Quantum communication via noisy channels

Inaugural-Dissertation

zur
Erlangung des Doktorgrades der
Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von
Zahra Shadman
aus Quchan/Iran

January 2011

Aus dem Institut für Theoretische Physik, Lehrstuhl III

der Heinrich-Heine Universität Düsseldorf

Gedruckt mit der Genehmigung der

Mathematisch-Naturwissenschaftlichen Fakultät der

Heinrich-Heine-Universität Düsseldorf

Referent: Prof. Dr. D. Bruß

Koreferent: Prof. Dr. R. Egger

Tag der mündlichen Prüfung: 25.01.2011

To my family

Abstract

Quantum communication is at the heart of the quantum information theory. We study two types of quantum communication protocols over noisy transmission channels, i.e. super dense coding and cryptography protocols. In the first part of this thesis, for various scenarios, it is discussed how the super dense coding capacity is influenced by noisy quantum channels. The case of memoryless channels as well as those channels with memory which are modelled by uncorrelated and correlated noise, respectively, are considered. Explicitly Pauli channels over arbitrary dimensions are treated and the super dense coding capacity for some resource states is derived. For the qubit depolarizing channel, when noise is uncorrelated, the super dense coding capacity with respect to the input state is also optimized. This illustrates a threshold value of the noise parameter below which the super dense coding capacity is optimized by a maximally entangled initial state, while above the threshold value the super dense coding capacity is optimized by a product state. For Pauli channels, with correlated noise, the case of non-unitary encoding is studied and the super dense coding capacity is derived. The first part of this thesis is concluded with an example for multipartite super dense coding.

In the second part of the thesis, the problem of optimal eavesdropping on noisy states in quantum key distribution is investigated. The case of the six state protocol, when the signal states are mixed with white noise, is considered. This situation may arise either when Alice deliberately adds noise to the signal states before they leave her lab, or, in a realistic scenario when the eavesdropper (referred to as Eve) is not assumed to replace the noisy quantum channel by a noiseless one. For individual attacks, we find Eve's optimal mutual information

with Alice as a function of the quantum bit error rate. Finally, the results illustrate that adding noise on the quantum level can make quantum key distribution more robust against eavesdropping.

Zusammenfassung

Die Quantenkommunikation ist ein zentrales Thema der Quanteninformationstheorie. Wir untersuchen zwei Typen von Quantenkommunikationsprotokollen in verrauschten Übertragungskanälen: superdichte Kodierung und Kryptographieprotokolle. Im ersten Teil der Arbeit wird für verschiedene Szenarien der Einfluss von verrauschten Quantenkanälen auf superdichte Verschlüsselungskapazität diskutiert. Es wird sowohl der Fall von gedächtnislosen Kanälen als auch der Fall von Kanälen, deren Gedächtnis durch unkorreliertes und korreliertes Rauschen modelliert wird, betrachtet. Explizit werden Paulikanäle in beliebigen Dimensionen behandelt und die superdichte Verschlüsselungskapazität für einige Ressourcenzustände hergeleitet. Für den Qubit-Depolarisierungskanal mit unkorreliertem Rauschen wird die superdichte Kodierungkapazität abhängig vom Eingangszustand optimiert. Sie illustriert, dass unter einem bestimmten Grenzwert des Rauschparameters die superdichte Kodierungskapazität durch einen maximal verschränkten Zustand und über dem Grenzwert durch Produktzustände optimiert wird. Für Paulikanäle mit korreliertem Rauschen wird der Fall von nicht unitärer Verschlüsselung untersucht und die superdichte Kodierungkapazität hergeleitet. Der erste Teil dieser Arbeit wird mit Beispielen für superdichte Kodierung in Vielteilchensystemen abgeschlossen.

Im zweiten Teil dieser Arbeit wird das Problem des optimalen Abhörens von verrauschten Zuständen im Bereich der Quantenschlüsselübertragung untersucht. Der Fall des Sechs-Zustands-Protokolls, in dem die Signalzustände mit weißem Rauschen gemischt werden, wird betrachtet. Diese Situation kann auftreten, wenn Alice bewusst weißes Rauschen zu ihren Signalzuständen mischt, bevor sie ihr Labor verlassen, oder, in einem realistischen Fall, wenn für den Abhörer nicht angenommen wird, dass er den verrauschten Kanal durch einen rauschfreien

Kanal ersetzt. Für individuelle Attacken finden wir Eves und Alices optimale Transinformation als Funktion der Quantenbitfehlerrate. Abschließend zeigen die Resultate, dass das Hinzufügen von Rauschen im Quantenbereich die Quantenschlüsselübertragung robuster gegen Abhören machen kann.

Contents

List of Figures

Chapter 1

Preliminaries

1.1 Classical information theory

Classical information theory, sometimes referred to as information theory, provides a structure for any type of communication and information processing. It is generally believed that the information theory as a modern discipline has been developed by Claude E. Shannon in his landmark article *A mathematical theory of communication* [47]. In 1963, he expanded the ideas of this article in a book with Warren Weaver [48]. Shannon's theory provided a mathematical model for communication based on probability theory and statistics. He focused on engineering-type problems of communication and gave the engineers a way to determine the capacity of a channel. However, the theory immediately absorbed the interest of mathematicians and other scientists. Some of the applied areas of information theory include communication theory, and coding theory which can be subdivided to source coding and channel coding, signal analysis and cryptography.

Information theory is a well developed subject. It provides many of tools and concepts used in quantum information theory. Therefore, it is important to consider classical information theory before studying how it needs to be adjusted so that quantum effects can be taken into account.

The fundamental concept of the theory is *information*. It is defined as *a message or a sequence of messages to be communicated to the receiving terminal* [47]. Based on the binary logarithm, the unit of information is the *bit*. To convey

information from one location to another, a communication system is essential. As depicted by Shannon, any communication system consists of five important parts. An *information source* which produces a message and a *destination* to whom the message is transmitted. Moreover, a *transmitter* and a *receiver* that converts each message into some physical signals and vice versa. They can also be called an encoder and a decoder. Finally, a *channel* is the medium over which the physical signal is transmitted.

A key measure of information is known as *entropy*. It is closely related to the thermodynamic entropy as defined by physicists. Entropy is a good measure of the uncertainty associated with a random variable. For a random variable with entropy of α, when a receiver (say Bob) at the other end of the channel gets a signal that tells the value of that random variables, he gets rid of the uncertainty α, or in the other words, Bob gains information α.

Besides the entropy, another important quantity in the theory is the *mutual information*. In a simple communication language it measures the amount of information shared between input and output of a noisy channel.

The rest of this section is dedicated to an introductory over classical entropies as well as mutual information and classical channels.

1.1.1 Classical entropy

Shannon entropy

For a discrete random variable X with a sequence of messages $x_1, x_2, ..., x_n$ and the probability distribution $p(x_1), p(x_2), ..., p(x_n)$, the Shannon entropy is defined by

$$H(X) = -\sum_x p(x) \log p(x), \tag{1.1}$$

where logarithms are taken to base 2. The Shannon entropy is a non-negative concave function. That is, for two random variables X and Y with discrete probability distributions $\{p(x)\}$ and $\{q(x)\}$, the entropy of any average of these two distributions is greater than the average of the entropies. i.e.

$$\gamma H(X) + (1-\gamma)H(Y) \leqslant H(\gamma X + (1-\gamma)Y), \tag{1.2}$$

where $0 \leqslant \gamma \leqslant 1$ is the weight of each distribution.

Conditional entropy

Conditional entropy measures, on average, the amount of uncertainty of Y conditional on knowing X. In the language of the communication, for noisy channels without memory, and for given $p(x)$, the output of the channel depends on the input by the conditional probability $p(y|x)$. The conditional entropy or conditional uncertainty is then given by

$$H(Y|X) = -\sum_{x,y} p(x)p(y|x)\log p(y|x). \tag{1.3}$$

Relative entropy

Relative entropy gives a measure of something like the distance between two different probability distributions. For two random variables X and Y with the probability distributions $\{p(x)\}$ and $\{q(x)\}$, the relative entropy is defined by

$$H(p(x)\|q(x)) = \sum_{x} p(x)\log\frac{p(x)}{q(x)}. \tag{1.4}$$

Relative entropy is non-symmetrical and non-negative. It takes the zero value if, and only if, the two distributions are the same, and increases as the distributions diverge.

1.1.2 Mutual information

For two random variables X and Y, the mutual information measures the amount of information that can be obtained about one of the random variables by observing another. It is defined as

$$I(X:Y) = \sum_{x,y} p(x,y)\log p(y|x) - \sum_{y} p(y)\log p(y). \tag{1.5}$$

Mutual information is non-negative and symmetrical. It takes the zero value if, and only if, two random variables are statistically independent.

1.1.3 Classical channel

Channels have important roles in any communication system. A pair of wires, a coaxial cable, a band of radio frequencies, a beam of light, etc., are examples of classical channels. A channel can assumed to be noiseless, which is far from reality, or it can be noisy. Mathematically speaking, a channel is a stochastic map modeling the effect of the noise experienced by the classical message on its way from the sender to the remote receiver. A *binary symmetric channel* is a simple example for a classical channel. It can transmit only one of two symbols (usually called 0 and 1). The effect of noise in this channel is to flip the bit being transmitted with probability p, while with probability $1-p$, the bit is transmitted without error.

An important concept related to any communication channel is the capacity. In the asymptotic limit of many uses of the channel, and by applying appropriate error correction, the channel capacity is defined to be the maximum possible rate at which information can be reliably transmitted through a noisy classical channel. For a discrete channel, it is given by the maximum of the mutual information $I(X:Y)$ with respect to all possible probability distributions $\{p(x)\}$. i.e.,

$$C = \max_{\{P(x)\}} I(X:Y) = \max_{\{P(x)\}} \left(H(X) - H(X|Y)\right). \tag{1.6}$$

1.2 Quantum information theory

Quantum information theory is the result of the generalization of the classical information theory to the quantum world, where the information is carried by quantum systems. Analogous to the *bit*, the unit in quantum information is the quantum bit (*qubit*), a two-level quantum system. An example of a qubit is the direction of the spin of an electron. Unlike classical states (which are discrete), a quantum system can be in a superposition of states. This property is one of the magnificent aspects of the quantum mechanics. The statistical state of a quantum system is described by a density matrix ρ. It is a positive semidefinite matrix with trace one. The operator that is represented by the density matrix is

called the density operator. A *pure* state density matrix has the form $\rho = |\psi\rangle\langle\psi|$, while a density matrix not expressible in this form is in a *mixed* state.

Another important nature of quantum world is the possibility of entangling quantum systems. A state is said to be entangled if it cannot be written as a convex combination of tensor products of its subsystems. Entanglement can be used to perform tasks that are not possible with classical states. Teleportation and super dense coding protocols are two phenomena based on entanglement.

Some concepts like capacity, entropy and channel in classical information theory have been generalized for quantum information theory. In the following, a short overview of some concepts in quantum world is provided.

1.2.1 Unitary evolution

A postulate in quantum mechanics expresses that the state of a closed (isolated) quantum system at two different times are related by a unitary operator. That is, if the state of a quantum system at time t_1 is given by $|\psi(t_1)\rangle$, at a later time of t_2 it goes to $|\psi(t_2)\rangle = U(t_1, t_2)|\psi(t_1)\rangle$. The unitary transformation $U(t_1, t_2)$ is then defined by the Schrödinger equation,

$$i\hbar\frac{d|\psi(t)\rangle}{dt} = H|\psi(t)\rangle, \tag{1.7}$$

where H is the Hamiltonian of the system and $\hbar$ is the Planck's constant. Using the Schrödinger equation (1.7), for a time-independent Hamiltonian H, the unitary transformation $U(t_1, t_2)$ is determined by $U(t_1, t_2) = \exp(-\frac{iH(t_2-t_1)}{\hbar})$.

1.2.2 Measurement

The act of acquiring information about a physical system unavoidably disturbs the state of the quantum system. There is no counterpart of this limitation in classical physics. In quantum physics, a *measurement* is an interaction between a quantum system and a measuring device. This interaction leaves the device in a state representing the outcome of the measurement. The tradeoff between acquiring information and creating disturbance is related to quantum randomness. It is because the outcome of a measurement has a random element that we are unable to infer the initial state of the system from the outcome.

A postulate in quantum mechanics associates to each observable a self-adjoint operator. Another postulate states that the measurement of a physical observable corresponds to an action of the respective operator (say A) on the state of the system. For a quantum system being described by a pure state $|\Psi\rangle$, the probability of the outcome $|\Phi_k\rangle$ is given by

$$p_k = |\langle\Phi_k|\Psi\rangle|^2. \tag{1.8}$$

The process of measurement then causes a collapse of the quantum state $|\Psi\rangle$ to the state $|\Phi_k\rangle$. The set of the outcomes $\{|\Phi_k\rangle\}$ are described by the set of the measurement operators $\{A_k\}$ which act on the initial state $|\Psi\rangle$,

$$|\Phi_k\rangle = \frac{A_k|\Psi\rangle}{\sqrt{p_k}}. \tag{1.9}$$

The measurement operators A_k satisfy the completeness relation,

$$\sum_k A_k^\dagger A_k = \mathbb{1}. \tag{1.10}$$

For a quantum system being described by a mixed state ρ the probability of the outcome $|\Phi_k\rangle$ is given by

$$p_k = \langle\Phi_k|\rho|\Phi_k\rangle. \tag{1.11}$$

Von Neumann measurement

In this part, we explain an special type of general measurement, i.e. von Neumann measurement. For an observable represented by an operator A, when the outcomes of the measurement are the eigenvalues (say α) of the operator with the eigenvectors $|\alpha\rangle$ ($A|\alpha\rangle = \alpha|\alpha\rangle$), the quantum measurement is the so called von Neumann or projective measurement. The observable A can then be represented by

$$A = \sum_\alpha \alpha|\alpha\rangle\langle\alpha| = \sum_\alpha \alpha P_\alpha, \tag{1.12}$$

where $P_\alpha = |\alpha\rangle\langle\alpha|$ is the projection operator with the properties $P_\alpha^2 = P_\alpha$, $P_\alpha^\dagger = P_\alpha$, and $P_\alpha P_{\tilde{\alpha}} = \delta_{\alpha,\tilde{\alpha}} P_\alpha$. The projection operators also satisfy the completeness relation,

$$\sum_\alpha P_\alpha^\dagger P_\alpha = \sum_\alpha P_\alpha^2 = \sum_\alpha P_\alpha = \mathbb{1}. \tag{1.13}$$

1.2.3 Quantum entropy

Von Neumann entropy

Quantum entropy is not a single concept but rather a family of notions. It starts with von Neumann entropy which is an analogue of the Shannon entropy. Von Neumann entropy, quantitatively, measures the information contained within a quantum system. Instead of being defined on probability distributions, it is defined on density matrices ρ by

$$S(\rho) = -\operatorname{tr}(\rho \log \rho). \tag{1.14}$$

To compute the von Neumann entropy, one should find the spectrum of ρ. For λ_i being the eigenvalues of ρ, it is computed by

$$S(\rho) = -\sum_i \lambda_i \log \lambda_i. \tag{1.15}$$

Some of the important properties of von Neumann entropy are given in following.
• The von Neumann entropy is a non-negative function, $S(\rho) \geqslant 0$.
• For a Hilbert space with dimension d, $S(\rho)$ takes the maximum value of $\log d$ for the maximally mixed state $\mathbb{1}/d$ and the zero value for a pure state.
• The von Neumann entropy is invariant under unitary transformation U, i.e. $S(\rho) = S(U\rho U^\dagger)$.
• The von Neumann entropy is a concave function, i.e. for non-negative real numbers α_i and density operators ρ_i the von Neumann entropy satisfies the inequality

$$\sum_i \alpha_i S(\rho_i) \leqslant S\left(\sum_i \alpha_i \rho_i\right), \tag{1.16}$$

where $\sum_i \alpha_i = 1$.
• The von Neumann entropy is additive. Given two density operators ρ and σ in different Hilbert spaces A and B, we have $S(\rho \otimes \sigma) = S(\rho) + S(\sigma)$.

Quantum relative entropy

Another type of quantum entropy is the quantum relative entropy, which is a measure of the closeness between two quantum states. For two density matrices ρ and σ, the quantum relative entropy of ρ with respect to σ is defined by

$$S(\rho \parallel \sigma) = \operatorname{tr} \rho \left(\log \rho - \log \sigma\right). \tag{1.17}$$

Some properties of relative entropy are:
• $S(\rho \parallel \sigma) \geqslant 0$, with equality if and only if $\rho = \sigma$.
• $S(\rho \parallel \sigma) \leqslant \infty$ if and only if $\operatorname{supp}(\rho) \subseteq \operatorname{supp}(\sigma)$.
• $S(\rho \parallel \sigma)$ is invariant under unitary transformations U, i.e. $S(U\rho U^\dagger \parallel U\sigma U^\dagger) = S(\rho \parallel \sigma)$
• Let $\mathcal{H}$ and $\mathcal{L}$ be two Hilbert spaces and $\mathcal{T}$ a completely positive trace preserving (CPTP) map from $\mathcal{H}$ to $\mathcal{L}$. Then, $S(\mathcal{T}(\rho) \parallel \mathcal{T}(\sigma)) \leqslant S(\rho \parallel \sigma)$. Physically speaking, the completely positive trace preserving map cannot increase the closeness of two quantum states.

1.2.4 Quantum channel

Transmitting both known and unknown quantum states from one location to another is usually an important task. It is crucial to have physical systems which serve as quantum channels. Optical fibers and an unmodulated spin chain [7] are examples of quantum channels which are suitable for long distance and short distance quantum communication, respectively.

