n} \equiv \frac{1}{M}\sum_{m=1}^{M}\sum_{l=1}^{\infty} p_{L|\hat{M}}(l|m)|l\rangle\langle l|_L \otimes |m\rangle\langle m|_{\hat{M}} \otimes [\mathcal{U}^{\mathcal{N}}]^{\otimes n}(|\psi^{l,m}\rangle\langle\psi^{l,m}|_{A^n}). \tag{2.160}$$

We can also define the state after the decoding measurement acts as

$$\sigma_{L\hat{M}M'E^n} \equiv \frac{1}{M}\sum_{m,m'=1}^{M}\sum_{l=1}^{\infty} p_{L|\hat{M}}(l|m)|l\rangle\langle l|_L \otimes |m\rangle\langle m|_{\hat{M}} \otimes \mathrm{Tr}_{B^n}\{\Lambda_{B^n}^{m'}[\mathcal{U}^{\mathcal{N}}]^{\otimes n}(|\psi^{l,m}\rangle\langle\psi^{l,m}|_{A^n})\} \otimes |m'\rangle\langle m'|_{M'}. \tag{2.161}$$

Let $\overline{\rho}_A$ denote the average single-channel input state, defined as

$$\overline{\rho}_A \equiv \frac{1}{Mn}\sum_{m=1}^{M}\sum_{i=1}^{n} \mathrm{Tr}_{A^n\setminus A_i}\{\rho_{A^n}^m\}. \tag{2.162}$$

Applying the partial trace and the assumption in (2.25), it follows that

$$\mathrm{Tr}\{G\overline{\rho}_A\} = \frac{1}{M}\sum_{m=1}^{M}\mathrm{Tr}\{\overline{G}_n\rho_{A^n}^m\} \leq P. \tag{2.163}$$

Let $\overline{\sigma}_B$ denote the average single-channel output state:

$$\overline{\sigma}_B \equiv \mathcal{N}(\overline{\rho}_A) = \frac{1}{n}\sum_{i=1}^{n}\mathrm{Tr}_{B^n\setminus B_i}\{\sigma_{B^n}\}, \tag{2.164}$$

and let $\overline{\sigma}_E$ denote the average single-channel environment state:

$$\overline{\sigma}_E \equiv \hat{\mathcal{N}}(\overline{\rho}_A) = \frac{1}{n}\sum_{i=1}^{n} \mathrm{Tr}_{E^n \setminus E_i}\{\sigma_{E^n}\}. \tag{2.165}$$

It follows from non-negativity, subadditivity of entropy, concavity of entropy, (2.163), and the assumption that G is a Gibbs observable that

$$0 \leq H\left(\frac{1}{M}\sum_{m=1}^{M} \rho_{A^n}^m\right) \tag{2.166}$$

$$\leq \sum_{i=1}^{n} H\left(\frac{1}{M}\sum_{m=1}^{M} \mathrm{Tr}_{A^n \setminus A_i}\{\rho_{A^n}^m\}\right) \tag{2.167}$$

$$\leq nH(\overline{\rho}_A) < \infty. \tag{2.168}$$

Similar reasoning but applying Condition 2.10 implies that

$$0 \leq H(B^n)_\sigma \leq \sum_{i=1}^{n} H(B_i)_\sigma \leq nH(B)_{\overline{\sigma}} < \infty. \tag{2.169}$$

Similar reasoning but applying Lemma 2.1 implies that

$$0 \leq H(E^n)_\sigma \leq \sum_{i=1}^{n} H(E_i)_\sigma \leq nH(E)_{\overline{\sigma}} < \infty. \tag{2.170}$$

Furthermore, the entropy $H(\hat{M})_\sigma = \log_2 M$ because the reduced state σ_M is maximally mixed with dimension equal to M.

Our analysis makes use of several other entropic quantities, each of which we need to argue is finitely bounded from above and below and thus can be added or subtracted at will in our analysis. The quantities involved are as follows, along with bounds for them

[17, 20, 21]:

$$0 \leq I(\hat{M}; B^n)_\sigma \leq \min\{\log_2 M, nH(B)_{\overline{\sigma}}\}, \tag{2.171}$$

$$0 \leq I(\hat{M}; E^n)_\sigma \leq \min\{\log_2 M, nH(E)_{\overline{\sigma}}\}, \tag{2.172}$$

$$0 \leq H(\hat{M}|E^n)_\sigma \leq \log_2 M, \tag{2.173}$$

as well as

$$0 \leq I(\hat{M}L; B^n)_\sigma,\ I(L; B^n|\hat{M})_\sigma, H(B^n|L\hat{M})_\sigma \leq nH(B)_{\overline{\sigma}}, \tag{2.174}$$

and

$$0 \leq I(\hat{M}L; E^n)_\sigma,\ I(L; E^n|\hat{M})_\sigma, H(E^n|L\hat{M})_\sigma \leq nH(E)_{\overline{\sigma}}. \tag{2.175}$$

We now proceed with the converse proof:

$$\log_2 M = H(\hat{M})_\sigma \tag{2.176}$$

$$= I(\hat{M}; M')_\sigma + H(\hat{M}|M')_\sigma \tag{2.177}$$

$$\leq I(\hat{M}; M')_\sigma + h_2(\varepsilon) + \varepsilon \log_2(M-1) \tag{2.178}$$

$$\leq I(\hat{M}; B^n)_\sigma + h_2(\varepsilon) + \varepsilon \log_2 M. \tag{2.179}$$

The first equality follows because the entropy of a uniform distribution is equal to the logarithm of its cardinality. The second equality is an identity. The first inequality follows from applying Fano's inequality in (1.32) to the condition in (2.26). The second inequality follows from applying the Holevo bound [76, 77]. The direct sum property of the trace distance and the security condition in (2.27) imply that

$$\frac{1}{2}\left\|\sigma_{\hat{M}E^n} - \pi_{\hat{M}} \otimes \omega_{E^n}\right\|_1 = \frac{1}{M}\sum_{m=1}^{M} \frac{1}{2}\left\|\hat{\mathcal{N}}^{\otimes n}(\rho_{A^n}^m) - \omega_{E^n}\right\|_1 \leq \varepsilon, \tag{2.180}$$

which, by the AFW inequality in (1.37) for classical–quantum states, means that

$$\left| H(\hat{M}|E^n)_{\pi\otimes\omega} - H(\hat{M}|E^n)_{\sigma} \right| \leq \varepsilon \log_2(M) + g(\varepsilon). \tag{2.181}$$

But

$$\begin{aligned} & H(\hat{M}|E^n)_{\pi\otimes\omega} - H(\hat{M}|E^n)_{\sigma} \\ & = H(\hat{M})_{\pi} - H(\hat{M}|E^n)_{\sigma} && (2.182) \\ & = H(\hat{M})_{\sigma} - H(\hat{M}|E^n)_{\sigma} && (2.183) \\ & = I(\hat{M};E^n)_{\sigma}, && (2.184) \end{aligned}$$

so then

$$I(\hat{M};E^n)_{\sigma} \leq \varepsilon \log_2(M) + g(\varepsilon). \tag{2.185}$$

Returning to (2.179) and inserting (2.185), we find that

$$\log_2 M \leq I(\hat{M};B^n)_{\sigma} - I(\hat{M};E^n)_{\sigma} + 2\varepsilon \log_2 M + h_2(\varepsilon) + g(\varepsilon). \tag{2.186}$$

We now focus on bounding the term $I(\hat{M};B^n)_\sigma - I(\hat{M};E^n)_\sigma$:

$$\begin{aligned}
&I(\hat{M};B^n)_\sigma - I(\hat{M};E^n)_\sigma \\
&= I(\hat{M}L;B^n)_\sigma - I(L;B^n|\hat{M})_\sigma \\
&\quad - \left[I(\hat{M}L;E^n)_\sigma - I(L;E^n|\hat{M})_\sigma\right] && (2.187)\\
&= I(\hat{M}L;B^n)_\sigma - I(\hat{M}L;E^n)_\sigma \\
&\quad - \left[I(L;B^n|\hat{M})_\sigma - I(L;E^n|\hat{M})_\sigma\right] && (2.188)\\
&\leq I(\hat{M}L;B^n)_\sigma - I(\hat{M}L;E^n)_\sigma && (2.189)\\
&= H(B^n)_\sigma - H(B^n|L\hat{M})_\sigma \\
&\quad - \left[H(E^n)_\sigma - H(E^n|L\hat{M})_\sigma\right] && (2.190)\\
&= H(B^n)_\sigma - H(B^n|L\hat{M})_\sigma \\
&\quad - \left[H(E^n)_\sigma - H(B^n|L\hat{M})_\sigma\right] && (2.191)\\
&= H(B^n)_\sigma - H(E^n)_\sigma. && (2.192)
\end{aligned}$$

The first equality follows from the chain rule for mutual information. The second equality follows from a rearrangement. The first inequality follows from the assumption of degradability of the channel, which implies that Bob's mutual information is never smaller than Eve's: $I(L;B^n|\hat{M})_\sigma \geq I(L;E^n|\hat{M})_\sigma$. The third equality follows from definitions. The fourth equality follows because the marginal entropies of a pure state are equal, i.e.,

$$\begin{aligned}
&H(B^n|L\hat{M})_\sigma \\
&= \frac{1}{M}\sum_{l,m} p_{L|\hat{M}}(l|m) H(\mathrm{Tr}_{E^n}\{[\mathcal{U}^{\mathcal{N}}]^{\otimes n}(|\psi^{l,m}\rangle\langle\psi^{l,m}|_{A^n})\}) \\
&= \frac{1}{M}\sum_{l,m} p_{L|\hat{M}}(l|m) H(\mathrm{Tr}_{B^n}\{[\mathcal{U}^{\mathcal{N}}]^{\otimes n}(|\psi^{l,m}\rangle\langle\psi^{l,m}|_{A^n})\}) \\
&= H(E^n|L\hat{M})_\sigma. && (2.193)
\end{aligned}$$

Continuing, we have that

$$
\begin{aligned}
(2.192) &= H(B_1)_\sigma - H(E_1)_\sigma + H(B_2 \cdots B_n)_\sigma \\
&\quad - H(E_1 \cdots E_n)_\sigma \\
&\quad - [I(B_1; B_2 \cdots B_n)_\sigma - I(E_1; E_2 \cdots E_n)_\sigma] && (2.194) \\
&\leq H(B_1)_\sigma - H(E_1)_\sigma \\
&\quad + H(B_2 \cdots B_n)_\sigma - H(E_1 \cdots E_n)_\sigma && (2.195) \\
&\leq \sum_{i=1}^{n} H(B_i)_\sigma - H(E_i)_\sigma && (2.196) \\
&\leq n \left[H(B)_{\mathcal{U}(\overline{\rho})} - H(E)_{\mathcal{U}(\overline{\rho})} \right] && (2.197) \\
&\leq n \left[\sup_{\rho : \mathrm{Tr}\{G\rho\} \leq P} H(\mathcal{N}(\rho)) - H(\hat{\mathcal{N}}(\rho)) \right]. && (2.198)
\end{aligned}
$$

The first equality follows by exploiting the definition of mutual information. The first inequality follows from the assumption of degradability, which implies that $I(B_1; B_2 \cdots B_n)_\sigma \geq I(E_1; E_2 \cdots E_n)_\sigma$. The second inequality follows by iterating the argument. The third inequality follows from the concavity of the coherent information for degradable channels (Proposition 2.3), with $\overline{\rho}_A$ defined as in (2.162) and satisfying (2.163). Thus, the final inequality follows because we can optimize the coherent information with respect all density operators satisfying the energy constraint.

Putting everything together and assuming that $\varepsilon < 1/2$, we find the following bound for all $(n, M, G, P, \varepsilon)$ private communication codes:

$$
(1 - 2\varepsilon) \frac{1}{n} \log_2 M - \frac{1}{n} \left[h_2(\varepsilon) + g(\varepsilon) \right] \leq \sup_{\rho : \mathrm{Tr}\{G\rho\} \leq P} H(\mathcal{N}(\rho)) - H(\hat{\mathcal{N}}(\rho)). \tag{2.199}
$$

Now taking the limit as $n \to \infty$ and then as $\varepsilon \to 0$, we can conclude the inequality in (2.157). This concludes the proof. ■

2.8 APPLICATION TO GAUSSIAN QUANTUM CHANNELS

We first shown that, under certain conditions, the energy-constrained coherent information is optimized by a thermal state.

Theorem 2.13 *Let G be a Gibbs observable and $P \in [0, \infty)$. Let $\mathcal{N} : \mathcal{T}(\mathcal{H}_A) \to \mathcal{T}(\mathcal{H}_B)$ be a degradable quantum channel satisfying Condition 2.10. Let θ_β denote the thermal state of G, as in* (1.60)*, satisfying $\mathrm{Tr}\{G\theta_\beta\} = P$ for some $\beta > 0$. Suppose that $\mathcal{N}$ and a complementary channel $\hat{\mathcal{N}} : \mathcal{T}(\mathcal{H}_A) \to \mathcal{T}(\mathcal{H}_E)$ are Gibbs preserving, in the sense that there exist $\beta_1, \beta_2 > 0$ such that*

$$\mathcal{N}(\theta_\beta) = \theta_{\beta_1}, \qquad \hat{\mathcal{N}}(\rho_\beta) = \theta_{\beta_2}. \tag{2.200}$$

Set

$$P_1 \equiv \mathrm{Tr}\{G\mathcal{N}(\theta_\beta)\}, \qquad P_2 \equiv \mathrm{Tr}\{G\hat{\mathcal{N}}(\theta_\beta)\}. \tag{2.201}$$

Suppose further that $\mathcal{N}$ and $\hat{\mathcal{N}}$ are such that, for all input states ρ such that $\mathrm{Tr}\{G\rho\} = P$, the output energies satisfy

$$\mathrm{Tr}\{G\mathcal{N}(\rho)\} \leq P_1, \qquad \mathrm{Tr}\{G\hat{\mathcal{N}}(\rho)\} \geq P_2. \tag{2.202}$$

Then the function

$$\sup_{\mathrm{Tr}\{G\rho\}=P} H(\mathcal{N}(\rho)) - H(\hat{\mathcal{N}}(\rho)), \tag{2.203}$$

is optimized by the thermal state θ_β.

Proof. Let $\mathcal{D} : \mathcal{T}(\mathcal{H}_B) \to \mathcal{T}(\mathcal{H}_E)$ be a degrading channel such that $\mathcal{D} \circ \mathcal{N} = \hat{\mathcal{N}}$. Consider a state ρ such that $\mathrm{Tr}\{G\rho\} = P$. The monotonicity of quantum relative entropy with respect

to quantum channels (see (1.40)) implies that

$$D(\mathcal{N}(\rho)\|\mathcal{N}(\theta_\beta)) \geq D((\mathcal{D}\circ\mathcal{N})(\rho)\|(\mathcal{D}\circ\mathcal{N})(\theta_\beta)) \tag{2.204}$$

$$= D(\hat{\mathcal{N}}(\rho)\|\hat{\mathcal{N}}(\theta_\beta)). \tag{2.205}$$

By the assumption of the theorem, this means that

$$D(\mathcal{N}(\rho)\|\theta_{\beta_1}) \geq D(\hat{\mathcal{N}}(\rho)\|\theta_{\beta_2}), \tag{2.206}$$

where β_1 and β_2 are such that $\mathrm{Tr}\{G\theta_{\beta_1}\} = P_1$ and $\mathrm{Tr}\{G\theta_{\beta_2}\} = P_2$. After a rewriting using definitions, the inequality above becomes

$$\mathrm{Tr}\{\hat{\mathcal{N}}(\rho)\log\theta_{\beta_2}\} - \mathrm{Tr}\{\mathcal{N}(\rho)\log\theta_{\beta_1}\} \geq H(\mathcal{N}(\rho)) - H(\hat{\mathcal{N}}(\rho)). \tag{2.207}$$

Set $Z_1 \equiv \mathrm{Tr}\{e^{-\beta_1 G}\}$ and $Z_2 \equiv \mathrm{Tr}\{e^{-\beta_2 G}\}$. We can then rewrite the upper bound as

$$\begin{aligned}
&\mathrm{Tr}\{\hat{\mathcal{N}}(\rho)\log\theta_{\beta_2}\} - \mathrm{Tr}\{\mathcal{N}(\rho)\log\theta_{\beta_1}\} \\
&= \mathrm{Tr}\{\hat{\mathcal{N}}(\rho)\log\left[e^{-\beta_2 G}/Z_2\right]\} \\
&\quad - \mathrm{Tr}\{\mathcal{N}(\rho)\log\left[e^{-\beta_1 G}/Z_1\right]\} && (2.208)\\
&= \log\left[Z_1/Z_2\right] - \beta_2\,\mathrm{Tr}\{G\hat{\mathcal{N}}(\rho)\} + \beta_1\,\mathrm{Tr}\{G\mathcal{N}(\rho)\} && (2.209)\\
&\leq \log\left[Z_1/Z_2\right] - \beta_2 P_2 + \beta_1 P_1. && (2.210)
\end{aligned}$$

Thus, we have established a uniform upper bound on the coherent information of states subject to the constraints given in the theorem:

$$H(\mathcal{N}(\rho)) - H(\hat{\mathcal{N}}(\rho)) \leq \log\left[Z_1/Z_2\right] - \beta_2 P_2 + \beta_1 P_1. \tag{2.211}$$

This bound is saturated when we choose the input $\rho = \theta_\beta$, where β is such that $\mathrm{Tr}\{G\theta_\beta\} = P$, because

$$\log\left[Z_1/Z_2\right] - \beta_2 P_2 + \beta_1 P_1 = H(\mathcal{N}(\theta_\beta)) - H(\hat{\mathcal{N}}(\theta_\beta)). \tag{2.212}$$

This concludes the proof. ■

Remark 2.14 *Note that we can also conclude that $P_1 \geq P_2$ for channels satisfying the hypotheses of the above theorem because the channel is degradable, implying that $H(\theta_{\beta_1}) \geq H(\theta_{\beta_2})$, and the entropy of a thermal state is a strictly increasing function of the energy (and thus invertible) [13, Proposition 10].*

The main result of this section is the following theorem, which gives an explicit expression for the energy-constrained capacities of all phase-insensitive degradable Gaussian channels that satisfy the conditions of Theorem 2.13 for all $\beta > 0$:

Theorem 2.15 *Let $\mathcal{N}_{X,Y}$ be a phase-insensitive degradable Gaussian channel, having a dilation of the form in* (1.71)*. Suppose that $\mathcal{N}_{X,Y}$ satisfies the conditions of Theorem 2.13 for all $\beta > 0$. Then its energy-constrained capacities $Q(\mathcal{N}_{X,Y}, \hat{E}_m, P)$, $E(\mathcal{N}_{X,Y}, \hat{E}_m, P)$, $P(\mathcal{N}_{X,Y}, \hat{E}_m, P)$, and $K(\mathcal{N}_{X,Y}, \hat{E}_m, P)$ are equal and given by the following formula:*

$$g(X^T V^{\theta_\beta} X + Y) - g(X_E^T V^{\theta_\beta} X_E + Y_E), \tag{2.213}$$

where θ_β is a thermal state of mean photon number P.

Proof. Since the channel is degradable, satisfies Condition 2.10, and $\hat{E}_m$ is a Gibbs observable, Theorem 2.12 applies and these capacities are given by the following formula:

$$\sup_{\rho:\mathrm{Tr}\{\hat{E}_m\rho\}\leq P} H(\mathcal{N}_{X,Y}(\rho)) - H(\hat{\mathcal{N}}_{X_E,Y_E}(\rho)). \tag{2.214}$$

By assumption, the channel satisfies the conditions of Theorem 2.13 as well for all $\beta > 0$, so that the following function is optimized by a thermal state θ_β of mean photon number P:

$$\sup_{\rho:\mathrm{Tr}\{\hat{E}_m \rho\}=P} H(\mathcal{N}_{X,Y}(\rho)) - H(\hat{\mathcal{N}}_{X_E,Y_E}(\rho)) = H(\mathcal{N}_{X,Y}(\theta_\beta)) - H(\hat{\mathcal{N}}_{X_E,Y_E}(\theta_\beta)). \tag{2.215}$$

It thus remains to prove that $H(\mathcal{N}_{X,Y}(\theta_\beta)) - H(\hat{\mathcal{N}}_{X_E,Y_E}(\theta_\beta))$ is increasing with decreasing β. This follows from the covariance property in (1.70), the concavity of coherent information in the input for degradable channels (Proposition 2.3), and the fact that thermal states can be realized by random Gaussian displacements of thermal states with lower temperature. Consider that

$$\begin{aligned}
&H(\mathcal{N}_{X,Y}(\theta_{\beta'})) - H(\hat{\mathcal{N}}_{X_E,Y_E}(\theta_{\beta'})) \\
&= \int d^{2m}\xi \; q(\xi) \; \left[H(\mathcal{N}_{X,Y}(\theta_{\beta'})) - H(\hat{\mathcal{N}}_{X_E,Y_E}(\theta_{\beta'})) \right] && (2.216) \\
&= \int d^{2m}\xi \; q(\xi) \; \Big[H(D(X\xi)\mathcal{N}_{X,Y}(\theta_{\beta'})D^\dagger(X\xi)) \\
&\qquad - H(D(X_E\xi)\hat{\mathcal{N}}_{X_E,Y_E}(\theta_{\beta'})D^\dagger(X_E\xi)) \Big] && (2.217) \\
&= \int d^{2m}\xi \; q(\xi) \; \Big[H(\mathcal{N}_{X,Y}(D(\xi)\theta_{\beta'}D^\dagger(\xi))) \\
&\qquad - H(\hat{\mathcal{N}}_{X_E,Y_E}(D(\xi)\theta_{\beta'}D^\dagger(\xi))) \Big] && (2.218) \\
&\leq H(\mathcal{N}_{X,Y}(\theta_\beta)) - H(\hat{\mathcal{N}}_{X_E,Y_E}(\theta_\beta)). && (2.219)
\end{aligned}$$

The first equality follows by placing a probability distribution in front, and the second follows from the unitary invariance of quantum entropy. The third equality follows from the covariance property of quantum Gaussian channels, given in (1.70). The inequality follows because degradable channels are concave in the input state (Proposition 2.3) and from (1.63). ■

Special cases: Pure-loss and quantum-limited amplifier channels

We can now discuss some special cases of the above result, some of which have already been known in the literature. Suppose that the channel is a single-mode pure-loss channel $\mathcal{L}_\eta$, where $\eta \in [1/2, 1]$ characterizes the average fraction of photons that make it through the channel from sender to receiver [1]. In this case, the channel has $X = \sqrt{\eta} I_2$ and $Y = (1-\eta) I_2$. We take the Gibbs observable to be the photon-number operator $\hat{a}^\dagger \hat{a}$ and the energy constraint to be $N_S \in [0, \infty)$. Such a channel is degradable [78] and was conjectured [53] to have energy-constrained quantum and private capacities equal to

$$g(\eta N_S) - g((1-\eta) N_S). \tag{2.220}$$

This conjecture was proven for the quantum capacity in [49, Theorem 8], and the present paper establishes the statement for private capacity. This was argued by exploiting particular properties of the g function (established in great detail in [79]) to show that the thermal state input is optimal for any fixed energy constraint. Here we can see this latter result as a consequence of the more general statements in Theorems 2.13 and 2.15, which are based on the monotonicity of relative entropy and other properties of this channel, such as covariance and degradability. Taking the limit $N_S \to \infty$, the formula in (2.220) converges to

$$\log_2(\eta/[1-\eta]), \tag{2.221}$$

which is consistent with the formula stated in [34].

Suppose that the channel is a single-mode quantum-limited amplifier channel $\mathcal{A}_\kappa$ of gain $\kappa \geq 1$. In this case, the channel has $X = \sqrt{\kappa} I_2$ and $Y = (\kappa - 1) I_2$. Again we take the energy operator and constraint as above. This channel is degradable [78] and was

[1] We do not consider transmissivities $\eta \in [0, 1/2]$ because the quantum capacity vanishes in this range since the channel becomes antidegradable.

recently proven [29] to have energy-constrained quantum and private capacity equal to

$$g(\kappa N_S + \kappa - 1) - g([\kappa - 1]\,[N_S + 1]). \tag{2.222}$$

The result was established by exploiting particular properties of the g function in addition to other arguments. However, we can again see this result as a consequence of the more general statements given in Theorems 2.13 and 2.15. Taking the limit $N_S \to \infty$, the formula converges to

$$\log_2(\kappa/\,[\kappa - 1]), \tag{2.223}$$

which is consistent with the formula stated in [34].

Remark 2.16 *Ref. [34] has been widely accepted to have provided a complete proof of the unconstrained quantum capacity formulas given in* (2.221) *and* (2.223)*. The important developments of [34] were to identify that it suffices to optimize coherent information of these channels with respect to a single channel use and Gaussian input states. The issue is that [34] relied on an "optimization procedure carried out in" [33] in order to establish the infinite-energy quantum capacity formula given there (see just before [34, Eq. (12)]). However, a careful inspection of [33, Section V-B] reveals that no explicit optimization procedure is given there. The contentious point is that it is necessary to show that, among all Gaussian states, the thermal state is the input state optimizing the coherent information of the quantum-limited attenuator and amplifier channels. This point is not argued or in any way justified in [33, Section V-B] or in any subsequent work or review on the topic [80, 81, 82, 58]. As a consequence, we have been left to conclude that the proof from [34] features a gap which was subsequently closed in [49, Section III-G-1] and [29].*

Our results from Theorems 2.13 and 2.15 allow for making more general statements, applicable to broadband scenarios considered in prior works for other capacities [52, 83, 84]. For details please refer to Sec. IX. A in [28].

2.9 Appendix: Minimum fidelity and minimum entanglement fidelity

The following proposition states that a quantum code with good minimum fidelity implies that it has good minimum entanglement fidelity with negligible loss in parameters. This was first established in [60] and reviewed in [61]. Here we follow the proof available in [63], which therein established a relation between trace distance and diamond distance between an arbitrary channel and the identity channel.

Proposition 2.17 *Let $\mathcal{C} : \mathcal{T}(\mathcal{H}) \to \mathcal{T}(\mathcal{H})$ be a quantum channel with finite-dimensional input and output. Let $\mathcal{H}'$ be a Hilbert space isomorphic to $\mathcal{H}$. If*

$$\min_{|\phi\rangle \in \mathcal{H}} \langle \phi | \mathcal{C}(|\phi\rangle\langle\phi|) | \phi \rangle \geq 1 - \varepsilon, \tag{2.224}$$

then

$$\min_{|\psi\rangle \in \mathcal{H}' \otimes \mathcal{H}} \langle \psi | (\mathrm{id}_{\mathcal{H}'} \otimes \mathcal{C})(|\psi\rangle\langle\psi|) | \psi \rangle \geq 1 - 2\sqrt{\varepsilon}, \tag{2.225}$$

where the optimizations are with respect to state vectors.

Proof. The inequality in (2.224) implies that the following inequality holds for all state vectors $|\phi\rangle \in \mathcal{H}$:

$$\langle \phi | \left[|\phi\rangle\langle\phi| - \mathcal{C}(|\phi\rangle\langle\phi|) \right] |\phi\rangle \leq \varepsilon. \tag{2.226}$$

By the inequalities in (1.31), this implies that

$$\| |\phi\rangle\langle\phi| - \mathcal{C}(|\phi\rangle\langle\phi|) \|_1 \leq 2\sqrt{\varepsilon}, \tag{2.227}$$

for all state vectors $|\phi\rangle \in \mathcal{H}$. We will show that

$$\left| \langle \phi | \left[|\phi\rangle\langle\phi^\perp| - \mathcal{C}(|\phi\rangle\langle\phi^\perp|) \right] |\phi^\perp\rangle \right| \leq 2\sqrt{\varepsilon}, \tag{2.228}$$

for every orthonormal pair $\{|\phi\rangle, |\phi^{\perp}\rangle\}$ of state vectors in $\mathcal{H}$. Set

$$|w_k\rangle \equiv \frac{|\phi\rangle + i^k |\phi^{\perp}\rangle}{\sqrt{2}} \tag{2.229}$$

for $k \in \{0, 1, 2, 3\}$. Then it follows that

$$|\phi\rangle\langle\phi^{\perp}| = \frac{1}{2} \sum_{k=0}^{3} i^k |w_k\rangle\langle w_k|. \tag{2.230}$$

Consider now that

$$\begin{aligned}
&\left|\langle\phi| \left[|\phi\rangle\langle\phi^{\perp}| - \mathcal{C}(|\phi\rangle\langle\phi^{\perp}|)\right] |\phi^{\perp}\rangle\right| \\
&\leq \left\| |\phi\rangle\langle\phi^{\perp}| - \mathcal{C}(|\phi\rangle\langle\phi^{\perp}|) \right\|_{\infty} && (2.231) \\
&\leq \frac{1}{2} \sum_{k=0}^{3} \left\| |w_k\rangle\langle w_k| - \mathcal{C}(|w_k\rangle\langle w_k|) \right\|_{\infty} && (2.232) \\
&\leq \frac{1}{4} \sum_{k=0}^{3} \left\| |w_k\rangle\langle w_k| - \mathcal{C}(|w_k\rangle\langle w_k|) \right\|_{1} && (2.233) \\
&\leq 2\sqrt{\varepsilon}. && (2.234)
\end{aligned}$$

The first inequality follows from the characterization of the operator norm as $\|A\|_{\infty} = \sup_{|\phi\rangle,|\psi\rangle} |\langle\phi|A|\psi\rangle|$, where the optimization is with respect to state vectors $|\varphi\rangle$ and $|\psi\rangle$. The second inequality follows from substituting (2.230) and applying the triangle inequality and homogeneity of the ∞-norm. The third inequality follows because the ∞-norm of a traceless Hermitian operator is bounded from above by half of its trace norm [85, Lemma 4]. The final inequality follows from applying (2.227).