Mathematically, a quantum channel is a completely positive trace preserving map which projects the quantum state ρ acting on the Hilbert space $\mathcal{H}_1$ to the Hilbert space $\mathcal{H}_2$. That is, for Λ to be the quantum channel, $\Lambda : \mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$. Quantum channels are classified into channels with and without memory (they are also called memory and memoryless channels). When noise acts independently

on each use of the channel we deal with memoryless channels. In comparison to channels with memory, memoryless channels have received a great deal of attention. However, since correlation between errors are common in the physical systems, channels with memory are more realistic [9; 15; 29; 37; 38]. The spin chain channel proposed by Bose [7] provides a physical example for a channel with memory. It uses an unmodulated and unmeasured spin chain to transmit quantum information. The state to be transmitted is placed by Alice on the nearest spin to her. This state acts both as an input state and also part of memory state for further uses of the channel. Afterwards, Bob who is at the opposite end of the chain receives this state with some fidelity on the nearest spin to him.

Both classical and quantum information can be carried over a quantum channel. The classical capacity theorem was independently proved by Holevo [23] and Schumacher and Westmoreland [42] while the quantum capacity theorem was originally stated by Lloyd [34]. In the asymptotic limit of many uses of the channel, the classical (or quantum) capacity measures the maximum rate of information in bits (or qubits) that can be faithfully transmitted per channel use.

The notions of channel capacity and super dense coding capacity are closely related. The slight difference is that: the channel capacity is based on sending the optimal states (including separable ones) through the channel, while, providing a shared entangled state between Alice and Bob, the super dense coding capacity is reached by the optimal local encoding on Alice's side of the shared state (more details are presented in chapter 3).

Chapter 2

Introduction

A realistic quantum system usually suffers from unwanted interactions with the outside world. It interacts with the environment via user interaction, power supply, and also more fundamental processes like thermodynamic and electromagnetic interaction with objects outside the system. Any of these unwanted effects are assumed as noise. Quantum noises such as thermal, shot, and Heisenberg uncertainty, all effect systems in various ways. Some are fundamental and unavoidable, while others are due to quantum interactions with the environment. Statistical prediction of the behavior of noise can be done by considering it as quantum operations. From here we can built a picture of how a quantum state changes when noise is introduced.

In quantum information schemes, noise does not always play a negative role. Some positive aspects of noise have been already recognized in cryptography and quantum biological systems. It is therefore necessary to understand the sources and properties of quantum noise in order to design technological devices that can avoid or possibly make use of them.

In this thesis, noise functions in both destructive and constructive directions. In chapter 3, noise shows its destroying face in the context of super dense coding, while in chapter 3, an improving aspect of quantum noise for a problem in cryptography is presented.

2.1 Super dense coding

Entanglement is a completely quantum characteristic which has given rise to a number of phenomena such as super dense coding. In the original super dense coding protocol, by making use of shared entanglement, it is possible to communicate two bits of classical information by sending one qubit only. It is possible due to the nonlocal properties of quantum entanglement.

In a realistic super dense coding scenario, noise is unavoidably present in the system. The central theme of chapter 3 is the question: How does noise in the transmission quantum channel affect the super dense coding capacity (the optimal information transfer)? In our approach, noise can influence the super dense coding protocol in two stages. Firstly, during the process of distributing the entangled state between the sender and the receiver. Secondly, when the sender (Alice) sends her part of the encoded state through the noisy channel to the receiver (Bob). The notion of *one-sided* channel then regards to the special case that noise only influences the Alice's subsystem after encoding. It is assumed no noise affects the Bob's side. The *two-sided* channel refers to the case where both Alice's and Bob's subsystems experience noisy channels. For *two-sided* channels, noise can then be either uncorrelated or correlated. A Pauli channel in arbitrary dimensions as one type of noise, and Bell states and Werner states as resources for super dense coding, are widely used in this chapter.

In chapter 3, we discuss bipartite and multipartite super dense coding scenarios, taken with the consideration of noisy channels and (non)unitary encoding. We focus on the optimization problem of the Holevo quantity in order to find the super dense coding capacity. chapter 3 is organized as follows:

Section 3.1 gives an introduction over super dense coding protocol and the Holevo bound as a key concept in finding the super dense coding capacity is discussed. An overview over different super dense coding scenarios in the presence of noiseless channels is also provided. Section 3.2 is devoted to super dense coding in the presence of uncorrelated noise and using unitary encoding. In section 3.2.1, we introduce a certain condition on the von Neumann entropy and we derive the super dense coding capacity for those cases where this condition is fulfilled. In 3.2.2 and 3.2.3, we give examples of initial states and channels

for which this condition on the von Neumann entropy is satisfied, and calculate their optimal super dense coding capacity explicitly. Section 3.2.4 provides a comparison between the super dense coding capacities in the presence of a *one-sided* or *two-sided* 2-dimensional depolarizing channel, and the classical capacity of a 2-dimensional depolarizing channel. In section 3.3, we discuss the super dense coding capacity in the presence of a *correlated* Pauli channel, considering unitary and non-unitary encoding. In 3.3.2 and 3.3.3, we give examples of correlated channels and initial states for which we can explicitly find the capacity. Section 3.4 is devoted to multipartite super dense coding. It provides a generalization of the example discussed in 3.2.2. We will consider the case of *one-sided* Pauli channels which influence Alices part of the state. We suppose that Bobs system does not experience any noise. Section 3.5 provides an conclusion to chapter 3. We show that in comparison to the super dense coding capacity with noiseless channels, noise causes some degradation in information transform.

2.2 Cryptography

Cryptography is used to transmit a message from a sender to a receiver without leaking useful information to any unauthorized party. It is a part of broader field of cryptology, which also includes cryptanalysis, the art of code breaking. Classical cryptography relies on computational complexity. It may be broken by an effective algorithm or a powerful computer. In contrast to classical cryptography, quantum cryptographical protocols provide an unconditional secure key, the security of which is inherent in the laws of quantum mechanics.

In a typical implementation in quantum cryptosystems, polarized photons are sent from Alice to Bob through an optical fiber. An eavesdropper (Eve) is usually assumed to have every possible power that is compatible with the laws of quantum mechanics. This power is not necessarily realistic, with respect to existing tools and technology. In particular, Eve is supposed to be able to replace the quantum channel (i.e. the optical fiber in our example given above), which in reality always introduces some noise, by a noiseless fiber. Even though this assumption is compatible with quantum mechanics, it is too restrictive from a realist's point of view.

Chapter 4 is concerned with an eavesdropping problem in quantum key distribution. It is assumed that Eve does not possess a noiseless fiber. The quantum states thus will experience noise in transit from Alice to Bob. We will consider white noise, where the noise parameter p describes the best existing fiber that Eve can possibly get hold of. Our results also hold for the scenario where Alice deliberately adds noise to the signal states before sending them to Bob, and the quantum channel is noiseless. We focus in particular on the six state protocol, for which signals are mixed with white noise. In this scenario, Alice sends the signal states in three bases with equal probability through the channel where eavesdropper is assumed to have access to it. Eve's strategy is base to attack the signal states individually. She attaches to each signal an ancilla state and apply a unitary operator. She then attacks the signal states in a way that introduces the same quantum bit error rate for different directions. That is because a basis-dependent quantum bit error rate indicates the presence of an eavesdropper, which Alice and Bob can easily test it.

The aim of chapter 4 is to present an intuitive understanding for the counter-intuitive fact that *additional* noise on the quantum level may help the trusted parties to improve the performance of a six state protocol. Chapter 4 is structured as follows:

Section 4.1 provides an introduction on quantum key distribution protocols and also two security theorems. In section 4.2, we discuss the eavesdropping on noisy signals and we calculate the mutual information between Alice & Eve. We will then in section 4.3 and section 4.4 derive the optimal mutual information that Eve can obtain, when using individual attacks on noisy quantum signals, and compare it to the mutual information achievable by eavesdropping on pure states. We also obtain the mutual information between Alice & Bob. One expects that Alice and Bob, but also Eve, will lose some information, due to the additional noise. It is not evident, however, how the relation between the two mutual information curves (Alice & Bob versus Alice & Eve) changes, when the noise increases. By using Csiszár and Körner theorem we then in section 4.5, discuss the security proof.

Chapter 3

Super dense coding in the presence of noise

3.1 Introduction

Super dense coding as a communication protocol was introduced in 1992 by Bennet and Wiesner [6]. It is one of the notable areas in which quantum entanglement plays an essential role. Crucial to this communication protocol is an entangled initial state that is shared between sender(s) and receiver(s), together with the property that an entangled state can be transformed by the sender into another state via a *local* operation, taken from some set of operations. The protocol of super dense coding works as follows. Alice and Bob share a maximally entangled pure state of two spin-1/2 particles of the form $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|0_A 0_B\rangle + |1_A 1_B\rangle)$ where, $0_{A,B}$ and $1_{A,B}$ denote the spin-up and spin-down states, respectively. Alice has qubit A and Bob has qubit B. For convenience, we drop labels A and B. Alice locally applies one of the four unitary operations σ_0, σ_1, σ_2 and σ_3 to her single qubit, where σ_0 is the identity operator and the other three are the Pauli operators

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.1}$$

Depending on Alice's choice of transformation (up to some phases) the original state $|\phi^+\rangle$ transforms into one of the four mutually orthogonal states $|\phi^\pm\rangle$ and

$|\psi^{\pm}\rangle$

$$(\sigma_0 \otimes \mathbb{1})|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \equiv |\phi^+\rangle, \tag{3.2}$$

$$(\sigma_1 \otimes \mathbb{1})|\phi^+\rangle = \frac{1}{\sqrt{2}}(|10\rangle + |01\rangle) \equiv |\psi^+\rangle, \tag{3.3}$$

$$(i\sigma_2 \otimes \mathbb{1})|\phi^+\rangle = \frac{1}{\sqrt{2}}(|10\rangle - |01\rangle) \equiv |\psi^-\rangle, \tag{3.4}$$

$$(\sigma_3 \otimes \mathbb{1})|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle) \equiv |\phi^-\rangle. \tag{3.5}$$

These four states are the so-called Bell states. Subsequent to the unitary rotation, Alice sends her single qubit to Bob, who is now in possession of the total state. Since Bell states are orthogonal, they all can perfectly be distinguished by Bob via a suitable von Neumman measurement. If each of these states represents one piece of information, Alice has managed to send one out of four messages (i.e. two bits of classical information) by transmitting only a single qubit. This is impossible without entanglement since the amount of classical information conveyed by an isolated qubit cannot exceed one bit. These two classical bits of information, as we precisely define later, is the super dense coding capacity. For the first time the increased capacity of a quantum information channel by super dense coding was experimentally verified in [39]. Until recently, experiments have only been able to distinguish three out four of these states [39]. However, more complicated schemes that make use of so-called hyper-entanglement can now distinguish all four states [1].

The super dense coding capacity is defined to be the maximal amount of classical information that can be reliably transmitted to the receiver for a given entangled state. A crucial element in finding the super dense coding capacity is the Holevo bound [22], which is a universal upper bound on classical information that can be decoded from a quantum ensemble. Below, for noiseless and noisy channels, we discuss the bound, and subsequently define the super dense coding capacity for both cases.

3.1.1 Holevo bound

A theorem stated by Gordon [20] and Levitin [32] and proved by Holevo [22] states that the amount of accessible classical information (I_{acc}) contained in an ensemble $\{\rho_i, p_i\}$ is upper bounded by the so-called Holevo quantity $\chi(\{\rho_i, p_i\})$. This upper bound is regardless to the measurement that can be performed on the system, and is given by

$$I_{acc} \leq \chi(\{\rho_i, p_i\}) \equiv S(\overline{\rho}) - \sum_i p_i S(\rho_i), \tag{3.6}$$

where $\overline{\rho} = \sum_i p_i \rho_i$ is the average ensemble state and $S(\eta) = -\operatorname{tr}(\eta \log \eta)$ is the von Neumann entropy of η. From the concavity of the von Neumann entropy $S(\rho)$ it follows that, the Holevo quantity is non-negative. In the limit of sending long strings of the input quantum states ρ_i the Holevo bound is reachable [23; 42].

3.1.2 Super dense coding capacity

The first attentions, after proposing the super dense coding protocol, were given to various scenarios over noiseless channels and unitary encoding [11; 12; 21]. For ρ to be a shared mixed state between the sender and the receiver, Alice performs a local unitary operation W_i on ρ to encode classical information through the state ρ_i

$$\rho_i = (W_i \otimes \mathbb{1})\rho(W_i^\dagger \otimes \mathbb{1}). \tag{3.7}$$

Subsequently, she sends her subsystem of dimension d to Bob (ideally via a noiseless channel) with a probability p_i. The ensemble that Bob receives is $\{\rho_i, p_i\}$. As we mentioned in the previous section, the amount of classical information transmitted via a quantum channel in this process is measured by the Holevo quantity (3.6). Then for a given shared state ρ, the super dense coding *capacity* has been defined to be the maximal amount of the Holevo quantity $\chi\{\rho_i, p_i\}$ with respect to the probability p_i and the unitary operator W_i

$$C = \max_{\{p_i, W_i\}} (\chi\{\rho_i, p_i\}) \equiv \max_{\{p_i, W_i\}} \left(S(\overline{\rho}) - \sum_i p_i S(\rho_i) \right). \tag{3.8}$$

Several lines of argument are possible tackle the optimization procedure [21; 53]. For instance, the approach given in [53] has introduced an upper bound on the Holevo quantity (3.8)

$$\chi(\{\rho_i, p_i\}) \leqslant \log d + S(\rho_b) - S(\rho), \tag{3.9}$$

where ρ is the initial resource state shared between Alice and Bob. Here, d is the dimension of Alice's subsystem, ρ_b is Bob's reduced density operator with $\rho_b = \mathrm{tr}_a\, \rho$ and $S(\rho) = -\mathrm{tr}(\rho \log \rho)$ is the von Neumann entropy. The Holevo quantity $\chi\{\rho_i, p_i\}$ is a function of the resource state ρ and the encoding $\{W_i, p_i\}$. It has been proven that the upper bound (3.9) can be achieved by any complete set of orthogonal unitary operators $\{V_i\}$ which satisfy $\frac{1}{d}\sum_{i=1}^{d^2} V_i \Xi V_i^\dagger = \mathrm{tr}[\Xi]\mathbb{1}$ for any operator Ξ, while the probability of choosing the unitary operators are equal, i.e. $p_i = \frac{1}{d^2}$. Then the super dense coding capacity for noiseless channels and unitary encoding, is given by

$$C = \log d + S(\rho_b) - S(\rho)\ . \tag{3.10}$$

Without the additional resource of entangled states, a d-dimensional quantum state can be used to transmit the information $\log d$. Hence, quantum states for which $S(\rho_b) - S(\rho) > 0$, i.e. those which are more mixed locally than globally, are the useful states for super dense coding and are the so called *dense-codeable* states. For instance, any pure entangled bipartite state satisfies this inequality and therefore, is useful for super dense coding while separable states always violate it [28]. There exist a class of states which are entangled but not *distillable*, i.e. using many copies of the state, it is impossible to obtain a maximally entangled state from them by local operation and classical communication (LOCC). These states are the so called *bound entangled* states [24]. Any bipartite bound entangled state also violates $S(\rho_b) - S(\rho) > 0$ [11; 26] and therefore, like separable states are useless for super dense coding. In general, relation $S(\rho_b) - S(\rho) > 0$ cannot hold for quantum states with positive partial transpose [11]. Therefore, states that are useful for super dense coding always have a non-positive partial transpose (NPT). However, the converse is not true: There exist states which are NPT, but which are not useful for super dense coding. One can also classify bipartite states according to their usefulness for super dense coding [12]. A simple

classification of bipartite quantum states according to their *dense-codeability* is coming in the following (see also Figure 3.1).
1-Separable state: These states are not useful for super dense coding.
2-PPT entangled states: These states are entangled with positive partial transpose but cannot be used for super dense coding. The entanglement of these states cannot be detected by the partial transposition criterion.
3-NPT non-DC states: These states are entangled with non-positive partial transpose but still not useful for super dense coding. However, their entanglement can be detected by the partial transposition criterion.
4- NPT DC states. These states are entangled with non-positive partial transpose and they are useful for super dense coding.
A generalization of this classification for the multipartite quantum states has been considered in [11; 12]. Besides the case of a single sender and receiver sharing an initial entangled state and using unitary encoding some other scenarios also have been discussed: many senders and either one or two receivers, non-unitary encoding, continuous variables, etc. [12; 27; 33].

3.1.3 Holevo quantity in the presence of noise

As we defined before, a quantum channel is a communication channel which can transmit quantum information. Physically, a noisy quantum channel is a process that arises through interaction with the environment. Mathematically, a noisy quantum channel can be described as a completely positive trace preserving (CPTP) map acting on the quantum state. We consider $\Lambda : \rho_i \to \Lambda(\rho_i)$ to be a CPTP map that acts on the encoded state ρ_i (3.7). For $\{\Lambda(p_i, \rho_i)\}$ to be the ensemble that Bob receives, the Holevo quantity is given by

$$\chi\{\Lambda(\rho_i, p_i)\} = S\left(\overline{\Lambda(\rho)}\right) - \sum_i p_i S\left(\Lambda(\rho_i)\right) = \sum_i p_i S\left(\Lambda(\rho_i) \| \overline{\Lambda(\rho)}\right), \quad (3.11)$$

where $\overline{\Lambda(\rho)} = \sum_i p_i \Lambda(\rho_i)$ is the average state and $S(\rho \parallel \sigma) = \operatorname{tr} \rho (\log \rho - \log \sigma)$ is the relative entropy (1.17). Note that $\chi\{\Lambda(\rho_i, p_i)\}$ is a function of the resource state ρ, the encoding $\{W_i, p_i\}$ and the channel Λ. For brevity of notation we will

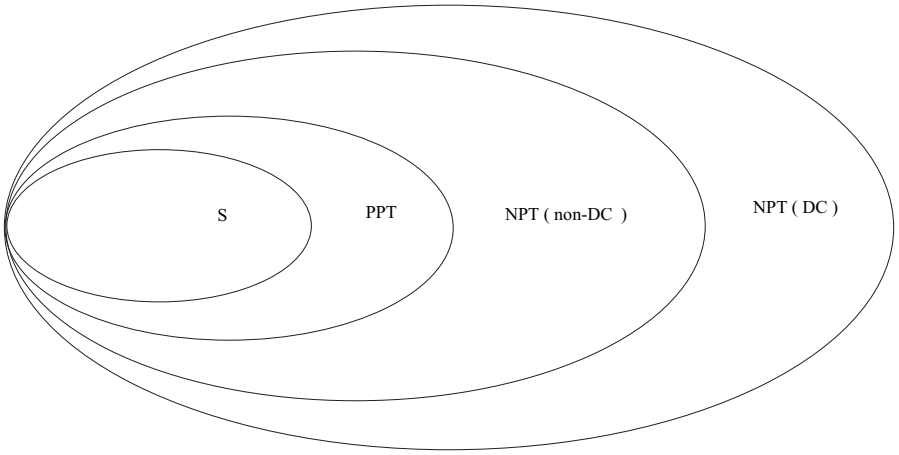

Figure 3.1: Classification of bipartite quantum states, according to their usefulness for super dense coding [12]. S stands for separable states. PPT, NPT (non-DC) are the set of PPT entangled states and NPT entangled states which, despite of being entangled, cannot be useful for super dense coding. NPT (DC) are the set of NPT entangled states that can be used for super dense coding.

not write explicitly these arguments of χ. Likewise to the noiseless channel, the super dense coding *capacity* C for a given resource state ρ and the noisy channel Λ is defined to be the maximum of the Holevo quantity $\chi\{\Lambda(p_i, \rho_i)\}$ with respect to $\{W_i, p_i\}$

$$C = \max_{\{W_i, p_i\}} (\chi\{\Lambda(p_i, \rho_i)\}) \equiv \max_{\{W_i, p_i\}} \left(S\left(\overline{\Lambda(\rho)}\right) - \sum_i p_i S\left(\Lambda(\rho_i)\right) \right). \qquad (3.12)$$

The rest of the present chapter discusses different super dense coding scenarios, taken with the consideration of noisy channels. For these scenarios, we focus on the optimization problem of the Holevo quantity in order to find the super dense coding capacity.

3.2 Super dense coding with uncorrelated noise

A majority of research related to quantum channels has focused on *memoryless* channels. As we defined in preliminary chapter, a channel is called memoryless if the output of the channel just depends on the corresponding input and not on any previous inputs. Uncorrelated noise is a model of such a channel. In this section, we focus on uncorrelated noise, the case of a single sender and a single receiver, and assumes unitary encoding. We study two different scenarios of noisy channels. Firstly, we will assume that the sender Alice and the receiver Bob share already a bipartite quantum state ρ (it could for instance have been distributed to them by a third party). After Alice's local encoding operation, she sends her part of the quantum state to Bob via the noisy channel, described by the map Λ_a, see Figure 3.2. We call this the case of a *one-sided* channel. Secondly, we consider the case where Alice prepares the bipartite state ρ and sends one part of it via a noisy channel, described by the map Λ_b, to Bob, thus establishing the shared resource state for super dense coding. When the two parties want to use this resource, Alice does the local encoding and then sends her part of the state via the channel Λ_a to Bob, see Figure 3.3. We call this case a *two-sided* channel.

3.2.1 Optimal super dense coding capacity

In order to obtain the capacity in the presence of uncorrelated noise, we consider bipartite systems, where each subsystem has finite dimension d. A general density matrix on $C^d \otimes C^d$ in the Hilbert-Schmidt representation can be conveniently decomposed as

$$\begin{aligned} \rho &= \frac{1}{d^2}\left(\mathbb{1}\otimes\mathbb{1} + \sum_{j=1}^{d^2-1} r_j\lambda_j\otimes\mathbb{1} + \mathbb{1}\otimes\sum_{j=1}^{d^2-1} s_j\lambda_j + \sum_{j,k=1}^{d^2-1} t_{jk}\lambda_j\otimes\lambda_k\right) \\ &= \mathbb{1}\otimes\frac{\rho_b}{d} + \frac{1}{d^2}\left(\sum_{j=1}^{d^2-1} r_j\lambda_j\otimes\mathbb{1} + \sum_{j,k=1}^{d^2-1} t_{jk}\lambda_j\otimes\lambda_k\right), \end{aligned} \tag{3.13}$$

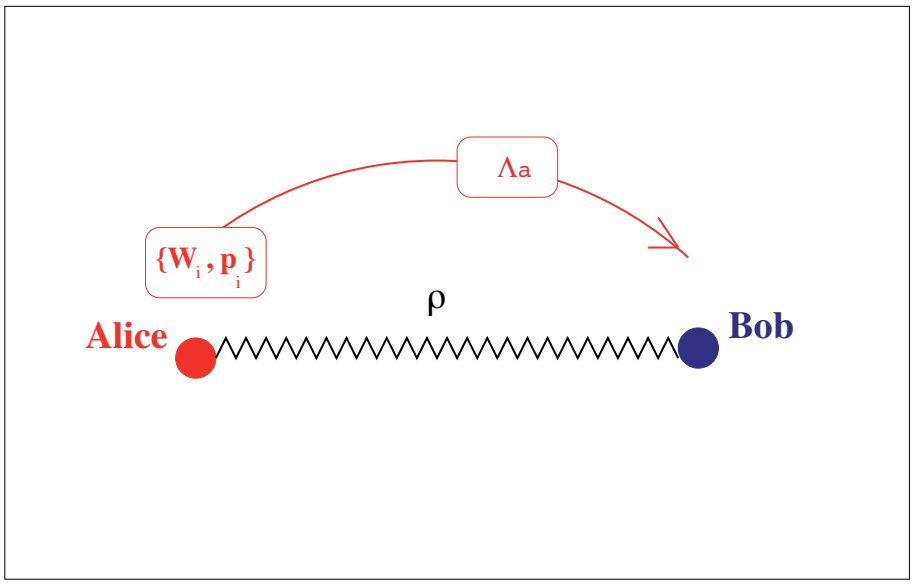

Figure 3.2: *One-sided* noise: Bipartite super dense coding with an initially entangled state ρ, shared between Alice and Bob. Alice applies the unitary operator W_i, taken from a set $\{W_i\}$ with probability $\{p_i\}$, on her part of the entangled state ρ. She sends the encoded state with probability p_i over a noisy channel, described by the map Λ_a, to Bob. In the first approach we assume that Λ_a just affects Alice's subsystem, but that there is no noise on Bob's side.

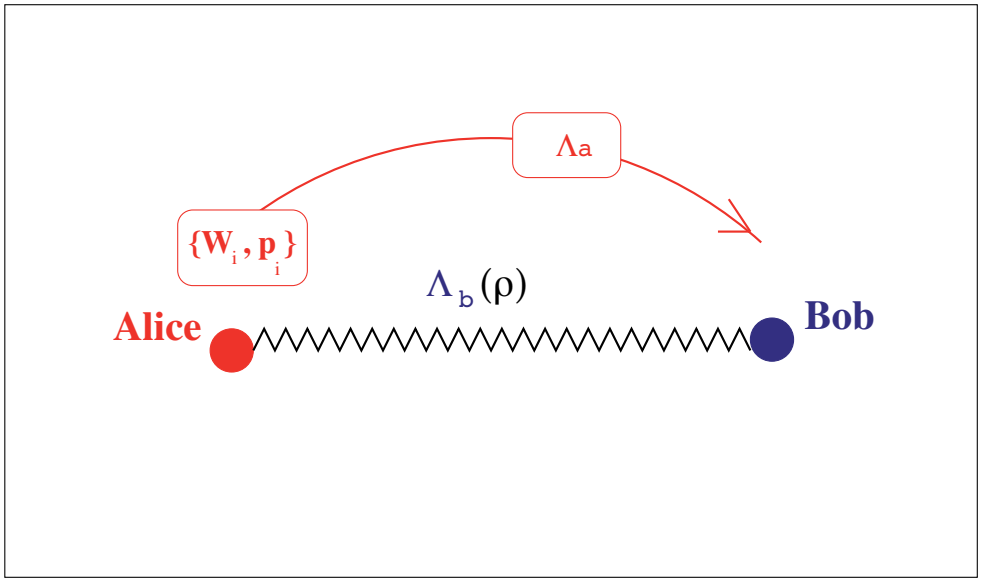

Figure 3.3: *Two-sided* noise: Bipartite super dense coding with an initially entangled state ρ, shared between Alice and Bob. In the second approach, the noisy channel Λ_a influences Alice's subsystem after encoding while the noisy channel Λ_b has already affected Bob's side in the distribution step of the initial state ρ.

where $\rho_b = \mathrm{tr}_a\,\rho$ represents Bob's reduced density operator and λ_j are the generators of the SU(d) algebra with $\mathrm{tr}\,\lambda_j = 0$. The parameters r_j, s_j, t_{jk} are real coefficients. We introduce the set of unitary operators $\{V_i\}$, defined as

$$V_{i=(m,n)}|j\rangle = e^{(\frac{2\pi i n j}{d})}|j + m(\mathrm{mod}\,d)\rangle. \tag{3.14}$$

These operators satisfy the condition $d^{-1}\,\mathrm{tr}(V_i V_j^\dagger) = \delta_{ij}$. Integers m and n run from 0 to $d-1$ such that we have d^2 unitary operators V_i. We will also consider in the following the case of unital noisy channels acting on Alice's and Bob's subsystems, namely channels described by the completely positive trace preserving map

$$\Lambda(\rho) = \sum_m K_m \rho K_m^\dagger\ , \quad \sum_m K_m^\dagger K_m = \mathbb{1}\ , \quad \sum_m K_m K_m^\dagger = \mathbb{1}\ , \tag{3.15}$$

where K_m are Kraus operators. Here, the first condition on the Kraus operators corresponds to trace preservation, and the second condition guarantees the unital property $\Lambda(\mathbb{1}) = \mathbb{1}$. We will show that for unital memoryless noisy quantum channels and certain initial resource states, the set of unitary operators $\{V_i\}$ in equation (3.14) with equal probabilities is the optimum encoding and leads to the maximum of the Holevo quantity.

We will first prove in *Lemma 1* some properties that hold for the specific encoding $\{V_i\}$. In the following the symbol τ_i will denote the resource state after encoding with V_i, whereas τ will denote the resource state after encoding with an arbitrary unitary operation U. The ensemble average after the specific encoding with $\{V_i\}$, the probability distribution $p_i = 1/d^2$ and after action of the channel will be denoted as $\tilde{\rho}$.

Lemma 1. Let $\Lambda_a(\rho_1) = \sum_m A_m \rho_1 A_m^\dagger$ and $\Lambda_b(\rho_2) = \sum_{\tilde{m}} B_{\tilde{m}} \rho_2 B_{\tilde{m}}^\dagger$ be any two unital channels which act on Alice's and Bob's side, respectively. For an initial resource state ρ shared between Alice and Bob, the global channel Λ_{ab} then acts as

$$\Lambda_{ab}(\rho) = \sum_{m,\tilde{m}} (A_m \otimes B_{\tilde{m}})\,\rho\left(A_m^\dagger \otimes B_{\tilde{m}}^\dagger\right). \tag{3.16}$$

Then the following statements hold:

1-a) For $\tau_i = (V_i \otimes \mathbb{1})\rho(V_i^\dagger \otimes \mathbb{1})$, with V_i being defined in (3.14), the average $\tilde{\rho}$ of

the ensemble $\{p_i = \frac{1}{d^2}, \Lambda_{ab}(\tau_i)\}$ takes the form $\tilde{\rho} = \mathbb{1} \otimes \Lambda_b(\frac{\rho_b}{d})$.
1-b) For $\tau = (U \otimes \mathbb{1})\, \rho \left(U^\dagger \otimes \mathbb{1}\right)$ with U being any unitary operator acting on Alice's system, $\mathrm{tr}\left(\Lambda_{ab}(\tau) \log \tilde{\rho}\right) = -S(\tilde{\rho})$.
1-c) The relative entropy between $\Lambda_{ab}(\tau)$ and $\tilde{\rho}$ can be expressed as $S\left(\Lambda_{ab}(\tau) \| \tilde{\rho}\right) = S\left(\tilde{\rho}\right) - S\left(\Lambda_{ab}(\tau)\right)$.

Proof 1-a): In [21] it was shown that the average of the ensemble $\{p_i = \frac{1}{d^2}, \tau_i\}$ is

$$\begin{aligned}\sum_i \frac{1}{d^2}\tau_i &= \frac{1}{d^2}\sum_i (V_i \otimes \mathbb{1})\rho(V_i^\dagger \otimes \mathbb{1}) \\ &= \frac{1}{d^2}\sum_i (V_i \otimes \mathbb{1}) \left[\mathbb{1} \otimes \frac{\rho_b}{d} + \frac{1}{d^2}\left(\sum_{j=1}^{d^2-1} r_j \lambda_j \otimes \mathbb{1} + \sum_{j,k=1}^{d^2-1} t_{jk}\lambda_j \otimes \lambda_k\right)\right](V_i^\dagger \otimes \mathbb{1}).\end{aligned}$$

By using $\frac{1}{d}\sum_i (V_i \lambda_j V_i^\dagger) = \mathrm{tr}[\lambda_j]\mathbb{1}$ and the Hilbert-Schmidt decomposition for ρ (3.13) we get

$$\begin{aligned}&= \mathbb{1} \otimes \frac{\rho_b}{d} + \frac{1}{d^2}\sum_{j=1}^{d^2-1} r_j \sum_i (V_i \lambda_j V_i^\dagger) \otimes \mathbb{1} + \frac{1}{d^2}\sum_{j,k=1}^{d^2-1} t_{jk} \sum_i (V_i \lambda_j V_i^\dagger) \otimes \lambda_k \\ &= \mathbb{1} \otimes \frac{\rho_b}{d} + \frac{1}{d}\sum_{j=1}^{d^2-1} r_j \underbrace{\mathrm{tr}[\lambda_j]}_{0}(\mathbb{1} \otimes \mathbb{1}) + \frac{1}{d}\sum_{j,k=1}^{d^2-1} t_{j,k} \underbrace{\mathrm{tr}[\lambda_j]}_{0}(\mathbb{1} \otimes \lambda_k) \\ &= \mathbb{1} \otimes \frac{\rho_b}{d} \qquad (3.17)\end{aligned}$$

Then the average of the ensemble $\{p_i = \frac{1}{d^2}, \Lambda_{ab}(\tau_i)\}_{i=0}^{d^2-1}$ by using (3.17), the linearity of the channel and its unital property is

$$\begin{aligned}\tilde{\rho} &= \sum_i \frac{1}{d^2}\Lambda_{ab}(\tau_i) = \Lambda_{ab}(\sum_i \frac{1}{d^2}\tau^i) = \Lambda_{ab}(\mathbb{1} \otimes \frac{\rho_b}{d}) \\ &= \sum_{m,\tilde{m}} \left(A_m \otimes B_{\tilde{m}}\right)\left(\mathbb{1} \otimes \frac{\rho_b}{d}\right)\left(A_m^\dagger \otimes B_{\tilde{m}}^\dagger\right) \\ &= \sum_m A_m A_m^\dagger \otimes \sum_{\tilde{m}} B_{\tilde{m}} \frac{\rho_b}{d} B_{\tilde{m}}^\dagger \\ &= \mathbb{1} \otimes \Lambda_b(\frac{\rho_b}{d}). \qquad (3.18)\end{aligned}$$

□

Proof 1-b): In Lemma (1-a) we showed that $\tilde{\rho} = \mathbb{1} \otimes \Lambda_b(\frac{\rho_b}{d})$ and hence, $\log \tilde{\rho} =$

$\mathbb{1} \otimes \log \Lambda_b(\frac{\rho_b}{d})$. Therefore:

$$\begin{aligned} \mathrm{tr}\left(\Lambda_{ab}(\tau)\log\tilde{\rho}\right) &= \mathrm{tr}\left[\left(\sum_{m,\tilde{m}}(A_mU\otimes B_{\tilde{m}})\rho(U^\dagger A_m^\dagger\otimes B_{\tilde{m}}^\dagger)\right)\left(\mathbb{1}\otimes\log\Lambda_b(\frac{\rho_b}{d})\right)\right] \\ &= \mathrm{tr}\left[\sum_{m,\tilde{m}}(A_mU\otimes B_{\tilde{m}})\rho\left(U^\dagger A_m^\dagger\otimes B_{\tilde{m}}^\dagger\log\Lambda_b(\frac{\rho_b}{d})\right)\right]. \end{aligned}$$

By using the decomposition (3.13) for ρ we have

$$\begin{aligned} \mathrm{tr}\left(\Lambda_{ab}(\tau)\log\tilde{\rho}\right) &= \mathrm{tr}\left[\left(\sum_m A_mUU^\dagger A_m^\dagger\right)\otimes\left(\sum_{\tilde{m}}B_{\tilde{m}}\frac{\rho_b}{d}B_{\tilde{m}}^\dagger\log\Lambda_b(\frac{\rho_b}{d})\right)\right. \\ &+\frac{1}{d^2}\left(\sum_{j=1}^{d^2-1}r_j\sum_m A_mU\lambda_jU^\dagger A_m^\dagger\right)\otimes\left(\sum_{\tilde{m}}B_{\tilde{m}}B_{\tilde{m}}^\dagger\log\Lambda_b(\frac{\rho_b}{d})\right) \\ &\left.+\frac{1}{d^2}\sum_{j,k=1}^{d^2-1}t_{jk}\left(\sum_m A_mU\lambda_jU^\dagger A_m^\dagger\right)\otimes\left(\sum_{\tilde{m}}B_{\tilde{m}}\lambda_kB_{\tilde{m}}^\dagger\log\Lambda_b(\frac{\rho_b}{d})\right)\right]. \quad (3.19) \end{aligned}$$

By using linearity of the trace and the relations

$$\mathrm{tr}[\sum_m A_mUU^\dagger A_m^\dagger] = \mathrm{tr}[\sum_m A_mA_m^\dagger] = \mathrm{tr}[\mathbb{1}]\,, \quad (3.20a)$$

$$\mathrm{tr}[\sum_m A_mU\lambda_jU^\dagger A_m^\dagger] = \mathrm{tr}[U\lambda_jU^\dagger\sum_m A_m^\dagger A_m] = \mathrm{tr}[U\lambda_jU^\dagger] = 0, \quad (3.20b)$$

we can write

$$\begin{aligned} &\mathrm{tr}\left(\Lambda_{ab}(\tau)\log\tilde{\rho}\right) \\ &= \mathrm{tr}_a\,\mathrm{tr}_b\left[\sum_{m,\tilde{m}}\mathbb{1}\otimes\left(B_{\tilde{m}}\frac{\rho_b}{d}B_{\tilde{m}}^\dagger\log\Lambda_b(\frac{\rho_b}{d})\right)\right] \\ &= \mathrm{tr}_b\left[\left(\sum_{\tilde{m}}B_{\tilde{m}}\rho_bB_{\tilde{m}}^\dagger\right)\log\Lambda_b(\frac{\rho_b}{d})\right] \\ &= \mathrm{tr}_b\left[\Lambda_b(\rho_b)\log\Lambda_b(\frac{\rho_b}{d})\right] = -S(\tilde{\rho}). \quad (3.21) \end{aligned}$$

□

Proof 1-c): Using the definition of the relative entropy $S(\sigma\|\rho) = \mathrm{tr}(\sigma\log\sigma -$

$\sigma \log \rho$) and the result of Lemma (1-b) we can write

$$\begin{aligned} S(\Lambda_{ab}(\tau)\|\tilde{\rho}) &= \mathrm{tr}(\Lambda_{ab}(\tau)\log\Lambda_{ab}(\tau) - \Lambda_{ab}(\tau)\log\tilde{\rho}) \\ &= \mathrm{tr}(\Lambda_{ab}(\tau)\log\Lambda_{ab}(\tau)) - \mathrm{tr}(\Lambda_{ab}(\tau)\log\tilde{\rho}) \\ &= S(\tilde{\rho}) - S(\Lambda_{ab}(\tau)). \end{aligned} \tag{3.22}$$

□

We now show that for resource states with a certain symmetry property, namely for those states where the von Neumann entropy after the channel action is independent of the unitary encoding, the encoding with the equally probable operators $\{V_i\}$, as given in (3.14), is optimal. Our proof follows the line of argument developed in [21].