Let $|\psi\rangle \in \mathcal{H}' \otimes \mathcal{H}$ be an arbitrary state vector. All such state vectors have a Schmidt

decomposition of the following form:

$$|\psi\rangle = \sum_x \sqrt{p(x)}|\zeta_x\rangle \otimes |\varphi_x\rangle, \tag{2.235}$$

where $\{p(x)\}_x$ is a probability distribution and $\{|\zeta_x\rangle\}_x$ and $\{|\varphi_x\rangle\}_x$ are orthonormal sets, respectively. Then consider that

$$\begin{aligned}
&1 - \langle\psi|(\mathrm{id}_{\mathcal{H}'} \otimes \mathcal{C})(|\psi\rangle\langle\psi|)|\psi\rangle \\
&= \langle\psi|(\mathrm{id}_{\mathcal{H}'} \otimes \mathrm{id}_{\mathcal{H}} - \mathrm{id}_{\mathcal{H}'} \otimes \mathcal{C})(|\psi\rangle\langle\psi|)|\psi\rangle \\
&= \langle\psi|(\mathrm{id}_{\mathcal{H}'} \otimes [\mathrm{id}_{\mathcal{H}} - \mathcal{C}])(|\psi\rangle\langle\psi|)|\psi\rangle \\
&= \sum_{x,y} p(x)p(y)\langle\varphi_x| \left[|\varphi_x\rangle\langle\varphi_y| - \mathcal{C}(|\varphi_x\rangle\langle\varphi_y|)\right] |\varphi_y\rangle.
\end{aligned} \tag{2.236}$$

Now applying the triangle inequality and (2.228), we find that

$$\begin{aligned}
&1 - \langle\psi|(\mathrm{id}_{\mathcal{H}'} \otimes \mathcal{C})(|\psi\rangle\langle\psi|)|\psi\rangle \\
&= \left| \sum_{x,y} p(x)p(y)\langle\varphi_x| \left[|\varphi_x\rangle\langle\varphi_y| - \mathcal{C}(|\varphi_x\rangle\langle\varphi_y|)\right] |\varphi_y\rangle \right| \\
&\leq \sum_{x,y} p(x)p(y) \left|\langle\varphi_x| \left[|\varphi_x\rangle\langle\varphi_y| - \mathcal{C}(|\varphi_x\rangle\langle\varphi_y|)\right] |\varphi_y\rangle\right| \\
&\leq 2\sqrt{\varepsilon}.
\end{aligned} \tag{2.237}$$

This concludes the proof. ■

Chapter 3
Capacities of Quantum-limited Amplifier Channels

3.1 Introduction

In this Chapter we will focus on the application of energy-constrained quantum Shannon theory to specific Bosonic quantum channels. Specifically we would like to calculate the exact capacity formula for different optical channels. Although the energy-constrained capacities are established for many communication protocols [35, 37, 28] as we mentioned in Sec. 2.1.1, to calculate the capacity for specific bosonic channel is still challenging due to the maximization over quantum states. Therefore, we first need to guess a good bosonic input state, which is called the achievability part. This part only involves a direct calculation of entropic quantities thus is relative simple. The real difficult part is the converse theorem, in which we need to prove that the bosonic input state we choose is indeed the optimal one. It is in this converse part that the *minimum output-entropy* (MOE) conjectures play a vital role. There are many different forms of MOE conjectures. MOE conjecture with more general form is more difficult to prove. And the calculation of capacity for more sophisticated protocol usually requires more general MOE conjecture. We review the different versions of MOE conjectures and their relationships in Appendix A.

The quantum-limited amplifier channel [86, 87] is a fundamental building block of bosonic Gaussian channel, given that it can be decomposed as the serial concatenation of a quantum-limited attenuator followed by a quantum-limited amplifier or its phase conjugate [88, 89]. Interestingly, the Bogoliubov transformation governing spontaneous parametric down-conversion in a nonlinear optical system [90] also describes a variety of

different physical processes, such as the dynamical Casimir effect [91], the Unruh effect [92] and Hawking radiation [93]. For example, the gain of a quantum amplifier channel is directly related to the acceleration of an observer in the setting of the Unruh effect. By employing Einstein's equivalence principle, the Unruh effect has a correspondence in the setting of Hawking radiation, in which the amplifier gain plays the role of the surface gravity of the black hole. For a review on the close relationship between the above phenomena, see Ref. [94]. Related, several papers have studied quantum communication in situations where relativistic effects cannot be ignored [95, 96]. Thus, the importance of quantum amplifier channels in various different fields of physics suggests that studying its communication capacities has both practical and theoretical relevance.

In this chapter, we first determine communication trade-offs for a quantum-limited amplifier channel in which a sender has access to the input of the amplifier and a receiver to its output. The information trade-off problem is one of the most general information-processing tasks that one can consider for a point-to-point quantum communication channel. It allows the sender and receiver to simultaneously generate or consume any of the three fundamental information resources: classical information, quantum information, and shared entanglement. The protocol from [97, 98, 99] (see also [14, Chapter 25]) establishes an achievable rate region, which yields remarkable gains over the naive strategy of time sharing, as discussed in [98, 99]. In this work we prove that this achievable rate region is optimal, which establishes the capacity region for this setting. In order to do so, we establish some new mathematical properties of the entropy of the bosonic thermal state (see Appendix B.1), a function which is of physical interest in a variety of contexts. We suspect that these established properties could have application in the analysis of other communication problems and in studies of quantum thermodynamics, but this is more appropriate to remain as the topic of future work.

We also consider the trade-off between public classical bits, private classical bits, and

secret key bits [100], and we establish the capacity region in this setting as well. This capacity region clearly has relevance when using a channel for the communication of secret information in addition to ordinary, public classical information.

Beyond the point-to-point setup, we also determine the capacity region for the single-sender, two-receiver broadcast channel induced by a unitary dilation of the quantum-limited amplifier channel. We do so by first giving a rate region achieved by inputting coherent states [101] to a noisy amplifier channel. We find that this rate region improves upon those achieved using traditional strategies such as coherent homodyne or heterodyne detection. We also prove that this rate region is optimal for quantum-limited amplifier channels by employing similar techniques that we use for the first two scenarios mentioned above. These techniques are different when compared to those used in previous works [98, 102] for the setting of the pure-loss channel.

This chapter is organized as follows. In Section 3.2, we review the main result of [103], which establishes a minimum output-entropy theorem essential for our developments here. In Section 3.3, we consider the communication trade-off for a quantum-limited amplifier channel. After briefly reviewing the characterization of the trade-off capacity region and the achievable rate region established in [99], we prove that this rate region is optimal. We then show that the trade-off capacity region outperforms that achievable with a naive time-sharing strategy. We also find that capacities decrease with increasing amplifier gain. We then consider the unitary dilation of the quantum-limited amplifier channel as a quantum broadcast channel in Section 3.4. In the first part of Section 3.4, we determine an achievable rate region for two receivers by using coherent-state encoding. In the second part of Section 3.4, we prove that this achievable rate region is optimal. In the third part of Section 3.4, we show that the capacity region outperforms those achieved by using homodyne and heterodyne detection. In Section 3.5, we consider the trade-off between public and private classical communication. We determine these trade-off

capacities for quantum-limited amplifier channels by employing techniques similar to those from Sections 3.3 and 3.4. Finally, we discuss the relationship between entropy conjectures and capacities of bosonic Gaussian channels in Section 3.6.

3.2 MINIMUM OUTPUT-ENTROPY THEOREM

All of our converse proofs in this work rely on the following minimum output-entropy theorem (single-mode version of Conjecture A.2, see Appendix A for more details), which holds for a single-mode, phase-insensitive quantum-limited amplifier (and its weak conjugate [104]) channel with a given input entropy constraint [103]. We restate this result as the following theorem:

Theorem 3.1 ([103]) *Consider a single-mode, phase-insensitive amplifier channel $\mathcal{N}_{A\to B}$. Let $H_0 > 0$ be a positive constant. For any input state ρ_A such that $H(\rho_A) \geq H_0$, the output von Neumann entropy $H(\mathcal{N}_{A\to B}(\rho_A))$ is minimized when ρ_A is a thermal state with mean photon number $g^{-1}(H_0)$, where*

$$g(x) \equiv (x+1)\log_2(x+1) - x\log_2 x\,. \tag{3.1}$$

is the entropy of a thermal state with mean photon number x. The same is true for the quantum-limited weak conjugate amplifier [30] (the complementary channel of $\mathcal{N}_{A\to B}$ [104]).

Theorem 3.1 provides lower bounds for certain terms in the capacity regions in (3.3)–(3.5) and (3.39)–(3.40), which are crucial for our converse proofs. Due to additivity issues of capacity regions in quantum information theory, proofs of converses generally require a multi-mode version of Conjecture A.2. However, a quantum-limited amplifier channel, the complementary channel of which is entanglement-breaking, is a Hadamard channel [105, 106]. It is known that the capacity regions of both the information trade-off and broadcast problems are single-letter for Hadamard channels [107, 108, 100, 109].

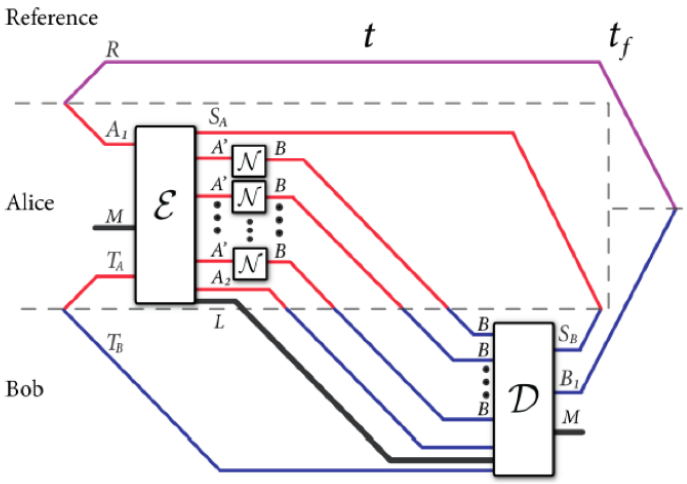

Figure 3.1: Triple trade-off protocol (taken from [3]).M is the classical register of Alice, T_AT_B is her pre-shared entanglement with Bob and A_1 is the quantum system she would like to transmit. After the encoding $\mathcal{E}$, Alice holds quantum system S_A and classical register L. She send A'^n to Bob via the quantum channel $\mathcal{N}$. After Bob's decoding, we still have a classical output M, a quantum output B_1 and an entangled state on S_AS_B.

Our first result concerns the transmission (or consumption) of classical bits, quantum bits, and shared entanglement along with the consumption of many independent uses of a quantum-limited amplifier channel. The communication trade-off is characterized by *rate triples* (C, Q, E), where C is the net rate of classical communication, Q is the net rate of quantum communication, and E is the net rate of entanglement generation. See Figure 3.1 for the most general protocol. The triple trade-off capacity region of a quantum channel $\mathcal{N}$ is the regularization of the union of regions of the following form [108] (see also [3, Chapter 25]):

$$\begin{aligned} C + 2Q &\leq H(\mathcal{N}(\rho)) + \sum_x p_X(x) \left[H(\rho_x) - H(\mathcal{N}^c(\rho_x)) \right], \\ Q + E &\leq \sum_x p_X(x) \left[H(\mathcal{N}(\rho_x)) - H(\mathcal{N}^c(\rho_x)) \right], \\ C + Q + E &\leq H(\mathcal{N}(\rho)) - \sum_x p_X(x) H(\mathcal{N}^c(\rho_x)), \end{aligned} \tag{3.2}$$

where the union is with respect to all possible input ensembles $\{p_X(x), \rho_x\}$ and $\rho \equiv \sum_x p_X(x)\rho_x$. Here $\mathcal{N}^c$ is a complementary channel of $\mathcal{N}$ [14].

Achievable rate region

The achievability part of the capacity theorem for the quantum-limited amplifier channel was already established in [98, 99]. The coding strategy is to employ an input ensemble of Gaussian-distributed phase-space displacements of the two-mode squeezed vacuum. We restate this result as the following theorem, which is given as Theorem 3 in [99]:

Theorem 3.2 *An achievable rate region for a quantum-limited amplifier channel with amplifier gain $\kappa \geq 1$ is given by the union of regions of the following form:*

$$C + 2Q \leq g(\lambda N_S) + g(\kappa N_S + \bar{\kappa}) - g(\bar{\kappa}[\lambda N_S + 1]) \, , \tag{3.3}$$

$$Q + E \leq g(\kappa\lambda N_S + \bar{\kappa}) - g(\bar{\kappa}[\lambda N_S + 1]) \, , \tag{3.4}$$

$$C + Q + E \leq g(\kappa N_S + \bar{\kappa}) - g(\bar{\kappa}[\lambda N_S + 1]) \, , \tag{3.5}$$

where $\lambda \in [0, 1]$ is a photon-number-sharing parameter and g(x) is defined in (3.1)*. The parameter $\bar{\kappa} \equiv \kappa - 1$ denotes the mean number of photons generated by the channel when the vacuum is input.*

Outer bound for the capacity region

Our contribution here is to prove that the rate region in Theorem 3.2 is equal to the capacity region.

Theorem 3.3 *The triple trade-off capacity region for a quantum-limited amplifier channel with amplifier gain $\kappa \geq 1$ is equal to the rate region given in Theorem 3.2.*

Proof. We first recall that the capacity region of a quantum limited amplifier channel is

single-letter [107] due to the fact that a quantum-limited amplifier channel is a Hadamard channel [105, 106]. Thus, there is no need to consider the regularization in (3.2). To give an upper bound on the single-letter capacity region of the quantum-limited amplifier channel, we prove that for all input ensembles $\{p_X(x), \rho_x\}$, obeying the energy constraint $\mathrm{Tr}\{\hat{n} \sum_x p(x)\rho_x\} \leq N_S$, there exists a $\lambda \in [0, 1]$ such that the following four inequalities hold

$$H(\mathcal{N}(\rho)) \leq g(\kappa N_S + \kappa - 1) \, , \tag{3.6}$$

$$\sum_x p_X(x) H(\rho_x) \leq g(\lambda N_S) \, , \tag{3.7}$$

$$\sum_x p_X(x) H(\mathcal{N}(\rho_x)) \leq g(\kappa\lambda N_S + \kappa - 1) \, , \tag{3.8}$$

$$\sum_x p_X(x) H(\mathcal{N}^c(\rho_x)) \geq g((\kappa - 1)(\lambda N_S + 1)) \, . \tag{3.9}$$

We start by establishing the inequality in (3.6):

$$H(\mathcal{N}(\rho)) \leq g(\kappa N_S + \kappa - 1) \, . \tag{3.10}$$

This inequality follows from the facts that the output state has mean photon number no larger than $\kappa N_S + \kappa - 1$ when the input mean photon number is no larger than N_S and because the thermal state of mean photon number $\kappa N_S + \kappa - 1$ realizes the maximum entropy at the output.

We now argue the inequalities in (3.7) and (3.8). Consider that concavity of entropy and that the thermal state realizes the maximum entropy imply the following bound:

$$\sum_x p_X(x) H(\rho_x) \leq H(\rho) \leq g(N_S) \, . \tag{3.11}$$

Since $g(x)$ is monotonically increasing, there exists a $\lambda' \in [0,1]$ such that

$$\sum_x p_X(x) H(\rho_x) = g(\lambda' N_S) \, . \tag{3.12}$$

From concavity of entropy and (3.10), we find that

$$\sum_x p_X(x) H(\mathcal{N}(\rho_x)) \leq H(\mathcal{N}(\rho)) \tag{3.13}$$

$$\leq g(\kappa N_S + \kappa - 1) \, . \tag{3.14}$$

Due to the fact that the vacuum-state input realizes the minimum output entropy for any phase-insensitive quantum Gaussian channel [106], the following lower bound applies

$$H(\mathcal{N}(\rho_x)) \geq g(\kappa - 1) \, . \tag{3.15}$$

Since $g(x)$ is monotonically increasing and since we have shown that

$$g(\kappa - 1) \leq \sum_x p_X(x) H(\mathcal{N}(\rho_x)) \leq g(\kappa N_S + \kappa - 1), \tag{3.16}$$

there exists $\lambda \in [0,1]$ such that

$$\sum_x p_X(x) H(\mathcal{N}(\rho_x)) = g(\lambda \kappa N_S + \kappa - 1) \, . \tag{3.17}$$

However, λ and λ' are different in general. But we can use Theorem B.2 in Appendix B.1 to establish that $\lambda' \leq \lambda$. To use it we need to know the entropy of the input state. Supposing that the mean photon number of ρ_x is $N_{S,x}$, we have that

$$H(\rho_x) \leq g(N_{S,x}) \, . \tag{3.18}$$

Therefore there exists $\lambda'_x \in [0,1]$ such that

$$H(\rho_x) = g(\lambda'_x N_{S,x}) \, . \tag{3.19}$$

Now employing Theorem 3.1 for the quantum-limited amplifier channel, we have that

$$H(\mathcal{N}(\rho_x)) \geq g(\kappa\lambda'_x N_{S,x} + \kappa - 1) \, , \tag{3.20}$$

which in turn implies that

$$\begin{aligned} & g(\lambda\kappa N_S + \kappa - 1) \\ & \quad = \sum_x p_X(x) H(\mathcal{N}(\rho_x)) & (3.21) \\ & \quad \geq \sum_x p_X(x) g(\kappa\lambda'_x N_{S,x} + \kappa - 1) \, . & (3.22) \end{aligned}$$

Together with $\sum_x p_X(x) g(\lambda'_x N_{S,x}) = g(\lambda' N_S)$, and using Theorem B.2 in Appendix B.1 with $q = \kappa$ we find that

$$\sum_x p_X(x) g(\kappa\lambda'_x N_{S,x} + \kappa - 1) \geq g(\kappa\lambda' N_S + \kappa - 1) \, , \tag{3.23}$$

which, by combining (3.22) and (3.23), implies that

$$g(\lambda\kappa N_S + \kappa - 1) \geq g(\kappa\lambda' N_S + \kappa - 1) \, . \tag{3.24}$$

Since g is monotonically increasing and it has a well-defined inverse function, we find that

$$\lambda \geq \lambda' \, , \tag{3.25}$$

which, after combining with (3.12) and the monotonicity of $g(x)$, implies that

$$\sum_x p_X(x) H(\rho_x) \leq g(\lambda N_S) \ . \tag{3.26}$$

This concludes the proof of the inequalities in (3.7) and (3.8).

To prove the last bound in (3.9), by (3.15) and

$$H(\mathcal{N}(\rho_x)) \leq g(\kappa N_{S,x} + \kappa - 1) \ , \tag{3.27}$$

we can conclude that there exists $\lambda_x \in [0,1]$ such that the following equality holds

$$H(\mathcal{N}(\rho_x)) = g(\lambda_x \kappa N_{S,x} + \kappa - 1) \ . \tag{3.28}$$

The quantum-limited amplifier channel $\mathcal{N}$ is degradable [104], and its degrading channel $\mathcal{D}_{B\to C}$ is the weakly-conjugate channel of the quantum-limited amplifier with $\kappa' = (2\kappa - 1)/\kappa$ [104]. The main property of this degrading channel that we need is that an input thermal state of mean photon number K leads to an output thermal state of mean photon number $(\kappa' - 1)(K+1)$. Theorem 3.1 applied to this case gives that for given input entropy $g(K)$, the minimum output entropy of $\mathcal{D}_{B\to C}$ is equal to $g((\kappa' - 1)(K+1))$. By applying it, we find that

$$\begin{aligned}
&\sum_x p_X(x) H(\mathcal{N}^c(\rho_x)) \\
&\quad \geq \sum_x p_X(x) g((\kappa' - 1)(\lambda_x \kappa N_{S,x} + \kappa)) &(3.29)\\
&\quad = \sum_x p_X(x) g((\kappa - 1)\lambda_x N_{S,x} + \kappa - 1). &(3.30)
\end{aligned}$$

Since $\sum_x p_X(x) g(\kappa \lambda_x N_{S,x} + \kappa - 1) = g(\lambda \kappa N_S + \kappa - 1)$, using Theorem B.1 in Appendix B.1

with $q = (\kappa - 1)/\kappa$ and $C = (\kappa - 1)/\kappa$, we find that

$$\begin{aligned}
&\sum_x p_X(x) H(\mathcal{N}^c(\rho_x)) \\
&\geq g(q(\lambda\kappa N_S + \kappa - 1) + (\kappa - 1)/\kappa) \, , \qquad (3.31) \\
&= g((\lambda N_S + 1)(\kappa - 1)) \, . \qquad (3.32)
\end{aligned}$$

This concludes our proof for the four bounds in (3.6)–(3.9). Together with the achievability part in [99] (recalled as Theorem 3.2), this concludes the proof that the union of regions given by (3.3)–(3.5) is equal to the quantum dynamic capacity region for the quantum-limited amplifier channel. ■

Returning to our discussion from the introduction, we note that Theorem 3.3 completely characterizes the communication abilities of any phase-insensitive quantum-limited amplifier channel, particular examples of this channel occurring in a number of scenarios of physical interest, including spontaneous parametric down-conversion in a nonlinear optical system [90], the dynamical Casimir effect [91], the Unruh effect [92], and Hawking radiation [93]. That is, if one desires to use any such channel for sending classical and quantum information along with the assistance of shared entanglement, then Theorem 3.3 sets the ultimate limits for such a task. Theorem 3.3 thus subsumes and places a capstone on much previous literature in quantum information having to do with capacities of phase-insensitive, quantum-limited amplifier channels.

Comparison with time-sharing strategy and large κ limit

Figure 3.2 displays two special cases of the capacity region in (3.3)–(3.5). We consider a quantum-limited amplifier channel with gain $\kappa = 2$ and choose the mean input photon number to be $N_S = 200$. In Figure 3.2(a), we plot the trade-off between classical and quantum communication without entanglement assistance. The maximum quantum

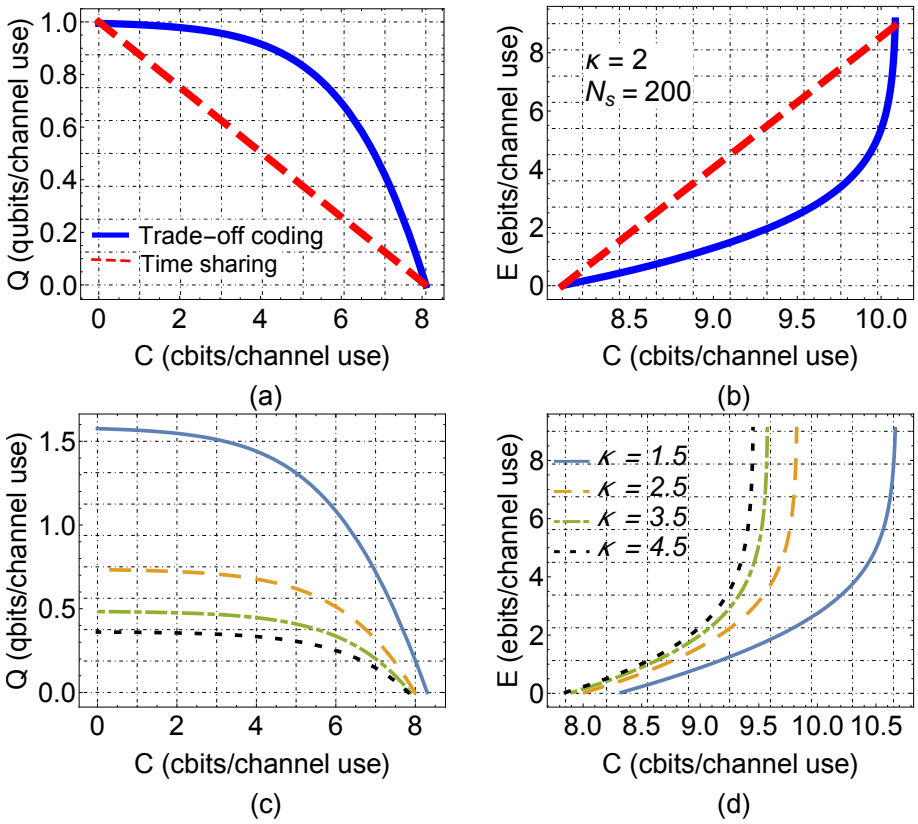

Figure 3.2: We consider a quantum-limited amplifier channel with $\kappa = 2$ and mean photon number constraint $N_S = 200$. In (a), we plot the (C, Q) trade-off. The maximum quantum capacity is equal to $\log_2(2) - \log_2(1) = 1$ qubit per channel use. A trade-off coding strategy shows an improvement compared to time sharing, wherein we see that the classical data rate can be boosted while still maintaining a high quantum transmission rate. In (b) we plot the (C, E) trade-off. The sender and the receiver share entanglement, and the sender would like to transmit classical information while minimizing the consumption of entanglement. As can be seen, with trade-off coding, the sender can significantly reduce the consumption of entanglement while still keeping the classical communication rate near to its maximum value. In (c) and (d) we plot the capacity region for the (C, Q) and (C, E) trade-off with amplifier gain $\kappa = 1.5$, 2.5, 3.5, and 4.5. Each capacity region shrinks as the amplifier gain κ increases.

transmission rate is $\log_2(\kappa/\bar{\kappa}) = 1$ qubits per channel use, established jointly in [110, 50] (see discussion in [28]). This result also follows from the results of the present paper by considering that the bound in (3.4) for $\lambda = 1$ gives the finite-energy quantum capacity of the quantum limited amplifier channel:

$$g(\kappa N_S + \bar{\kappa}) - g(\bar{\kappa}[N_S + 1]). \tag{3.33}$$

Taking the infinite-energy limit, we recover the formula established in [110, 50]:

$$\lim_{N_S \to \infty} g(\kappa N_S + \bar{\kappa}) - g(\bar{\kappa}[N_S + 1]) = \log_2(\kappa/\bar{\kappa}). \tag{3.34}$$

Around 200 photons per channel use is large enough to approximate this quantum capacity well for the above parameter choices. The figure indicates a remarkable improvement over a time-sharing strategy, in which the sender transmits classical information for some fraction of the channel uses and transmits quantum information for the other fraction. By using a trade-off coding strategy, lowering the quantum data rate by about 0.1 qubits per channel use allows for sending roughly three extra classical bits per channel use. However, if a time-sharing strategy is adopted, lowering the quantum data rate by the same amount gives only one additional bit per channel use.

In Figure 3.2(b), we plot the trade-off between entanglement-assisted and unassisted classical communication. Again, a trade-off coding strategy gives a dramatic improvement over time sharing. In this figure, we take the convention that positive E corresponds to entanglement consumption. With mean photon number $N_S = 200$, the sender can reliably transmit a maximum of around 10.2 classical bits per channel use by consuming around 9.1 entangled bits per channel use [111, 112]. By using trade-off coding, the sender can reduce the consumption of entanglement to around 4 entangled bits per channel use, while still being able to transmit classical data at around 9.8 bits per channel

use.

One trend we see for the quantum-limited amplifier channel is that a large amplifier gain κ compromises its communication ability, as shown in Figures 3.2(c) and 3.2(d). For the (C, Q) trade-off, as κ increases, the quantum capacity decreases for a fixed classical rate. For the (C, E) trade-off, not only the maximum classical rate is reduced, but the savings of entangled bits for a constant classical rate are also diminished. This effect results from the fact that a quantum-limited amplifier channel with large κ generates more photons from the vacuum, and thus injects more noise into the transmitted quantum signal. Mathematically the shrinkage of the capacity region is due to the term $g(\bar{\kappa}[\lambda N_S+1])$ appearing in all of the inequalities in (3.3)–(3.5), which increases with increasing amplifier gain.

3.4 QUANTUM BROADCAST AMPLIFIER CHANNEL

Our next result concerns the classical capacity of a quantum broadcast channel induced by a unitary dilation of the quantum amplifier channel. We consider the single-sender, two-receiver case in which Alice simultaneously transmits classical data to Bob (B) via the amplifier channel and to Charlie (C) via its complementary channel. The full Bogoliubov transformation for this setup is given by

$$\begin{aligned} \hat{b} &= \sqrt{\kappa}\hat{a} + \sqrt{\kappa - 1}\hat{e}^\dagger \, , \\ \hat{c}^\dagger &= \sqrt{\kappa - 1}\hat{a} + \sqrt{\kappa}\hat{e}^\dagger \, , \end{aligned} \tag{3.35}$$

where $\hat{a}$, $\hat{b}$, $\hat{c}$, and $\hat{e}$ are the field-mode annihilation operators corresponding to the sender Alice's input mode, the receiver Bob's output mode, the receiver Charlie's output mode, and an environmental input, respectively. Here we consider a general amplifier channel with thermal noise, in which the input state represented by $\hat{e}$ is a thermal state with mean photon number N_B. Such a channel could model information propagation to two ob-

servers, one outside and one beyond the event horizon of a black hole [96]. This channel could also model information propagation from an inertial observer to two constantly accelerated complementary observers moving with opposite accelerations in two causally disconnected regions of Rindler spacetime, if we take the convention that the inertial observer can encode information into Unruh modes, which arguably allows for computing estimates for an upper bound of channel capacities between inertial and relativistically accelerating observers [113].

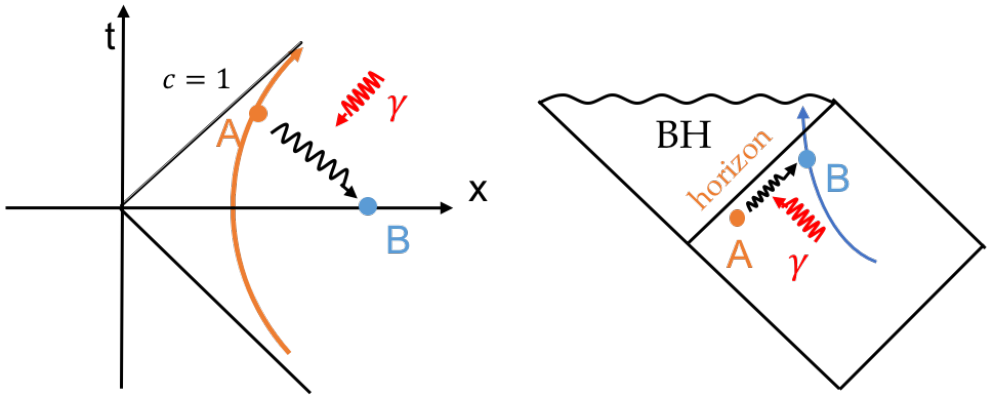

Figure 3.3: Communication in Rindler space-time and near the horizon of a black hole. (Left) A spacetime diagram. Alice travels with constant proper acceleration and Bob stays at rest. Alice will see thermal noise due to the Uhruh effect. This communication line can be modeled as a quantum amplifier channel. (Right) A Penrose diagram of a Schwarzchild black hole. Bob is free falling in to the black hole while Alice is in a spaceship staying just outside the horizon. Due to the equivalence principle, at the region near the horizon, the physics in the two subfigures are equivalent to each other.

The most general classical communication protocol over quantum broadcast channel is shown in Figure 3.4.

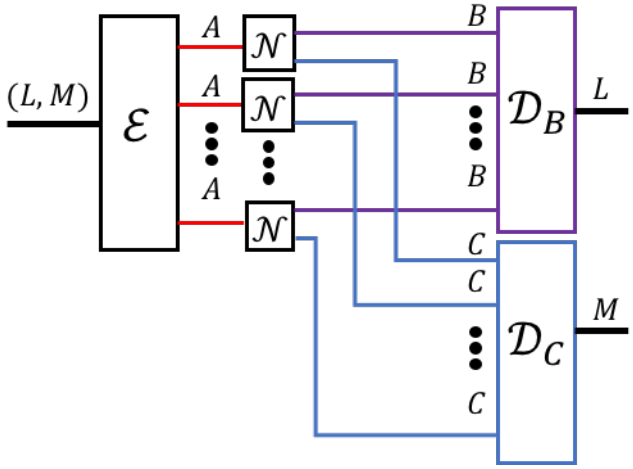

Figure 3.4: Classical communication over quantum broadcast channel. Alice encodes message l, m into quantum states by $\mathcal{E}$. She then send A^n to Bob and Charlie via quantum broadcast channel $\mathcal{N}$. Bob and Charlie try to decode their own messages from their decoders $\mathcal{D}_B,\ .D_C$.

The classical capacity region of the two-user degraded quantum broadcast channel was derived in [114] (see also [115] for the achievability part) and found to be equal to the regularization of the union of the following rate regions:

$$\begin{aligned} R_B &\leq \sum_x p_X(x)\left[H(\mathcal{N}(\rho_x)) - \sum_y p_{Y|X}(y|x)H(\mathcal{N}(\rho_y))\right] , \\ R_C &\leq H(\mathcal{N}^c(\rho)) - \sum_x p_X(x)H(\mathcal{N}^c(\rho_x)) , \end{aligned} \tag{3.36}$$

where the union is with respect to input ensembles $\{p_X(x)p_{Y|X}(y|x), \rho_y\}$ with

$$\rho_x \equiv \sum_y p_{Y|X}(y|x)\rho_y , \tag{3.37}$$

$$\rho \equiv \sum_x p_X(x)\rho_x . \tag{3.38}$$

In the following we first give an achievable rate region for an amplifier channel with

thermal noise. We then prove that this rate region is optimal if the multi-mode version of Theorem 3.1 is true. For the case in which the amplifier channel is quantum-limited, the capacity region is single-letter [109] and therefore Theorem 3.1 implies the broadcast capacity region for the quantum-limited amplifier channel.

Achievable rate region by coherent-state encoding

Theorem 3.4 *Consider a quantum broadcast amplifier channel as given in* (3.35) *with amplifier gain* $\kappa \geq 1$ *and environmental thermal-state input with mean photon number* N_B*. Suppose that the mean input photon number for each channel use is no larger than* N_S*. Then the following rate region for Bob and Charlie*

$$R_B \leq g(\kappa\lambda N_S + \bar{\kappa}(N_B + 1)) - g(\bar{\kappa}(N_B + 1)) \, , \tag{3.39}$$

$$R_C \leq g(\bar{\kappa}(N_S + 1) + \kappa N_B) - g(\bar{\kappa}(\lambda N_S + 1) + \kappa N_B) \, , \tag{3.40}$$

with $\lambda \in [0, 1]$ *is achievable by using coherent-state encoding according to the following ensemble:*

$$\{p(t)p(\alpha|t), |\alpha\rangle\langle\alpha|\} \, , \tag{3.41}$$

where

$$p(t) = \frac{1}{\pi N_S} \exp\left(-\frac{|t|^2}{N_S} \right) \, , \tag{3.42}$$

$$p(\alpha|t) = \frac{1}{\pi\lambda N_S} \exp\left(-\frac{|\sqrt{1-\lambda}t - \alpha|^2}{\lambda N_S} \right) \, . \tag{3.43}$$

Here α *and* t *are complex variables and* $\bar{\lambda} = 1 - \lambda$*.*

Proof. Using (3.42) and (3.43), we find that

$$\begin{aligned}\rho_t &= \int d^2\alpha\, p(\alpha|t)\, |\alpha\rangle\langle\alpha| \\ &= \int d^2\gamma\, \frac{1}{\pi\lambda N_S}\, \exp\left(-\frac{|\gamma|^2}{N_S\lambda}\right)\, |\gamma+\sqrt{\bar{\lambda}}t\rangle\langle\gamma+\sqrt{\bar{\lambda}}t| \\ &= D(\sqrt{\bar{\lambda}}t)\rho^{\rm th}_{\lambda N_S} D^\dagger(\sqrt{\bar{\lambda}}t)\ . \end{aligned} \tag{3.44}$$

In the above, $D(\alpha)$ is a displacement operator [101] and $\rho^{\rm th}_{\lambda N_S}$ denotes a thermal state of mean photon number λN_S. The overall average input state is

$$\begin{aligned}\rho &= \int d^2t\, p(t)\, \rho_t\ , \\ &= \int d^2t'\, \frac{1}{\pi\bar{\lambda}N_S}\, \exp\left(-\frac{|t'|^2}{\bar{\lambda}N_S}\right) D(t')\rho^{\rm th}_{\lambda N_S} D^\dagger(t') \\ &= \rho^{\rm th}_{N_S}\ , \end{aligned} \tag{3.45}$$

which is just a thermal state with mean photon number N_S, in agreement with the energy constraint. There are four entropies we need to evaluate in (3.39) and (3.40). The first one is

$$\begin{aligned}&\int d^2t\, p(t)\, H(\mathcal{N}(\rho_t)) \\ &= \int d^2t\, p(t)\, H(\mathcal{N}(D(\sqrt{\bar{\lambda}}t)\rho^{\rm th}_{\lambda N_S} D^\dagger(\sqrt{\bar{\lambda}}t))) \\ &= \int d^2t\, p(t)\, H(\mathcal{N}(\rho^{\rm th}_{\lambda N_S})) = H(\mathcal{N}(\rho^{\rm th}_{\lambda N_S})) \\ &= g(\kappa\lambda N_S + (\kappa-1)(N_B+1))\ . \end{aligned} \tag{3.46}$$

The second equality follows because the amplifier channel is covariant with respect to displacement operators and the fact that entropy is invariant with respect to a unitary transformation.

Since the output state is unitarily related to a thermal state with mean photon number $(\kappa - 1)(N_B + 1)$ when Alice sends a coherent state into an amplifier channel, the second term in (3.39) is given by

$$\int d^2t \; d^2\alpha \; p(t) \; p(\alpha|t) \; H(\mathcal{N}(|\alpha\rangle\langle\alpha|)) = g((\kappa - 1)(N_B + 1)) \; . \tag{3.47}$$

Now similarly for (3.40), the first term is

$$H(\mathcal{N}^c(\rho^{\text{th}}_{N_S})) = g((\kappa - 1)(N_S + 1) + \kappa N_B) \; . \tag{3.48}$$

The last term can be calculated as follows:

$$\begin{aligned}
&\int d^2t \; p(t) \; H(\mathcal{N}^c(\rho_t)) \\
&= \int d^2t \; p(t) \; H(\mathcal{N}^c(D(\sqrt{\bar{\lambda}}t)\rho^{\text{th}}_{\lambda N_S} D^\dagger(\sqrt{\bar{\lambda}}t))) \\
&= \int d^2t \; p(t) \; H(\mathcal{N}^c(\rho^{\text{th}}_{\lambda N_S})) \\
&= g((\kappa - 1)(\lambda N_S + 1) + \kappa N_B) \; .
\end{aligned} \tag{3.49}$$

We use the facts that a gauge-contravariant bosonic Gaussian channel is contravariant with respect to displacement operators and that entropy is invariant with respect to a unitary transformation. Combining the above results, we conclude that the rate region in (3.39) and (3.40) is achievable. ■

Outer bound for the capacity region

We first prove that the rate region in (3.39) and (3.40) is optimal if a multi-mode version of Theorem 3.1 is true. To do so we need to show that it is also an outer bound for the capacity region.

Theorem 3.5 *Consider a quantum amplifier channel with amplifier gain $\kappa \geq 1$ and environ-*

mental thermal-state input with mean photon number N_B. Suppose that the mean input photon number for each channel use is no larger than N_S. Suppose that a multi-mode version of Theorem 3.1 is true. Then the region given by (3.39) *and* (3.40) *is an outer bound for the broadcast capacity region.*

Proof. Since a general quantum amplifier channel with thermal noise is not a Hadamard channel, we need to consider the n-letter version of (3.39) and (3.40). Specifically, we need to prove that for all input ensembles $\{p_X(x)p_{Y|X}(y|x), \rho_y\}$ for n uses of the channel, there exists $\lambda \in [0, 1]$ such that the following four bounds hold

$$\sum_x p_X(x) H(\mathcal{N}^{\otimes n}(\rho_x)) \leq ng(\kappa\lambda N_S + \bar{\kappa}(N_B + 1)) \, , \tag{3.50}$$

$$H((\mathcal{N}^c)^{\otimes n}(\rho)) \leq ng(\bar{\kappa}(N_S + 1) + \kappa N_B) \, , \tag{3.51}$$

$$\sum_x \sum_y p_X(x) p_{Y|X}(y|x) H(\mathcal{N}^{\otimes n}(\rho_y)) \geq ng(\bar{\kappa}(N_B + 1)) \, , \tag{3.52}$$

$$\sum_x p_X(x) H((\mathcal{N}^c)^{\otimes n}(\rho_x)) \geq ng(\bar{\kappa}(\lambda N_S + 1) + \kappa N_B) \, . \tag{3.53}$$

The second inequality holds because

$$H((\mathcal{N}^c)^{\otimes n}(\rho)) \leq \sum_{j=1}^{n} H(\rho_C^j) \tag{3.54}$$

$$\leq ng((\kappa - 1)(N_S + 1) + \kappa N_B) \, . \tag{3.55}$$

The first inequality follows from the subadditivity of quantum entropy. The second inequality follows from the fact that each output state at C has mean photon number $(\kappa - 1)(N_S + 1) + \kappa N_B$ and the thermal state maximizes the entropy.

Since the vacuum minimizes the output entropy for any phase-insensitive Gaussian channel [106], we find that $H(\mathcal{N}^{\otimes n}(\rho_y)) \geq ng((\kappa - 1)(N_B + 1))$, which leads to the third

bound:

$$\sum_x \sum_y p_X(x) p_{Y|X}(y|x) H(\mathcal{N}^{\otimes n}(\rho_y)) \geq ng((\kappa - 1)(N_B + 1)) \, . \tag{3.56}$$

Now we prove the first bound. From the concavity of quantum entropy, we have that

$$H\left(\sum_y p_{Y|X}(y|x) \mathcal{N}^{\otimes n}(\rho_y)\right) \geq \sum_y p_{Y|X}(y|x) H(\mathcal{N}^{\otimes n}(\rho_y)) \, . \tag{3.57}$$

Thus we have

$$\begin{aligned} &\sum_x p_X(x) H(\mathcal{N}^{\otimes n}(\rho_x)) \\ &\geq \sum_{x,y} p_X(x) p_{Y|X}(y|x) H(\mathcal{N}^{\otimes n}(\rho_y)) \\ &\geq ng((\kappa - 1)(N_B + 1)) \, . \end{aligned} \tag{3.58}$$

On the other hand, we have that

$$\begin{aligned} \sum_x p_X(x) H(\mathcal{N}^{\otimes n}(\rho_x)) &\leq H(\mathcal{N}^{\otimes n}(\rho)) \\ &\leq ng(\kappa N_S + (\kappa - 1)(N_B + 1)) \, . \end{aligned} \tag{3.59}$$

Together with (3.58) and the fact that $g(x)$ is monotonic, there exists $\lambda \in [0, 1]$ such that

$$\sum_x p_X(x) H(\mathcal{N}^{\otimes n}(\rho_x)) = ng(\kappa \lambda N_S + (\kappa - 1)(N_B + 1)) \, .$$

To prove the last bound, we use the fact that the weakly degrading channel of the amplifier channel is the weakly-conjugate of an amplifier channel with $\kappa' = (2\kappa - 1)/\kappa > 1$

[104]. We first calculate the entropy of the output state:

$$\begin{aligned} H(\mathcal{N}^{\otimes n}(\rho_x)) &= H(\rho_{B,x}) \ , \\ &\leq \sum_{j=1}^{n} H(\rho_{B,x}^{j}) \ , \\ &\leq n \sum_{j=1}^{n} \frac{1}{n} g(\kappa N_{S,x_j} + (\kappa - 1)(N_B + 1)) \ , \\ &\leq n g(\kappa N_{S,x} + (\kappa - 1)(N_B + 1)) \ . \end{aligned} \tag{3.60}$$

The first inequality follows from subadditivity of quantum entropy. Letting N_{S,x_j} be the mean photon number for the jth symbol of ρ_x, the second inequality follows because the thermal state maximizes the entropy. Letting $N_{S,x} = \sum_j N_{S,x_j}/n$, the last inequality follows from concavity of $g(x)$. Since we also have that

$$H(\mathcal{N}^{\otimes n}(\rho_x)) \geq n g((\kappa - 1)(N_B + 1)), \tag{3.61}$$

there exists $\lambda_x \in [0, 1]$ such that

$$H(\mathcal{N}^{\otimes n}(\rho_x)) = n g(\kappa \lambda_x N_{S,x} + (\kappa - 1)(N_B + 1)) \ . \tag{3.62}$$

Using the multi-mode version of Theorem 1 for the degrading channel, we find that

$$\begin{aligned} &\sum_x p_X(x) H((\mathcal{N}^c)^{\otimes n}(\rho_x)) \\ \geq &\sum_x p_X(x) n g((\kappa' - 1)[\kappa \lambda_x N_{S,x} + \bar{\kappa}(N_B + 1) + 1] + \kappa' N_B) \\ = &\sum_x p_X(x) n g((\kappa - 1)(\lambda_x N_{S,x} + 1) + \kappa N_B) \ . \end{aligned} \tag{3.63}$$

Together with

$$\sum_x p_X(x) g(\kappa \lambda_x N_{S,x} + \bar{\kappa}(N_B + 1)) = g(\kappa \lambda N_S + \bar{\kappa}(N_B + 1)) \, , \tag{3.64}$$

we can invoke Theorem B.1 in Appendix B.1 with $q = (\kappa - 1)/\kappa$ and $C = \frac{2\kappa - 1}{\kappa}(N_B + 1) - 1$ to find that

$$\sum_x p_X(x) H((\mathcal{N}^c)^{\otimes n}(\rho_x)) \\ \geq n g((\kappa - 1)(\lambda N_S + 1) + \kappa N_B) \, . \tag{3.65}$$

This concludes our proof. Together with the achievability of (3.39)–(3.40), we establish it as the capacity region for the quantum broadcast amplifier channel, provided that the multi-mode version of Theorem 1 is true. ■

Now let us consider the quantum-limited amplifier channel. Since the broadcast capacity region for Hadamard channels is single-letter [109], by setting $n = 1$ and $N_B = 0$ in the above proof, we establish the following:

Corollary 3.6 *For a quantum-limited amplifier broadcast channel,* (3.39)–(3.40) *with* $N_B = 0$ *is equal to the capacity region.*

Coherent-detection and large κ limit

To evaluate the performance of the capacity region given by (3.39) and (3.40) with $N_B = 0$, we compare it with what can be achieved by conventional, coherent-detection strategies [116, 117]. When Alice inputs a coherent state $|\alpha\rangle$, Bob receives a displaced thermal state $D(\sqrt{\kappa}\alpha)\rho_{\bar{\kappa}}^{\text{th}} D^\dagger(\sqrt{\kappa}\alpha)$, where $D(\sqrt{\kappa}\alpha)$ denotes a displacement operator and $\rho_{\bar{\kappa}}^{\text{th}}$ the density operator corresponding to a thermal state with mean photon number $\bar{\kappa}$ [101]. When Bob employs homodyne or heterodyne detection [101], his measurement

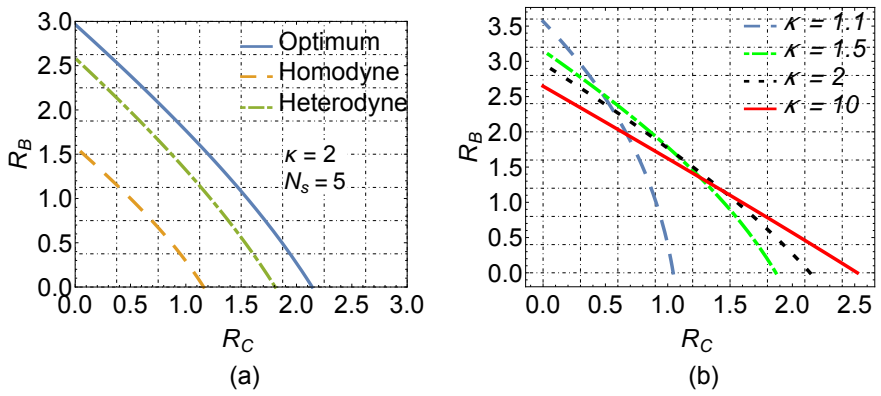

Figure 3.5: In (a) we consider a quantum-limited broadcast amplifier channel with $N_S = 5$ and $N_B = 0$. We compare the capacity region obtained by homodyne detection ((3.66) with $\xi = 1/2$), heterodyne detection ((3.66) with $\xi = 1$) and the optimal measurement ((3.39)–(3.40) with $N_B = 0$). In (b) we plot the large κ limit of the rate region. At $\kappa = 10$, it is indistinguishable with the limit in (3.67).

outcomes have particular Gaussian distributions, and similarly for Charlie. The quantum broadcast channel then reduces to a classical Gaussian channel with additive noise [118]. Using known results for classical Gaussian broadcast channels [118, 117, 71], we find that coherent-detection strategies lead to the following capacity regions:

$$\begin{aligned} R_B &\leq \xi \log_2\left(1 + \frac{\lambda\kappa N_S}{\xi(\xi + \bar{\kappa})}\right) , \\ R_C &\leq \xi \log_2\left(1 + \frac{(1-\lambda)\bar{\kappa} N_S}{\xi(\xi + \bar{\kappa}) + \lambda\bar{\kappa} N_S}\right) , \end{aligned} \tag{3.66}$$

where $\xi = 1/2$ for homodyne detection and $\xi = 1$ for heterodyne detection. See Appendix B.2 for a detailed derivation.

In Figure 3.5(a), we compare these strategies with the optimal strategy for a quantum-limited amplifier with $\kappa = 2$ and $N_S = 5$. As we can see, the capacity region we find in (3.39) and (3.40) outperforms both coherent detection schemes. For relatively high mean photon number, heterodyne detection outperforms homodyne detection as expected from prior results [102].

Notice that in the first equation of (3.66), the amplifier gain κ happens to cancel out in

the case of heterodyne detection ($\xi = 1$). This indicates that amplifying will both boost and hurt the transmission rate, so that there should exist a 'balanced point'. Actually, if we consider the large κ limit, (3.39) and (3.40) reduce to a gain-independent linear trade-off:

$$R_B + R_C \leq \log_2(N_S/[N_B + 1] + 1) \, . \tag{3.67}$$

Physically, although a large amplifier gain will amplify the input energy power and thus potentially increase the capacity, it is balanced out by the increasing noise generated from amplifying the vacuum, manifested by the negative terms in (3.39) and (3.40). With mean photon numbers $N_S = 5$ and $N_B = 0$, the maximum classical capacity of Bob and Charlie converges to around $\log_2(6) \approx 2.58$ bits per channel use. In Figure 3.5(b), we plot the rate region for amplifier gain κ increasing from 1.1 to 10. The capacity region converges to (3.67) very quickly. The maximum capacities for both receivers approach around 2.6 bits per channel use, as expected from the reasoning above.

3.5 TRADING PUBLIC AND PRIVATE RESOURCES

Here we briefly argue that we obtain the private dynamic capacity region [100] of quantum-limited amplifier channels. The techniques for establishing this result are similar to those from previous sections, so we merely state the result rather than going through all the details.

The information-theoretic task is similar to the triple trade-off discussed previously, but the resources involved are different. Here we are concerned with the transmission (or consumption) of public classical bits, private classical bits, and secret key along with the consumption of many independent uses of a quantum-limited amplifier channel. The communication trade-off is characterized by *rate triples* (R, P, S), where R is the net rate of public classical communication, P is the net rate of private classical communication,

and S is the net rate of secret key generation.

Since the quantum-limited amplifier channel is a Hadamard channel, the private dynamic capacity region of a quantum channel $\mathcal{N}$ is given by the union of regions of the following form [100]:

$$\begin{aligned} R + P &\leq H(\mathcal{N}(\rho)) - \sum_{x,y} p_X(x) p_{Y|X}(y|x) H(\mathcal{N}(\psi_{x,y})), \\ P + S &\leq \sum_x p_X(x) \left[H(\mathcal{N}(\rho_x)) - H(\mathcal{N}^c(\rho_x)) \right], \\ R + P + S &\leq H(\mathcal{N}(\rho)) - \sum_x p_X(x) H(\mathcal{N}^c(\rho_x)), \end{aligned} \tag{3.68}$$

where the union is with respect to all possible pure-state input ensembles $\{p_X(x)p_{Y|X}(y|x), \psi_{x,y}\}$,

$$\rho_x \equiv \sum_y p_{Y|X}(y|x)\psi_{x,y}, \tag{3.69}$$

$$\rho \equiv \sum_x p_X(x)\rho_x, \tag{3.70}$$

and $\mathcal{N}^c$ is a complementary channel of $\mathcal{N}$. To give an upper bound on the single-letter private dynamic capacity region of the quantum-limited amplifier channel, we need to show that for all input ensembles $\{p_X(x)p_{Y|X}(y|x), \psi_{x,y}\}$, there exists a $\lambda \in [0,1]$ such that the following four inequalities hold

$$H(\mathcal{N}(\rho)) \leq g(\kappa N_S + \bar{\kappa}) , \tag{3.71}$$

$$\sum_x \sum_y p_X(x) p_{Y|X}(y|x) H(\mathcal{N}(\psi_{x,y})) \geq g(\bar{\kappa}) , \tag{3.72}$$

$$\sum_x p_X(x) H(\mathcal{N}(\rho_x)) \leq g(\kappa\lambda N_S + \bar{\kappa}) , \tag{3.73}$$

$$\sum_x p_X(x) H(\mathcal{N}^c(\rho_x)) \geq g(\bar{\kappa}(\lambda N_S + 1)) . \tag{3.74}$$

We can establish these bounds using methods from the previous sections. Thus, we find that the private dynamic capacity region of the quantum-limited amplifier channel is as follows:

$$R + P \leq g(\kappa N_S + \bar{\kappa}) - g(\bar{\kappa}), \tag{3.75}$$

$$P + S \leq g(\kappa\lambda N_S + \bar{\kappa}) - g(\bar{\kappa}(\lambda N_S + 1)), \tag{3.76}$$

$$R + P + S \leq g(\kappa N_S + \bar{\kappa}) - g(\bar{\kappa}(\lambda N_S + 1))). \tag{3.77}$$

This rate region is achievable as well, as shown in [98, 99], and so the union of (3.75)–(3.77) with respect to $\lambda \in [0, 1]$ is equal to the private dynamic capacity region.

3.6 Discussion

Theorem 3.1 from [103] plays an important role in our proof of the capacity regions for the information trade-off and quantum broadcast settings. For a long time now, thermal states have been conjectured to minimize the output entropy for pure-loss channels with an input entropy constraint [102]. The authors of [103] established this result for all single-mode phase-insensitive bosonic Gaussian channels, going well beyond the original conjecture and including it as a special case. The special case for $H_0 = 0$ was proved for all multi-mode phase-insensitive Gaussian channels [106, 89]. After that, de Palma *et al.* first reduced the optimizer problem to the set of all possible passive states [119] using the technique of majorization [120, 121, 122] and subsequently proved the conjecture for single-mode pure-loss channels [123]. The multi-mode generalization of the results in Ref. [103], which would determine capacity regions for pure-loss channels [102, 98], is still unsolved.

The strongest conjecture proposed so far is the Entropy Photon number Inequality (EPnI) [124] which takes on a role analogous to Shannon's entropy power inequality [125]. The truth of the EPnI subsumes all minimum output entropy conjectures. Although the

EPnI has not been proved yet, a different quantum analog of EPI, quantum EPI (qEPI) has been proved recently for a multi-mode lossy channel [126]. Although the qEPI does not imply the truth of the EPnI, the lower bounds given by the two inequalities are extremely close for a large range of parameters [126]. This fact strongly suggests the truth of the multi-mode EPnI. We give an upper bound in Appendix B.3 for the capacity region of information trade-off over the pure-loss channel by using the qEPI. This represents the first application of the qEPI to the information trade-off problem. The bound given by the qEPI is extremely close to the upper bound, if we assume the multi-mode minimum output entropy conjecture is true. Therefore, it is safe to say that the achievable rate region found in Ref. [98] is the optimal capacity region for all practical purposes.

Chapter 4
Review of Dynamical Decoupling

4.1 Introduction

One of the major difficulties in the realization of quantum computing and quantum information processing is to protect a quantum state from decoherence. Quantum error correction protocols were developed to meet this challenge [127, 128, 129]. However, fault-tolerant quantum computation requires the fidelity of quantum gate at the physical layer is be at least 99.99%. The recent development of surface codes implemented on superconducting quibt systems has lifted this threshold to 99.9% [130]. Therefore it is urgent to reduce the error at physical layer below the threshold. To meet this challenge, dynamical decoupling (DD) has been proposed as a way to counteract the interaction between a quantum system and the environment by an open-loop control field. The idea of using pulse sequences, to protect nuclear spins from classical decoherence, dates back to 1950, when the spin-echo method was developed [131]. Since then, many pulse methods have been developed in nuclear magnetic resonance spectroscopy [132]. In 1998, it was first pointed out that a similar technique, periodical dynamical decoupling (PDD), can be applied to open quantum systems [133]. By using a control field, DD coherently averages out the unwanted system-environment interaction through the application of tailored sequences of (ideally, instantaneous) pulses.