Lemma 2. Let τ_i denote the resource state after encoding with V_i, given in (3.14). Let

$$\tilde{\chi} = S(\tilde{\rho}) - \frac{1}{d^2}\sum_{i}^{d^2-1} S(\Lambda_{ab}(\tau_i)) \tag{3.23}$$

be the Holevo quantity for the ensemble $\{p_i = \frac{1}{d^2}, \Lambda_{ab}(\tau_i)\}$, where $\tilde{\rho}$ is the average state of this ensemble and $\Lambda_{ab}(\cdot)$ is defined in (3.16). For all the channels Λ_{ab} and all initial states ρ for which

$$S(\Lambda_{ab}(\tau)) = \frac{1}{d^2}\sum_{i}^{d^2-1} S(\Lambda_{ab}(\tau_i)) \tag{3.24}$$

holds, $\tilde{\chi}$ is the super dense coding capacity. Here $\tau = (U \otimes \mathbb{1})\,\rho\,\left(U^\dagger \otimes \mathbb{1}\right)$, as we defined already above, with U being any unitary operator.

Proof: Let us consider an arbitrary encoding, leading to an ensemble $\{p_i, \Lambda_{ab}(\rho_i)\}$. We will show that its Holevo quantity χ cannot be higher than $\tilde{\chi}$ in (3.23), if the condition (3.24) is fulfilled.

If $S(\Lambda_{ab}(\tau)) = \frac{1}{d^2}\sum_i S(\Lambda_{ab}(\tau^i))$, then from (3.23) and Lemma (1-c),

$$\begin{aligned} \tilde{\chi} &= S(\tilde{\rho}) - S(\Lambda_{ab}(\tau)) \\ &= S(\Lambda_{ab}(\tau)\|\tilde{\rho}). \end{aligned} \tag{3.25}$$

Since this equation holds for any τ that fulfills (3.24), it specially holds for arbitrary encoding ρ_i, i.e.

$$\tilde{\chi} = S(\Lambda_{ab}(\rho_i)\|\tilde{\rho}) = \sum_i p_i S(\Lambda_{ab}(\rho_i)\|\tilde{\rho}). \tag{3.26}$$

Using Donald's identity, see [17], the right hand side of the above equation can be decomposed as

$$\sum_i p_i S(\Lambda_{ab}(\rho_i)\|\tilde{\rho}) = \sum_i p_i S(\Lambda_{ab}(\rho_i)\|\overline{\Lambda_{ab}(\rho)}) + S(\overline{\Lambda_{ab}(\rho)}\|\tilde{\rho}) \tag{3.27}$$

with $\overline{\Lambda_{ab}(\rho)} = \sum_i p_i \Lambda_{ab}(\rho_i)$. The first term on the right hand side is the Holevo quantity for any arbitrary ensemble $\{p_i, \Lambda_{ab}(\rho_i)\}$. Hence,

$$\tilde{\chi} = \chi + S(\overline{\Lambda_{ab}(\rho)}\|\tilde{\rho}). \tag{3.28}$$

Since the relative entropy $S(\overline{\Lambda_{ab}(\rho)}\|\tilde{\rho})$ is always positive or zero we can say that $\tilde{\chi}$ is always bigger or equal than χ and hence, $\tilde{\chi}$ is the super dense coding capacity. □

From Lemma 2 we find that

$$\tilde{\chi} = S(\tilde{\rho}) - S(\Lambda_{ab}(\tau)). \tag{3.29}$$

Since the above equation holds for $\tau = (U \otimes \mathbb{1})\,\rho\,\left(U^\dagger \otimes \mathbb{1}\right)$ with any unitary U, it especially holds for $\tau = \rho$. Hence, whenever the condition (3.24) is true, the super dense coding capacity is given by

$$C = \tilde{\chi} = S(\tilde{\rho}) - S(\Lambda_{ab}(\rho)), \tag{3.30}$$

where $\tilde{\rho}$ is the average of the ensemble after encoding with the specific (and equally probable) unitaries $\{V_i\}$ and after the channel action, as introduced in Lemma 1. As an interpretation of this formula, note that the action of a noisy channel typically will increase the entropy of a given state, and therefore will decrease the super dense coding capacity of the original resource state.

In the next two sub-sections we will study examples of channels and bipartite states satisfying the condition (3.24), and evaluate explicitly the corresponding super dense coding capacities.

3.2.2 One-sided d-dimensional Pauli channel

A d-dimensional Pauli channel [18] that acts just on Alice's side is defined by

$$\Lambda_a^P(\rho_i) = \sum_{m,n=0}^{d-1} q_{mn}(V_{mn} \otimes \mathbb{1})\rho_i(V_{mn}^\dagger \otimes \mathbb{1}) \,, \tag{3.31}$$

where q_{mn} are probabilities (i.e. $q_{mn} \geq 0$ and $\sum_{mn} q_{mn} = 1$). The operators V_{mn}, defined in (3.14) with a slightly different notation for the indices, can be expressed as

$$V_{mn} = \sum_{k=0}^{d-1} \exp\left(\frac{2i\pi kn}{d}\right) |k\rangle\langle k + m(\mathrm{mod}\, d)| \,. \tag{3.32}$$

They satisfy $\mathrm{tr}\, V_{mn} = d\delta_{m0}\delta_{n0}$ and $V_{mn}V_{mn}^\dagger = \mathbb{1}$, and have the properties

$$V_{mn}V_{\tilde{m}\tilde{n}} = \exp\left(\frac{2i\pi\tilde{n}m}{d}\right) V_{m+\tilde{m}(mod\, d),n+\tilde{n}(mod\, d)}, \tag{3.33}$$

$$\mathrm{tr}[V_{mn}V_{\tilde{m}\tilde{n}}^\dagger] = d\delta_{m\tilde{m}}\delta_{n\tilde{n}}, \tag{3.34}$$

$$V_{mn}V_{\tilde{m}\tilde{n}} = \exp\left(\frac{2i\pi(\tilde{n}m - n\tilde{m})}{d}\right) V_{\tilde{m}\tilde{n}}V_{mn}. \tag{3.35}$$

As the Kraus operators of *one-sided* Pauli channel (3.31) are unitary it is a unital channel.

Bell states

A Bell state in $d \times d$ dimensions is defined as

$$|\Phi_{00}\rangle = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} |j\rangle \otimes |j\rangle. \tag{3.36}$$

The set of all maximally entangled Bell states is then obtained by $|\Phi_{mn}\rangle = (V_{mn} \otimes \mathbb{1})|\Phi_{00}\rangle$, for $m, n = 0, 1, ..., d-1$. For $d = 2$, we use the notation $|\Phi^+\rangle \equiv |\Phi_{00}\rangle$. We will show that for a Bell state shared between Alice and Bob, and with a *one-sided* d-dimensional Pauli channel, the condition (3.24) is fulfilled. To do so, we will first prove the following Lemma.

Lemma 3. Let us define $\pi_{mn} := (V_{mn}U \otimes \mathbb{1})\rho_{00}(U^\dagger V_{mn}^\dagger \otimes \mathbb{1})$, where U is a unitary operator, $\rho_{00} = |\Phi_{00}\rangle\langle\Phi_{00}|$ and V_{mn} is defined in (3.32). For $m \neq \tilde{m}$ or $n \neq \tilde{n}$,

$$\pi_{mn}\pi_{\tilde{m}\tilde{n}} = 0 \tag{3.37}$$

holds.

Proof:

First, we prove that $\rho_{00}(U^\dagger V_{mn}^\dagger V_{\tilde{m}\tilde{n}}U \otimes \mathbb{1})\rho_{00} = 0$. We start with proving $\langle\Phi_{00}|(U^\dagger V_{mn}^\dagger V_{\tilde{m}\tilde{n}}U \otimes \mathbb{1})|\Phi_{00}\rangle = 0$, from which the previous statement follows. Due to (3.34) for $m \neq \tilde{m}$ or $n \neq \tilde{n}$ the expression $V_{mn}^\dagger V_{\tilde{m}\tilde{n}}$ is traceless and $\{V_{jk}\}$ form a complete set. We can thus expand $V_{mn}^\dagger V_{\tilde{m}\tilde{n}} = \sum_{(j,k)\neq(0,0)} \beta_{jk}V_{jk}$ with expansion coefficients β_{jk}. Therefore,

$$\begin{aligned}
&\langle\Phi_{00}|(U^\dagger V_{mn}^\dagger V_{\tilde{m}\tilde{n}}U \otimes \mathbb{1})|\Phi_{00}\rangle \\
&= \sum_{(j,k)\neq(0,0)} \beta_{jk}\langle\Phi_{00}|(U^\dagger V_{jk}U \otimes \mathbb{1})|\Phi_{00}\rangle \\
&= \frac{1}{d}\sum_{(j,k)\neq(0,0)}\sum_{m,n=0}^{d-1} \beta_{jk}\langle mm|(U^\dagger V_{jk}U \otimes \mathbb{1})|nn\rangle \\
&= \frac{1}{d}\sum_{(j,k)\neq(0,0)}\sum_{m,n=0}^{d-1} \beta_{jk}\langle m|U^\dagger V_{jk}U|n\rangle \underbrace{\langle m|n\rangle}_{\delta_{mn}} \\
&= \frac{1}{d}\sum_{(j,k)\neq(0,0)} \beta_{jk}\,\mathrm{tr}[U^\dagger V_{jk}U] = \frac{1}{d}\sum_{(j,k)\neq(0,0)} \beta_{jk}\,\mathrm{tr}[V_{jk}] = 0.
\end{aligned}$$

Since $\rho_{00} = |\Phi_{00}\rangle\langle\Phi_{00}|$, we arrive at

$$\rho_{00}(U^\dagger V_{mn}^\dagger V_{\tilde{m}\tilde{n}}U \otimes \mathbb{1})\rho_{00} = 0. \tag{3.38}$$

which completes this part of the proof. Therefore, by using (3.38), for $m \neq \tilde{m}$ or $n \neq \tilde{n}$ we have,

$$\pi_{mn}\pi_{\tilde{m}\tilde{n}} = (V_{mn}U \otimes \mathbb{1})\underbrace{\rho_{00}(U^\dagger V_{mn}^\dagger V_{\tilde{m}\tilde{n}}U \otimes \mathbb{1})\rho_{00}}_{0}(U^\dagger V_{\tilde{m}\tilde{n}}^\dagger \otimes \mathbb{1}) = 0.$$

□

By using the orthogonality property (3.37) and the purity of the density operators

π_{mn}, we can write

$$\begin{aligned} S(\Lambda_a^P(\tau)) &= S\left(\Lambda_a^P\left((U\otimes\mathbb{1})\rho_{00}(U^\dagger\otimes\mathbb{1})\right)\right) \\ &= S\left(\sum_{m,n=0}^{d-1} q_{mn}\underbrace{(V_{mn}U\otimes\mathbb{1})\rho_{00}(U^\dagger V_{mn}^\dagger\otimes\mathbb{1})}_{:=\pi_{mn}}\right) \\ &= H(\{q_{mn}\})\,, \end{aligned} \tag{3.39}$$

where $H(\{q_{mn}\}) = -\sum_{m,n} q_{mn}\log q_{mn}$ is the Shannon entropy defined in (1.1). We note that the von Neumann entropy $S(\Lambda_a^P(\tau))$ is independent of the unitary encoding U. Consequently, for a *one-sided* d-dimensional Pauli channel with an initial Bell state, the condition (3.24) is satisfied. The super dense coding capacity (3.30) for an initial Bell state and a *one-sided* Pauli channel in d dimensions takes the form

$$\begin{aligned} C_{\text{Bell}}^{\text{one-sided P}_\text{d}} &= S(\frac{\mathbb{1}}{d}\otimes\rho_b) - H(\{q_{mn}\}) \\ &= S(\frac{\mathbb{1}}{d}) + S(\rho_b) - H(\{q_{mn}\}) = \log d^2 - H(\{q_{mn}\}) \end{aligned} \tag{3.40}$$

where $S(\rho_b) = \log d$ for a d-dimensional Bell state and $m, n = 0, 1, ..., d-1$. Using (3.10) we notice that the super dense coding capacity of a $d\times d$-dimensional Bell state in the noiseless case is given by $\log d^2$. Thus, in the presence of a *one-sided* Pauli channel the super dense coding capacity is reduced by the amount $H(\{q_{mn}\})$ with respect to the noiseless case - i.e. the channel noise is simply subtracted from the super dense coding capacity with noiseless channels.

We notice that the same capacity is achieved also for any maximally entangled state, i.e. for any $|\Phi\rangle = U_a\otimes U_b|\Phi_{00}\rangle$. Actually, Lemma 3 still holds in this case and therefore also the derivation of the capacity (3.40).

Werner states

For ρ_{00} to be a Bell state in d dimensions with $\rho_{00} = |\Phi_{00}\rangle\langle\Phi_{00}|$, we say ρ_W is a Werner state in d dimensions if $\rho_W = \frac{1-\eta}{d^2}\mathbb{1} + \eta\rho_{00}$ with $0\leq\eta\leq 1$. Sometimes its also called the noisy Bell state or a pseudo pure Bell state. For $d = 2$ we use the notation $\rho_W = \frac{1-\eta}{4}\mathbb{1} + \eta\rho^+$ for a Werner state with $\rho^+ = |\Phi^+\rangle\langle\Phi^+|$. We will now evaluate the super dense coding capacity for an input Werner state.

The Werner state ρ_W in the presence of a *one-sided* d-dimensional Pauli channel provides another example of states and channels that satisfy (3.24).
Using (3.39), $\{q_{mn}\}$ is the set of eigenvalues of $\Lambda_a^P\left[(U\otimes\mathbb{1})\rho_{00}(U^\dagger\otimes\mathbb{1})\right]$. The Pauli channel is a linear and unital map. Expressing the identity matrix $\mathbb{1}$ in a suitable basis, we arrive at

$$\begin{aligned}
&S\left(\Lambda_a^P\left((U\otimes\mathbb{1})\rho_W(U^\dagger\otimes\mathbb{1})\right)\right)\\
&=S\left(\eta\Lambda_a^P\left[(U\otimes\mathbb{1})\rho_{00}(U^\dagger\otimes\mathbb{1})\right]+\frac{1-\eta}{d^2}\mathbb{1}\right)\\
&=S\left(\eta\mathrm{diag}\left(q_{00},...,q_{d-1d-1}\right)+\frac{1-\eta}{d^2}\mathbb{1}\right)\\
&=S\left(\mathrm{diag}\left(\eta q_{00}+\frac{1-\eta}{d^2},...,\eta q_{d-1,d-1}+\frac{1-\eta}{d^2}\right)\right)\\
&=H\left(\{\eta q_{mn}+\frac{1-\eta}{d^2}\}\right).
\end{aligned}\tag{3.41}$$

From (3.41) it is apparent that the output channel entropy is independent of the unitary encoding. Consequently, the super dense coding capacity, according to (3.30), is given by

$$C_{\text{Werner}}^{\text{one-sided P}_\text{d}}=\log d^2-H(\{\frac{1-\eta}{d^2}+\eta q_{mn}\}).\tag{3.42}$$

The above capacity is also achieved by any other state with the form $U_a\otimes U_b\rho_W U_a^\dagger\otimes U_b^\dagger$.

3.2.3 Two-sided d-dimensional depolarizing channel.

In (3.31) we introduced the concept of a *one-sided* d-dimensional Pauli channel. A *two-sided* d-dimensional Pauli channel is then defined by

$$\Lambda_{ab}^{\text{P}}(\rho_i)=\sum_{m,n,\tilde{m},\tilde{n}=0}^{d-1}q_{mn}q_{\tilde{m}\tilde{n}}(V_{mn}\otimes V_{\tilde{m}\tilde{n}})\rho_i(V_{mn}^\dagger\otimes V_{\tilde{m}\tilde{n}}^\dagger).\tag{3.43}$$

The d-dimensional depolarizing channel is a special case of a d-dimensional Pauli channel, with probability parameters

$$q_{mn}=\begin{cases}1-p+\frac{p}{d^2}, & m=n=0\\ \frac{p}{d^2}, & \text{otherwise,}\end{cases}\tag{3.44}$$

with the noise parameter p, $0 \leq p \leq 1$, and $m, n = 0, ..., d-1$. In a $d-$dimensional depolarizing channel with probability p, the quantum state is replaced with a completely mixed state $\frac{\mathbb{1}}{d}$ and with the probability $1-p$ left unchanged. In the following Lemma we make the statement that the von Neumann entropy of a state that was sent through the *two-sided* depolarizing channel is independent of any local unitary transformations that were performed before the action of the channel.

Lemma 4. Let $\Lambda_{ab}^{\text{dep}}$ denote a *two-sided* d-dimensional depolarizing channel. For a state ρ and bilateral unitary operator $U_a \otimes U_b$, we have

$$S\left(\Lambda_{ab}^{\text{dep}}\left((U_a \otimes U_b)\,\rho(U_a^\dagger \otimes U_b^\dagger)\right)\right) = S(\Lambda_{ab}^{\text{dep}}(\rho)). \tag{3.45}$$

Proof: Considering Λ_a^{dep} and Λ_b^{dep} to be the d-dimensional depolarizing channels that act on Alice's and Bob's system, respectively, it is straightforward to verify that

$$\Lambda_a^{\text{dep}}(\lambda_j) = (1-p)\lambda_j\ , \tag{3.46}$$

(where λ_j are as before the generators of $SU(d)$), and analogously for Bob's system. We also show that for $\Lambda_a^{dep}(U\lambda_j U^\dagger)$ a similar expression holds. By the fact that λ_j is traceless, we have

$$\Lambda_a^{dep}(U\lambda_j U^\dagger) \;=\; \Lambda_a^{dep}\left(\sum_{m,n\neq(0,0)}^{d-1} \gamma_{mn} V_{mn}\right).$$

Here, $\Lambda_a^{dep}(\cdot)$ as we defined before is a linear map that is given by $\Lambda_a^{dep}(\cdot) = \sum_{\tilde{m},\tilde{n}=0}^{d-1} q_{\tilde{m}\tilde{n}} V_{\tilde{m}\tilde{n}}(\cdot)V_{\tilde{m}\tilde{n}}^\dagger$. Then we can write

$$\begin{aligned}\Lambda_a^{dep}(U\lambda_j U^\dagger) &= \sum_{m,n\neq(0,0)}^{d-1} \gamma_{mn}\Lambda_a^{dep}(V_{mn}) \\ &= \sum_{\tilde{m},\tilde{n}=0}^{d-1}\sum_{m,n\neq(0,0)}^{d-1} \gamma_{mn} q_{\tilde{m}\tilde{n}} V_{\tilde{m}\tilde{n}} V_{mn} V_{\tilde{m}\tilde{n}}^\dagger.\end{aligned}$$

By using (3.35) and unitarity of V_{mn}, we have

$$\begin{aligned}\Lambda_a^{dep}(U\lambda_j U^\dagger) &= \sum_{\tilde{m},\tilde{n}=0}^{d-1}\sum_{m,n\neq(0,0)}^{d-1}\gamma_{mn}q_{\tilde{m}\tilde{n}}\exp\left(\frac{2i\pi(n\tilde{m}-\tilde{n}m)}{d}\right)V_{mn}\\ &= \sum_{m,n\neq(0,0)}^{d-1}\gamma_{mn}V_{mn}\sum_{\tilde{m},\tilde{n}=0}^{d-1}q_{\tilde{m}\tilde{n}}\exp\left(\frac{2i\pi(n\tilde{m}-\tilde{n}m)}{d}\right).\end{aligned}$$

For $q_{\tilde{m}\tilde{n}}$ we replace the expression of (3.44) and we can write

$$\begin{aligned}\Lambda_a^{dep}(U\lambda_j U^\dagger) &= \sum_{m,n\neq(0,0)}^{d-1}\gamma_{mn}V_{mn}\left(1-p+\frac{p}{d^2}\underbrace{\sum_{\tilde{m},\tilde{n}=0}^{d-1}\exp\left(\frac{2i\pi(n\tilde{m}-\tilde{n}m)}{d}\right)}_{\delta_{0,m}\delta_{0,n}}\right)\\ &= (1-p)\sum_{m,n\neq(0,0)}^{d-1}\gamma_{mn}V_{mn} = (1-p)U\lambda_j U^\dagger,\end{aligned}$$

which completes this part of proof. Therefore,

$$\Lambda_a^{\rm dep}(U_a\lambda_j U_a^\dagger) = (1-p)U_a\lambda_j U_a^\dagger\,. \tag{3.47}$$

In the next step we show the following covariance property of the *two-sided* depolarizing channel, namely

$$\Lambda_{ab}^{\rm dep}\Big((U_a\otimes U_b)\,\rho(U_a^\dagger\otimes U_b^\dagger)\Big) = (U_a\otimes U_b)\left[\Lambda_{ab}^{\rm dep}(\rho)\right](U_a^\dagger\otimes U_b^\dagger), \tag{3.48}$$

holds. To prove the above expression, we use the decomposition (3.13) for ρ and

(3.47), and get

$$\begin{aligned}
&\Lambda_{ab}^{\mathrm{dep}}\left((U_a\otimes U_b)\,\rho(U_a^\dagger\otimes U_b^\dagger)\right)=\frac{1}{d^2}\Lambda_{ab}^{\mathrm{dep}}\left[\left(\mathbb{1}\otimes\mathbb{1}+\sum_{j=1}^{d^2-1}r_jU_a\lambda_jU_a^\dagger\otimes\mathbb{1}\right.\right.\\
&\left.\left.+\mathbb{1}\otimes\sum_{j=1}^{d^2-1}s_jU_b\lambda_jU_b^\dagger+\sum_{j,k=1}^{d^2-1}t_{jk}U_a\lambda_jU_a^\dagger\otimes U_b\lambda_kU_b^\dagger\right)\right]\\
&=\frac{1}{d^2}\left[\mathbb{1}\otimes\mathbb{1}+\sum_{j=1}^{d^2-1}r_j\Lambda_a^{\mathrm{dep}}(U_a\lambda_jU_a^\dagger)\otimes\mathbb{1}+\mathbb{1}\otimes\sum_{j=1}^{d^2-1}s_j\Lambda_b^{\mathrm{dep}}(U_b\lambda_jU_b^\dagger)\right.\\
&\left.+\sum_{j,k=1}^{d^2-1}t_{jk}\Lambda_a^{\mathrm{dep}}\left(U_a\lambda_jU_a^\dagger\right)\otimes\Lambda_b^{\mathrm{dep}}\left(U_b\lambda_kU_b^\dagger\right)\right]\\
&=\frac{1}{d^2}\left[\mathbb{1}\otimes\mathbb{1}+(1-p)\left(\sum_{j=1}^{d^2-1}r_jU_a\lambda_jU_a^\dagger\otimes\mathbb{1}+\mathbb{1}\otimes\sum_{j=1}^{d^2-1}s_jU_b\lambda_jU_b^\dagger\right)\right.\\
&\left.+(1-p)^2\sum_{j,k=1}^{d^2-1}t_{jk}U_a\lambda_jU_a^\dagger\otimes U_b\lambda_kU_b^\dagger\right]\\
&=(U_a\otimes U_b)\frac{1}{d^2}\left[\mathbb{1}\otimes\mathbb{1}+(1-p)\left(\sum_{j=1}^{d^2-1}r_j\lambda_j\otimes\mathbb{1}+\mathbb{1}\otimes\sum_{j=1}^{d^2-1}s_j\lambda_j\right)\right.\\
&\left.+(1-p)^2\sum_{j,k=1}^{d^2-1}t_{jk}\lambda_j\otimes\lambda_k\right](U_a^\dagger\otimes U_b^\dagger)\\
&=(U_a\otimes U_b)\left[\Lambda_{ab}^{\mathrm{dep}}(\rho)\right](U_a^\dagger\otimes U_b^\dagger).
\end{aligned}$$

Since the von Neumann entropy is invariant under unitary transformations, the proof of Lemma 4 is complete. □

As a consequence of Lemma 4 we can conclude that for a *two-sided* d-dimensional depolarizing channel the entropy for a given initial state ρ is independent of the unitary encoding, namely

$$S\left(\Lambda_{ab}^{\mathrm{dep}}\left((U\otimes\mathbb{1})\,\rho\left(U^\dagger\otimes\mathbb{1}\right)\right)\right)=S\left(\Lambda_{ab}^{\mathrm{dep}}(\rho)\right). \tag{3.49}$$

Therefore, (3.24) holds and, according to (3.30), the super dense coding capacity for a given general resource state ρ, with a *two-sided* d-dimensional depolarizing

channel is given by

$$\begin{aligned} C^{\text{two-sided dep}_\text{d}}(\rho) &= S\left(\frac{\mathbb{1}}{d}\otimes\Lambda_b^{\text{dep}}(\rho_b)\right) - S\left(\Lambda_{ab}^{\text{dep}}(\rho)\right) \\ &= S\left(\frac{\mathbb{1}}{d}\right) + S\left(\Lambda_b^{\text{dep}}(\rho_b)\right) - S\left(\Lambda_{ab}^{\text{dep}}(\rho)\right) \\ &= \log d + S\left(\Lambda_b^{\text{dep}}(\rho_b)\right) - S\left(\Lambda_{ab}^{\text{dep}}(\rho)\right). \end{aligned} \tag{3.50}$$

Notice that since Lemma 4 holds for any local unitary $U_a \otimes U_b$, the capacity (3.50) depends only on the degree of entanglement of the input state ρ. In other words, all input states with the same degree of entanglement have the same super dense coding capacity.

Comparing the above expression (3.50) with the one for the noiseless case, given by $C = \log d + S(\rho_b) - S(\rho)$, one realizes that in the case of *two-sided* noise the channel that affects Bob's subsystem enters twice, both in the von Neumann entropies for the local and the global density matrix.

Super dense coding capacity and optimal initial state

In (3.50) we obtained the super dense coding capacity of an arbitrary given initial resource state ρ for the *two-sided* d-dimensional depolarizing channel. In this subsection we perform the optimization of the super dense coding capacity over the initial state of two qubits for the *two-sided* 2-dimensional depolarizing channel. Thus, we derive the optimal value of the super dense coding capacity, if Alice and Bob have a depolarizing channel available for the transfer of 2-dimensional quantum states and can choose the initial resource state.

A pure state of two qubits $|\vartheta_\alpha\rangle$ can be written in the Schmidt bases $\{|u_i\rangle\}, \{|v_i\rangle\}$ as $|\vartheta_\alpha\rangle = \sqrt{1-\alpha}|u_1v_1\rangle + \sqrt{\alpha}|u_2v_2\rangle$ with $0 \leq \alpha \leq 1/2$. Two local unitaries V_a and V_b convert the computational bases to the Schmidt bases. Therefore, $|\vartheta_\alpha\rangle$ in computational bases can be written as $|\vartheta_\alpha\rangle = V_a \otimes V_b(\sqrt{1-\alpha}|00\rangle + \sqrt{\alpha}|11\rangle)$. In (3.45) we showed that the output von Neumann entropy of the *two-sided* depolarizing channel is invariant under previous local unitary transformations. Therefore $|\vartheta_\alpha\rangle$ and $|\varphi_\alpha\rangle = \sqrt{1-\alpha}|00\rangle + \sqrt{\alpha}|11\rangle$ lead to the same dense coding capacity. We can thus parametrize a pure initial state as a function of a single real parameter, namely as the state $|\varphi_\alpha\rangle$, and follow the approach of Ref. [13]. The super dense

coding capacity (3.50) of a pure state of two qubits as a function of α and the noise parameter p is given by

$$\begin{aligned} C_{\alpha}^{\text{two-sided dep}_2}\left(|\varphi_\alpha\rangle\langle\varphi_\alpha|\right) &= 1 - \xi_1 \log \xi_1 - \xi_2 \log \xi_2 + \gamma_1 \log \gamma_1 + \gamma_2 \log \gamma_2 \\ &+ 2\gamma_3 \log \gamma_3 \ , \end{aligned} \tag{3.51}$$

where γ_i (with $i = 1, 2, 3, 4$) are the eigenvalues of $\Lambda_{ab}^{\text{dep}}\left(|\varphi_\alpha\rangle\langle\varphi_\alpha|\right)$ and ξ_s (with $s = 1, 2$) are the eigenvalues of $\Lambda_{b}^{\text{dep}}\left(\rho_{b,\alpha}\right)$, where $\rho_{b,\alpha} = \text{tr}_a(|\varphi_\alpha\rangle\langle\varphi_\alpha|)$. The eigenvalues γ_i and ξ_s are explicitly given by

$$\begin{aligned} \gamma_{1,2} &= \frac{1}{2}\left(1 - p(1 - \frac{p}{2}) \pm (1-p)\sqrt{1 - 4p\alpha(2-p)(1-\alpha)}\right) \ , \\ \gamma_3 = \gamma_4 &= \frac{p}{2}(1 - \frac{p}{2}) \ , \\ \xi_1 &= \alpha - p\alpha + \frac{p}{2} \ , \\ \xi_2 &= 1 - \alpha + p\alpha - \frac{p}{2} \ . \end{aligned} \tag{3.52}$$

We can now maximize expression (3.51) with respect to the variable α, for a given noise parameter p, and find interesting results. They are illustrated in Figure 3.4, where we plot the super dense coding capacity in (3.51) as a function of the noise parameter p, for various values α. We find that there is a threshold value $p_t \approx 0.345$, where two curves cross each other: for $0 \leq p \leq 0.345$ the value $\alpha = 1/2$ leads to the highest super dense coding capacity, i.e. the optimal initial resource state is a Bell state. For $p \geq 0.345$, the optimal choice is $\alpha = 0$, i.e. product states are best for super dense coding. As shown graphically in the close-up of Figure 3.4 , the curves for intermediate values of α are always lower than $\alpha = 1/2$ or $\alpha = 0$. In order to prove this claim, we also evaluated $C_{\alpha=1/2}^{\text{two-sided dep}_2} - C_{\alpha}^{\text{two-sided dep}_2}$ in the range of $0 \leq p \leq 0.345$ and $C_{\alpha=0}^{\text{two-sided dep}_2} - C_{\alpha}^{\text{two-sided dep}_2}$ in the range of $0.345 \leq p \leq 1$ as functions of the parameters α and p. We found that these two functions are positive or zero. Thus, for pure initial states it is always best to either use maximally entangled states or product states, depending on the noise level.

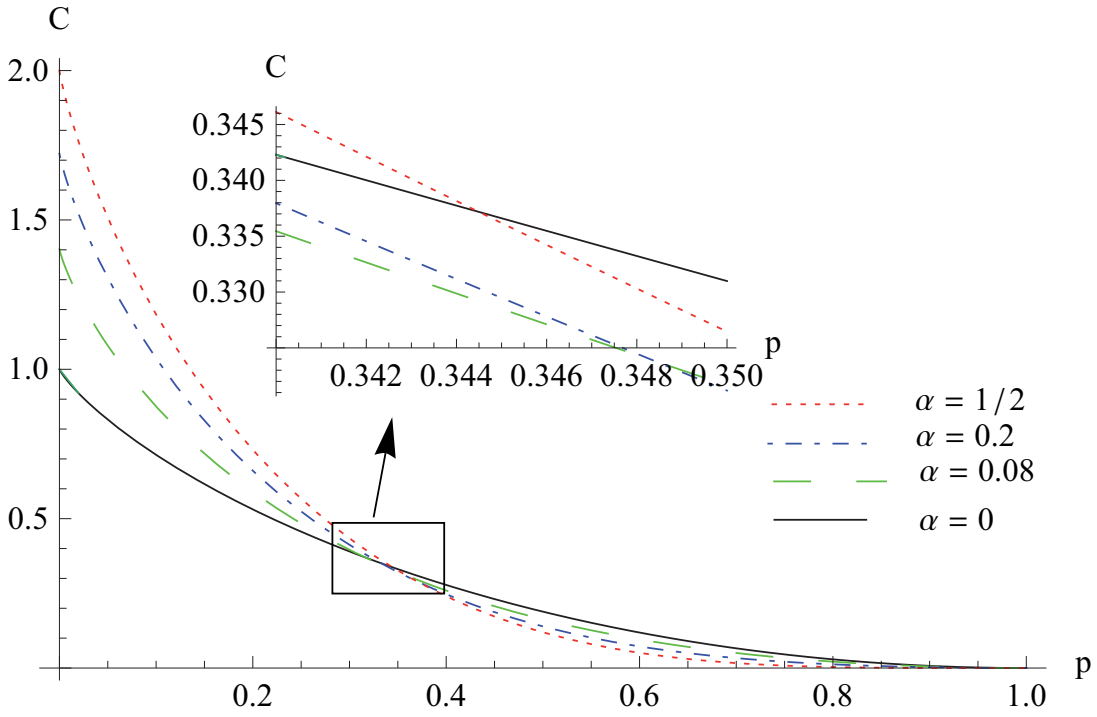

Figure 3.4: The super dense coding capacity for the *two-sided* depolarizing channel in 2 dimensions, $C_\alpha^{\text{two-sided dep}_2}$, as function of the noise parameter p, for $\alpha = 0$, $\alpha = 0.08$, $\alpha = 0.2$ and $\alpha = 1/2$. For the definition of α see main text. For $0 \leq p \leq 0.345$ a Bell state, i.e. $\alpha = 1/2$, leads to the optimal capacity, while for $0.345 \leq p \leq 1$ the optimal initial state is a product state ($\alpha = 0$).

In the following we call the super dense coding capacity of an initial Bell state $|\varphi_{1/2}\rangle$ in the presence of a *two-sided* 2-dimensional depolarizing channel $C_{\text{Bell}}^{\text{two-sided dep}_2}$. Using (3.51) with $\alpha = 1/2$, this capacity is given by

$$\begin{aligned} C_{\text{Bell}}^{\text{two-sided dep}_2} &= 2 + \frac{1+3(1-p)^2}{4} \log \frac{1+3(1-p)^2}{4} \\ &+ 3\frac{1-(1-p)^2}{4} \log \frac{1-(1-p)^2}{4} \,. \end{aligned} \tag{3.53}$$

The super dense coding capacity with an initial product state $|\varphi_0\rangle$ in the presence of a *two-sided* 2-dimensional depolarizing channel is denoted in the following as $C^{\text{ch dep}_2}$. From (3.51) with $\alpha = 0$ it follows that

$$C^{\text{ch dep}_2} = 1 + \frac{p}{2} \log \frac{p}{2} + \frac{2-p}{2} \log \frac{2-p}{2}. \tag{3.54}$$

Note that (3.54) is identical to the classical channel capacity of the depolarizing channel for qubits [30].