Later Viola and Zanardi showed that this average-out-effect can be understood as a symmetrizing procedure over the Pauli group space [134, 135]. However, the finite switching time in real experimental conditions makes the symmetrizing imperfect. Therefore, concatenated dynamical decoupling (CDD) was proposed to eliminate higher-order

errors in the interaction Hamiltonian [136]. Another advantage of CDD is its robustness against pulse errors. In real experiments, the imperfection of pulses, such as the rotation angle error and the finite width, is unavoidable. The concatenation of pulse sequences not only suppressed the unwanted interaction, but also the pulse errors to higher order [136, 137].

In 2008, Uhrig proposed a dynamical decoupling scheme now called Uhrig DD (UDD), in which he considered pulses applied at arbitrary time points [138]. In this case, not only we can optimize the performance of DD over different pulses at each time point, but also have the freedom to choose arbitrary pulse intervals. Uhrig solved this optimization problem and found that UDD is efficient, in the sense that the decoupling order is linear with the number of pulses applied. People even started to call UDD the optimal DD. Later the result was generalized to general decoherence model: two UDD sequences with different pulse directions are embedded into each other, form the so-called quadratic DD (QDD) [139]. However, it was soon realized that irrationally-timed DD sequences are more sensitive to both the form of the spectral cutoff and to inevitable pulse errors [140, 141, 142], while being less amenable to the additional compensation steps (e.g., via phase-shifts or composite pulses) that are needed to mitigate these errors for arbitrary input states [143, 144, 145, 146].

Although it turns out that UDD may not be practical at all, its development actually stimulated two important directions in the field of DD over the past decade. In Uhrig's seminal paper [147], instead of working in the time domain like before, he wrote the final fidelity as an integral in frequency domain. In this way, DD sequences could be considered as a high-pass filter and the fidelity as just the overlap between the noise spectrum and the DD filter. Later on, several works have been devoted to develop a systematic framework for DD in the frequency domain [148, 149, 150, 151]. Analyzing DD in the frequency domain not only provides a theoretical alternative, but it also connects DD

to fields like noise spectroscopy and pulse sequence engineering. It is the consideration of reducing pulse sequence complexity that led the attention back to the 'old' digital DD, where the pulse intervals are integer multiplies of some constant. The digtial timing structure is compatible with the digital circuitry and digital clock in a quantum computer. The most important progress in this direction is the Walsh DD (WDD) [152, 153, 154, 155], where many digital DD sequences can be simply understood as Walsh functions. It is well-known how to efficiently generate such shapes in the lab [156].

In the rest of this chapter, we review the mathematical framework of DD in time domain upon which following chapters are built. For the framework of DD in the frequency domain, please refer to [148, 149, 150, 151]. It involves different mathematical tools and offers a complementary understanding of DD.

4.2 Dynamical decoupling in the time domain

Control framework

We consider a single-qubit system S coupled to an uncontrollable quantum environment (bath) B, which forms a closed system on the Hilbert space $\mathcal{H}_S \otimes \mathcal{H}_B$. The system and the bath are coupled via an arbitrary interaction; that is, we let the joint evolution in the absence of a control to be generated by a Hamiltonian of the form

$$H \equiv H_S \otimes \mathbb{1}_B + H_{SB} + \mathbb{1}_S \otimes H_B \ , \tag{4.1}$$

where H_S and $H_B \equiv B_0$ are, respectively, the internal Hamiltonian for S and B alone.

We assume the interaction takes the general linear form,

$$H_{SB} = \sigma_x \otimes B_x + \sigma_y \otimes B_y + \sigma_z \otimes B_z \ , \tag{4.2}$$

The bath operators B_0, B_u with $u \in \{x, y, z\}$ are assumed to be bounded but other-

wise arbitrary (possibly unknown). In what follows, we shall use $\beta \equiv ||B_0||$ and $J \equiv \max_{u\in\{x,y,z\}}\{||B_u||\}$ to quantify the strength of the internal-bath dynamics vs. the system-bath interaction, with $||\cdot||$ being the operator norm.

DD is implemented via a control action on S alone, generated by a control Hamiltonian of the form $H_c(t) \otimes \mathbb{1}_B$. Let

$$U_c(t) \equiv \mathcal{T} \exp[-i \int_0^t H_c(t')dt'] \tag{4.3}$$

be the *control propagator*, with $\hbar = 1$ and $\mathcal{T}$ denoting time-ordering. The effect of the control field is easy to understand in the *toggling frame*, in which the modulated Hamiltonian is given by

$$\tilde{H}_e(t) = U_c^\dagger(t) H_e U_c(t) \ . \tag{4.4}$$

Since the DD objective is to achieve an identity gate on S, all the evolution induced by H contributes to unwanted error dynamics [157], whereby $H \equiv H_e$. The effect of H_e may be isolated by expressing the propagator $U(T)$, for evolution under $H(t) \equiv H_e + H_c(t)$ over time T, as

$$U(T) = U_c(T)\, \mathcal{T} \exp[-i \int_0^t \tilde{H}_e(t')dt'] \ , \tag{4.5}$$

where $U_c(T) = \mathbb{1}_S$ for DD.

We can think of the full evolution as generated by some effective time-independent Hamiltonian,

$$U(T) \equiv e^{-iH^{\text{eff}}(T)T} \equiv e^{-i\Omega_e(T)} \tag{4.6}$$

Sometimes it is more convenient to define the error action operator $\Omega_e(T)$ [157]. The

norm of $\Omega_e(T)$, up to pure-bath terms that do not enter the reduced dynamics, quantifies the achievable error per gate (EPG). Specifically, $\Omega_e(T)$ and the associate effective Hamiltonian may be obtained via a perturbative Magnus expansion,

$$\Omega_e(T) \equiv [H_{SB}^{\text{eff}}(T) + H_B^{\text{eff}}(T)]T \tag{4.7}$$

$$= \exp[\sum_{m=1}^{\infty} \Omega_e^{(m)}(T)] \,, \tag{4.8}$$

where $\Omega_e^{(m)}(T)$ is a time-ordered integral involving m^{th}-order nested commutators. The Magnus expansion converges as long as $||H||T < \pi$ [158].

The DD performance in the time domain is then characterized by the order of error suppression, or *cancellation order* (CO). CO is given by the leading order of the error action operator [159, 150]. That is, we say a DD sequence achieves α-order decoupling, if the following is true at control time T,

$$\text{EPG} \equiv ||\text{mod}_B(\Omega_e(T))|| = ||TH_{SB}^{\text{eff}}(T)|| = \mathcal{O}(T^{\alpha+1}) \,. \tag{4.9}$$

Here mod means we do not include the pure bath term since it does not contribute to the evolution of the qubit.

Digital pulse sequence

A idea pulse sequence is specified by pulse timings and the π-rotations, namely, $\{t_j, \sigma_j\}$. Thus, for an ideal pulse sequence specified by a total of N pulses over a running time T, the control Hamiltonian is given by

$$H_c(t) = \frac{\pi}{2} \sum_{j=1}^{N} \sigma_j \delta(t - t_j) \,, \tag{4.10}$$

where we let $t_0 \equiv 0, t_N \equiv T$, and $\sigma_j \in \{\sigma_u\}, \forall j$.

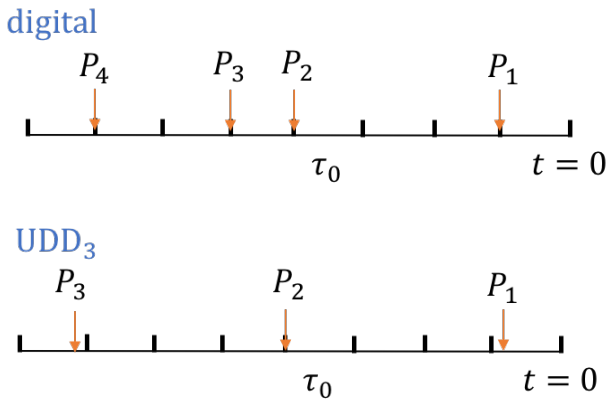

Figure 4.1: Digital and non-digital pulse sequences. The upper plot is an illustration of a digital pulse sequence, where the pulse intervals are integer multiplies of τ_0. The lower plot is UDD_3 in which the pulse timing are given by irrational numbers.

Crucially, a digital pulse sequence is a pulse sequence such that all inter-pulse separations obey $t_j - t_{j-1} \equiv n_j \tau_0$, with $n_j \in \mathbb{N}$. Here $\tau_0 > 0$ is the the *minimum pulse interval* determined by hardware limitations. Another convenient representation we shall use for the above sequence is

$$P_N \mathtt{f}_{n_N \tau_0} \ldots P_2 \mathtt{f}_{n_2 \tau_0} P_1 \mathtt{f}_{n_1 \tau_0} \ ,$$

where the $P_j \in \{X, Y, Z\}$ represent different π pulses and $\mathtt{f}_{n_j \tau_0}$ denotes free evolution between P_{j-1} and P_j. In Figure 4.1 we give examples of digital and non-digital pulse sequences.

Chapter 5
Concatenated-projection Dynamical Decoupling

As we mentioned in Chapter 4, by optimizing the pulse interval to suppress the low-frequency region of the noise spectrum, Uhrig dynamical decoupling (UDD) can achieve the same cancellation order (CO) with exponentially fewer pulses, compared to CDD [147]. Although it has superior performance, UDD is very sensitive to pulse errors, due to the fact that it only uses single-axis rotations. However, to protect unknown states, digital DD with multi-axis rotations can compensate errors due to its symmetric structure [141, 160, 140, 143].

Another development in dynamical decoupling is to use random pulses, instead of deterministic schemes, for sufficiently long sequences [161, 162]. Instead of trying to suppress the interaction to arbitrarily high orders, random dynamical decoupling schemes improves the time dependence of the error accumulation from quadratic to linear [161, 162].

In this chapter we only consider deterministic digital dynamical decoupling for two reasons: (1) it is easy to implement in experiments, and (2) it is robust against pulse errors as compared to non-digital ones. Besides analytical calculations, recently many other DD sequences have been found using genetic algorithms to optimize the CO [163], some of which even achieve the same CO with fewer pulses than CDD. This result suggests that there are undiscovered digital pulse sequences with performance better than those which are the best known so far. We confirm this conjecture by establishing a unified framework that encompasses almost all known digital DD schemes.

In this work, we propose concatenated-projection dynamical decoupling (CPDD) to unify all known uniform DD schemes. Our framework gives a way to construct new dig-

ital pulse sequences and to calculate their CO. In Sec. 5.1, we first define our projection pulse sequence and explain its effect as an 'atomic' projection. Then we use concatenated projections along different directions to construct more complex pulse sequences with arbitrary CO in Sec. 5.2. These results comprise the two cornerstones of our theory of CPDD. In Sec. 5.3 we formally introduce CPDD by defining the CPDD equivalence classes, which are specified by three integers. We first develop a series of properties of CPDD. We then design a deterministic scheme to construct an optimized pulse sequence for each CO. A table of known DD schemes is given as well to show how these known schemes fit into our CPDD framework. In Sec. 5.4, we discuss why, intuitively, some CPDD sequences are superior than CDD. We also point out a typo in Ref. [163], which is easily detected within the framework of CPDD.

5.1 PROJECTION PULSE SEQUENCE

Now consider that we apply K digital ideal π pulses with total control time T. Thus the control field takes the form of Eq. (4.10). In the limit of $\tau_0 \to 0$, we have a continuous pulse sequence, and only the zeroth-order term in the Magnus expansion survives,

$$H^{\text{eff}(0)}(T) = \frac{1}{T}\int_0^T U_c^\dagger(t) H_e U_c(t) dt. \tag{5.1}$$

Also, since the pulses are ideal, $U_c(t)$ is a piece-wise constant function. The first-order average Hamiltonian reduces to

$$H^{\text{eff}(0)} = \frac{1}{K}\sum_{j=1}^{K} U_c^\dagger(t_j) H_e U_c(t_j). \tag{5.2}$$

If we choose pulses such that $U_c(t_j)$ go through each element of a certain group such as $\mathcal{G} = \{I, X, Y, Z\}$, Eq. (5.2) is just a symmetrizing procedure which projects H_e onto the commutant of the group algebra [164, 135].

A more intuitive way to view the effect of a pulse sequence is to look at it as a combination of basic projections [165]. The simplest pulse sequence is P_iP_i, which is similar to the CPMG pulse sequence [166, 167]. Hereafter we call it the projection pulse sequence and use the notation $\mathtt{p}_i = P_iP_i$. To explain its projecting effect, we consider the spin-boson model, which induces the longitudinal decay of the spin. The interaction takes the form,

$$H_{SB} = \sum_k (g_k \sigma^+ \otimes b_k + g_k^* \sigma^- \otimes b_k^\dagger), \tag{5.3}$$

where b_k is the annihilation operator of a photon with momentum k, and g_k is the coupling strength between photon with mode k and the spin. Here $\sigma^\pm$ is the creation (annihilation) operator, $\sigma^\pm = \sigma_x \pm i\sigma_y$.

If we apply $\mathtt{p}_z = ZZ$ to the system, then $U_c(t_j) \in \mathcal{G} = \{\mathbb{1}, Z\}$. Using Eq. (5.2), the interaction term will be completely removed in the continuous limit due to the fact $\sigma_z \sigma_{x(y)} \sigma_z = -\sigma_{x(y)}$. Therefore, geometrically the pulse sequence $\mathtt{p}_z$ projects the Hamiltonian along the z direction.

However, in real experimental conditions, there is an upper limit of the pulse switching rate; thus τ_0 is finite. Although the Magnus expansion is still valid, so long as $||H||T << \pi$, the projection is not exact anymore due to higher-order corrections.

Theorem 5.1 *Assume the interaction between a single qubit and bath takes the form of Eq. (4.2). After applying the projection pulse sequence* $\mathtt{p}_j$ *(*$j = x, y, z$*) with pulse interval* τ_0*, the zeroth order effective Hamiltonian is given by,*

$$H_{SB}^{\mathit{eff}(0)} \equiv \pi_j^{(0)} H_e = \sigma_j \otimes B_j. \tag{5.4}$$

Thus the full error average Hamiltonian is given by,

$$H_{SB}^{\mathit{eff}} = \sigma_j \otimes [B_j + k_j^{(1)}(\tau_0)] + \sum_{i \perp j} \sigma_i \otimes [B_i^{(1)} + k_i^{(2)} o(\tau_0^2)], \tag{5.5}$$

where we use $\pi_j^{(0)}$ to represent the mapping from H_e to $H_{SB}^{\mathit{eff}(0)}$ that is induced by the projection pulse sequence $\mathtt{p}_j$. The symbol $\perp$ represents the directions orthogonal to direction j and $B_i^{(1)} \sim k_i^{(1)} o(\tau_0)$. Here $k_i^{(n)}$ is some combination of commutators of the bath operators with dimension of $[H]^n$.

Proof. Without loss of generality, we consider the pulse sequence is ZZ, where $Z = -i\sigma_z$. The transformed Hamiltonian $\tilde{H}(t)$ is given by

$$\tilde{H}_e(t) \equiv \begin{cases} H_1, \ 0 < t < \tau_0 \\ H_2, \ \tau_0 < t < 2\tau_0 \ , \end{cases} \tag{5.6}$$

where

$$H_1 = \mathbb{1} \otimes B_0 + \sigma_x \otimes B_x + \sigma_y \otimes B_y + \sigma_z \otimes B_z, \tag{5.7}$$

$$H_2 = \mathbb{1} \otimes B_0 - \sigma_x \otimes B_x - \sigma_y \otimes B_y + \sigma_z \otimes B_z. \tag{5.8}$$

After applying the pulse sequence $P_j P_j$, the first and second order expansion of the average Hamiltonian is given by,

$$H^{\text{eff}(0)} = \frac{1}{2} \sum_{i=1}^{2} H_i, \tag{5.9}$$

$$H^{\text{eff}(1)} = \frac{-i}{2} \tau_c \sum_{i<j} [H_i, H_j]. \tag{5.10}$$

Using Eq. (5.6), we have

$$H^{\text{eff}(0)} = \mathbb{1} \otimes B_0 + \sigma_z \otimes B_z, \tag{5.11}$$

$$H^{\text{eff}(1)} = \tau_0[B_0, B_x] \otimes \sigma_x + \tau_0[B_0, B_y] \otimes \sigma_y + 2\tau_0[B_y, B_x] \otimes \sigma_z. \tag{5.12}$$

Since the commutator between bath operators is not zero in general, we have

$$H_{SB}^{\text{eff}(0)} = \sigma_z \otimes [B_z + k_z^{(1)} o(\tau_0)] + \sum_{i \perp z} \sigma_i \otimes [B_i^{(1)} + k_i^{(2)} o(\tau_0^2)], \tag{5.13}$$

where $B_i^{(1)} = \tau_0[B_0, B_i]$. The same calculation gives similar results for $j = x, y$. ■

This projection point of view gives a geometrical and intuitive way to understand the effect of the pulse sequence $P_j P_j$. In summary, we have shown that the effect of the projection pulse sequence $\mathrm{p}_i = P_i P_i$ is to project the Hamiltonian along i direction up to first order.

5.2 Concatenation of cyclic pulse sequences as successive projections

The higher-order terms remaining in H_{SB}^{eff}, after applying the projections, will coherently add up with time. To achieve higher-order suppression, we need to project the Hamiltonian along different directions successfully. We will show in this section that the effect of the concatenation of cyclic pulse sequences is to apply successively the projections induced by each pulse sequence.

A pulse sequence is called cyclic when the generated evolution operator is periodic with period T up to a phase factor,

$$U_c(nT) = e^{i\phi} U_c(T), \tag{5.14}$$

where $n = 0, 1, 2 \ldots$ and ϕ is an arbitrary phase.

An equivalent definition of a cyclic pulse sequence is that the product of all the pulses is equal to one, up to an arbitrary phase,

$$\prod_{i=1}^{K} P_i = e^{i\phi} \mathbb{1}, \tag{5.15}$$

which follows directly from Eq. (5.14) when $n = 0$. We will see that this property is necessary for the proof of the equivalence between concatenation and successive projections.

Another useful property of the cyclic pulse sequence is that the concatenation of two cyclic pulse sequences is still cyclic.

Lemma 5.1 *The concatenation of two cyclic pulse sequences is still cyclic.*

Proof. Consider two cyclic pulse sequences A and B. From Eq. (5.15) we have

$$\prod_{i=1}^{K_A} P_i^A = e^{i\phi_A} \mathbb{1}, \tag{5.16}$$

$$\prod_{i=1}^{K_B} P_i^B = e^{i\phi_B} \mathbb{1}. \tag{5.17}$$

The pulse sequence C is constructed by concatenating A and B, and thus

$$\begin{aligned} C &= A[B] \\ &\equiv P_1^A (P_1^B \ldots P_{K_B}^B) P_2^A (P_1^B \ldots P_{K_B}^B) \ldots P_{K_A}^A (P_1^B \ldots P_{K_B}^B) \end{aligned} \tag{5.18}$$

Therefore the product of all pulses of sequence C is

$$\begin{aligned}\prod_{i=1}^{K_C} P_i^C &= P_1^A(\prod_{i=1}^{K_B} P_i^B)....P_{K_A}^A(\prod_{i=1}^{K_B} P_i^B) \\ &= P_1^A e^{i\phi_B}...P_{K_A}^A e^{i\phi_B} \\ &= e^{iK_B\phi_B}\prod_{i=1}^{K_A} P_i^A \\ &= e^{i(K_B\phi_B+\phi_A)}\mathbb{1} \\ &= e^{i\phi_C}\mathbb{1} \end{aligned} \tag{5.19}$$

where we define $\phi_c = K_B\phi_B + \phi_A$. Therefore, the pulse sequence C=$A[B]$ is also cyclic. ■

Having the definition of cyclic pulse sequence, we now prove the second basic theorem of our CPDD scheme.

Lemma 5.2 *Consider two pulse sequences A and B, $P_K^i...P_1^i$, where i = A, B, with the same pulse interval. The first pulse sequence $A = P_jP_j$ ($j \in \{x, y, z\}$), which is a projection pulse sequence, and sequence B is concatenated from multiple projection pulse sequences. A third pulse sequence C is constructed by concatenating A and B, $C = A[B] \equiv P_jBP_jB$. The following relationship holds,*

$$\pi_C^0 = \pi_B^0\pi_A^0, \tag{5.20}$$

where the mapping π_i^0 induced by applying the pulse sequence i is defined in Theorem 5.1.

Proof. The concatenated sequence C is given by

$$(P_iP_K^B)P_{K-1}^B...P_1^B(P_iP_K^B)...P_1^B, \tag{5.21}$$

where the bracket means that there is no free evolution in between the two pulses inside the bracket. Since the projection pulse sequence is cyclic by definition, and sequence B is

also cyclic by Theorem 5.1,

$$\prod_{i=1}^{K} P_i^B = e^{i\phi}\mathbb{1}. \tag{5.22}$$

The $2K_B$ evolution operator of the control field, U_m^C is given by,

$$U_m^C = \prod_{j\leq m} P_j^C \tag{5.23}$$

To construct $\pi_A^{(0)}$ and $\pi_B^{(0)}$, we group U_m^C and $U_{m+K_B}^C$ together($m \leq K_B$).

For $1 \leq m < K_B$,

$$U_m^C \;=\; \prod_{j\leq m} P_j^B = U_m^B, \tag{5.24}$$

and,

$$\begin{aligned} U_{m+K_B}^C &= \Big(\prod_{j\leq m} P_j^B\Big) P_i \Big(\prod_{j>m}^{K_B} P_j^B\Big)\Big(\prod_{j\leq m} P_j^B\Big) \\ &= e^{i\phi} P_i \prod_{j\leq m} P_j^B \\ &= e^{i\phi} P_i U_m^B\,, \end{aligned} \tag{5.25}$$

where we have used the commutativity of Pauli matrices and the cyclic property Eq. (5.22) of pulse sequence B.

Now adding the action of U_m^C and $U_{m+K_B}^C$ on H_{SB} together, we have

$$\begin{aligned} & U_m^{C\dagger} H_{SB} U_m^C + U_{m+K_B}^{C\dagger} H_{SB} U_{m+K_B}^C \\ =\;& U_m^{B\dagger} H_{SB} U_m^B + U_m^{B\dagger} P_i^\dagger H_{SB} P_i U_m^B \\ =\;& U_m^{B\dagger} (2\pi_A^{(0)} H_{SB}) U_m^B \end{aligned} \tag{5.26}$$

If $m = K_B$, we have $U^C_{K_B}$ and $U^C_{2K_B}$, which are

$$\begin{aligned} U^C_{K_B} &= P_i P^B_K \prod_{j<K_B} P^B_j \\ &= P_i \prod_{j\leq K_B} P^B_j \\ &= e^{i\phi} P_i U^B_{K_B} \end{aligned} \tag{5.27}$$

$$\begin{aligned} U^C_{2K_B} &= P_i\Big(\prod_{j\leq K_B} P^B_j\Big) P_i\Big(\prod_{j\leq K_B} P^B_j\Big) \\ &= e^{i\phi} U^B_{K_B}. \end{aligned} \tag{5.28}$$

Now using Eqs. (5.26, 5.27, 5.28), the first order of average Hamiltonian after applying sequence C is given by

$$\begin{aligned} \bar{H}^{(0)}_C &= \frac{1}{2K_B}\sum_{m=1}^{2K_B} U^{C\dagger}_m H_{SB} U^C_m \\ &= \frac{1}{2K_B}\sum_{m=1}^{K_B}\Big(U^{C\dagger}_m H_{SB} U^C_m + U^{C\dagger}_{m+K_B} H_{SB} U^C_{m+K_B}\Big) \\ &= \frac{1}{K_B}\sum_{m=1}^{K_B} U^{B\dagger}_m\Big(\frac{1}{2}\sum_{l=1}^{K_A} U^{A\dagger}_l H_{SB} U^A_l\Big) U^B_m \\ &= \pi^{(0)}_B \pi^{(0)}_A H_{SB}\ . \end{aligned} \tag{5.29}$$

■

Lemma 5.2 is the theoretical cornerstone of this work. It explains why concatenation can increase the CO, which is not so obvious. The cyclic properties and the changeability of different pulses (up to an irrelevant phase factor) are necessary for the proof.

5.3 CONCATENATED PROJECTIONS DYNAMICAL DECOUPLING

What really distinguishes our work from the CDD scheme is that we chose the projection pulse sequence as the basic element of concatenation. Motivated by Theorems 5.1 and 5.3, we define our concatenated-projection dynamical decoupling (CPDD) as a new way to construct pulse sequences by applying projections along different directions successively. Since each projection kills the interaction terms orthogonal to it by one more order, by appropriately combining different projections, our CPDD can achieve arbitrarily high CO.

Definition 5.2 *A CPDD pulse sequence is specified by an ordered series $i_N, i_{N-1}, ..., i_1$. It is constructed by concatenating N projection pulse sequences successively, $A = \mathtt{p}_{i_N}[\mathtt{p}_{i_{N-1}}[...[\mathtt{p}_{i_1}\underbrace{]...]}_{N}$, where $i_j \in \{x, y, z\}$ and $1 \leq j \leq N$.*

The suppressing effect of the CPDD sequence on the Hamiltonian follows immediately from the combination of the effects of projections and concatenation.

Theorem 5.3 *: Consider a CPDD pulse sequence A specified by $i_N, i_{N-1}, ..., i_1$. After applying pulse sequence A, the average interaction Hamiltonian is given by*

$$H_{SB}^{\mathit{eff}} = \sum_{i=x,y,z} \sigma_i \otimes [B_i^{(d_i)} + k_i^{(d_i+1)} o(\tau_0^{d_i+1})], \tag{5.30}$$

where

$$d_i = \sum_{j \perp i} n_j, \tag{5.31}$$

and n_j is the number of $\mathtt{p}_j$ sequences.

Proof. Repeatedly using Lemma 5.2, the leading order of the error average Hamiltonian after applying sequence A is given by

$$\pi_A^{(0)} H_e = \pi_{i_1}^{(0)} \pi_{i_2}^{(0)} ... \pi_{K_A}^{(0)} H_e \, . \tag{5.32}$$

From Theorem 5.1, each projection removes the first order term in the perpendicular direction. Therefore,

$$H_{SB}^{\text{eff}(0)} = \sum_{i=x,y,z} \sigma_i \otimes B_i^{(d_i)}, \tag{5.33}$$

where

$$d_i = \sum_{j \perp i} n_j. \tag{5.34}$$

■

From Theorem 5.3 we can see that the effect of n_i p_is is to suppress the error Hamiltonian along the i direction to n_ith order. From Eq. (5.34) we also notice that the order of how different projection pulse sequences is concatenated does not affect the leading order of the average Hamiltonian along each direction. Therefore we can define an equivalence relationship between different CPDD pulse sequences,

Definition 5.4 *Consider two pulse sequences A and A'. The leading order of the error Hamiltonians induced by each of them are $H_{SB}^{\text{eff}(0)} = \sum_{i=x,y,z} \sigma_i \otimes B_i^{(d_i)}$ and $H_{SB}^{\text{eff}(0)'} = \sum_{i=x,y,z} \sigma_i \otimes B_i^{(d_i')}$. We define A and A' to be equivalent to each other*

$$A \sim A', \tag{5.35}$$

if the leading order of the average error Hamiltonians are the same along each direction, namely $n_i = n_i'$.

It can be easily proved that the relationship defined above satisfies the three properties of an equivalence relationship.

Therefore, for a CPDD sequence specified by sequence $a = i_N, i_{N-1}, ..., i_1$, all CPDD sequences specified by a's permutations $A\{i_N, i_{N-1}, ..., i_1\}$ form an equivalence class. By virtue of the equivalence class, only three numbers n_x, n_y, n_z are needed to completely

specify a CPDD class.

Definition 5.5 *: A CPDD class is defined as an equivalence class with equivalence relationship defined in Definition 5.4, specified by three integers,* $\{n_x, n_y, n_z\}$*. The structure of the pulse sequence can be generated by concatenating all* n_i $\mathtt{p}_i$ *sequences* ($i = x, y, z$) *in arbitrary order.*

From the definition of CPDD and Theorem 5.3, we derive a series of properties satisfied by CPDD sequences and their equivalence classes.

Properties of CPDD

Due to the way concatenation connects two pulse sequences, we can derive two properties that the structure of each CPDD sequence must satisfy.

1. *For an arbitrary CPDD pulse sequence, each odd site has the same kind of* π *pulse.*

We prove this by induction. Consider a pulse sequence $A = P_K P_{K-1} ... P_1$, which is concatenated from N projection pulse sequences. Pulse sequences A_n are concatenated from the first n ($1 \leq n \leq N$) of them. The following relations are satisfied,

$$A_n = \mathtt{p}_i[A_{n-1}], \tag{5.36}$$

where $\mathtt{p}_i$ is the nth projection pulse sequence. Assume for subsequence A_{n-1} the pulses are the same for each odd site,

$$A_{n-1} = P_{2n-2} P_0 ..... P_2 P_0. \tag{5.37}$$

Let's examine the pulse sequence A_n,

$$\begin{aligned} A_n &= \mathtt{p}_i[A_{n-1}] \\ &= (P_i P_{2n-2}) P_0 ... P_2 P_0 (P_i P_{2n-2}) P_0 ... P_2 P_0. \end{aligned} \tag{5.38}$$

Therefore, the fact that all the pulses at each odd site are the same still holds.

Since $A_2 = \mathtt{p}_i[\mathtt{p}_j] = (P_iP_j)P_j(P_iP_j)P_j$, which also has the same kind of pulse on its odd sites, by induction we have proved that for each odd sites of A_N the pulses are the same:

$$P_{2m+1} = P_0, \quad m = 0, 1, ... \tag{5.39}$$

2. *For an arbitrary CPDD pulse sequence, the first half and the second half subsequences are the same.*

Again this property follows from the definition of concatenation. Using Eq. (5.36) for $n = N$ we have

$$A = \mathtt{p}_{i_n}[A_{n-1}]. \tag{5.40}$$

Assume $A_{n-1} = P_{2N-2}P_{2N-3}...P_1$, we have

$$A = (P_{i_n}P_{2N-2})P_{2N-3}...P_1(P_{i_n}P_{2N-2})P_{2N-3}...P_1. \tag{5.41}$$

Obviously sequence A is composed the same two copies of A_{N-1}, or more precisely

$$P_m = P_{m-N/2}, \quad m = 1, 2, ..., N. \tag{5.42}$$

3. *For CPDD class $\{n_x, n_y, n_z\}$, the number of pulses or sequence length K is given by*

$$K_{n_x,n_y,n_z} = 2^{n_x+n_y+n_z}. \tag{5.43}$$

The proof is straightforward. At first, each basic projection is induced by the two same pulses. Secondly, the length of two concatenated pulse sequences, A and B, is equal to the product of the length of each two, $K_{A[B]} = K_A K_B$. Therefore, for a pulse sequence composed of n_i pairs of (P_i, P_i), the total pulse number K is given by Eq. (5.43).

4. *For the CPDD class $\{n_x, n_y, n_z\}$, the CO achieved is given by,*

$$\mathrm{CO}_{n_x,n_y,n_z} = \min\{n_y + n_z, n_x + n_z, n_x + n_y\}. \tag{5.44}$$

From Theorem 5.3, the leading order of the error Hamiltonian induced by any pulse sequence in the CPDD class $\{n_x, n_y, n_z\}$ is given by

$$\bar{H}^{(0)}_{err} = \sum_{i=x,y,z} \sigma_i \otimes B_i^{(d_i)}, \tag{5.45}$$

where $d_i = \sum_{j \perp i} n_j$. Since the CO α is defined as the leading order of $\bar{H}$, $\alpha = \min\{d_x, d_y, d_z\}$.

From the expression of CO, Eq. (5.44), we can see that simply increasing pulse numbers (number of projections) does not necessarily increase the CO. Actually, in the framework of CPDD, only for pulse sequences with certain pulse numbers CO could increase.

5. *For a given CO α, the minimum number of pulses, K_{min}, required to achieve such CO is given by*

$$\log_2(K_{\min}) = \frac{1}{2}[3\alpha + \frac{1}{2}(1 - 1^{\oplus\alpha})], \tag{5.46}$$

where $1^{\oplus\alpha} = \underbrace{1 \oplus 1 \oplus \ ... \ \oplus 1}_{\alpha}$.

To achieve CO α, 3α terms in the interaction Hamiltonian need to be eliminated due to

the form of H_{SB}, Eq. (4.2). However, each basic projection π_i requires two pulses, which implies that only an even number of terms can be eliminated for a given CPDD sequence. Therefore, we need to add one more pulse depending on whether α is odd or not. Dividing the total number of eliminated terms by two we have,

$$\sum_{i=x,y,z} n_i = \frac{1}{2}[3\alpha + \frac{1}{2}(1 + (-1)^{\alpha+1}]. \tag{5.47}$$

Using the results of property 3, we have Eq. (5.46). The mysterious series $4, 8, 32, 64, 256...$ was first found by a genetic algorithm in Ref. [163] which is now understood, thanks to our unifying framework of CPDD.

Optimized digital dynamical decoupling

The pulse sequences corresponding to the pulse number in Eq. (5.46) uses the minimum number of pulses at each CO. This optimized digital dynamical decoupling (OUDD) scheme can be represented using the CPDD indexes as following

$$\text{ODDD}_k : \frac{1}{2}\{k - 1^{\oplus k}, k + 1^{\oplus k}, k + 1^{\oplus k}\}. \tag{5.48}$$

The CO of ODDD_k is $\alpha_k = k$ and the sequence length is given by $K_{\min}$ in Eq. (5.46).

As we can see, some of the pulse sequences are particular levels of CDD_l and GA8_l. To compare with ODDD and other known DD schemes, we list the corresponding CPDD indexes of known DD schemes in Table 5.1 below.

Table 5.1: Known DD schemes represented as CPDD

$\{n_x, n_y, n_z\}$	Name	Pulse sequence	K	N
$\{0, 0, 1\}$	Projection	$P_i P_i$	2	0
$\{0, 1, 1\}$	PDD(CDD$_1$)	$P_i P_j P_i P_j$	4	1
$\{1, 1, 1\}$	GA8$_a$	$IP_iP_jP_iIP_iP_jP_i$	8	2
$\{0, l, l\}$	CDD$_l$	CDD[CDD$_{l-1}$]	4^l	l
$\{l, l, l\}$	GA8$_l$	GA8$_a$[GA8$_{l-1}$]	8^l	$2l$

5.4 Discussion

Although CDD also relies on concatenation, the fact that our CPDD uses basic projections as building blocks makes finding more efficient pulse sequences possible. To make this clear, we consider the CDD from the view point of our CPDD. CDD$_1$ can be considered as the concatenation of two different projections, CDD$_1 = X(YY)X(YY) = ZYZY$ [137]. Therefore the effect of CDD$_1$ is to apply projections along the y and x direction successively,

$$\begin{aligned} \pi^{(0)}_{CDD_1} H_{SB} &\equiv \pi^{(0)}_y \pi^{(0)}_x H_0 \\ &= \pi^{(0)}_y [\sigma_x \otimes B_x + \sigma_y \otimes B^{(1)}_y + \sigma_z \otimes B^{(1)}_z] \\ &= \sigma_x \otimes B^{(1)}_x + \sigma_y \otimes B^{(1)}_y + \sigma_z \otimes B^{(2)}_z \\ &\sim k^{(1)} o(\tau_0^1). \end{aligned} \tag{5.49}$$

As we can see, CDD$_1$ completely removes the zeroth order interaction terms, thus achieving CO$_{\text{CDD}_1} = 1$.

Now consider CDD$_2$, which is the concatenation of two CDD$_1$ sequences: we write explicitly the process of the successive projections ,

$$
\begin{aligned}
\pi^0_{CDD_2} H_{SB} &= \pi^{(0)}_y \pi^{(0)}_x \pi^{(0)}_y \pi^{(0)}_x H_0 \\
&= \pi^{(0)}_y \pi^{(0)}_x [\sigma_x \otimes B^{(1)}_x + \sigma_y \otimes B^{(1)}_y + \sigma_z \otimes B^{(2)}_z] \\
&= \sigma_x \otimes B^{(2)}_x + \sigma_y \otimes B^{(2)}_y + \sigma_z \otimes B^{(4)}_z \\
&\sim k^{(2)} o(\tau_0^2). \qquad (5.50)
\end{aligned}
$$

As we can see from above, a total of eight eliminations (each projection pulse sequence eliminates two terms in the orthogonal directions) are used to completely remove the first two orders of the interaction H_{SB}. However, the two additional eliminations of B_z does not contribute to further increasing of the CO. To avoid this, we consider projecting along each direction exactly once, namely $\pi_x \pi_y \pi_z$, which belong to the CPDD class $\{1, 1, 1\}$. Translating the projection back to corresponding pulse sequence according to the rule of concatenation, we have

$$
\begin{aligned}
\pi^{(0)}_x \pi^{(0)}_y \pi^{(0)}_z &: \mathtt{p}_x[\mathtt{p}_y[\mathtt{p}_z]] \\
&= \mathtt{p}_x[Y(ZZ)Y(ZZ)] \\
&= X(XZXZ)X(XZXZ) \\
&= IZXZIZXZ, \qquad (5.51)
\end{aligned}
$$

which only uses eight pulses and six projections. This was first found by a genetic algorithm and called GA8$_a$ in Ref. [163].

To achieve CO of $\alpha = 2$, CDD$_2$ needs 16 pulses while GA8a sequences only requires 8 pulses. This efficiency of using pulses comes from the very fact that GA8a uses basic

projections π_i as building blocks while CDD_2 uses composed projections $\pi_i\pi_j$ ($i \neq j$) as building blocks.

In Ref. [163], the author claimed to find another 8 pulse sequence GA8_b=$Z(XYXY)Z(XYXY)$ which also achieved CO of $\alpha = 2$. From the structure of GA8_b, we know the projections induced by it is,

$$\pi^{(0)}_{GA8_b} = \pi^{(0)}_z \pi^{(0)}_z \pi^{(0)}_y, \tag{5.52}$$

which belong to CPDD class $\{0, 1, 2\}$. Using the results from Eq. (5.44), the CO of GA8_b is equal to 1. Therefore the claim in Ref. [163] is a typo.

To double check our results, we also use the multi-precision package *mpmath* [1] to compute the CO of both GA8_a and GA8_b for a 5-spin model with random coupling constants. Here the distance D is defined as the distance between an actual evolution operator and the unit operator [163],

$$D(U, \mathbb{1}_S) = \sqrt{1 - \frac{1}{d_{\mathcal{H}}}||\Gamma||_1}, \tag{5.53}$$

where $d_{\mathcal{H}}$ is the dimension of the Hilbert space, $\Gamma = \text{Tr}_S\{U\}$ and $||\cdot||_1$is the trace norm. An upper bound of D can be calculated[163],

$$D \lesssim O[\tau_0^{\alpha+1}]. \tag{5.54}$$

Therefore, we can extract the CO by plotting D versus τ_0 in the log-log diagram. As we can see in Figure 5.1, GA_{8a} achieves higher CO than GA_{8b} which is in agreement with the argument from the view point of our CPDD.

The main advantage of using UDD is that the pulse number needed scales linearly

[1] F. Johansson et al. MPMATH: a Python library for arbitraryprecision floating-point arithmetic (version 0.18), December 2013; http://mpmath.org/

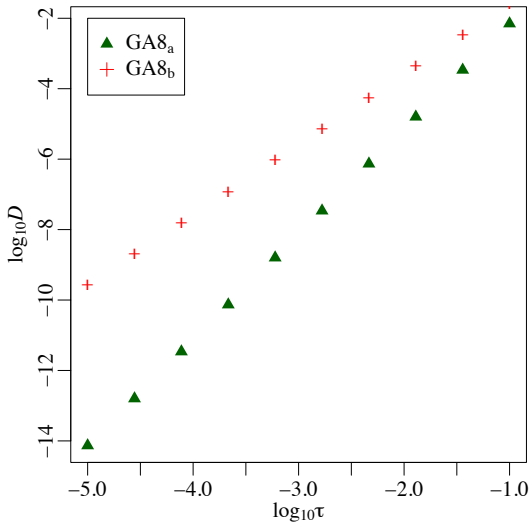

Figure 5.1: Comparing the cancellation order of $GA8_a$ and $GA8_b$. D is the dimensionless distance between actual evolution operator and the unit operator, defined in [163], and τ_0 is the pulse interval with unit second. The cancellation order α is defined by the relation $D \sim O(\tau_0^{\alpha+1})$. We consider the parameters J in the range of $J\tau_0 \in [10^{-6}, 10^{-1}]$, where J is the norm of the interaction Hamiltonian.

with the CO, $N_{UDD} \sim \mathcal{O}(K)$, which is much more efficient than the exponentially dependence in CPDD. However, UDD is subject to several difficulties. Firstly it is valid only for environmental spectrum with a hard cut-off [168, 169], and secondly it is very sensitive to pulse errors [141, 160, 170, 143]. However, for our CPDD, especially the OUDD class, some pulse sequences have rotation symmetry thus making them robust against pulse error. Although we have not given a rigorous bound analysis for CPDD here, the results and the calculation should be similar to those in Ref. [137]: the distance between the actual state and the desired state goes to zero as the concatenation level goes to infinity. More importantly, pulse errors are suppressed along with series of concatenations as

long as the error is not too large.

Chapter 6
General Walsh Dynamical Decoupling

As we mentioned in Chapter 4, digital DD sequences are highly compatible with hardware constraints stemming from digital sequencing circuitry and clocking, which makes them attractive in terms of minimizing sequencing complexity, as ultimately demanded for large-scale implementations. In this chapter, we provide a general framework for constructing digital dynamical decoupling sequences based on Walsh modulation — applicable to arbitrary qubit decoherence scenarios [32].

Control modulation based on Walsh functions [156], has been proposed as a unifying approach for generating digital-efficient protocols, for both dynamically corrected quantum storage and gates [152, 153, 154, 155]. Walsh DD (WDD) has been shown to naturally incorporate existing digital sequences as special instances (including concatenated DD for both single- and multi-axis decoherence [171]), and to provide a restricted search space for numerical sequence optimization and analytic performance analysis under finite timing resources. For dephasing noise on a qubit, concatenated DD sequences based on single-axis control are provably optimal, among the Walsh suite, in the sense of guaranteeing a desired order of error suppression with minimum total pulse number [152].

In this chapter, we identify optimal single-qubit WDD sequences capable of canceling out the simultaneous dephasing and relaxation effects that arise from arbitrary environmental couplings. The key step is to generalize existing sequence constructions of WDD based on multi-axis control, and establish formal equivalence of the resulting general WDD formalism with the concatenated-projection DD (CPDD) approach proposed in Ref. [31] (see Chapter 5 for details). By leveraging this equivalence, we explicitly char-

acterize the error-suppression capabilities of *any* general WDD sequence, along with its complexity in terms of the required control time slots. We show that, unlike in the dephasing scenario, concatenated DD is no longer optimal, and identify a large family of *optimal* WDD (OWDD) schemes, whose complexity is exponentially smaller for the same order of suppression. While the performance of different OWDD sequences depends additionally on the specific control path, our analysis indicates that OWDD can substantially improve over existing digital schemes in relevant parameter regimes.

The rest of this chapter is organized as follows. In Sec. 6.1 we first review single-axis WDD and CPDD formalism. We then give the formal definition of GWDD and show its equivalence to CPDD. We analyze the error suppression capability of GWDD sequences in Sec. 6.2. Specifically we 1) derive a formula to calculate the cancellation order; 2) provide a analytical procedure to calculate an upper bound on the decoupling error and 3) illustrate the control sensitivity and give a intuitive understanding of it.

6.1 General Walsh dynamical decoupling

We first review single-axis Walsh DD formalism and a recently proposed digital DD framework called CPDD. We then generalize WDD to include multi-axis pulses such that it protects qubit against arbitrary general decoherence. Then we establish the equivalence between newly defined GWDD and CPDD. The equivalence established here will be crucial in our later analysis on the performance of OWDD sequences.

Walsh vs. concatenated-projection DD formalism

The Walsh functions are a well-known family of binary-valued piecewise-constant functions orthonormal over $[0, 1]$, which may be naturally employed to describe digital DD sequences [156, 152]. For dephasing noise, single-axis control via π-pulses around

(say) the x-axis suffices in the ideal case, resulting in a control propagator of the form

$$U_c(t) \equiv \sigma_x^{[x(t)+1]/2} , \tag{6.1}$$

where the control switching function $x(t)$ toggles between the values ± 1 at instants corresponding to the applied pulse timings. Let the Walsh function of *Paley order* n be defined as

$$W_n(x) \equiv \prod_{j=1}^{m} R_j(x)^{b_j}, \quad x \in [0,1], \tag{6.2}$$

where $\{b_j\}$ is the binary representation of n, namely $n = \sum_{j=1}^{m} b_j 2^{j-1}$, and $R_j(x) \equiv \text{sgn}[\sin(2^j \pi x)]$ is the *Rademacher function*, which switches between ± 1 with frequency 2^{j-1}.

A *WDD$_n$ sequence* is then defined as the pulse sequence with switching function

$$x(t) = -W_n(t/T),\ t \in [0,T] . \tag{6.3}$$

If $r \equiv \sum_m b_m$ is the *Hamming weight* of n (hence the number of Rademacher functions used to construct $W_n(x)$), the corresponding WDD$_n$ protocol achieves CO $= r$ [152] .

For a single qubit exposed to multi-axis decoherence, Ref. [152] also defines two-axis WDD protocols by allowing for the control propagator $U_c(t)$ to involve *two* switching functions, say, for π-pulses along the x and y directions, with the form $x(t) = R_{j_1} R_{j_3} \ldots R_{j_{2r-1}}$, $y(t) = R_{j_2} R_{j_4} \ldots R_{j_{2r}}$. In this way, for $n = 4^r - 1$, the resulting WDD$_n$ protocol reproduces concatenated DD (CDD) of level r, again achieving CO $= r$ for this general error model [171].

A different approach to digital DD design is provided by CPDD [31], whereby pulse sequences are built by concatenating projection sequences. There are four such sequences, $\mathtt{p}_0 \equiv I\mathtt{f}_{\tau_0} I\mathtt{f}_{\tau_0}$, $\mathtt{p}_x \equiv X\mathtt{f}_{\tau_0} X\mathtt{f}_{\tau_0}$, and similarly for $\mathtt{p}_y$ and $\mathtt{p}_z$. Applying $\mathtt{p}_u$, with $u \in \{x,y,z\}$,

suppresses the interaction along perpendicular directions, to the first order, that is, with corresponding EPG $= \mathcal{O}(\tau_0^2||H||^2)$ [172]. Given two pulse sequences A and B, their concatenation may be defined as $A[B] \equiv P_{N_A}^A(B)\ldots P_2^A(B)P_1^A$. The new pulse sequence constructed in this way inherits the suppression capabilities from each of the original pulse sequences. Concatenating a pulse sequence with $\mathtt{p}_0$ corresponds to simply repeating the sequence twice.

A *CPDD$_s$ sequence* is then specified by an ordered string $s \equiv s_m s_{m-1}\ldots s_1$, with $s_j \in \{0, x, y, z\}$, with each symbol labeling a projection sequence. To construct the corresponding pulse sequence, projection sequences are concatenated according to the specified string, namely,

$$\text{CPDD}_s \equiv \mathtt{p}_{s_1}[\ldots[\mathtt{p}_{s_{m-1}}[\mathtt{p}_{s_m}]]] \,. \tag{6.4}$$

For example, in this notation $\text{CDD}_r = \text{CPDD}_{(xy)^r}$.

General WDD and its equivalence to CPDD

Our first result is a generalization of multi-axis WDD beyond the existing one. Unlike the construction in [152], we start by expressing the control propagator in terms of *three* distinct switching functions:

$$U_c(t) = \sigma_x^{[x(t)+1]/2}\sigma_y^{[y(t)+1]/2}\sigma_z^{[z(t)+1]/2}. \tag{6.5}$$

We define general WDD (GWDD) sequences as follows:

Definition 6.1 *A GWDD$_{\vec{n}}$ sequence is specified by an integer vector consisting of three Paley orders, $\vec{n} \equiv (n_x, n_y, n_z)$, subject to the constraint $\sum_{u=x,y,z} b_j^u \leq 1$, $1 \leq j \leq m_u$. Here, b_j^u is the jth digit in the binary representation of n_u, where $n_u = \sum_{j=1}^{m_u} b_j^u 2^{j-1}$. The switching function for*

control along direction u in Eq. (6.5) is

$$u(t) = -WDD_{n_u}(t/T) = \prod_{j=1}^{m} R_j^u(t/T)^{b_j^u},\ t \in [0,T], \tag{6.6}$$

with $m \equiv \max\{m_x, m_y, m_z\}$ and $b_j^u \equiv 0$ for $m_u < j \leq m$.

Since any π-pulse can be obtained as the product of two π-pulses along orthogonal directions, the constraint on the coefficients b_j^u is necessary to avoid redundant sequences, by allowing at most one non-zero digit among all three digits at each binary location. Clearly, the above definition recovers the one in Refs. [152, 154], where a single integer suffices to specify a two-axis WDD$_n$, due to the assumed particular structure. For instance, r^{th}-order CDD corresponds to a GWDD$_{\vec{n}}$ with

$$\vec{n} = \left(2\frac{4^r-1}{4-1}, \frac{4^r-1}{4-1}, 0\right) .$$

Here the single above-mentioned Paley order $n = 4^r - 1$ being the sum of three Paley orders in our definition.

Crucially, the above GWDD definition is instrumental to both establish equivalence with the CPDD formalism, and uncover optimal GWDD sequences not accounted for otherwise. To demonstrate the equivalence, note that each non-zero digit b_j^u, in the binary representation of n_u in a GWDD sequence, may be associated to a projection p_u in the equivalent CPDD sequence. When $b_j^u = 0$ for all $u \in \{x, y, z\}$, we have an identity projection p_0 in CPDD. Explicitly, the following conversion rules hold:

(i) *CPDD-to-GWDD.* Given a CPDD$_s$ with $s = s_m s_{m-1} \ldots s_1$, calculate $n_u = \sum_{j=1}^{m} b_j^u 2^{j-1}$ for $u \in \{x, y, z\}$, where $b_j^u = 1$ if $s_j = u$, otherwise $b_j^\mu = 0$. The corresponding GWDD sequence is GWDD$_{n_x,n_y,n_z}$.

(ii) *GWDD-to-CPDD.* Given a GWDD$_{n_x,n_y,n_z}$, first convert each Paley order to its binary

Table 6.1: Equivalence between single-axis WDD and CPDD.

WDD_n	CPDD_s
WDD_0	$\mathrm{CPDD}_0 = \mathrm{p}_0$
WDD_1	$\mathrm{CPDD}_x = \mathrm{p}_x$
$\mathrm{WDD}_2 = \mathrm{WDD}_{10}$	$\mathrm{CPDD}_{x0} = \mathrm{p}_0[\mathrm{p}_x]$
$\mathrm{WDD}_3 = \mathrm{WDD}_{11}$	$\mathrm{CPDD}_{xx} = \mathrm{p}_x[\mathrm{p}_x]$
$\mathrm{WDD}_4 = \mathrm{WDD}_{100}$	$\mathrm{CPDD}_{x00} = \mathrm{p}_0[\mathrm{p}_0[\mathrm{p}_x]]$

representation, $n_u = (b^\mu_{m_u} b^\mu_{m_u-1} \dots b^u_1)_2$. Second, leftpad the binary representations with zeros so that they all have the same length m. For the jth digit and $u \in \{x, y, z\}$, set $s_j = u$ if $b^u_j = 1$; else, if all $b^u_j = 0$, set $s_j = 0$. The corresponding CPDD sequence is CPDD_s with $s = s_m s_{m-1} \dots s_1$.

The resulting correspondence is illustrated in Table 6.1 for single-axis sequences. For multi-axis control, we use the so-called GA_{8r} sequences as an example. The latter is obtained from concatenation of a basic six-pulse, 2nd-order GA_8 sequence, $I\mathtt{f}X\mathtt{f}Y\mathtt{f}X\mathtt{f}I\mathtt{f}X\mathtt{f}Y\mathtt{f}X\mathtt{f}$, found by a genetic search algorithm in Ref. [163]. In the CPDD framework, $\mathrm{GA}_{8r} = \mathrm{CPDD}_{(zyx)^r}$. By using the above rules, we have $n_x = (100\dots100)_2$, $n_y = (010\dots010)_2$, $n_z = (001\dots001)_2$. Accordingly, the corresponding GWDD sequence is $\mathrm{GWDD}_{\vec{n}}$, where $\vec{n}$ is given by

$$\vec{n} = \left(4\frac{1-2^{3r}}{1-2^3}, 2\frac{1-2^{3r}}{1-2^3}, \frac{1-2^{3r}}{1-2^3}\right) . \tag{6.7}$$

6.2 Performance analysis of GWDD

We first give the formula of cancellation for arbitrary GWDD sequence, which is trivial due to the equivalence between GWDD and CPDD. However, the CO itself cannot fully characterize the decoupling error, therefore we show how to directly calculate a upper bound on the decoupling error. Finally we observe that different orders of concatenation will accumulate different prefactors in the decoupling error. We give a intuitive way to

understand this control path sensitivity by considering several explicit examples.

Cancellation order

Cancellation order is a well-defined quantity to measure the decoupling ability of a DD sequences in the weak-interaction limit. The equivalence with the CPDD formalism makes it possible to easily obtain the CO of an arbitrary GWDD sequence. As shown in Ref. [31] (also see Chapter 5), the CO of CPDD is given by

$$\alpha = \min\{r_y + r_z, r_x + r_z, r_x + r_y\} \,, \tag{6.8}$$

where r_u is the number of projections along the u-axis. From the above rules, we see that each such projection implies a non-zero bit in the binary representation of the corresponding Paley order. Therefore, the CO of GWDD is still given by the above equation, but with $\{r_u = \sum_j b_j^u\}$ now being Hamming weights. It follows that GWDD/CPDD sequences with the same CO are highly non-unique: permuting the order of projections will produce a different GWDD sequence, but leave the CO unchanged. Accordingly, we may think of GWDD sequences specified by $(r_{\mathcal{P}(x)}, r_{\mathcal{P}(y)}, r_{\mathcal{P}(z)})$, where $\mathcal{P} \in \mathcal{S}_3$ is any permutation, as forming an equivalence class with respect to CO.

Upper bound on error per gate

The EPG provides an appropriate performance measure for control since it upper-bounds the trace-norm distance between the intended and the actual final states of the system, say,

$$\Delta(\rho_S^0(T), \rho_S(T)) \equiv \|\rho_S^0(T) - \rho_S(T)\|_1 \,, \tag{6.9}$$

where $\rho_S^0(T) = \rho_S^0(0) \equiv |\psi\rangle\langle\psi|$ for DD [173, 157]. That is, $\Delta(\rho_S^0(T), \rho_S(T)) \leq ||TH_{SB}^{\text{eff}}(T)||$ independently of $|\psi\rangle$, which in turn allows us to bound experimentally accessible fideli-

ties as $1-\Delta \leq F \leq \sqrt{1-\Delta^2}$ (here, $F(\rho, |\psi\rangle\langle\psi|) \equiv \mathrm{tr}\sqrt{\sqrt{\rho}\,|\psi\rangle\langle\psi|\sqrt{\rho}}$).

Thanks to the equivalence between GWDD and CPDD, we only need to calculate the upper bound for the corresponding CPDD sequence. The geometrical picture of projections makes CPDD the natural framework to use. Specifically, we first show how the norm of the relevant interaction Hamiltonian is renormalized by a single projection sequence. Since every CPDD sequence arises from concatenation of a series of projections, we can then apply the result of a single projection recursively, to establish the desired upper bound.

1. Bath renormalization by a single projection

   Consider first the effect of a single projection sequence, say $\mathtt{p_x}$. The resulting toggling-frame error Hamiltonian is

$$\tilde{H}_e(t) = \begin{cases} H, & 0 \leq t \leq \tau_0, \\ XHX, & \tau_0 \leq t \leq 2\tau_0. \end{cases} \tag{6.10}$$

   Since $\tilde{H}_e(t)$ is a piece-wise constant function, the first three orders of the Magnus series expansion may be easily computed as

$$\Omega_e^{(1)} = \tau_0(H + XHX), \tag{6.11}$$

$$\Omega_e^{(2)} = -\frac{i}{2}\tau_0^2[H, XHX], \tag{6.12}$$

$$\Omega_e^{(3)} = \frac{1}{3!}\tau_0^3[XHX[H, XHX]]. \tag{6.13}$$

   By using the explicit form of $H = \sum_{\mu=0,x,y,z} \sigma_\mu \otimes B_\mu$, together with Eq. (6.10) above,

the first two contributions become

$$\Omega_e^{(1)} = 2\tau_0(\mathbb{1} \otimes B_0 + B_x \otimes X), \tag{6.14}$$

$$\Omega_e^{(2)} = Y \otimes \tau_0^2(i[B_0, B_y] + \{B_z, B_x\}) + Z \otimes \tau_0^2(i[B_0, B_z] + \{B_y, B_x\}), \tag{6.15}$$

with a corresponding norm

$$||\Omega_e^{(1)}|| = O(\tau_0(\beta + J)), \qquad ||\Omega_e^{(2)}|| = O(\tau_0^2 J(\beta + J)).$$

Although the Magnus expansion converges as long as $||H||T < \pi$, care is needed in discarding higher-order terms. The norm of the third-order term is found to be

$$||\Omega_e^{(3)}|| = O(\tau_0^3 \beta J(\beta + J)) + O(\tau_0^3 J^3). \tag{6.16}$$

Accordingly, it is not possible in general to ignore this contribution as it is not clear which term in Eq. (6.16) dominates. Following the analysis in [171], we proceed by addressing separately two limiting regimes:

(a) When $J \ll \beta$, we have $||\Omega_e^{(1)}|| = O(\tau_0\beta)$, $||\Omega_2^{(2)}|| = O(\tau_0\beta J)$ and $||\Omega_e^{(i)}|| = O(\tau_0^i \beta^{i-1} J)$. Therefore, we have

$$||\Omega_e^{(1)}|| < ||\Omega_e^{(2)}|| \ll ||\Omega_e^{(i\geq 3)}||, \tag{6.17}$$

as long as the condition $\beta\tau_0 \ll 1$ is obeyed.