We now show that using mixed initial states as a resource cannot increase the super dense coding capacity, i.e. $|\varphi_{1/2}\rangle$ and $|\varphi_0\rangle$ are the optimal input states for the range of noise parameter $0 \leq p \leq 0.345$ and $0.345 \leq p \leq 1$, respectively. To show this claim we first write the super dense coding capacity (3.50) in the form of the relative entropy

$$\begin{aligned} C^{\text{two-sided dep}_\text{d}}(\rho) &= \log d + S\left(\Lambda_b^{\text{dep}}(\rho_b)\right) - S\left(\Lambda_{ab}^{\text{dep}}(\rho)\right) \\ &= -\operatorname{tr}_b\left[\Lambda_b^{\text{dep}}(\rho_b)\left(\log \Lambda_b^{\text{dep}}(\rho_b) - \log d\right)\right] - S\left(\Lambda_{ab}^{\text{dep}}(\rho)\right) \\ &= -\operatorname{tr}_b\left[\Lambda_b^{\text{dep}}(\rho_b)\left(\log \Lambda_b^{\text{dep}}\left(\frac{\rho_b}{d}\right)\right)\right] - S\left(\Lambda_{ab}^{\text{dep}}(\rho)\right) \end{aligned}$$

By using $\operatorname{tr}_a \Lambda_{ab}^{\text{dep}}(\rho) = \Lambda_b^{\text{dep}}(\rho_b)$ and also the property $\mathbb{1} \otimes \log \sigma = \log(\mathbb{1} \otimes \sigma)$, we arrive at the relative entropy form

$$\begin{aligned} C^{\text{two-sided dep}_\text{d}}(\rho) &= -\operatorname{tr}_a \operatorname{tr}_b\left[\Lambda_{ab}^{\text{dep}}(\rho)\left(\mathbb{1} \otimes \log \Lambda_b^{\text{dep}}\left(\frac{\rho_b}{d}\right)\right)\right] - S\left(\Lambda_{ab}^{\text{dep}}(\rho)\right) \\ &= -\operatorname{tr}\left[\Lambda_{ab}^{\text{dep}}(\rho) \log\left(\mathbb{1} \otimes \Lambda_b^{\text{dep}}\left(\frac{\rho_b}{d}\right)\right)\right] - S\left(\Lambda_{ab}^{\text{dep}}(\rho)\right) \\ &= \operatorname{tr}\left[\Lambda_{ab}^{\text{dep}}(\rho) \log \Lambda_{ab}^{\text{dep}}(\rho)\right] - \operatorname{tr}\left[\Lambda_{ab}^{\text{dep}}(\rho) \log\left(\mathbb{1} \otimes \Lambda_b^{\text{dep}}\left(\frac{\rho_b}{d}\right)\right)\right] \\ &= S(\Lambda_{ab}^{\text{dep}}(\rho) \| \frac{\mathbb{1}}{d} \otimes \Lambda_b^{\text{dep}}(\rho_b)) \,. \end{aligned} \tag{3.55}$$

Since any mixed state can be written as a convex combination of pure states ρ_k, i.e. $\rho_{mix} = \sum_k p_k \rho_k$, and $\rho_{b,mix} = \mathrm{tr}_a(\rho_{mix}) = \sum_k p_k \rho_{b,k}$, we can write

$$\begin{aligned} C_{\rho_{mix}} &= S(\Lambda_{ab}^{\mathrm{dep}}(\rho_{mix})\|\tilde{\rho}) \\ &= S(\Lambda_{ab}^{\mathrm{dep}}(\rho_{mix})\|\frac{\mathbb{1}}{d} \otimes \Lambda_b^{\mathrm{dep}}(\rho_{b,mix})) \\ &= S(\sum_k p_k \Lambda_{ab}^{\mathrm{dep}}(\rho_k)\|\frac{\mathbb{1}}{d} \otimes \sum_k p_k \Lambda_b^{\mathrm{dep}}(\rho_{b,k})) \\ &= S(\sum_k p_k \Lambda_{ab}^{\mathrm{dep}}(\rho_k)\|\sum_k p_k \frac{\mathbb{1}}{d} \otimes \Lambda_b^{\mathrm{dep}}(\rho_{b,k})) \\ &\leq \sum_k p_k S(\Lambda_{ab}^{\mathrm{dep}}(\rho_k)\|\frac{\mathbb{1}}{d} \otimes \Lambda_b^{\mathrm{dep}}(\rho_{b,k})) \ . \end{aligned} \tag{3.56}$$

In the above inequality we have used the subadditivity of the relative entropy, i.e. $S(\sum_i p_i r_i\| \sum_i q_i s_i) \leq \sum_i p_i S(r_i\|s_i) + H(p_i\|q_i)$, where $H(\cdot\|\cdot)$ is the Shannon relative entropy, defined as $H(p_i\|q_i) = \sum_i p_i \log \frac{p_i}{q_i}$ [40]. We showed before that the super dense coding capacity of a pure state for $0 \leq p \leq 0.345$ is upper bounded by the super dense coding capacity of a Bell state $|\varphi_{1/2}\rangle$, and for $0.345 \leq p \leq 1$ it is upper bounded by the product state $|\varphi_0\rangle$. Remembering that ρ_k is pure, and using (3.55), we find that for $0 \leq p \leq 0.345$

$$C_{\rho_{mix}} \leq \sum_k p_k S(\Lambda_{ab}^{\mathrm{dep}}(\rho_k)\|\frac{\mathbb{1}}{d} \otimes \Lambda_b^{\mathrm{dep}}(\rho_{b,k})) \leq C_{Bell}^{\text{two-sided dep}_2} \ , \tag{3.57}$$

and for $0.345 \leq p \leq 1$

$$C_{\rho_{mix}} \leq \sum_k p_k S(\Lambda_{ab}^{\mathrm{dep}}(\rho_k)\|\frac{\mathbb{1}}{d} \otimes \Lambda_b^{\mathrm{dep}}(\rho_{b,k})) \leq C^{\mathrm{ch\,dep}_2} \ , \tag{3.58}$$

which proves our claim.

It is interesting to note that the optimal capacity for the *two-sided* qubit depolarizing channel is a non-differentiable function of the noise parameter p, and that the optimal states are either maximally entangled or separable. In other words, there is a transition in the entanglement of the optimal input states at the particular threshold value of the noise parameter $p_t \approx 0.345$. Notice that a similar transition behavior in the entanglement of the optimal input states for transmission of classical information was found also for the qubit depolarizing channel

with correlated noise [36]. It is interesting that in the present context the transition behavior arises in a memoryless channel and is not related to correlations introduced via the noise process.

3.2.4 Super dense coding capacity versus channel capacity

In this section, we consider the question of whether or not it is reasonable in the presence of noise to use the super dense coding protocol for the transmission of classical information? To answer this question, we provide a comparison between the classical capacity of a 2-dimensional depolarizing channel and the super dense coding capacities of a *one-sided* and *two-sided* 2-dimensional depolarizing channel, for the resource of an initially shared Bell state. Since the depolarizing channel is a special form of a Pauli channel, according to (3.40) the super dense coding capacity for a *one-sided* 2-dimensional depolarizing channel for an initially shared Bell state is

$$C_{Bell}^{\text{one-sided dep}_2} = 2 + \frac{4-3p}{4}\log\frac{4-3p}{4} + 3\frac{p}{4}\log\frac{p}{4}. \tag{3.59}$$

The super dense coding capacity for a *two-sided* 2-dimensional depolarizing channel with a Bell state as resource is given in (3.53). The classical capacity $C^{\text{ch dep}_2}$ of the 2-dimensional depolarizing channel is achieved by an ensemble of pure states belonging to an orthonormal basis, say $\{|0\rangle, |1\rangle\}$ at the channel input, with equal probability $\frac{1}{2}$ and performing a complete von Neumann measurement in the same basis over the channel output [30]. Its expression is given explicitly in (3.54).

In Figure 3.5, we plot $C_{Bell}^{\text{one-sided dep}_2}$, $C_{Bell}^{\text{two-sided dep}_2}$, $C^{\text{ch dep}_2}$, and $C = 1$ in terms of the noise parameter p. As we expect, the first three capacities $C_{Bell}^{\text{one-sided dep}_2}$, $C_{Bell}^{\text{two-sided dep}_2}$ and $C^{\text{ch dep}_2}$ decrease as the noise increases. As expected, the super dense coding capacity of a *one-sided* 2-dimensional depolarizing channel $C_{Bell}^{\text{one-sided dep}_2}$ is greater than the classical capacity $C^{\text{ch dep}_2}$ for all values of p, as the additional resource of entanglement is used in super dense coding. The comparison between $C_{Bell}^{\text{two-sided dep}_2}$ and $C^{\text{ch dep}_2}$ illustrates that for $0.345 \leq p \leq 1$ the 2-dimensional depolarizing channel capacity is greater than the super dense coding capacity for a *two-sided* 2-dimensional depolarizing channel. This suggests

that for $0.345 \leq p \leq 1$ Alice and Bob do not win by sending classical information via a super dense coding protocol with unitary encoding. For this regime, the noise degrades the entanglement too much to be useful. Now we can answer the question posed at the beginning of this sub-section: super dense coding is not always a useful scheme for sending classical information in the presence of noise.

We notice also that $C_{Bell}^{\text{one-sided dep}_2}$ corresponds to the entanglement assisted capacity for the depolarizing channel [5]. According to (3.59) for $p = 0.252$ the super dense coding capacity for an initial Bell state via the *one-sided* 2-dimensional depolarizing channel is equal to *one*. The maximum information that can be transmitted by two-dimensional systems without any source of entangled states is $C = 1$. That is, for $p = 0.252$ the super dense coding capacity reaches the classical limit, as can be seen in Figure 3.5. It was shown in [25] that the classical limit of the quantum teleportation protocol, when using a Bell state and distributing one subsystem of it via a depolarizing channel, is reached at $p = 1/3$. In the absence of noise, quantum teleportation and super dense coding are two equivalent protocols [49]. According to our results this is not true in the presence of noise, as we have shown explicitly for the depolarizing channel: here, the quantum/classical boundary for super dense coding occurs at a different noise value than for quantum teleportation.

We point out that the expression (3.40) for the super dense coding capacity of a Bell state provides a lower bound to the entanglement-assisted capacity of a general Pauli channel.

We also note that the discussed issues in the present section have been mainly published in [45].

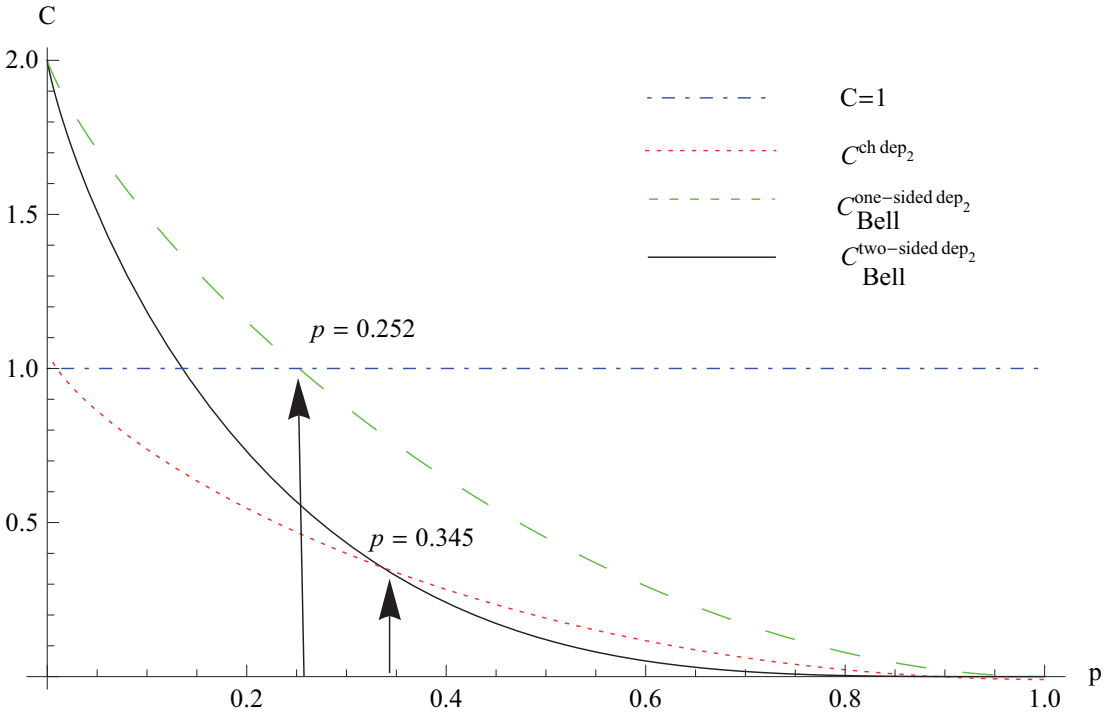

Figure 3.5: The classical capacity $C^{\text{ch dep}_2}$ of the 2-dimensional depolarizing channel and the super dense coding capacities for an initial Bell state in the presence of a *one-sided* and *two-sided* 2-dimensional depolarizing channel, $C_{Bell}^{\text{one-sided dep}_2}$ and $C_{Bell}^{\text{two-sided dep}_2}$, respectively, as functions of the noise parameter p.

3.3 Super dense coding in the presence of a correlated Pauli channel

In many real world applications the assumption of having uncorrelated noisy channels is far from reality and memory effects need to be taken into account. We already in the preliminaries chapter, briefly, explained the notion of the channels with memory. In this section, we use the correlated noise as a model for a channel with memory. In this model, noise in the consecutive uses of the channel is correlated. One of the aspects of such channels is the possibility to enhance the capacity. Here, we specifically consider the Pauli channel (3.31). Let Λ_a^P and Λ_b^P being two d-dimensional Pauli channels which act on Alice's and Bob's subsystems, respectively. These two channels are correlated with the definition,

$$\Lambda_{ab}^P(\cdot) = \sum_{m,n,\tilde{m},\tilde{n}=0}^{d-1} q_{mn\tilde{m}\tilde{n}}(V_{mn} \otimes V_{\tilde{m}\tilde{n}})(\cdot)(V_{mn}^\dagger \otimes V_{\tilde{m}\tilde{n}}^\dagger). \tag{3.60}$$

where the probability $q_{mn\tilde{m}\tilde{n}}$ is given by $q_{mn\tilde{m}\tilde{n}} = (1-\mu)q_{mn}q_{\tilde{m}\tilde{n}} + \mu q_{mn}\delta_{m,\tilde{m}}\delta_{n,\tilde{n}}$ with the correlation degree $0 \leq \mu \leq 1$. For $\mu = 0$ the two channels Λ_a^P and Λ_b^P are uncorrelated and for $\mu = 1$ they are fully correlated.

We name (3.60) a *correlated* Pauli channel. In the present section, for a single sender and a single receiver, a *correlated* Pauli channel as well as unitary and non unitary encoding, we obtain two explicit expressions for super dense coding capacity. We show that both unitary and non-unitary encoding problems reduce to the problem of finding a single CPTP map (in the case of unitary encoding this is a specific unitary transformation) that minimizes the output von Neumann entropy after applying it and the channel to the input state ρ. For the case of unitary encoding we find examples that the single unitary operator is explicitly defined. The rest of the present section will explain in detail the above statements.

3.3.1 Unitary encoding

This subsection treats the optimization of the Holevo quantity having a *correlated* Pauli channel and unitary encoding. We introduce an upper bound on the Holevo quantity and we show that this upper bound is reachable and thus is the super

dense coding capacity. This procedure phrased in the following Lemma.