(b) When $J \gg \beta$, we have $||\Omega_e^{(i)}|| = O(\tau_0^i J^i)$. Thus, the same relation given in Eq. (6.17) holds, as long as $J\tau_0 \ll 1$.

In summary, when $J \gg \beta$ or $J \ll \beta$, provided that $\tau_0||H|| \ll 1$, it suffices to retain the first two orders of the Magnus expansion, giving an approximate expression for the error action operator as

$$\Omega_e(2\tau_0) \approx \Omega^{(1)}(2\tau_0) + \Omega^{(2)}(2\tau_0) \tag{6.18}$$

$$\equiv 2\tau_0 \bar{H}^x \tag{6.19}$$

$$= 2\tau_0 \sum_{\mu=0,x,y,z} \sigma_\mu \otimes B_\mu^x, \tag{6.20}$$

where in the last equality we have defined the average Hamilton associated with p_x and the relevant renormalized bath operators. From Eqs. (6.14) and (6.15), we can read them off as

$$\begin{aligned} B_0^x &= B_0, \\ B_x^x &= B_x, \\ B_y^x &= \frac{\tau_0}{2}(i[B_0, B_y] + \{B_z, B_x\}), \\ B_z^x &= \frac{\tau_0}{2}(i[B_0, B_z] + \{B_y, B_x\}). \end{aligned} \tag{6.21}$$

Similar equations hold for projections along the y or z directions. When the strength of the system-bath interaction and the pure bath dynamics are of the same order of magnitude, $J \sim \beta$, the calculation depends on the specific value of J and β, and no general analytic error bound may be established. From now on, we thus assume that the system is in either of the two regimes mentioned above.

2. Bath renormalization in an arbitrary CPDD sequence

Consider a CPDD sequence specified by an ordered string s_0, with total running

time T_{s_0}. Let the relevant effective Hamiltonian be denoted by $\bar{H}^{s_0}$. We now construct a new CPDD sequence by concatenating it with a projection sequence, say, $\mathtt{p}_x$, obtaining a CPDD$_{s_0x}$, whose renormalized effective Hamiltonian $\bar{H}^{s_0x}$ we wish to determine.

The evolution propagator of the system under the control of CPDD$_{s_0x}$ is

$$X(e^{-i\bar{H}^{s_0}T_{s_0}})X(e^{-i\bar{H}^{s_0}T_{s_0}}) = e^{-iXe^{-i\bar{H}^{s_0}T_{s_0}}X}e^{-ie^{-i\bar{H}^{s_0}T_{s_0}}}.$$

Therefore, the toggling-frame error Hamiltonian $\bar{H}^{s_0x}(t)$ is still a piece-wise constant function,

$$\bar{H}^{s_0x}(t) = \begin{cases} \bar{H}^{s_0} \ \ , 0 \le t \le T_{s_0}, \\ X\bar{H}^{s_0}X \ \ , T_{s_0} \le t \le 2T_{s_0}, \end{cases}$$

which makes it possible to use the same analysis used in the previous section. Accordingly, in the two regimes where $J \ll \beta$ or $J \gg \beta$, the renormalized bath operators are given by

$$\begin{aligned} B_0^{s_0x} &= B_0^{s_0}, \\ B_x^{s_0x} &= B_x^{s_0}, \\ B_y^{s_0x} &= \frac{T_{s_0}}{2}\left(i[B_0, B_y^{s_0}] + \{B_z^{s_0}, B_x^{s_0}\}\right), \\ B_z^{s_0x} &= \frac{T_{s_0}}{2}\left(i[B_0, B_z^{s_0}] + \{B_y^{s_0}, B_x^{s_0}\}\right), \end{aligned} \tag{6.22}$$

where $B_\mu^{s_0}, \mu \in \{0, x, y, z\}$ are the effective bath operators of CPDD$_{s_0}$. As we can see, $\mathtt{p}_x$ leaves $B_x^{s_0}$ unchanged, but renormalizes $B_y^{s_0}$ and $B_z^{s_0}$ to the next order. Similar renormalization relations hold for $\mathtt{p}_y$ and $\mathtt{p}_z$.

Applying standard operator-norm inequalities to the renormalized bath operators in Eq. (6.22), in particular, $|||[A,B]|| \leq 2||A||\,||B||$, $|||A+B|| \leq ||A|| + ||B||$, and $||AB|| \leq ||A||\,||B||$, we have

$$||B_0^{s_0 x}|| = ||B_0^{s_0}||, \quad ||B_x^{s_0 x}|| = ||B_x^{s_0}||, \tag{6.23a}$$

$$||B_y^{s_0 x}|| \leq T_{s_0}(\beta||B_y^{s_0}|| + ||B_z^{s_0}||\,||B_x^{s_0}||), \tag{6.23b}$$

$$||B_z^{s_0 x}|| \leq T_{s_0}(\beta||B_z^{s_0}|| + ||B_y^{s_0}||\,||B_x^{s_0}||), \tag{6.23c}$$

This is the main result of this section. Along with the definition of the EPG, this yields the desired result for CPDD$_s$,

$$\text{EPG} \leq T_s \sum_{u=x,y,z} ||B_u^s||.$$

Control path sensitivity

As remarked in previous section, any permutation of the order of concatenation in building CPDD sequences will leave the CO invariant. We expect that pulse sequence with a different control path will give different performance, since the EPG (or fidelity) do not solely depends on the CO. In the context of GWDD, control path sensitivity may be understood by comparing the upper bounds of the EPG generated by different control paths. As we can see from Eq. (6.23), the EPG of CPDD sequences generated by permutations of a sequence s, have the same scaling behavior on τ_0, but produce different prefactors. In this section, we first present a concrete example to demonstrate how the information about the control path is "encoded" into the prefactors of the relevant EPG. We then provide a more convenient way to calculate the prefactor for any GWDD/CPDD sequences.

- Switches in the control path are good for error suppression

  The simplest non-trivial example we may consider is to compare $\text{CDD}_2 = \text{CPDD}_{xyxy}$ with the CPDD sequence generated by a permutation of $xyxy$, denoted by $\overline{\text{CDD}}_2 = \text{CPDD}_{xxyy}$. To simplify the calculation and to focus on prefactor, we assume the regime $\beta \gg J$. Applying the renormalization given in Eq. (6.23) repeatedly, we have for CDD_2

$$||B_x^{xyxy}|| \leq 2^3 \cdot 2^1 \tau_0^2 \beta^2 J, \tag{6.24}$$

$$||B_y^{xyxy}|| \leq 2^2 \cdot 2^0 \tau_0^2 \beta^2 J, \tag{6.25}$$

  where at each step we only keep the leading-order terms. The bound for $||B_z||$ is always higher order than the other two directions since both $\mathtt{p}_x$ and $\mathtt{p}_y$ suppress B_z. In the above equations we also see explicitly how the prefactors are accumulated. Similarly, for $\overline{\text{CDD}}_2$, we apply Eq. (6.23) repeatedly but with a different order, obtaining

$$||B_x^{xxyy}|| \leq 2^3 \cdot 2^2 \tau_0^2 \beta^2 J, \tag{6.26}$$

$$||B_y^{xxyy}|| \leq 2^1 \cdot 2^0 \tau_0^2 \beta^2 J. \tag{6.27}$$

  As we can see, the upper bounds of CDD_2 and $\overline{\text{CDD}}_2$ have the same scaling over τ_0, consistent with the fact that both of them achieve $\text{CO} = 2$. However, CDD_2 has a smaller prefactor than $\overline{\text{CDD}}_2$:

$$\text{EPG}^{xyxy} \leq 20\, \tau_0^2 \beta^2 JT \tag{6.28}$$

$$\text{EPG}^{xxyy} \leq 34\, \tau_0^2 \beta^2 JT. \tag{6.29}$$

  This can be qualitatively explained as follows. When $\mathtt{p}_x$ is applied, the upper bound

for $||B_y||$ will start to accumulate a prefactor. If we continue applying $\mathtt{p}_x$, like in $\overline{\text{CDD}}$, the prefactor for $||B_y||$ will grow exponentially since the length of the sequence is exponentially increasing. However, if the direction of the projection sequence is changed at a certain point, say to $\mathtt{p}_y$, then the prefactor for $||B_y||$ will stop increasing. Therefore, CPDD sequences, with a large number of switches in the direction of the corresponding projections, tend to have lower error and better performance. For sufficiently large CO we may write

$$\text{EPG}^{\text{CDD}_\alpha} \leq 2^{\alpha^2}(\tau_0\beta)^\alpha JT, \tag{6.30}$$

$$\text{EPG}^{\overline{\text{CDD}}_\alpha} \leq 2^{\frac{1}{2}(3\alpha^2-\alpha)}(\tau_0\beta)^\alpha JT. \tag{6.31}$$

The above conclusions remain unchanged if we work in the opposite regime, $\beta \ll J$, since the prefactor only depends on the order of concatenations.

- Calculating prefactors for GWDD/CPDD sequences

  The method we described above to calculate the prefactors for GWDD sequences relies upon the geometric picture of CPDD. However, the calculation is tedious, especially for long pulse sequences. Here we present an alternative method to directly calculate the prefactor for any GWDD/CPDD sequence.

  Consider a pulse sequence CPDD$_s$. Then:

  1. Define s' to be the sequence of letters in the reverse order of s, namely, $s' \equiv s_1 \dots s_m$. Construct a $3 \times |s|$ matrix, denoted by $\mathcal{L}$, according to the following rule:

$$\mathcal{L}_{\mu j} \equiv \begin{cases} 1, \text{ if } s_j = \mu \\ 0, \text{ otherwise} \end{cases} . \tag{6.32}$$

where we use $\mu \in \{x, y, z\}$ to label the 1^{st}, the 2^{nd} and the 3^{rd} row of $\mathcal{L}$.

2. The prefactor in the upper bound on $||B_\mu^s||$ is then given by

$$\prod_{\substack{j=1 \\ \bar{\mathcal{L}}_{\mu j} \neq 0}}^{|s|} \bar{\mathcal{L}}_{\mu j} 2^{j-1} \,, \tag{6.33}$$

where the matrix $\bar{\mathcal{L}}$ is the logical negation of $\mathcal{L}$.

3. If we assume $\beta \gg J$ and ignore higher-order contributions, we have the following upper bound

$$||B_\mu^s|| \leq \left(\prod_{\substack{j=1 \\ \bar{\mathcal{L}}_{\mu j} \neq 0}}^{|s|} \bar{\mathcal{L}}_{\mu j} 2^{j-1} \right) (\tau_0 \beta)^{\sum_{k=1}^{|s|} \bar{\mathcal{L}}_{\mu k}} J \,. \tag{6.34}$$

We illustrate the above procedure by considering a simple example, namely, the second level of OWDD sequences, CPDD_{xyz}. From the definition of CPDD sequence, $s' = xyz$, hence the matrix $\mathcal{L}$ is given by

$$\begin{array}{c} s' \\ x \\ y \\ z \end{array} \begin{array}{c} \begin{array}{ccc} x & y & z \end{array} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{array} .$$

Here, the row indexes represent different directions while the column indexes are

specified by s'. Applying Eq. (6.33) and Eq. (6.34), and assuming $\beta \gg J$, we get

$$||B_x^{xyz}|| \leq 2^2 \cdot 2^1 (\tau_0 \beta)^2 J,$$
$$||B_y^{xyz}|| \leq 2^2 \cdot 2^0 (\tau_0 \beta)^2 J,$$
$$||B_z^{xyz}|| \leq 2^1 \cdot 2^0 (\tau_0 \beta)^2 J.$$

6.3 Optimal Walsh Dynamical decoupling

Since the resources in the real control experiment are always limited, pulse sequences which achieve high performance with few applied pulses are practically important. In this section we identity optimal sequences within OWDD framework. We also analyze its performance and show its superiority over well-know digital DD scheme.

Definition of OWDD

In the presence of a realistic constraint, $\tau_0 > 0$, achieving higher CO comes at the price of either increasing the total number of pulses N for fixed storage time T — until the maximum CO compatible with the constraints is accommodated; or of increasing both N and T — until perturbative error suppression breaks down and, again, a maximum CO is reached beyond which no further improvement occur [155]. This motivates defining *optimal WDD* (OWDD) sequences by demanding that they guarantee a desired CO with minimum pulse number or, equivalently, minimum number of time slots, N_T, each slot having duration τ_0. Within single-axis WDD, CDD sequences are provably optimal [152]. However, this is no longer true for multi-axis GWDD. The optimal GWDD can be inferred from the CPDD framework.

Definition 6.2 *For CO $= \alpha$, let $\overline{\alpha} = \pm 1$ denote the parity of α. Then, all GWDD$_{\vec{n}}$ satisfying the*

Table 6.2: Number of control time slots and applied pulses for OWDD vs. CDD

	OWDD			CDD		
CO	(n_x, n_y, n_z)	N_T	N	(n_x, n_y, n_z)	N_T	N
1	(2,1,0)	4	4	(2,1,0)	4	4
2	(4,2,1)	8	6	(10,5,0)	16	14
3	(18,5,4)	32	32	(42,21,0)	64	60
4	(36,18,9)	64	42	(170,85,0)	256	238

following two conditions are optimal and define an equivalence class referred to as OWDD$_\alpha$:

$$\sum_{u=x,y,z} b_j^u = 1, \quad 1 \leq j \leq m, \tag{6.35}$$

$$(r_{\mathcal{P}(x)}, r_{\mathcal{P}(y)}, r_{\mathcal{P}(z)}) = \frac{1}{2}\left(\alpha - \overline{\alpha}, \alpha + \overline{\alpha}, \alpha + \overline{\alpha}\right). \tag{6.36}$$

Eq. (6.35) ensures that the pulse sequence does not expend any pulse on repetition, whilst Eq. (6.36) gives the Hamming weights (number of projections) needed to suppress decoherence up to the required CO. From the equivalence between CPDD and GWDD, and the analysis in Ref. [31], it follows that OWDD$_\alpha$ uses a number of time slots given by $\log_2(N_T) = \frac{1}{2}\left(3\alpha + \overline{\alpha}\right)$. A comparison between OWDD$_\alpha$ and CDD$_\alpha$ is included in Table **??**. If the CO is sufficiently large, OWDD$_\alpha$ is *exponentially more efficient* than CDD, since

$$N_T^{\text{OWDD}_\alpha} / N_T^{\text{CDD}_\alpha} \approx 2^{\frac{3}{2}\alpha} / 2^{2\alpha} = 2^{-\alpha/2}. \tag{6.37}$$

Performance analysis

Since for each α there is an equivalence class of OWDD$_\alpha$ sequences, and the fact that different control paths result different error suppression ability, we need to choose specific OWDD sequences before trying to evaluate their performance.

Inspired by the analysis in Sec. 6.2.3, we consider OWDD sequences with the maximum number of switches, OWDD$_\alpha^h \equiv \{$CPDD$_{xy}$, CPDD$_{xyz}$, CPDD$_{xyzxy}$, CPDD$_{(xyz)^2}$, ...$\}$,

as well as sequences where this number is minimized and the lattice-path trajectory has long straight segments: OWDD$^l_\alpha \equiv$ {CPDD$_{xy}$,CPDD$_{xyz}$, CPDD$_{xxyyz}$, CPDD$_{x^2y^2z^2}$, ...} for $\alpha = 1, 2, 3, 4$. The first two orders of OWDD sequences are the same for any choice of OWDD. Thus, in order to illustrate the control path sensitivity, below we explicitly calculate the upper bound of EPG for the next two levels of OWDD^h and OWDD^l, corresponding to CO $= 3, 4$, respectively.

To calculate the upper bound for OWDD$^h_3$ = CPDD$_{xyzxy}$, we first write down the $\mathcal{L}$ matrix according to Eq.(6.32),

$$\mathcal{L} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} . \tag{6.38}$$

Calculate its negation $\bar{\mathcal{L}}$ and then the upper bound according to Eq.(6.34):

$$\begin{aligned} ||B_x^{xyzxy}|| &\leq 2^4 \cdot 2^2 \cdot 2^1 (\tau_0\beta)^3 J, \\ ||B_y^{xyzxy}|| &\leq 2^3 \cdot 2^2 \cdot 2^0 (\tau_0\beta)^3 J, \\ ||B_z^{xyzxy}|| &\leq 2^4 \cdot 2^3 \cdot 2^1 \cdot 2^0 (\tau_0\beta)^4 J. \end{aligned}$$

By discarding the higher-order contribution from $||B_z||$, we have EPG$^{\text{OWDD}^h_3} \leq 5{\cdot}2^5(\tau_0\beta)^3 JT$, where $T = 2^5\tau_0$. Following the same procedure for OWDD$^h_4$, we have

$$\begin{aligned} ||B_x^{xyzxyz}|| &\leq 2^5 \cdot 2^4 \cdot 2^2 \cdot 2^1 (\tau_0\beta)^4 J, \\ ||B_y^{xyzxyz}|| &\leq 2^5 \cdot 2^3 \cdot 2^2 \cdot 2^0 (\tau_0\beta)^4 J, \\ ||B_z^{xyzxyz}|| &\leq 2^4 \cdot 2^3 \cdot 2^1 \cdot 2^0 (\tau_0\beta)^4 J, \end{aligned}$$

the EPG is dominated by the x direction and we have EPG$^{\text{OWDD}^h_4} \leq 2^{12} (\tau_0\beta)^4 JT$, with

$T = 2^6\tau_0$.

Now we calculate the upper bounds for OWDD^l_α. To calculate the upper bound of OWDD^l_3, we write down the $\mathcal{L}$ matrix first, namely,

$$\mathcal{L} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} . \tag{6.39}$$

Then the upper bounds of bath operators are given by Eq. (6.34),

$$\begin{aligned} ||B_x^{xxyyz}|| &\leq 2^4 \cdot 2^3 \cdot 2^2 (\tau_0\beta)^3 J, \\ ||B_y^{xxyyz}|| &\leq 2^4 \cdot 2^1 \cdot 2^0 (\tau_0\beta)^3 J, \\ ||B_z^{xxyyz}|| &\leq 2^3 \cdot 2^2 \cdot 2^1 \cdot 2^0 (\tau_0\beta)^4 J. \end{aligned}$$

Therefore, $\text{EPG}^{\text{OWDD}^l_3} \leq 2^4 \cdot 2^5 (\tau_0\beta)^3 JT$, with $T = 2^5\tau_0$. Similarly, the upper bounds for OWDD^l_4 are given by

$$\begin{aligned} ||B_x^{xxyyzz}|| &\leq 2^5 \cdot 2^4 \cdot 2^3 \cdot 2^2 (\tau_0\beta)^4 J, \\ ||B_y^{xxyyzz}|| &\leq 2^5 \cdot 2^4 \cdot 2^1 \cdot 2^0 (\tau_0\beta)^4 J, \\ ||B_z^{xxyyzz}|| &\leq 2^3 \cdot 2^2 \cdot 2^1 \cdot 2^0 (\tau_0\beta)^4 J, \end{aligned}$$

whereby $\text{EPG}^{\text{OWDD}^l_4} \leq 2^{14} (\tau_0\beta)^4 JT$, with $T = 2^6\tau_0$.

As one can see from the above calculations, at $\text{CO} = 3$ the EPG upper bound of OWDD^h is about 3.2 times smaller than the one for OWDD^l, and becomes four times smaller at $\text{CO} = 4$. Therefore, an increasingly larger benefit is expected also in terms of fidelity from using OWDD^h with larger CO. Geometrically, if one visualizes the implemented sequence of projections in terms of a lattice path starting at the origin in

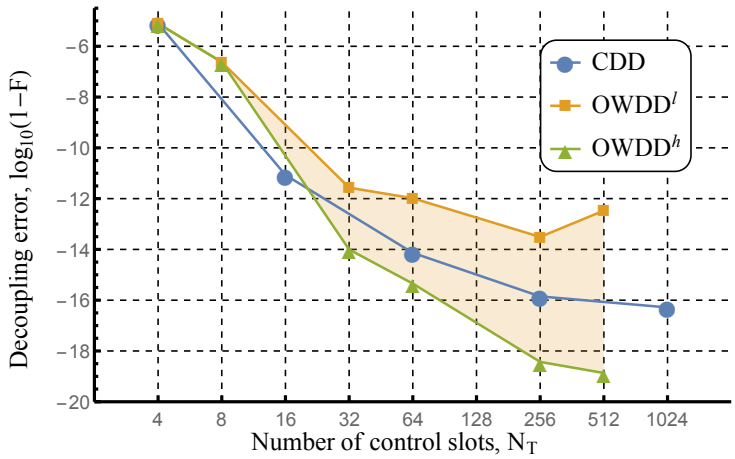

Figure 6.1: Fidelity loss vs. N_T for two choices of OWDD$_\alpha$ protocols with same CO and CDD. The shaded area marks the performance spread expected for all OWDD$_\alpha$ protocols in the same equivalence class. A toy model consisting of three bath spins is used, with an initial joint state of the form $|\Psi\rangle_{SB} \equiv |\psi\rangle \otimes |z_1 z_2 z_3\rangle$, where $|\psi\rangle$ is a random qubit state and each bath spin is randomly chosen over $z_i \in \{0, 1\}$. Results are averaged over 500 realizations. We choose parameters β = 10kHz, J = 1MHz, and $\tau_0 = 0.1\mu$s, suitable for qualitatively describing GaAs quantum dots [175].

$\mathbb{N}^3$, OWDD^h maximizes the number of switches in direction as compared to OWDD^l. That avoiding control path repetitions is generally useful in slowing down coherent error build-up, has been emphasized in the context of randomized DD design [174], and we conjecture that a similar intuition may be key for further optimizing OWDD against path variations.

We conclude by comparing in Fig. 6.1 the performance of OWDD and CDD directly in terms of average fidelity loss, by resorting to an exact numerical simulation of a low-dimensional spin-bath model, which mimics the basic features of hyperfine-induced decoherence of an electron spin qubit in a quantum dot [176, 175]. The bath operators are

$$B_\mu = \sum_{i \neq j} \sum_{\alpha,\beta} c^\mu_{\alpha\beta} (\sigma^\alpha_i \otimes \sigma^\beta_j) \,, \tag{6.40}$$

where i, j index the bath qubits, $\mu, \alpha, \beta \in \{0, x, y, z\}$, and $c^\mu_{\alpha\beta}$ are uniformly random coupling constants in $[0, 1]$. We assume a fixed minimum pulse interval $\tau_0 = 0.1\mu$s. At large

N_T, the performance tends to plateau (or even deteriorate) due to the fact that convergence breaks down for long T. In addition, the fact that OWDD_2 underperforms CDD_2 in this simulation may be accounted for by the effect of the prefactor (see Sec. 6.2.3). Remarkably, if OWDD_α^h is used, comparable performance to CDD_α is found for smaller N_T, whereas for same N_T, the fidelity of OWDD can be higher than the one of CDD by *up to two orders of magnitude*.

References

[1] George Gamow. *Thirty years that shook physics: The story of quantum theory*. Courier Corporation, 1985.

[2] Michael A Nielsen and Isaac Chuang. *Quantum computation and quantum information*. AAPT, 2002.

[3] Mark M. Wilde. From classical to quantum Shannon theory. *arXiv preprint arXiv:1106.1445*, 2011.

[4] Shuang Cong. *Control of quantum systems: theory and methods*. John Wiley & Sons, 2014.

[5] Peter W Shor. Algorithms for quantum computation: Discrete logarithms and factoring. In *Foundations of Computer Science, 1994 Proceedings., 35th Annual Symposium on*, pages 124–134. IEEE, 1994.

[6] Nicolas Gisin, Grégoire Ribordy, Wolfgang Tittel, and Hugo Zbinden. Quantum cryptography. *Review of Modern Physics*, 74:145–195, 2002.

[7] Charles H Bennett, Gilles Brassard, Sandu Popescu, Benjamin Schumacher, John A Smolin, and William K Wootters. Purification of noisy entanglement and faithful teleportation via noisy channels. *Physical Review Letters*, 76(5):722, 1996.

[8] Bei Zeng, Xie Chen, Duan-Lu Zhou, and Xiao-Gang Wen. Quantum information meets quantum matter. *arXiv preprint arXiv:1508.02595*, 2015.

[9] Jens Eisert. Entanglement and tensor network states. *arXiv preprint arXiv:1308.3318*, 2013.

[10] Armin Uhlmann. The "transition probability" in the state space of a *-algebra. *Reports on Mathematical Physics*, 9(2):273–279, 1976.

[11] Christopher A. Fuchs and Jeroen van de Graaf. Cryptographic distinguishability measures for quantum mechanical states. *IEEE Transactions on Information Theory*, 45(4):1216–1227, 1998. arXiv:quant-ph/9712042.

[12] Robert Alicki and Mark Fannes. Continuity of quantum conditional information. *Journal of Physics A: Mathematical and General*, 37(5):L55–L57, 2004. arXiv:quant-ph/0312081.

[13] Andreas Winter. Tight uniform continuity bounds for quantum entropies: conditional entropy, relative entropy distance and energy constraints. *Communications in Mathematical Physics*, 347(1):291–313, 2016. arXiv:1507.07775.

[14] Mark M. Wilde. From classical to quantum Shannon theory. 2016. arXiv:1106.1445v7.

[15] Alfred Wehrl. Three theorems about entropy and convergence of density matrices. *Reports on Mathematical Physics*, 10(2):159 – 163, 1976.

[16] Valentina Baccetti and Matt Visser. Infinite Shannon entropy. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(04):P04010, 2013.

[17] Göran Lindblad. Entropy, information and quantum measurements. *Communications in Mathematical Physics*, 33(4):305–322, 1973.

[18] Göran Lindblad. Completely positive maps and entropy inequalities. *Communications in Mathematical Physics*, 40(2):147–151, 1975.

[19] Alexander S. Holevo and Maksim E. Shirokov. Mutual and coherent information for infinite-dimensional quantum channels. *Problems of Information Transmission*, 46(3):201–218, 2010. arXiv:1004.2495.

[20] Anna A. Kuznetsova. Conditional entropy for infinite-dimensional quantum systems. *Theory of Probability & Its Applications*, 55(4):709–717, 2011. arXiv:1004.4519.

[21] Maksim E. Shirokov. Measures of quantum correlations in infinite-dimensional systems. *Sbornik: Mathematics*, 207(5):724, 2015. arXiv:1506.06377.

[22] Maksim E. Shirokov. Squashed entanglement in infinite dimensions. *Journal of Mathematical Physics*, 57(3):032203, 2016. arXiv:1507.08964.

[23] F Caruso, J Eisert, V Giovannetti, and AS Holevo. Multi-mode bosonic gaussian channels. *New Journal of Physics*, 10(8):083030, 2008.

[24] Christian Weedbrook, Stefano Pirandola, Raul Garcia-Patron, Nicolas J. Cerf, Timothy C. Ralph, Jeffrey H. Shapiro, and Seth Lloyd. Gaussian quantum information. *Reviews of Modern Physics*, 84(2):621–669, 2012. arXiv:1110.3234.

[25] John Williamson. On the algebraic problem concerning the normal forms of linear dynamical systems. *American Journal of Mathematics*, 58(1):141–163, 1936.

[26] Mark M. Wilde, Marco Tomamichel, Seth Lloyd, and Mario Berta. Gaussian hypothesis testing and quantum illumination. 2016. arXiv:1608.06991.

[27] Filippo Caruso, Jens Eisert, Vittorio Giovannetti, and Alexander S. Holevo. Multi-mode bosonic Gaussian channels. *New Journal of Physics*, 10:083030, 2008. arXiv:0804.0511.

[28] Mark M Wilde and Haoyu Qi. Energy-constrained private and quantum capacities of quantum channels. *arXiv preprint arXiv:1609.01997*, 2016.

[29] Haoyu Qi and Mark M. Wilde. Capacities of quantum amplifier channels. 2016. arXiv:1605.04922.

[30] Haoyu Qi, Mark M Wilde, and Saikat Guha. Thermal states minimize the output entropy of single-mode phase-insensitive gaussian channels with an input entropy constraint. *arXiv preprint arXiv:1607.05262*, 2016.

[31] Haoyu Qi and Dowling Jonathan P. Method for generating all uniform π-pulse sequences used in deterministic dynamical decoupling. *Physical Review A*, 92:032303, 2015.

[32] Haoyu Qi, Jonathan P Dowling, and Lorenza Viola. Optimal digital dynamical decoupling for general decoherence via walsh modulation. *arXiv preprint arXiv:1702.05533*, 2017.