Lemma 5. Let

$$\chi = S\left(\overline{\Lambda_{ab}^P(\rho)}\right) - \sum_i p_i S\left(\Lambda_{ab}^P(\rho_i)\right) \tag{3.61}$$

be the Holevo quantity with $\rho_i = (W_i \otimes \mathbb{1})\rho(W_i{}^\dagger \otimes \mathbb{1})$, the average state $\overline{\Lambda_{ab}^P(\rho)} = \sum_i p_i \Lambda_{ab}^P(\rho_i)$ and Λ_{ab}^P to be the *correlated* Pauli channel defined via (3.60). Let U_{min} be the unitary operator that minimizes the von Neumann entropy after application of this unitary operator and the channel Λ_{ab}^P to the initial state ρ, i.e. U_{min} minimizes the expression $S\left(\Lambda_{ab}^P((U_{min} \otimes \mathbb{1})\rho(U_{min}^\dagger \otimes \mathbb{1}))\right)$. Then the super dense coding capacity C_{un}^P is given by

$$C_{un}^P = \log d + S\left(\Lambda_b^P(\rho_b)\right) - S\left[\Lambda_{ab}^P\left((U_{min} \otimes \mathbb{1})\rho(U_{min}^\dagger \otimes \mathbb{1})\right)\right]. \tag{3.62}$$

Proof: We start with introducing an upper bound on the Holevo quantity (3.61). Since U_{min} is a unitary operator that leads to the minimum of the output von Neumann entropy, for χ we have

$$\begin{aligned}\chi &= S\left(\overline{\Lambda_{ab}^P(\rho)}\right) - \sum_i p_i S\left(\Lambda_{ab}^P(\rho_i)\right) \\ &\leq S\left(\overline{\Lambda_{ab}^P(\rho)}\right) - S\left[\Lambda_{ab}^P\left((U_{min} \otimes \mathbb{1})\rho(U_{min}^\dagger \otimes \mathbb{1})\right)\right].\end{aligned}$$

The von Neumann entropy is subadditive and the maximum entropy of a d-dimensional system is $\log d$. Therefore,

$$\begin{aligned}\chi &\leq S\left(\mathrm{tr}_b\,\overline{\Lambda_{ab}^P(\rho)}\right) + S\left(\mathrm{tr}_a\,\overline{\Lambda_{ab}^P(\rho)}\right) - S\left[\Lambda_{ab}^P\left((U_{min} \otimes \mathbb{1})\rho(U_{min}^\dagger \otimes \mathbb{1})\right)\right] \\ &\leq \log d + S\left(\mathrm{tr}_a\,\overline{\Lambda_{ab}^P(\rho)}\right) - S\left[\Lambda_{ab}^P\left((U_{min} \otimes \mathbb{1})\rho(U_{min}^\dagger \otimes \mathbb{1})\right)\right].\end{aligned} \tag{3.63}$$

Since $\mathrm{tr}_a\,\overline{\Lambda_{ab}^P(\rho)} = \Lambda_b^P(\rho_b)$ it follows that

$$\chi \leq \log d + S\left(\Lambda_b^P(\rho_b)\right) - S\left[\Lambda_{ab}^P\left((U_{min} \otimes \mathbb{1})\rho(U_{min}^\dagger \otimes \mathbb{1})\right)\right]. \tag{3.64}$$

The upper bound (3.64) is achievable. To show this claim, we consider the ensemble $\{\tilde{p}_i = \frac{1}{d^2}, \tilde{U}_i = V_i U_{min}\}$ with V_i being defined in (3.14). The Holevo quantity for this ensemble is denoted by $\tilde{\chi}$ and is given by

$$\begin{aligned}\tilde{\chi} &= S\left[\sum_i \frac{1}{d^2}\Lambda_{ab}^P\left((\tilde{U}_i \otimes \mathbb{1})\rho(\tilde{U}_i{}^\dagger \otimes \mathbb{1})\right)\right] \\ &- \sum_i \frac{1}{d^2} S\left[\Lambda_{ab}^P\left((\tilde{U}_i \otimes \mathbb{1})\rho(\tilde{U}_i{}^\dagger \otimes \mathbb{1})\right)\right].\end{aligned} \tag{3.65}$$

By using (3.17), Lemma (1-a), and noting that U_{min} acts only on Alice's side, we find that the argument in the first term on the RHS of (3.65) is given by

$$\begin{aligned}
&\sum_i \frac{1}{d^2}\Lambda^P_{ab}\left(\tilde{U}_i\otimes\mathbb{1})\rho(\tilde{U_i}^\dagger\otimes\mathbb{1})\right)\\
=\;&\sum_i \frac{1}{d^2}\Lambda^P_{ab}\left((V_iU_{min}\otimes\mathbb{1})\rho(U^\dagger_{min}V_i^\dagger\otimes\mathbb{1})\right)\\
=\;&\Lambda^P_{ab}\left[\sum_i\frac{1}{d^2}(V_i\otimes\mathbb{1})(U_{min}\otimes\mathbb{1})\rho(U^\dagger_{min}\otimes\mathbb{1})(V_i^\dagger\otimes\mathbb{1})\right]\\
=\;&\Lambda^P_{ab}(\frac{\mathbb{1}}{d}\otimes\rho_b)=\frac{\mathbb{1}}{d}\otimes\Lambda^P_b(\rho_b).
\end{aligned}\tag{3.66}$$

Furthermore, the second term on the RHS of (3.65) can be expressed in terms of the unitary operator U_{min} and the channel. By inserting the action of the *correlated* Pauli channel, and using (3.35), from which it follows that $V_{i=jk}$ and V_{mn} commute up to a phase, we can write

$$\begin{aligned}
&\sum_i\frac{1}{d^2}S\left(\Lambda^P_{ab}\left(\tilde{U}_i\otimes\mathbb{1}\right)\rho(\tilde{U_i}^\dagger\otimes\mathbb{1})\right)\\
=\;&\frac{1}{d^2}\sum_i S\left(\sum_{m,n,\tilde{m},\tilde{n}}q_{mn\tilde{m}\tilde{n}}(V_{mn}\otimes V_{\tilde{m}\tilde{n}})\right.\\
&\left.\left[(V_iU_{min}\otimes\mathbb{1})\,\rho\left(U^\dagger_{min}V_i^\dagger\otimes\mathbb{1}\right)\right](V^\dagger_{mn}\otimes V^\dagger_{\tilde{m}\tilde{n}})\right)\\
=\;&\frac{1}{d^2}\sum_{i=k,j}S\left(\sum_{m,n,\tilde{m},\tilde{n}}q_{mn\tilde{m}\tilde{n}}(V_{kj}\otimes\mathbb{1})\,(V_{mn}\otimes V_{\tilde{m}\tilde{n}})\left[(U_{min}\otimes\mathbb{1})\,\rho\left(U^\dagger_{min}\otimes\mathbb{1}\right)\right]\right.\\
&\left.\left(V^\dagger_{mn}\otimes V^\dagger_{\tilde{m}\tilde{n}}\right)(V^\dagger_{kj}\otimes\mathbb{1})\right).
\end{aligned}\tag{3.67}$$

Since von Neumann entropy is invariant under unitary transformation we have

$$\begin{aligned}
&\sum_i \frac{1}{d^2} S\left(\Lambda_{ab}^P\left(\tilde{U}_i \otimes \mathbb{1}\right)\rho(\tilde{U_i}^\dagger \otimes \mathbb{1})\right) \\
= \;& \frac{1}{d^2}\sum_{kj} S\left((V_{kj}\otimes\mathbb{1})\left[\sum_{m,n,\tilde{m},\tilde{n}} q_{mn\tilde{m}\tilde{n}}\left(V_{mn}\otimes V_{\tilde{m}\tilde{n}}\right)\left(U_{min}\otimes\mathbb{1}\right)\rho\left(U_{min}^\dagger\otimes\mathbb{1}\right)\right.\right. \\
& \left.\left.\left(V_{mn}^\dagger\otimes V_{\tilde{m}\tilde{n}}^\dagger\right)\right](V_{kj}^\dagger\otimes\mathbb{1})\right) \\
= \;& \frac{1}{d^2}\sum_{kj} S\left(\sum_{m,n,\tilde{m},\tilde{n}} q_{mn\tilde{m}\tilde{n}}\left(V_{mn}\otimes V_{\tilde{m}\tilde{n}}\right)\left[\left(U_{min}\otimes\mathbb{1}\right)\rho\left(U_{min}^\dagger\otimes\mathbb{1}\right)\right]\right. \\
& \left.\left(V_{mn}^\dagger\otimes V_{\tilde{m}\tilde{n}}^\dagger\right)\right) \\
= \;& S\left[\Lambda_{ab}^P\left(\left(U_{min}\otimes\mathbb{1}\right)\rho\left(U_{min}^\dagger\otimes\mathbb{1}\right)\right)\right]
\end{aligned} \tag{3.68}$$

Inserting (3.66) and (3.68) into (3.65), one finds that the Holevo quantity $\tilde{\chi}$ is equal to the upper bound given in (3.64) and consequently, is the super dense coding capacity. □

By Lemma 5, we have proved that, in order to determine the super dense coding capacity , its enough to find an optimal U_{min} that minimizes the channel output von Neumann entropy $S\left[\Lambda_{ab}^P\left((U_{min}\otimes\mathbb{1})\rho(U_{min}^\dagger\otimes\mathbb{1})\right)\right]$. In the next subsection we give examples of the channels and initial states for which U_{min} is explicitly determined.

3.3.2 Correlated quasiclassical channel

A $d-$dimensional quasiclassical depolarizing channel (or simply quasiclassical channel) is a particular form of a $d-$dimensional Pauli channel. For this channel, the probabilities of the *displacement* operators V_{0n}, with *zero* mode ($m = 0$) and regardless to the phase shift n, are equal and they differ from the rest of the probabilities which are also equal

$$q_{mn} = \begin{cases} \frac{1-p}{d}, & m = 0 \\ \frac{p}{d(d-1)}, & \text{otherwise.} \end{cases} \tag{3.69}$$

The quasiclassical channel is characterized by a single probability parameter $0 \leq p \leq 1$. With the probability p, a *displacement* occurs and with the probability $1-p$, no *displacement* occurs to the quantum signal. Like in the classical case, p can also be seen as amount of the noise in the channel.

Bell states

Here, for a shared Bell state $\rho^+ = |\Phi^+\rangle\langle\Phi^+|$ with $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, and in the presence of a *Correlated* quasiclassical channel we find U_{min}. We investigate the case of $d = 2$. For two-dimensional systems the *displacement* operators are the identity and the three Pauli operators

$$\mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.70}$$

A two-dimensional *correlated* quasiclassical channel is then

$$\Lambda^Q_{ab}(\cdot) = \sum_{m,n} q_{mn}\sigma_m \otimes \sigma_n(\cdot)\sigma_m \otimes \sigma_n, \tag{3.71}$$

where $q_{mn} = (1-\mu)q_m q_n + \mu q_n \delta_{mn}$ with $q_0 = q_3 = \frac{1-p}{2}$ and $q_1 = q_2 = \frac{p}{2}$. According to find U_{min}, we start with the most general 2×2 unitary operator U

$$U = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}, \tag{3.72}$$

where a and b are complex variables which satisfy $|a|^2 + |b|^2 = 1$. The output of a *correlated* quasiclassical channel on an arbitrary state ρ is invariant under the unitary transformation $\sigma_3 \otimes \sigma_3$, i.e. $\Lambda^Q_{ab}(\rho) = \Lambda^Q_{ab}((\sigma_3 \otimes \sigma_3)\rho(\sigma_3 \otimes \sigma_3))$ [37]. The proof for this statement is coming in the following argument: The Pauli operators satisfy the property $\sigma_m\sigma_n = i\varepsilon_{mnk}\sigma_k + \delta_{mn}\sigma_0$. Therefore,

$$\begin{aligned} \Lambda^Q_{ab}((\sigma_3 \otimes \sigma_3)\rho(\sigma_3 \otimes \sigma_3)) &= \Sigma_{mn} q_{mn}(\sigma_m\sigma_3 \otimes \sigma_n\sigma_3)\rho(\sigma_m\sigma_3 \otimes \sigma_n\sigma_3) \\ &= q_{00}(\sigma_3 \otimes \sigma_3)\rho(\sigma_3 \otimes \sigma_3) + q_{01}(\sigma_3 \otimes \sigma_2)\rho(\sigma_3 \otimes \sigma_2) \\ &+ \ ... + q_{33}(\sigma_0 \otimes \sigma_0)\rho(\sigma_0 \otimes \sigma_0). \end{aligned} \tag{3.73}$$

Since $q_0 = q_3$ and $q_1 = q_2$, the probabilities $q_{mn} = (1-\mu)q_m q_n + \mu q_n \delta_{mn}$ fulfill $q_{00} = q_{33}, q_{01} = q_{32}, ..., q_{22} = q_{11}$. Therefore,

$$\Lambda^Q_{ab}((\sigma_3 \otimes \sigma_3)\rho(\sigma_3 \otimes \sigma_3)) = \Lambda^Q_{ab}(\rho). \tag{3.74}$$

□

By considering (3.74), we can use instead of $(U\otimes\mathbb{1})\rho^+(U^\dagger\otimes\mathbb{1})$, the expression $\frac{1}{2}\left[(U\otimes\mathbb{1})\rho^+(U^\dagger\otimes\mathbb{1})\right]+\frac{1}{2}\left[(\sigma_3U\otimes\sigma_3)\rho^+(U^\dagger\sigma_3\otimes\sigma_3)\right]$ which can be written as

$$\begin{aligned}
&\frac{1}{2}\left[(U\otimes\mathbb{1})\rho^+(U^\dagger\otimes\mathbb{1})\right]+\frac{1}{2}\left[(\sigma_3U\otimes\sigma_3)\rho^+(U^\dagger\sigma_3\otimes\sigma_3)\right]\\
&=\begin{pmatrix}\frac{aa^*}{2} & 0 & 0 & \frac{a^2}{2}\\ 0 & \frac{bb^*}{2} & \frac{-b^2}{2} & 0\\ 0 & \frac{-(b^*)^2}{2} & \frac{bb^*}{2} & 0\\ \frac{(a^*)^2}{2} & 0 & 0 & \frac{aa^*}{2}\end{pmatrix}\\
&=|a|^2|\Phi_1\rangle\langle\Phi_1|+|b|^2|\Phi_2\rangle\langle\Phi_2|,
\end{aligned}\tag{3.75}$$

with $|\Phi_1\rangle$ and $|\Phi_2\rangle$ to be

$$|\Phi_1\rangle=\frac{1}{\sqrt{2}}\left(\frac{a}{|a|}|00\rangle+\frac{a^*}{|a|}|11\rangle\right),\tag{3.76a}$$

$$|\Phi_2\rangle=\frac{1}{\sqrt{2}}\left(\frac{b}{|b|}|01\rangle-\frac{b^*}{|b|}|10\rangle\right).\tag{3.76b}$$

After applying the quasiclassical channel, using (3.74) and (3.75), and the concavity of the von Neumann entropy, for $S\left[\Lambda^Q_{ab}\left((U\otimes\mathbb{1})\rho^+(U^\dagger\otimes\mathbb{1})\right)\right]$ we arrive at

$$\begin{aligned}
&S\left[\Lambda^Q_{ab}\left((U\otimes\mathbb{1})\rho^+(U^\dagger\otimes\mathbb{1})\right)\right]\\
=\;&S\left[\Lambda^Q_{ab}\left(\frac{1}{2}(U\otimes\mathbb{1})\rho^+(U^\dagger\otimes\mathbb{1})+\frac{1}{2}(\sigma_3U\otimes\sigma_3)\rho^+(U^\dagger\sigma_3\otimes\sigma_3)\right)\right]\\
=\;&S\left[|a|^2\Lambda^Q_{ab}(|\Phi_1\rangle\langle\Phi_1|)+|b|^2\Lambda^Q_{ab}(|\Phi_2\rangle\langle\Phi_2|)\right]\\
\geq\;&|a|^2S\left[\Lambda^Q_{ab}(|\Phi_1\rangle\langle\Phi_1|)\right]+|b|^2S\left[\Lambda^Q_{ab}(|\Phi_2\rangle\langle\Phi_2|)\right]\\
\geq\;&S\left[\Lambda^Q_{ab}(|\Phi^+\rangle\langle\Phi^+|)\right],
\end{aligned}\tag{3.77}$$

where in the last line we have used that both $S\left[\Lambda^Q_{ab}(|\Phi_{1,2}\rangle\langle\Phi_{1,2}|)\right]$ are lower bounded by $S\left[\Lambda^Q_{ab}(|\Phi^+\rangle\langle\Phi^+|)\right]$. The proof for this statement comes in the following: A quantum state does not change up to a global phase and thus, we can rewrite $|\Phi_1\rangle$ (a similar argument holds for $|\Phi_2\rangle$) as

$$|\Phi_1\rangle=\frac{1}{\sqrt{2}}\left(|00\rangle+\exp(i\phi)|11\rangle\right).\tag{3.78}$$

After applying the *correlated* quasiclassical channel Λ^Q_{ab} to the state $|\Phi_1\rangle$, we arrive at

$$\Lambda^Q_{ab}(|\Phi_1\rangle\langle\Phi_1|) = \frac{1}{4}\left[\mathbb{1}\otimes\mathbb{1} + \left(\mu + (1-\mu)(1-2p)^2\right)\sigma_3\otimes\sigma_3 + \right.$$
$$\left.\mu\cos\phi(\sigma_1\otimes\sigma_1 - \sigma_2\otimes\sigma_2) + \mu(1-2p)\sin\phi(\sigma_1\otimes\sigma_2 + \sigma_2\otimes\sigma_1)\right]. \quad (3.79)$$

The von Neumann entropy of a quantum state is defined via its eigenvalues. The eigenvalues of (3.79) are

$$\gamma_{1,2} = (1-\mu)p(1-p)$$
$$\gamma_{3,4} = \frac{1}{2}\left(1 - 2(1-\mu)p(1-p) \pm \sqrt{\mu^2(1-4p(1-p)\sin^2\phi)}\right). \quad (3.80)$$

To minimize the von Neumann entropy $S(\Lambda^Q_{ab}(|\Phi_1\rangle\langle\Phi_1|)) = -\sum_i \gamma_i \log\gamma_i$, the eigenvalues should diverge as much as possible with respect to the parameter ϕ. The eigenvalues $\gamma_{1,2}$ are independent of ϕ. Thus, the most divergence happens when we maximize γ_3 while we minimize γ_4. This happens by $\phi = 0$ and it leads the state (3.78) to be a Bell state which proves the above statement. Therefore,

$$S\left[\Lambda^Q_{ab}(|\Phi_1\rangle\langle\Phi_1|)\right] \geq S\left[\Lambda^Q_{ab}(|\Phi^+\rangle\langle\Phi^+|)\right], \quad (3.81)$$

□

The lower bound on the von Neumann entropy (3.77) is reachable. It is not difficult to see that, the variables $a = 1$ and $b = 0$, which leads to the identity operator for U, reaches the bound. This ends successfully our attempt for fining a unitary operator which minimizes the output entropy. Therefore, the super dense coding capacity for a *correlated* quasiclassical channel, a Bell state, and $U_{min} = \mathbb{1}$, according to (3.62), is given by

$$C^{Q,B}_{un} = 2 - S\left(\Lambda^Q_{ab}\left(|\Phi^+\rangle\langle\Phi^+|\right)\right). \quad (3.82)$$