[33] Alexander S. Holevo and Reinhard F. Werner. Evaluating capacities of bosonic Gaussian channels. *Physical Review A*, 63(3):032312, 2001. arXiv:quant-ph/9912067.

[34] Michael M. Wolf, David Pérez-García, and Geza Giedke. Quantum capacities of bosonic channels. *Physical Review Letters*, 98(13):130501, 2007. arXiv:quant-ph/0606132.

[35] AS Holevo. Classical capacities of quantum channels with constrained inputs. *Probability Theory and Applications*, 48(2):359–374, 2003.

[36] Charles H Bennett, Peter W Shor, John A Smolin, and Ashish V Thapliyal. Entanglement-assisted classical capacity of noisy quantum channels. *Physical Review Letters*, 83(15):3081, 1999.

[37] Alexander S Holevo and Maksim E Shirokov. On classical capacities of infinite-dimensional quantum channels. *Problems of Information Transmission*, 49(1):15–31, 2013.

[38] ME Shirokov. Adaptation of the alicki-fannes-winter method for the set of states with bounded energy and its use. *arXiv preprint arXiv:1609.07044*, 2016.

[39] Debbie Leung and Graeme Smith. Continuity of quantum channel capacities. *Communications in Mathematical Physics*, 292(1):201–215, 2009.

[40] Maksim E Shirokov. Continuity bounds for information characteristics of quantum channels depending on input dimension and on input energy. *arXiv preprint arXiv:1610.08870*, 2016.

[41] Kunal Sharma, Mark M Wilde, Sushovit Adhikari, and Masahiro Takeoka. Bounding the energy-constrained quantum and private capacities of bosonic thermal channels. *arXiv preprint arXiv:1708.07257*, 2017.

[42] Valerio Scarani, Helle Bechmann-Pasquinucci, Nicolas J. Cerf, Miloslav Dušek, Norbert Lütkenhaus, and Momtchil Peev. The security of practical quantum key distribution. *Reviews of Modern Physics*, 81(3):1301–1350, 2009. arXiv:0802.4155.

[43] Igor Devetak. The private classical capacity and quantum capacity of a quantum channel. *IEEE Transactions on Information Theory*, 51(1):44–55, January 2005. arXiv:quant-ph/0304127.

[44] Benjamin Schumacher and Michael D. Westmoreland. Quantum privacy and quantum coherence. *Physical Review Letters*, 80(25):5695–5697, 1998. arXiv:quant-ph/9709058.

[45] Karol Horodecki, Michał Horodecki, Paweł Horodecki, and Jonathan Oppenheim. Secure key from bound entanglement. *Physical Review Letters*, 94(16):160502, 2005. arXiv:quant-ph/0309110.

[46] Karol Horodecki, Michal Horodecki, Pawel Horodecki, and Jonathan Oppenheim. General paradigm for distilling classical key from quantum states. *IEEE Transactions on Information Theory*, 55(4):1898–1929, 2009. arXiv:quant-ph/0506189.

[47] Karol Horodecki, Michał Horodecki, Paweł Horodecki, Debbie Leung, and Jonathan Oppenheim. Unconditional privacy over channels which cannot convey quantum information. *Physical Review Letters*, 100(11):110502, 2008. arXiv:quant-ph/0702077.

[48] Karol Horodecki, Michał Horodecki, Paweł Horodecki, Debbie Leung, and Jonathan Oppenheim. Quantum key distribution based on private states: Unconditional security over untrusted channels with zero quantum capacity. *IEEE Transactions on Information Theory*, 54(6):2604–2620, 2008. arXiv:quant-ph/0608195.

[49] Mark M. Wilde, Patrick Hayden, and Saikat Guha. Quantum trade-off coding for bosonic communication. *Physical Review A*, 86(6):062306, December 2012. arXiv:1105.0119.

[50] Stefano Pirandola, Riccardo Laurenza, Carlo Ottaviani, and Leonardo Banchi. Fundamental limits of repeaterless quantum communications. 2015. arXiv:1510.08863v5.

[51] Marco Tomamichel, Mario Berta, and Joseph M. Renes. Quantum coding with finite resources. *Nature Communications*, 7:11419, 2016. arXiv:1504.04617.

[52] Vittorio Giovannetti, Seth Lloyd, Lorenzo Maccone, and Peter W. Shor. Broadband channel capacities. *Physical Review A*, 68(6):062323, 2003. arXiv:quant-ph/0307098.

[53] Saikat Guha, Jeffrey H. Shapiro, and Baris I. Erkmen. Capacity of the bosonic wiretap channel and the entropy photon-number inequality. In *Proceedings of the IEEE International Symposium on Information Theory*, pages 91–95, Toronto, Ontario, Canada, 2008. arXiv:0801.0841.

[54] Mark M. Wilde, Patrick Hayden, and Saikat Guha. Information trade-offs for optical quantum communication. *Physical Review Letters*, 108(14):140501, 2012. arXiv:1105.0119.

[55] Ligong Wang, Jeffrey H. Shapiro, Nivedita Chandrasekaran, and Gregory W. Wornell. Private-capacity bounds for bosonic wiretap channels. 2012. arXiv:1202.1126.

[56] Giacomo De Palma, Andrea Mari, Seth Lloyd, and Vittorio Giovannetti. Multimode quantum entropy power inequality. *Physical Review A*, 91(3):032320, 2015. arXiv:1408.6410.

[57] Koenraad Audenaert, Nilanjana Datta, and Maris Ozols. Entropy power inequalities for qudits. *Journal of Mathematical Physics*, 57(5):052202, 2016. arXiv:1503.04213.

[58] Alexander S. Holevo. *Quantum Systems, Channels, Information*. de Gruyter Studies in Mathematical Physics (Book 16). de Gruyter, 2012.

[59] Alexander S. Holevo and Maksim E. Shirokov. On the entanglement-assisted classical capacity of infinite-dimensional quantum channels. *Problems of Information Transmission*, 49(1):15–31, 2013. arXiv:1210.6926.

[60] Howard Barnum, Emanuel Knill, and Michael A. Nielsen. On quantum fidelities and channel capacities. *IEEE Transactions on Information Theory*, 46(4):1317–1329, 2000. arXiv:quant-ph/9809010.

[61] Dennis Kretschmann and Reinhard F Werner. Tema con variazioni: quantum channel capacity. *New Journal of Physics*, 6(1):26, 2004. arXiv:quant-ph/0311037.

[62] Rochus Klesse. Approximate quantum error correction, random codes, and quantum channel capacity. *Physical Review A*, 75(6):062315, 2007. arXiv:quant-ph/0701102.

[63] John Watrous. *Theory of Quantum Information*. 2016.

[64] Alexander S. Holevo. Entanglement-assisted capacity of constrained channels. *Proceedings of SPIE, First International Symposium on Quantum Informatics*, 5128:62–69, 2003. arXiv:quant-ph/0211170.

[65] Alexander S. Holevo. Entanglement-assisted capacities of constrained quantum channels. *Theory of Probability & Its Applications*, 48(2):243–255, 2004. arXiv:quant-ph/0211170.

[66] Alexander S. Holevo and Maksim E. Shirokov. Continuous ensembles and the capacity of infinite-dimensional quantum channels. *Theory of Probability & Its Applications*, 50(1):86–98, 2006. arXiv:quant-ph/0408176.

[67] Alexander S. Holevo. The entropy gain of infinite-dimensional quantum evolutions. *Doklady Mathematics*, 82(2):730–731, 2010. arXiv:1003.5765.

[68] Rochus Klesse. A random coding based proof for the quantum coding theorem. *Open Systems & Information Dynamics*, 15(1):21–45, 2008. arXiv:0712.2558.

[69] Terence Tao. *Topics in Random Matrix Theory*, volume 132 of *Graduate Studies in Mathematics*. American Mathematical Society, 2012. See also http://terrytao.wordpress.com/2008/06/18/the-strong-law-of-large-numbers/.

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

[71] Abbas El Gamal and Young-Han Kim. *Network Information Theory*. Cambridge University Press, 2012. arXiv:1001.3404.

[72] Charles H. Bennett, Peter W. Shor, John A. Smolin, and Ashish V. Thapliyal. Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem. *IEEE Transactions on Information Theory*, 48(10):2637–2655, 2002. arXiv:quant-ph/0106052.

[73] Igor Devetak and Peter W. Shor. The capacity of a quantum channel for simultaneous transmission of classical and quantum information. *Communications in Mathematical Physics*, 256(2):287–303, 2005. arXiv:quant-ph/0311131.

[74] Graeme Smith. Private classical capacity with a symmetric side channel and its application to quantum cryptography. *Physical Review A*, 78(2):022306, 2008. arXiv:0705.3838.

[75] Jon Yard, Patrick Hayden, and Igor Devetak. Capacity theorems for quantum multiple-access channels: Classical-quantum and quantum-quantum capacity regions. *IEEE Transactions on Information Theory*, 54(7):3091–3113, 2008. arXiv:quant-ph/0501045.

[76] Alexander S. Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problems of Information Transmission*, 9:177–183, 1973.

[77] Horace P. Yuen and Masanao Ozawa. Ultimate information carrying limit of quantum systems. *Physical Review Letters*, 70(4):363–366, 1993.

[78] Filippo Caruso and Vittorio Giovannetti. Degradability of bosonic Gaussian channels. *Physical Review A*, 74(6):062307, December 2006. arXiv:quant-ph/0603257.

[79] Saikat Guha. *Multiple-User Quantum Information Theory for Optical Communication Channels*. PhD thesis, Massachusetts Institute of Technology, 2008.

[80] Alexander S. Holevo. One-mode quantum Gaussian channels: Structure and quantum capacity. *Problems of Information Transmission*, 43(1):1–11, 2007. arXiv:quant-ph/0607051.

[81] Filippo Caruso, Vittorio Giovannetti, and Alexander S. Holevo. One-mode bosonic Gaussian channels: a full weak-degradability classification. *New Journal of Physics*, 8(12):310, 2006. arXiv:quant-ph/0609013.

[82] Alexander S. Holevo and Vittorio Giovannetti. Quantum channels and their entropic characteristics. *Reports on Progress in Physics*, 75(4):046001, 2012. arXiv:1202.6480.

[83] Vittorio Giovannetti, Saikat Guha, Seth Lloyd, Lorenzo Maccone, Jeffrey H. Shapiro, and Horace P. Yuen. Classical capacity of the lossy bosonic channel: The exact solution. *Physical Review Letters*, 92(2):027902, January 2004. arXiv:quant-ph/0308012.

[84] Saikat Guha. Classical capacity of the free-space quantum-optical channel. Master's thesis, Massachusetts Institute of Technology, January 2004.

[85] Koenraad M. R. Audenaert and Jens Eisert. Continuity bounds on the quantum relative entropy. *Journal of Mathematical Physics*, 46(10):102104, 2005. arXiv:quant-ph/0503218.

[86] Hermann A. Haus and J. A. Mullen. Quantum noise in linear amplifiers. *Physical Review*, 128(5):2407–2413, 1962.

[87] Carlton M. Caves. Quantum limits on noise in linear amplifiers. *Physical Review D*, 26(8):1817–1839, 1982.

[88] Alexander S. Holevo. *Quantum systems, channels, information: A mathematical introduction*, volume 16. de Gruyter, 2013.

[89] Vittorio Giovannetti, Alexander S. Holevo, and Raul García-Patrón. A solution of Gaussian optimizer conjecture for quantum channels. *Communications in Mathematical Physics*, 334(3):1553–1571, 2014. arXiv:1312.2251.

[90] Aashish A. Clerk, Michel H. Devoret, Steven M. Girvin, Florian Marquardt, and Robert J. Schoelkopf. Introduction to quantum noise, measurement, and amplification. *Reviews of Modern Physics*, 82(2):1155, 2010. arXiv:0810.4729.

[91] Gerald T. Moore. Quantum theory of the electromagnetic field in a variable-length one-dimensional cavity. *Journal of Mathematical Physics*, 11(9):2679–2691, 1970.

[92] William G. Unruh. Notes on black-hole evaporation. *Physical Review D*, 14(4):870, 1976.

[93] Stephen W. Hawking. Black holes in general relativity. *Communications in Mathematical Physics*, 25(2):152–166, 1972.

[94] P. D. Nation, J. R. Johansson, M. P. Blencowe, and Franco Nori. Colloquium: Stimulating uncertainty: Amplifying the quantum vacuum with superconducting circuits. *Reviews of Modern Physics*, 84(1):1, 2012. arXiv:1103.0835.

[95] Kamil Brádler, Patrick Hayden, and Prakash Panangaden. Quantum communication in Rindler spacetime. *Communications in Mathematical Physics*, 312(2):361–398, 2012. arXiv:1007.0997.

[96] Kamil Brádler and Christoph Adami. Black holes as bosonic Gaussian channels. *Physical Review D*, 92(2):025030, 2015. arXiv:1405.1097.

[97] Min-Hsiu Hsieh and Mark M. Wilde. Entanglement-assisted communication of classical and quantum information. *IEEE Transactions on Information Theory*, 56(9):4682–4704, 2010. arXiv:0811.4227.

[98] Mark M. Wilde, Patrick Hayden, and Saikat Guha. Information trade-offs for optical quantum communication. *Physical Review Letters*, 108(14):140501, 2012. arXiv:1206.4886.

[99] Mark M. Wilde, Patrick Hayden, and Saikat Guha. Quantum trade-off coding for bosonic communication. *Physical Review A*, 86(6):062306, 2012. arXiv:1105.0119.

[100] Mark M. Wilde and Min-Hsiu Hsieh. Public and private resource trade-offs for a quantum channel. *Quantum Information Processing*, 11(6):1465–1501, 2012. arXiv:1005.3818.

[101] Christopher Gerry and Peter Knight. *Introductory Quantum Optics*. Cambridge University Press, 2004.

[102] Saikat Guha, Jeffrey H Shapiro, and Baris I Erkmen. Classical capacity of bosonic broadcast communication and a minimum output entropy conjecture. *Physical Review A*, 76(3):032303, 2007.

[103] Giacomo De Palma, Dario Trevisan, and Vittorio Giovannetti. Gaussian states minimize the output entropy of one-mode quantum Gaussian channels. *Physical Review Letters*, 118(16):160503, 2017. arXiv:1610.09970.

[104] Filippo Caruso and Vittorio Giovannetti. Degradability of bosonic Gaussian channels. *Physical Review A*, 74(6):062307, 2006. arXiv:quant-ph/0603257.

[105] Alexander S. Holevo. Entanglement-breaking channels in infinite dimensions. *Problems of Information Transmission*, 44(3):171–184, 2008. arXiv:0802.0235.

[106] Vittorio Giovannetti, Raúl García-Patrón, Nicholas J Cerf, and Alexander S Holevo. Ultimate classical communication rates of quantum optical channels. *Nature Photonics*, 8(10):796–800, 2014.

[107] Kamil Brádler, Patrick Hayden, Dave Touchette, and Mark M Wilde. Trade-off capacities of the quantum hadamard channels. *Physical Review A*, 81(6):062312, 2010.

[108] Mark M. Wilde and Min-Hsiu Hsieh. The quantum dynamic capacity formula of a quantum channel. *Quantum Information Processing*, 11(6):1431–1463, 2012.

[109] Qingle Wang, Siddhartha Das, and Mark M Wilde. Capacities of entanglement-breaking and Hadamard multiple-access and broadcast channels. *in preparation*, 2016.

[110] Michael M. Wolf, David Pérez-García, and Geza Giedke. Quantum capacities of bosonic channels. *Physical Review Letters*, 98(13):130501, March 2007. arXiv:quant-ph/0606132.

[111] Charles H. Bennett, Peter W. Shor, John A. Smolin, and Ashish V. Thapliyal. Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem. *IEEE Transactions on Information Theory*, 48(10):2637–2655, October 2002. arXiv:quant-ph/0106052.

[112] Vittorio Giovannetti, Seth Lloyd, Lorenzo Maccone, and Peter W. Shor. Entanglement assisted capacity of the broadband lossy channel. *Physical Review Letters*, 91(4):047901, 2003. arXiv:quant-ph/0304020.

[113] Eduardo Martín-Martínez, Dominic Hosler, and Miguel Montero. Fundamental limitations to information transfer in accelerated frames. *Physical Review A*, 86(6):062307, 2012. arXiv:1204.6271.

[114] Jon Yard, Patrick Hayden, and Igor Devetak. Quantum broadcast channels. *IEEE Transactions on Information Theory*, 57(10):7147–7162, 2011. arXiv:quant-ph/0603098.

[115] Ivan Savov and Mark M. Wilde. Classical codes for quantum broadcast channels. *IEEE Transactions on Information Theory*, 61(12):7017–7028, 2015. arXiv:1111.3645.

[116] Horace P. Yuen and Jeffrey H. Shapiro. Optical communication with two-photon coherent states–part iii: Quantum measurements realizable with photoemissive detectors. *IEEE Transactions on Information Theory*, 26(1):78–92, 1980.

[117] Carlton M. Caves and Peter D. Drummond. Quantum limits on bosonic communication rates. *Reviews of Modern Physics*, 66(2):481, 1994.

[118] Thomas M. Cover and Joy A. Thomas. *Elements of information theory*. John Wiley & Sons, 2012.

[119] Giacomo De Palma, Dario Trevisan, and Vittorio Giovannetti. Passive states optimize the output of bosonic Gaussian quantum channels. *IEEE Transactions on Information Theory*, 62(5):2895–2906, 2016.

[120] Albert W. Marshall, Ingram Olkin, and Barry C. Arnold. *Inequalities: Theory of Majorization and Its Applications*. Springer Series in Statistics. Springer, 2011.

[121] Vittorio Giovannetti, Saikat Guha, Seth Lloyd, Lorenzo Maccone, and Jeffrey H Shapiro. Minimum output entropy of bosonic channels: a conjecture. *Physical Review A*, 70(3):032315, 2004.

[122] Andrea Mari, Vittorio Giovannetti, and Alexander S. Holevo. Quantum state majorization at the output of bosonic Gaussian channels. *Nature communications*, 5:3826, 2014. arXiv:1312.3545.

[123] Giacomo De Palma, Dario Trevisan, and Vittorio Giovannetti. Gaussian states minimize the output entropy of the one-mode quantum attenuator. *IEEE Transactions on Information Theory*, 63(1):728–737, 2017. arXiv:1605.00441.

[124] Saikat Guha, Baris I Erkmen, and Jeffrey H Shapiro. The entropy photon-number inequality and its consequences. In *Information Theory and Applications Workshop, 2008*, pages 128–130. IEEE, 2008.

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

[126] Giacomo De Palma, Andrea Mari, and Vittorio Giovannetti. A generalization of the entropy power inequality to bosonic quantum systems. *Nature Photonics*, 8(12):958–964, 2014.

[127] Peter W Shor. Scheme for reducing decoherence in quantum computer memory. *Physical review A*, 52(4):R2493, 1995.

[128] Andrew M Steane. Error correcting codes in quantum theory. *Physical Review Letters*, 77(5):793, 1996.

[129] Emanuel Knill and Raymond Laflamme. Theory of quantum error-correcting codes. *Physical Review A*, 55(2):900, 1997.

[130] Rami Barends, Julian Kelly, Anthony Megrant, Andrzej Veitia, Daniel Sank, Evan Jeffrey, Ted C White, Josh Mutus, Austin G Fowler, Brooks Campbell, et al. Superconducting quantum circuits at the surface code threshold for fault tolerance. *Nature*, 508(7497):500–503, 2014.

[131] Erwin L Hahn. Spin echoes. *Physical review*, 80(4):580, 1950.

[132] Charles P Slichter. *Principles of magnetic resonance*, volume 1. Springer Science & Business Media, 2013.

[133] Lorenza Viola and Seth Lloyd. Dynamical suppression of decoherence in two-state quantum systems. *Physics Review A*, 58:2733–2744, Oct 1998.

[134] Lorenza Viola, Emanuel Knill, and Seth Lloyd. Dynamical decoupling of open quantum systems. *Physical Review Letters*, 82(12):2417, 1999.

[135] Paolo Zanardi. Symmetrizing evolutions. *Physics Letters A*, 258(2):77–82, 1999.

[136] K Khodjasteh and DA Lidar. Fault-tolerant quantum dynamical decoupling. *Physical Review Letters*, 95(18):180501, 2005.

[137] Kaveh Khodjasteh and Daniel A Lidar. Performance of deterministic dynamical decoupling schemes: Concatenated and periodic pulse sequences. *Physical Review A*, 75:062310, 2007.

[138] G.S. Uhrig. *Physics Review Letter*, 98(100504), 2007.

[139] Jacob R West, Bryan H Fong, and Daniel A Lidar. Near-optimal dynamical decoupling of a qubit. *Physical Review Letter*, 104(13):130501, 2010.

[140] CA Ryan, JS Hodges, and DG Cory. Robust decoupling techniques to extend quantum coherence in diamond. *Physical Review Letter*, 105(20):200402, 2010.

[141] Zhihao Xiao, Lewei He, and W. Wang. Efficiency of dynamical decoupling sequences in the presence of pulse errors. *Phys. Rev. A*, 83(3):032322, 2011.

[142] Ashok Ajoy, Gonzalo A. Álvarez, and Dieter Suter. Optimal pulse spacing for dynamical decoupling in the presence of a purely dephasing spin bath. *Physical Review A*, 83(3):032303, 2011.

[143] Alexandre M Souza, Gonzalo A Álvarez, and Dieter Suter. Robust dynamical decoupling. *Philos. T. R. Soc. A*, 370(1976):4748–4769, 2012.

[144] Zhi-Hui Wang, G De Lange, D Ristè, R Hanson, and VV Dobrovitski. Comparison of dynamical decoupling protocols for a nitrogen-vacancy center in diamond. *Phys. Rev. B*, 85(15):155204, 2012.

[145] Demitry Farfurnik, Andrey Jarmola, Linh M Pham, Zhi-Hui Wang, Viatcheslav V Dobrovitski, Ronald L Walsworth, Dmitry Budker, and Nir Bar-Gill. Optimizing a dynamical decoupling protocol for solid-state electronic spin ensembles in diamond. *Phys. Rev. B*, 92:060301, 2015.

[146] G. T. Genov, D. Schraft, N. V. Vitanov, and T. Halfmann, e-print arXiv:1609.09416.

[147] Götz S Uhrig. Keeping a quantum bit alive by optimized π-pulse sequences. *Physical Review Letter*, 98(10):100504, 2007.

[148] A. G. Kofman and G. Kurizki. Unified theory of dynamically suppressed qubit decoherence in thermal baths. *Physical Review Letter*, 93:130406, 2004.

[149] Todd J Green, Jarrah Sastrawan, Hermann Uys, and Michael J Biercuk. Arbitrary quantum control of qubits in the presence of universal noise. *New Journal of Physical*, 15(9):095004, 2013.

[150] G. A. Paz-Silva and L. Viola. *Physical Review Letter*, 113:250501, 2014.

[151] A. Soare, H. Ball, M. C. Jarratt, J. J. McLoughlin, X. Zhen, T. J. Green, and M. J. Biercuk. Experimental noise filtering by quantum control . *Nature Physics*, 10:825, 2014.

[152] David Hayes, Kaveh Khodjasteh, Lorenza Viola, and Michael J Biercuk. Reducing sequencing complexity in dynamical quantum error suppression by walsh modulation. *Phys. Rev. A*, 84(6):062323, 2011.

[153] D. Hayes, S. M. Clark, S. Debnath, D. Hucul, I. V. Inlek, K. W. Lee, Q. Quraishi, and C. Monroe. Coherent error suppression in multiqubit entangling gates. *Physical Review Letter*, 109:020503, 2012.

[154] H. Ball and M. J. Biercuk. Walsh-synthesized noise filters for quantum logic. *EPJ Quantum Tech.*, 2:1, 2015.

[155] K. Khodjasteh, J. Sastrawan, D. Hayes, T. J. Green, M. J. Biercuk, and L. Viola. *Nature Communication*, 4:2045, 2013.

[156] Kenneth George Beauchamp. *Walsh Functions and their Applications*, volume 3. Academic Press, 1975.

[157] K. Khodjasteh and L. Viola, Physical Review Letter **102**, 080501 (2009); K. Khodjasteh, D. A. Lidar and L. Viola, *ibid.* **104**, 090501 (2010).

[158] S. Blanes, F. Casas, J. A. Oteo, and J. Ros. The magnus expansion and some of its applications. *Physical Report*, 470:151, 2009.

[159] L. Viola and S. Lloyd, Physical Review A **58**, 2733 (1998); L. Viola, E. Knill, and S. Lloyd, Physical Review Letter. **82**, 2417 (1999).

[160] G de Lange, Z H Wang, D Rist, V V Dobrovitski, and R Hanson. *Science*, 330:60–63, 2010.

[161] Lorenza Viola and Emanuel Knill. Random decoupling schemes for quantum dynamical control and error suppression. *Physical Review Letters*, 94(6):060502, 2005.

[162] Lea F Santos and Lorenza Viola. Advantages of randomization in coherent quantum dynamical control. *New Journal of Physics*, 10(8):083009, 2008.

[163] G. Quiroz and D. A. Lidar. *Physical Review A*, 88:052306, 2013.

[164] L. Viola, E. Knill, and S. Lloyd. *Physical Review Letter*, 82:2417, 1999.

[165] Kaveh Khodjasteh and Daniel A. Lidar. Fault-tolerant quantum dynamical decoupling. *Physical Review Letter*, 95(18):180501, 2005.

[166] Herman Y Carr and Edward M Purcell. Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Physical Review*, 94(3):630, 1954.

[167] Nicolaas Bloembergen, Edward Mills Purcell, and Robert V Pound. Relaxation effects in nuclear magnetic resonance absorption. *Physical Review*, 73(7):679, 1948.

[168] S. Pasini and G. S. Uhrig. Optimized dynamical decoupling for power-law noise spectra. *Physical Review A*, 81:012309, Jan 2010.

[169] Götz S Uhrig and Daniel A Lidar. Rigorous bounds for optimal dynamical decoupling. *Physical Review A*, 82(1):012301, 2010.

[170] C. A. Ryan, J. S. Hodges, and D. G. Cory. Robust decoupling techniques to extend quantum coherence in diamond. *Physical Review Letter*, 105:200402, Nov 2010.

[171] Khodjasteh, Kaveh and Lidar, Daniel A., Physical Review Letter **95**, 180501 (2005); Phys. Rev. A **75**, 062310 (2007).

[172] Note that, with respect to Ref. [31], p_0 is introduced here for added generality.

[173] Daniel A. Lidar, Paolo Zanardi, and Kaveh Khodjasteh. Distance bounds on quantum dynamics. *Physical Review A*, 78(1):012308, 2008.

[174] L. F. Santos and L. Viola. *New Journal of Physics*, 10:083009, 2008.

[175] WA Coish, Vitaly N Golovach, J Carlos Egues, and Daniel Loss. Measurement, control, and decay of quantum-dot spins. *Physical Status Solidi B*, 243(14):3658–3672, 2006.

[176] Rogerio de Sousa and S Das Sarma. Theory of nuclear-induced spectral diffusion: Spin decoherence of phosphorus donors in si and gaas quantum dots. *Physical Review B*, 68(11):115322, 2003.

[177] Vittorio Giovannetti, Saikat Guha, Seth Lloyd, Lorenzo Maccone, Jeffrey H Shapiro, and Horace P Yuen. Classical capacity of the lossy bosonic channel: The exact solution. *Physical review letters*, 92(2):027902, 2004.

[178] Mark M. Wilde, Patrick Hayden, and Saikat Guha. Information trade-offs for optical quantum communication. *Physical Review Letters*, 108(140501), 2012.

[179] Andrea Mari, Vittorio Giovannetti, and Alexander S Holevo. Quantum state majorization at the output of bosonic gaussian channels. *arXiv preprint arXiv:1312.3545*, 2013.

[180] Saikat Guha. *Multiple-user quantum information theory for optical communication channels*. PhD thesis, MIT, 2008.

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

[182] Robert König and Graeme Smith. The entropy power inequality for quantum systems. *IEEE Transactions on Information Theory*, 60(3):1536–1548, 2014.

[183] Jeffrey H. Shapiro. Quantum optical communication course lecture notes (lecture 12). Available as MIT Course No. 6.453 on MIT OpenCourseware, 2008.

APPENDIX A

Minimum Output-entropy Conjectures

A.1 Introduction

The first capacity of bosonic channel solved exactly is the classical capacity of the pure-loss channel [177]. In this case none of MOE conjectures is needed, or put in other word, the MOE conjecture for pure-loss channel is trivial to prove, due to the fact that pure-loss channel preserves the vacuum state. Later the same authors realized that to prove the capacity for thermal channel, it requires a non-trivial inequality and they formally called it the MOE conjecture [121]. This conjecture states that it is the vacuum state that minimize the output entropy. Although intuitive as it seems, it takes ten years to prove this conjecture for all multi-mode phase-insensitive Gaussian channels [106]. In 2007, Guha and Shapiro showed that the proof of capacity for broadcast channel requires a MOE with constraint on the input entropy [102]. Later, Wilde et al showed that the triple trade-off capacity of pure-loss channel can be solved if the aforementioned MOE conjecture is true [178]. With recently development of the majorization techniques [179, 119], De Palma et al. proved this MOE conjecture for single-mode pure-loss channel. However, this result is not enough to solve the broadcast and triple trade-off capacity for pure-loss channel, since we always need a multi-mode MOE conjecture to be true in the converse proof due to the regularization issue. Later De Palma et al. generalized his result to all single-mode gauge-covariant channels [103]. Together with Guha and Wilde, I generalized this result to gauge-contravariant Gaussian channels [30] and thus solved the triple trade-off capacity and broadcast capacity for quantum-limited amplifier channels [29]. Our work is the first application of the constrained version of MOE conjecture.

A.2 MINIMUM OUTPUT-ENTROPY CONJECTURES

Von Neumann entropy characterizes the chaos or the noisiness of the corresponding quantum state. Therefore, the minimum output-entropy of a quantum channel measures the minimal amount of noise it injects, which directly relates to upper bound on channel capacities. This is the reason why MOE conjectures are so important in the converse proofs. Here we separate MOE conjectures into two classes depending on if there is a constraint on the input entropy. Therefore the numbering of MOE conjectures is different from those used in [180] and [30]. There, different MOE conjectures are sometimes classified by the existence of constraint and sometimes classified by the environment state (vacuum or thermal noise). Although the ordering in [102, 30] agrees with the historical development, the logic underneath is not satisfying. At the last we will briefly mention the entropy photon-number inequality (EPnI) [124], which subsumes all MOE conjectures and is considered as the holy grail in this topic.

The most elementary communication protocol in quantum Shannon theory is the classical communication over quantum channels. The capacity in this case is given by the regularization of the Holevo information of the channel

$$C(\mathcal{N}) = \lim_{k\to} \frac{1}{k}\chi(\mathcal{N}^{\otimes k}), \tag{A.1}$$

where

$$\chi(\mathcal{N}) = \max_{\rho_A:\mathrm{Tr}\{\hat{H}\rho_A\}\leq N_S} \left\{ H(\sum_x p(x)\rho_x^B) - \sum_x p(x) H(\rho_x^B) \right\} . \tag{A.2}$$

Here $\rho_x = \mathcal{N}(\psi_x)$ and $\rho_A = \sum_x p(x)\psi_x^A$.

It turns out that to calculate above capacity for phase-insensitive Gaussian channels, we need the following MOE theorem (since it has been proved):

Theorem A.1 ([106]) *Consider a n-mode phase insensitive Gaussian channel $\mathcal{N}_{A\to B}^{\otimes n}$. The output von Neumann entropy $H(\mathcal{N}^{\otimes n}(\rho_{A^n}))$ is minimized when ρ_{A^n} is the n-mode vacuum state:*

$$\min_{\rho} H(\mathcal{N}^{\otimes n}(\rho_{A^n})) = nH(\mathcal{N}(|0\rangle\langle 0|)). \tag{A.3}$$

For details of the proof please refer to [106]. Here I would like to illustrate how to use this MOE theorem to prove the classical capacity of a thermal channel. A thermal channel is given by the transformation

$$\hat{b} = \sqrt{\eta}\hat{a} + \sqrt{1-\eta}\hat{e}\,, \tag{A.4}$$

and the environment is a thermal state with mean photon number N_B. Here $\hat{a}, \hat{b}, \hat{e}$ are the annihilation operators of the input, output, and environment systems. We assume the input power is upper bounded by N_S. As we mentioned earlier, the first part of a capacity proof is to guess a good input ensemble $\{p(x), \psi_x\}$. Usually a good choice is the coherent states with complex Gaussian distribution,

$$\Big\{\frac{1}{\pi N_S} e^{-|\alpha|^2/2N_S}, |\alpha\rangle\Big\}\,. \tag{A.5}$$

It is not difficult to verify that the average input state is a thermal state with mean photon N_S. Therefore, we have the following lower bound on the energy-constrained capacity

$$C(\mathcal{N}) \geq g(\eta N_S + (1-\eta)N_B) - \int d\alpha^2 H(\mathcal{N}(D(\alpha)|0\rangle\langle 0|D^\dagger(\alpha))) \tag{A.6}$$

$$= g(\eta N_S + (1-\eta)N_B) - \int d\alpha^2 H(D(\sqrt{\eta}\alpha)(\mathcal{N}|0\rangle\langle 0|)D^\dagger(\sqrt{\eta}\alpha)) \tag{A.7}$$

$$= g(\eta N_S + (1-\eta)N_B) - g((1-\eta)N_B)\,. \tag{A.8}$$

The first step follows from that fact that an input thermal state with mean photon number

N_S is transformed to an output thermal state with mean photon number $\eta N_S + (1-\eta)N_B$ by the thermal channel. The first equality follows from the covariance of thermal channel. The last equality follows from the fact that entropy is invariant under unitaries.

Now comes the converse part. We need to consider the Holevo information of n-mode thermal channel. Since the mean photon number of the output state is at most $n(\eta N_S + (1-\eta)N_B)$, we know the corresponding tensor product thermal state maximizes the entropy $H(\mathcal{N}(\rho_{A^n}))$,

$$H(\mathcal{N}(\rho_{A^n})) \leq ng(\eta N_S + (1-\eta)N_B) \; . \tag{A.9}$$

Now from the MOE Theorem A.1 we know that

$$H(N^{\otimes n}(\rho_{A^n})) \geq H(\mathcal{N}^{\otimes n}(|0\rangle\langle 0|^{\otimes n})) \tag{A.10}$$

$$= ng((1-\eta)N_B) \; . \tag{A.11}$$

Together we have the following upper bound

$$C(\mathcal{N}^{\otimes n}) \leq ng(\eta N_S + (1-\eta)N_B) - ng((1-\eta)N_B) \; . \tag{A.12}$$

Combining with the lower bound we prove the following capacity for the thermal channel,

$$C(\mathcal{N}) = g(\eta N_S + (1-\eta)N_B) - g((1-\eta)N_B) \; . \tag{A.13}$$

As we can see, the MOE theorem is essential for proving the converse part. Similarly, applying Theorem A.1 to other phase-insensitive Gaussian channels, we can get the corresponding classical capacities [106].

Conjecture A.2 *Consider an n-mode phase insensitive Gaussian channel $\mathcal{N}_{A\to B}^{\otimes n}$. For any input state ρ_{A^n} such that $H(\rho_{A^n}) \geq nH_0$. The output von Neumann entropy $H(\mathcal{N}^{\otimes n}(\rho_{A^n}))$ is minimized when ρ_{A^n} is the tensor product of n thermal states, each with mean photon number $g^{-1}(H_0)$:*

$$\min_{\rho: H(\rho)\geq nH_0} H(\mathcal{N}^{\otimes n}(\rho_{A^n})) = nH(\mathcal{N}(\rho_{th})). \tag{A.14}$$

Here ρ_{th} is a thermal state such that $\mathrm{Tr}\{\hat{H}\rho_{th}\} = g^{-1}(H_0)$.

It is straightforward to see that Conjecture A.2 subsumes Theorem A.1 as a special case if it were true. Recently de Palma et al [103] proved Conjecture A.2 for single-mode gauge-covariant Gaussian channels, and we generalized their results to gauge-contravariant Gaussian channels, and thus completing the proof of Conjecture A.2 for $n = 1$. The n-mode MOE conjecture seems difficult to prove since it involves sorts of additivity of the output-entropy. So far, the only application of the single-mode theorem is our work on quantum-limited amplifier channels [29] (see Chapter 3).

Here we would like to briefly mention a more general form of the MOE conjectures. In his breakthrough paper 'A mathematical theory of communication' [181], Shannon not only laid down the foundation of communication theory, but also proved an elegant and fundamental inequality:

Theorem A.3 (Entropy power inequality (EPI)) *Let $\mathbf{X}$ and $\mathbf{Y}$ be two statistically independent, n dimensional, real-valued random vectors. Denote $h(\mathbf{X})$ and $h(\mathbf{Y})$ to be their differential entropies. Consider the following convex combination*

$$\mathbf{Z} = \sqrt{\lambda}\mathbf{X} + \sqrt{1-\lambda}\mathbf{Y}\ , \tag{A.15}$$

with $0 \leq \lambda \leq 1$. *In other word, the distribution of* $\mathbf{Z}$ *is given by*

$$p_{\mathbf{Z}}(\mathbf{z}) = \int d\mathbf{x} p_{\mathbf{X}}(\mathbf{x}) p_{\mathbf{Y}} \left(\frac{\mathbf{z} - \sqrt{\lambda}\mathbf{x}}{\sqrt{1-\lambda}} \right) . \tag{A.16}$$

Then the following inequality is true

$$e^{2h(\mathbf{Z})/n} \geq \lambda e^{2h(\mathbf{X})/n} + (1-\lambda) e^{2h(\mathbf{Y})/n} . \tag{A.17}$$

In the above theorem, the quantity $e^{2h(X)}/2\pi e$ is called the *entropy power* of a 1D random variable. The term 'power' comes from the fact that $e^{2h(X)}/2\pi e = \langle X^2 \rangle$ when X is a Gaussian variable. Therefore, the entropy power of a random variable is the mean-squared value of another variable which has the same entropy. The equality holds only when both $\mathbf{X}$ and $\mathbf{Y}$ are i.i.d. Gaussian distributed. The EPI has found many applications in converse theorem for additive Gaussian noise channels in both single and multi-user scenarios.

To seek a suitable form of EPI for bosonic system, we need first to figure out what is the correct analog of entropy power. It turns out that, for a n-mode density operator ρ, one natural analog of the entropy power for continuous random variable is the *entropy photon number*:

$$N(\rho) \equiv g^{-1}(H(\rho)/n) . \tag{A.18}$$

Therefore the entropy photon number of a bosonic state is equal to the mean photon number of the thermal state whose tensor product has the same entropy. Here the thermal state takes over the role of Gaussian distribution in EPI. Considering the specialty of thermal states in MOE Conjecture A.2, the entropy photon number is reasonable to be a natural choice. This point is clear after we give the formal conjecture of entropy photon-number

inequality (EPnI).

Conjecture A.4 (EPnI-1) *Let* $\hat{\mathbf{a}}$ *and* $\hat{\mathbf{b}}$ *be vectors of photon annihilation operators for* $2n$ *modes. Consider the following beam-splitter-like transformation with* $0 \leq \eta \leq 1$*,*

$$\hat{\mathbf{c}} = \sqrt{\eta}\,\hat{\mathbf{a}} + \sqrt{1-\eta}\,\hat{\mathbf{b}}\,. \tag{A.19}$$

Then the following is true

$$g^{-1}(H(\rho_{C^n})/n) \geq \eta g^{-1}(H(\rho_{A^n})/n) + (1-\eta)g^{-1}(H(\rho_{B^n})/n)\,, \tag{A.20}$$

for arbitrary input states ρ_{A^n} *and* ρ_{B^n}*.*

Consider the special case when the input B is a thermal state with mean photon number N_B; then the transformation becomes a thermal channel from A to C. Assume $H(\rho_{A^n}) \geq nH_0$ and from Conjecture A.4 we should have

$$\begin{aligned} g^{-1}(H(\rho_{C^n})/n) &\geq \eta g^{-1}(H(\rho_{A^n})/n) + (1-\eta)N_B & \text{(A.21)}\\ &\geq \eta g^{-1}(H_0) + (1-\eta)N_B\,, & \text{(A.22)} \end{aligned}$$

which is equivalent to

$$\begin{aligned} H(\mathcal{N}^{\otimes n}(\rho_{A^n})) &\geq ng\left(\eta g^{-1}(H_0) + (1-\eta)N_B\right) & \text{(A.23)}\\ &= nH(\mathcal{N}(\rho_{\text{th}})), & \text{(A.24)} \end{aligned}$$

where ρ_{th} is a thermal state with mean photon number $g^{-1}(H_0)$. Therefore we correctly reduce the EPnI to the MOE Conjecture A.2.

Since phase-insensitive Gaussian channels include both beam-splitters and amplifiers, we can have another EPnI conjecture for the amplifiers:

Conjecture A.5 (EPnI-2) *Let* $\hat{\mathbf{a}}$ *and* $\hat{\mathbf{b}}$ *be vectors of photon annihilation operators for* $2n$ *modes. Consider the following Bogoliubov transformation with* $\kappa > 1$*,*

$$\hat{\mathbf{c}} = \sqrt{\kappa}\,\hat{\mathbf{a}} + \sqrt{\kappa - 1}\,\hat{\mathbf{b}}^\dagger\,. \tag{A.25}$$

Then the following is true

$$g^{-1}(H(\rho_{C^n})/n) \geq \kappa g^{-1}(H(\rho_{A^n})/n) + (\kappa - 1)[g^{-1}(H(\rho_{B^n})/n) + 1]\,, \tag{A.26}$$

for arbitrary input states ρ_{A^n} *and* ρ_{B^n}*.*

It's not difficult to observe that this conjecture reduces to Conjecture A.2 for amplifier channels when one of the input state is thermal state. We can actually have another EPnI for the gauge-contravariant channels but we will stop here.

It worth to mention that, there is indeed a quantum EPI for bosonic system which was proposed and has been proved [182, 126]:

Theorem A.6 (qEPI) *Let* $\hat{\mathbf{a}}$ *and* $\hat{\mathbf{b}}$ *be vectors of photon annihilation operators for* $2n$ *modes. Consider the following beam-splitter-like transformation with* $0 \leq \eta \leq 1$*,*

$$\hat{\mathbf{c}} = \sqrt{\eta}\,\hat{\mathbf{a}} + \sqrt{1 - \eta}\,\hat{\mathbf{b}}\,. \tag{A.27}$$

Then the following is true

$$\exp(H(\rho_{C^n})/n) \geq \eta \exp(H(\rho_{A^n})/n) + (1 - \eta)\exp(H(\rho_{B^n})/n)\,, \tag{A.28}$$

for arbitrary input states ρ_{A^n} *and* ρ_{B^n}*.*

The proof follows a quantum generalization of the proof for classical EPI. Surprisingly the multi-mode seems not to be a issue at all in their proof but merely a parameter [182,

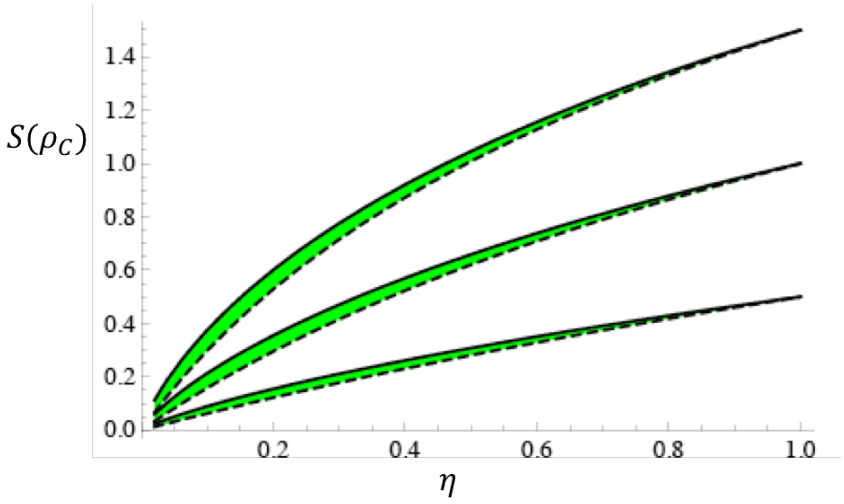

Figure A.1: Comparison between qEPI and EPnI (adopted from Ref. [126]).Set the input state at B system to vacuum so we have a pure-loss channel. We fix the entropic constraint on A to be $H_0 = 0.5, 1, 1.5$ The qEPI becomes $S(\rho_C) \geq \ln[\eta e^{H_0} + 1 - \eta]$ which is plotted in dotted lines. The lower bound achieved by the thermal input state at A is plotted in solid lines. The green area shows the tiny difference. This difference become negligible when H_0 is large.

126]. Although the qEPI does not reduce to MOE conjectures in special cases, authors in Ref. [126] shows that the difference between the lower bound given by the qEPI and those by EPnI is small, as we showed in Fig. A.1.

APPENDIX B

Extra Results For Chapter 3

B.1 TWO PROPERTIES OF $g(x)$

We first recall a property of the function

$$g(x) = (x+1)\log_2(x+1) - x\log_2 x, \tag{B.1}$$

which is helpful for our converse proofs. Recall that $g(x)$ is equal to the entropy of a thermal state with mean photon number x.

Theorem B.1 (Theorem A.3 of [180]) *Given $q \in [0,1]$ a probability distribution $p_X(x)$ and non-negative real numbers $\{y_x : 1 \leq x \leq n\}$, if*

$$\sum_{x=1}^{n} p_X(x) g(y_x) = g(y_0)\ , \tag{B.2}$$

then the following inequality holds for $C \geq 0$:

$$\sum_{x=1}^{n} p_X(x) g(q y_x + C) \geq g(q y_0 + C)\ . \tag{B.3}$$

As mentioned above, the above inequality is Theorem A.3 in Appendix C of [180]. Observe that Ref. [180] proved the inequality for $p_X(x)$ set to the uniform distribution. However, the argument there only relies on concavity of $g(x)$ and thus applies to an arbitrary distribution, as discussed later in [99].

Due to the requirement that $q \in [0,1]$, Theorem B.1 is not useful for the quantum

amplifier channel given that its amplifier gain $\kappa > 1$. To resolve this problem, we prove another property of $g(x)$:

Theorem B.2 *Given $q \in (1, +\infty)$, a probability distribution $p_X(x)$ and non-negative real numbers $\{y_x : 1 \leq x \leq n\}$, if*

$$\sum_{x=1}^{n} p_X(x) g(y_x) = g(y_0), \tag{B.4}$$

then

$$\sum_{x=1}^{n} p_X(x) g(q y_x + q - 1) \geq g(q y_0 + q - 1) \, . \tag{B.5}$$

Proof. The original proof of Theorem B.1 depends on the following inequality:

$$\log_2\left(1 + \frac{1}{qx + C}\right)(qx + C)(1 + qx + C) \geq \log_2\left(1 + \frac{1}{x}\right) qx(1 + x) \, , \tag{B.6}$$

which holds for $q \in [0, 1]$, $x \geq 0$, and $C \geq 0$. When considering $q > 1$, the above inequality does not generally hold. However, we can prove that it is true for $C = q - 1$. Substituting $C = q - 1$ in (B.6), we need to show that

$$(q(1 + x) - 1) \log_2 \frac{q(1 + x)}{q(1 + x) - 1} \geq x \log_2 \frac{1 + x}{x} \, . \tag{B.7}$$

Defining $h(x) = x \ln \frac{x+1}{x}$, then we can see that the above inequality is equivalent to the following one:

$$h(q(1 + x) - 1) \geq h(x) \, . \tag{B.8}$$

But

$$\lim_{x \to 0} h(x) = \lim_{t \to +\infty} \frac{\ln(1 + t)}{t} = \lim_{t \to +\infty} \frac{1}{1 + t} = 0 \, , \tag{B.9}$$

by L'Hospital's rule, and

$$h'(x) = \ln\left(1 + \frac{1}{x}\right) - \frac{1}{1+x} \, . \tag{B.10}$$

Since $h'(0) = +\infty$, $h'(+\infty) = 0$, and $h''(x) = -1/x(1+x)^2$, we have

$$h'(x) \geq 0 \, , \tag{B.11}$$

for $x \geq 0$, and the function $h(x)$ is non-negative and monotonically increasing for non-negative x. Now since $q(1+x) - 1 - x = (1+x)(q-1) \geq 0$, we find that

$$h(q(1+x) - 1) \geq h(x) \, . \tag{B.12}$$

This concludes the proof. ■

B.2 COHERENT-DETECTION SCHEMES

Although we have shown that (3.39)–(3.40) is achievable by using coherent-state encoding with a Gaussian distribution, implicitly we have also assumed that it is achieved by some fully quantum measurement scheme. If the two receivers use classical coherent detection instead, the problem reduces to a classical broadcast channel with Gaussian additive noise. We expect such schemes to be outperformed by those achieved with a fully quantum measurement.

One way to calculate the capacity region of the classical degradable broadcast channel is to use the formula from [114] with the same distribution as in (3.41). Another easier way is to first calculate the capacity of each classical channel to Bob and Charlie. Since each channel is Gaussian with additive noise, each capacity should have the following

form:

$$C \equiv C(\mathtt{snr}) = \frac{1}{2}\log_2(1+\mathtt{snr}_{\mathtt{B/C}}) \ , \tag{B.13}$$

where $\mathtt{snr}_{B/C}$ is the signal-to-noise ratio of the channel $A \to B/C$. Then we can use known results [118, 71] to directly get the capacity region for broadcast channel,

$$\begin{aligned} R_B &\leq C(\lambda \mathtt{snr}_B), \\ R_C &\leq C\left(\frac{(1-\lambda)\mathtt{snr}_C}{\lambda \mathtt{snr}_C + 1}\right) \ . \end{aligned} \tag{B.14}$$

For Bob, the channel could be modeled by the following transformation:

$$B = \sqrt{\kappa}A + Z \ . \tag{B.15}$$

If homodyne detection is employed, Bob is measuring one of the quadratures and B, A, and Z are scalar Gaussian random variables. The noise Z has distribution $Z \sim N(0, \frac{1}{4} + \frac{1}{2}\bar{\kappa})$, where the variance comes from both the vacuum itself and the thermal noise generated from the vacuum [183]. The capacity of the classical Gaussian channel is achieved by input with distribution $A \sim N(0, N_S)$, and therefore we have

$$\mathtt{snr}_B = \frac{\kappa N_S}{\frac{1}{4}(1+2\bar{\kappa})}. \tag{B.16}$$

When heterodyne detection is used, B, A, and Z are complex Gaussian random variables. The real part of the noise has distribution $\Re(Z) \sim N(0, \frac{1}{2} + \frac{1}{2}\bar{\kappa})$ and the same for the imaginary part [183]. The optimal input distribution for each part is $\Re(A) \sim N(0, N_S/2)$ and $\mathrm{Im}(A) \sim N(0, N_S/2)$ since the total input power is N_S. Thus for heterodyne detection

we have

$$\mathtt{snr}_B = \frac{\kappa N_S}{(1+\bar{\kappa})} = N_S. \tag{B.17}$$

Notice that we need to multiply the capacity formula by a factor of two, to take into account the contribution from each part of the complex variable. The channel to Charlie is modeled by

$$C = \sqrt{\bar{\kappa}}A + Z\ , \tag{B.18}$$

and all the analysis above for Bob still holds. We can write the capacity of each classical channel achieved by coherent detection in a unified way as

$$C_{A\to B} = \xi \log_2\left(1 + \frac{\kappa N_S}{\xi(\xi+\bar{\kappa})}\right)\ , \tag{B.19}$$

$$C_{A\to C} = \xi \log_2\left(1 + \frac{\bar{\kappa} N_S}{\xi(\xi+\bar{\kappa})}\right)\ , \tag{B.20}$$

where $\xi = \frac{1}{2}$ for homodyne detection and $\xi = 1$ for heterodyne detection.

Now using (B.14), we find the capacity region of coherent detection:

$$R_B \leq \xi \log_2\left(1 + \frac{\lambda\kappa N_S}{\xi(\xi+\bar{\kappa})}\right)\ , \tag{B.21}$$

$$R_C \leq \xi \log_2\left(1 + \frac{(1-\lambda)\bar{\kappa} N_S}{\xi(\xi+\bar{\kappa}) + \lambda\bar{\kappa}N_S}\right)\ , \tag{B.22}$$

thus giving (3.66).

B.3 Upper bound for trade-off capacity region of the pure-loss channel

The capacity region for the information trade-off over a pure-loss channel has been given in [98], provided that a multi-mode minimum output-entropy conjecture is true.

Although the multi-mode conjecture has not been proved yet, the recently proved quantum EPI (qEPI) can give a good upper bound [126, 182], holding for $\eta \in [1/2, 1]$. The qEPI is a direct translation of the classical EPI and is as follows:

$$2^{H(\rho_B)/n} \geq \lambda_A 2^{H(\rho_A)/n} + \lambda_E 2^{H(\rho_E)/n} , \tag{B.23}$$

where ρ_A is the input of one beamsplitter port, ρ_E is the input of the other beamsplitter port, ρ_B is the output of one port, and $\lambda_A = \eta, \lambda_E = 1 - \eta$ for a pure-loss channel with transmissivity η.

When we consider the case when the environment is in the vacuum state, we have $H(\rho_E) = 0$ and the following bound holds

$$H(\rho_B)/n \geq \log_2(\eta 2^{H(\rho_A)/n} + 1 - \eta) . \tag{B.24}$$

We will use this lower bound in what follows.

Recall the development in Eqs. (60)–(76) of [99]. Picking up from there, we have that

$$\sum_x p_X(x) H(\rho_x) = n g(\lambda' N_S) , \tag{B.25}$$

$$\sum_x p_X(x) H(\mathcal{N}(\rho_x)) = n g(\lambda \eta N_S) , \tag{B.26}$$

where $\lambda', \lambda \in [0, 1]$. Now instead of invoking the minimum output-entropy conjecture for a pure-loss channel, we use the lower bound given by the multi-mode qEPI in (B.24):

$$\begin{aligned} g(\lambda \eta N_S) &\geq \sum_x p_X(x) \log_2(\eta 2^{g(\lambda'_x N_{S,x})} + 1 - \eta) , \\ &\geq \log_2(\eta 2^{g(\lambda' N_S)} + 1 - \eta) . \end{aligned} \tag{B.27}$$

The last inequality follows from the fact that $f(x) = \log_2(\eta 2^x + 1 - \eta)$ is convex, and we have also used the equality $\sum_x p_X(x) g(\lambda'_x N_{S,x}) = g(\lambda' N_S)$. Rewriting this, we find that

$$\begin{aligned}
&\sum_x p_X(x) H(\rho_x) \\
&= n g(\lambda' N_S) && \text{(B.28)} \\
&\leq n \log_2 \left[\frac{1}{\eta} \left(2^{g(\lambda \eta N_S)} - (1-\eta) \right) \right] , && \text{(B.29)}
\end{aligned}$$

which replaces Eq. (60) in Ref. [99].

The lower bound given in Eq. (63) of [99] will be replaced by a new lower bound found by invoking the qEPI. Using (B.24) for a pure-loss channel with $\eta' = (1-\eta)/\eta$, we find that

$$\begin{aligned}
&\sum_x p_X(x) H(\mathcal{N}^c(\rho_x)) \\
&\geq \sum_x p_X(x) n \log_2(\eta' 2^{g(\eta \lambda_x N_{S,x})} + 1 - \eta') && \text{(B.30)} \\
&\geq n \log_2(\eta' 2^{\sum_x p(x) g(\eta \lambda_x N_{S,x})} + 1 - \eta') && \text{(B.31)} \\
&= n \log_2 \left(\frac{1-\eta}{\eta} 2^{g(\eta \lambda N_S + 1 - \eta)} + \frac{2\eta - 1}{\eta} \right) . && \text{(B.32)}
\end{aligned}$$

The two inequalities follow by invoking the qEPI and convexity of $f(x)$ as defined and used previously. In the last step, we have used $\sum_x p_X(x) g(\eta \lambda_x N_{S,x}) = g(\lambda \eta N_S)$.

In summary, an upper bound for the trade-off capacity region of the pure-loss channel,

derived from the qEPI, follows from the inequalities below:

$$\begin{aligned}
\frac{1}{n}H(\mathcal{N}^{\otimes n}(\rho)) &\leq g(\eta N_S)\ , \\
\frac{1}{n}\sum_x p_X(x)H(\rho_x) &\leq \log_2\left[\frac{1}{\eta}\left(2^{g(\lambda\eta N_S)}-(1-\eta)\right)\right]\ , \\
\frac{1}{n}\sum_x p_X(x)H(\mathcal{N}^{\otimes n}(\rho_x)) &\leq g(\eta\lambda N_S)\ , \\
\frac{1}{n}\sum_x p_X(x)H((\mathcal{N}^c)^{\otimes n}(\rho_x)) &\geq \log_2\left[\frac{1-\eta}{\eta}2^{g(\lambda\eta N_S)}+\frac{2\eta-1}{\eta}\right]\ .
\end{aligned} \tag{B.33}$$

For the broadcast capacity region of a pure-loss channel, the upper bound is given in Sec. IV.C of [126].

Vita

Haoyu Qi was born in Beijing, China, 1990. He attended Beijing Institute of Technology and graduated with a bachelor of science in physics in 2012. His thesis is about bouncing and cyclic cosmology models.

He joined the Department of Physics and Astronomy at Louisiana State University in 2012. He joined the QST group under the supervision of Dr. Jonathan P. Dowling and Dr. Mark M. Wilde. Since then he has been working on the theory of dynamical decoupling and quantum Shannon theory. He plans to receive his Ph.D. in physics in 2017.

Starting from December 2017, Haoyu Qi will join Xanadu AI in Toronto, Canada, as a research stuff. He will continue his research in optical quantum computation and quantum machine learning.