Werner states

For $d = 2$ the Werner state is $\rho_W = \eta|\Phi^+\rangle\langle\Phi^+| + \frac{1-\eta}{4}\mathbb{1}$. In order to find U_{min}, similar to the case of the Bell state in the previous part, we start by applying the most general unitary operator U on the Werner state ρ_W. Once more we use this property that the output of a quasiclassical channel is invariant under the

rotation $\sigma_3 \otimes \sigma_3$ on the input state (3.74), i.e. the channel outputs for the input states $(U \otimes \mathbb{1})\rho_W(U^\dagger \otimes \mathbb{1})$ and $(\sigma_3 U \otimes \sigma_3)\rho_W(U^\dagger \sigma_3 \otimes \sigma_3)$ are equal and therefore we can use the average of both states instead of only $(U \otimes \mathbb{1})\rho_W(U^\dagger \otimes \mathbb{1})$. Hence,

$$\begin{aligned}
&\frac{1}{2}(U \otimes \mathbb{1})\rho_W(U^\dagger \otimes \mathbb{1}) + \frac{1}{2}(\sigma_3 U \otimes \sigma_3)\rho_W(U^\dagger \sigma_3 \otimes \sigma_3) \\
= &\begin{pmatrix} \frac{\eta a a^*}{2} + \frac{1-\eta}{4} & 0 & 0 & \eta\frac{a^2}{2} \\ 0 & \eta\frac{bb^*}{2} + \frac{1-\eta}{4} & \eta\frac{-b^2}{2} & 0 \\ 0 & \eta\frac{-(b^*)^2}{2} & \eta\frac{bb^*}{2} + \frac{1-\eta}{4} & 0 \\ \eta\frac{(a^*)^2}{2} & 0 & 0 & \eta\frac{aa^*}{2} + \frac{1-\eta}{4} \end{pmatrix} \\
= &\ |a|^2\left(\eta|\Phi_1\rangle\langle\Phi_1| + \frac{1-\eta}{4}\mathbb{1}\right) + |b|^2\left(\eta|\Phi_2\rangle\langle\Phi_2| + \frac{1-\eta}{4}\mathbb{1}\right), \qquad (3.83)
\end{aligned}$$

where $|\Phi_1\rangle$ and $|\Phi_2\rangle$ are defined in (3.76a) and (3.76b). By applying the quasi-classical channel, by using (3.83), and also the concavity of the von Neumann entropy, for $S\left[\Lambda^Q_{ab}\left((U \otimes \mathbb{1})\rho_W(U^\dagger \otimes \mathbb{1})\right)\right]$ we find the upper bound

$$\begin{aligned}
&S\left[\Lambda^Q_{ab}\left((U \otimes \mathbb{1})\rho_W(U^\dagger \otimes \mathbb{1})\right)\right] \\
= &\ S\left[\Lambda^Q_{ab}\left(\frac{1}{2}(U \otimes \mathbb{1})\rho_W(U^\dagger \otimes \mathbb{1}) + \frac{1}{2}(\sigma_3 U \otimes \sigma_3)\rho_W(U^\dagger \sigma_3 \otimes \sigma_3)\right)\right] \\
= &\ S\left[|a|^2\Lambda^Q_{ab}\left(\eta|\Phi_1\rangle\langle\Phi_1| + \frac{1-\eta}{4}\mathbb{1}\right) + |b|^2\Lambda^Q_{ab}\left(\eta|\Phi_2\rangle\langle\Phi_2| + \frac{1-\eta}{4}\mathbb{1}\right)\right] \\
= &\ S\left[\Lambda^Q_{ab}\left(|a|^2\left(\eta|\Phi_1\rangle\langle\Phi_1| + \frac{1-\eta}{4}\mathbb{1}\right) + |b|^2\left(\eta|\Phi_2\rangle\langle\Phi_2| + \frac{1-\eta}{4}\mathbb{1}\right)\right)\right] \\
\geq &\ |a|^2 S\left[\Lambda^Q_{ab}\left(\eta|\Phi_1\rangle\langle\Phi_1| + \frac{1-\eta}{4}\mathbb{1}\right)\right] + |b|^2 S\left[\Lambda^Q_{ab}\left(\eta|\Phi_2\rangle\langle\Phi_2| + \frac{1-\eta}{4}\mathbb{1}\right)\right] \\
\geq &\ S\left[\Lambda^Q_{ab}\left(\eta|\Phi^+\rangle\langle\Phi^+| + \frac{1-\eta}{4}\mathbb{1}\right)\right]. \qquad (3.84)
\end{aligned}$$

In the last line we have used that both $S\left[\Lambda^Q_{ab}\left(\eta|\Phi_{1,2}\rangle\langle\Phi_{1,2}| + \frac{1-\eta}{4}\mathbb{1}\right)\right]$ are lower bounded by $S\left[\Lambda^Q_{ab}\left(\eta|\Phi^+\rangle\langle\Phi^+| + \frac{1-\eta}{4}\mathbb{1}\right)\right]$. The proof for this statement is similar to the Bell state case and is based on the eigenvalue computation. We let ν_i to be the spectrum of $\Lambda^Q_{ab}\left(\eta|\Phi_1\rangle\langle\Phi_1| + \frac{1-\eta}{4}\mathbb{1}\right)$. Since the channel is linear then $\Lambda^Q_{ab}\left(\eta|\Phi_1\rangle\langle\Phi_1| + \frac{1-\eta}{4}\mathbb{1}\right) = \eta\Lambda^Q_{ab}\left(|\Phi_1\rangle\langle\Phi_1|\right) + \frac{1-\eta}{4}\mathbb{1}$. Thus, its spectrum can be expressed in terms of γ_i (3.80) which are the eigenvalues of $\Lambda^Q_{ab}\left(|\Phi_1\rangle\langle\Phi_1|\right)$. Hence,

for ν_i we have

$$\begin{aligned}\nu_{1,2} &= \eta\gamma_{1,2} + \frac{1-\eta}{4},\\ \nu_{3,4} &= \eta\gamma_{3,4} + \frac{1-\eta}{4}.\end{aligned} \tag{3.85}$$

As we can see in (3.85) that the eigenvalues ν_i have the most divergence when γ_i are maximally divergent. As in the previous part this is reached by $\phi = 0$. Therefore, the minimum of the von Neumann entropy $S\left[\Lambda^Q_{ab}\left(\eta|\Phi_1\rangle\langle\Phi_1| + \frac{1-\eta}{4}\mathbb{1}\right)\right] = -\sum_i \nu_i \log \nu_i$ is obtained for $\phi = 0$ which leads to the entropy $S\Big[\Lambda^Q_{ab}\left(\eta|\Phi^+\rangle\langle\Phi^+| + \frac{1-\eta}{4}\mathbb{1}\right)\Big]$. Therefore, for any phase ϕ,

$$S\Big[\Lambda^Q_{ab}\left(\eta|\Phi_{1,2}\rangle\langle\Phi_{1,2}| + \frac{1-\eta}{4}\mathbb{1}\right)\Big] \geqslant S\Big[\Lambda^Q_{ab}\left(\eta|\Phi^+\rangle\langle\Phi^+| + \frac{1-\eta}{4}\mathbb{1}\right)\Big], \tag{3.86}$$

which completes this part of our proof. □

Now, we reach for $a = 1$ and $b = 0$ the lower bound (3.84) and thus $U_{min} = \mathbb{1}$. The super dense coding capacity for a shared Werner state and in the presence of a *correlated* quasiclassical channel is, according to (3.62), given by

$$C^{Q,W}_{un} = 2 - S\left(\Lambda^Q_{ab}\left(\rho_W\right)\right). \tag{3.87}$$

3.3.3 Fully correlated Pauli channel

In this section we give two more examples for which U_{min} is determined. Both examples are for *fully correlated* Pauli channel while the initial states are the Bell state and the Werner state. A *fully correlated* Pauli channel is a special form of a *correlated* Pauli channel (3.60) when $\mu = 1$. For $d = 2$ it is

$$\Lambda^f_{ab}(\cdot) \;=\; \sum_m q_m(\sigma_m \otimes \sigma_m)(\cdot)(\sigma_m \otimes \sigma_m). \tag{3.88}$$

where $\sum_m q_m = 1$ and σ_m are either identity or Pauli operators.

Bell states

As we defined the problem before, according to find the capacity (3.62), we should find U_{min}. To do so, we use the property that the Bell state ρ^+ is invariant

under the unitary transformation $(\sigma_m \otimes \sigma_m)$, i.e. $(\sigma_m \otimes \sigma_m)\rho^+(\sigma_m \otimes \sigma_m) = \rho^+$ and therefore is invariant under the action of a *fully correlated* Pauli channel. By using the invariance, the von Neumann entropy $S\left[\Lambda^f_{ab}\left((U \otimes \mathbb{1})\rho^+(U^\dagger \otimes \mathbb{1})\right)\right]$ takes a zero value by $U_{min} = \mathbb{1}$. Consequently, the super dense coding capacity, according to (3.62), for a Bell state and a *correlated* Pauli channel, is *two* bits. It is the maximum information transfer for $d = 2$. It shows that no information at all is lost to the environment and this class of channels behave like a noiseless one. This is a known result in the literatures of channel capacity.

Werner states

A *fully correlated* Pauli channel (3.88) with a Werner state ρ_W as a shared state between Alice and Bob is the last example for which we determine the operation U_{min}. To do so, we derive a lower bound on $S\left(\Lambda^f_{ab}\left((U \otimes \mathbb{1})\rho_W(U^\dagger \otimes \mathbb{1})\right)\right)$ with U to be an arbitrary unitary operator. By using the concavity of the von Neumann entropy and also by using the invariance of the von Neumann entropy under the unitary transformation, the lower bound on $S\left(\Lambda^f_{ab}\left((U \otimes \mathbb{1})\rho_W(U^\dagger \otimes \mathbb{1})\right)\right)$ is

$$\begin{aligned} & S\left(\Lambda^f_{ab}\left((U \otimes \mathbb{1})\rho_W(U^\dagger \otimes \mathbb{1})\right)\right) \\ = \; & S\left(\sum_m q_m(\sigma_m \otimes \sigma_m)(U \otimes \mathbb{1})(\eta\rho^+ + \frac{1-\eta}{4}\mathbb{1})(U^\dagger \otimes \mathbb{1})(\sigma_m \otimes \sigma_m)\right) \\ \geq \; & S\left(\eta\rho^+ + \frac{1-\eta}{4}\mathbb{1}\right). \end{aligned} \tag{3.89}$$

By using the invariance of a Bell state under the action of a *fully correlated* Pauli channel, i.e. $\Lambda^f_{ab}(\rho^+) = \rho^+$, it follows that the lower bound (3.89) is reachable by the identity operator. Then $U_{min} = \mathbb{1}$ and the super dense coding capacity, according to (3.62), is given by

$$C^{f,W}_{un} = 2 - S\left(\Lambda^f_{ab}\left(\rho_W\right)\right). \tag{3.90}$$

3.3.4 Non-unitary encoding

So far, we have assumed that the encoding in the super dense coding protocol is unitary. The super dense coding protocol with non-unitary encoding for noiseless

channels has been discussed by M. Horodecki et al. [26], M. Horodecki and Piani [27], and Winter [50]. In this part we consider the possibility of performing non-unitary encoding in the presence of a *correlated* Pauli channel. Let us consider Γ_i to be a completely positive trace preserving (CPTP) map. Alice applies the map Γ_i on her side of the shared state ρ, thereby encoding ρ as $\rho_i = [\Gamma_i \otimes \mathbb{1}](\rho) := \Gamma_i(\rho)$. The rest of the scheme is similar to the case of unitary encoding. Alice sends the encoded state $\rho_i = \Gamma_i(\rho)$ with the probability p_i to Bob through the *correlated* Pauli channel Λ^P_{ab}. Now, the question is: which ensemble of CPTP maps gives the super dense coding capacity? In other language, what is the optimum Holevo quantity with respect to the encoding Γ_i and p_i? To answer this question, first we give the expression for the super dense coding capacity with a *correlated* Pauli channel and non-unitary encoding:

$$C = \max_{\{\Gamma_i, p_i\}} \chi = \max_{\{\Gamma_i, p_i\}} \left(S\left[\sum_i p_i \Lambda^P_{ab}\left(\Gamma_i(\rho)\right)\right] - \sum_i p_i S\left[\Lambda^P_{ab}\left(\Gamma_i(\rho)\right)\right] \right), \quad (3.91)$$

where $\Lambda^P_{ab}(\rho)$ is defined via (3.60). Similar to the unitary encoding case in section 3.3.1, we find an upper bound on the Holevo quantity (3.91) and then we show that this upper bound is reachable by a pre-processing before unitary encoding. The above statement will be expressed in the following Lemma.

Lemma 6. Let χ be the Holevo quantity (3.91), and Let $\Gamma_{min}(\cdot) := [\Gamma_{min} \otimes \mathbb{1}](\cdot)$ be the map that minimizes the von Neumann entropy after application of this map and the channel Λ^P_{ab} to the initial state ρ, i.e. Γ_{min} minimizes the expression $S\left(\Lambda^P_{ab}(\Gamma_{min}(\rho))\right)$. Then the super dense coding capacity is given by

$$C^P_{non-un} = \log d + S\left(\Lambda^P_b(\rho_b)\right) - S\left(\Lambda^P_{ab}(\Gamma_{min}(\rho))\right). \quad (3.92)$$

where $\rho_b = tr_a \rho$ and Λ^P_b is the $d-$dimensional Pauli channel.

Proof: $\Gamma_{min}(\cdot)$ is a map that leads to the minimum of the entropy after applying it and the channel to the initial state ρ. Therefore,

$$\begin{aligned} \chi &= S\left(\sum_i p_i \Lambda^P_{ab}\left(\Gamma_i(\rho)\right)\right) - \sum_i p_i S\left(\Lambda^P_{ab}\left(\Gamma_i(\rho)\right)\right) \\ &\leq S\left(\sum_i p_i \Lambda^P_{ab}\left(\Gamma_i(\rho)\right)\right) - S\left(\Lambda^P_{ab}(\Gamma_{min}(\rho))\right). \end{aligned}$$

Since the von Neumann entropy is subadditive and since the maximum entropy of a d-dimensional system is $\log d$, we have

$$\chi \leq \log d + S\left(\mathrm{tr}_a\left(\sum_i p_i \Lambda_{ab}^P\left(\Gamma_i(\rho)\right)\right)\right) - S\left(\Lambda_{ab}^P\left(\Gamma_{min}(\rho)\right)\right).$$

By using $\mathrm{tr}_a \sum_i p_i \Lambda_{ab}^P\left(\Gamma_i(\rho)\right) = \Lambda_b^P(\rho_b)$, we find the upper bound

$$\chi \leq \log d + S\left(\Lambda_b^P(\rho_b)\right) - S\left(\Lambda_{ab}^P\left(\Gamma_{min}(\rho)\right)\right). \tag{3.93}$$

Now, we show that the ensemble $\{\tilde{p}_i, \tilde{\Gamma}_i(\rho)\}$ with $\tilde{p}_i = \frac{1}{d^2}$ and $\tilde{\Gamma}_i(\rho) = (V_i \otimes \mathbb{1})\Gamma_{min}(\rho)(V_i^\dagger \otimes \mathbb{1})$, where V_i is defined in (3.14) reaches the upper bound (3.93). In the other words, the optimal encoding consists of a fixed pre-processing with Γ_{min} and a subsequent unitary encoding. This is analogous to the case of noiseless channels, for which the same statement was shown in [27]. Below we prove the above claim.

The Holevo quantity of the ensemble $\{\tilde{p}_i, \tilde{\Gamma}_i(\rho)\}$ is

$$\tilde{\chi} = S\left(\sum_i \frac{1}{d^2}\Lambda_{ab}^P\left(\tilde{\Gamma}_i(\rho)\right)\right) - \sum_i \frac{1}{d^2} S\left[\Lambda_{ab}^P\left(\tilde{\Gamma}_i(\rho)\right)\right]. \tag{3.94}$$

By noting that Γ_{min} acts only on Alice's side, using (3.17), and Lemma (1-a), we find that the average of $\Lambda_{ab}^P\left(\tilde{\Gamma}_i(\rho)\right)$, i.e. the argument in the first term on the RHS of (3.94), is given by

$$\sum_i \frac{1}{d^2}\Lambda_{ab}^P\left(\tilde{\Gamma}_i(\rho)\right) = \frac{\mathbb{1}}{d} \otimes \Lambda_b^P(\rho_b). \tag{3.95}$$

Furthermore, the second term on the RHS of (3.94) is given by

$$\begin{aligned}
&\sum_i \frac{1}{d^2} S\left(\Lambda_{ab}^P\left(\tilde{\Gamma}_i(\rho)\right)\right) = \sum_i \frac{1}{d^2} S\left(\Lambda_{ab}^P\left(\left(V_i \otimes \mathbb{1}\right)\Gamma_{min}(\rho)\left(V_i^\dagger \otimes \mathbb{1}\right)\right)\right) \\
= &\frac{1}{d^2}\sum_i S\left(\left(V_i \otimes \mathbb{1}\right)\left[\sum_{m,n,\tilde{m},\tilde{n}=0}^{d-1} q_{mn\tilde{m}\tilde{n}}\left(V_{mn} \otimes V_{\tilde{m}\tilde{n}}\right)\Gamma_{min}(\rho)\left(V_{mn}^\dagger \otimes V_{\tilde{m}\tilde{n}}^\dagger\right)\right]\right. \\
&\left.\left(V_i^\dagger \otimes \mathbb{1}\right)\right) \\
= &\frac{1}{d^2}\sum_i S\left[\Lambda_{ab}^P\left(\Gamma_{min}(\rho)\right)\right] = S\left[\Lambda_{ab}^P\left(\Gamma_{min}(\rho)\right)\right],
\end{aligned} \tag{3.96}$$

where in the second line of the above equations we have inserted the action of the *correlated* Pauli channel, and we have used (3.35), from which it follows that V_i and V_{mn} commute up to a phase.

Inserting (3.95) and (3.96) into (3.94), one finds that the Holevo quantity $\tilde{\chi}$ is equal to the upper bound given in (3.93). Consequently, the Lemma has been proved and the super dense coding capacity with non-unitary encoding is determined by (3.92). □

Comparing (3.92) and (3.62) shows that applying the appropriate pre-processing Γ_{min} on the initial state ρ before the unitary encoding $\{V_i\}$ may increase the super dense coding capacity, with respect to only using unitary encoding for the case of a *correlated* Pauli channel. However, for some examples no better encoding than unitary encoding is possible. For instance, since *two* bits is the highest super dense coding capacity for $d = 2$, our results derived in section 3.3.3 for *fully correlated* Pauli channel and the Bell state provide an example where no pre-processing can improve the capacity.

Note that the present section is being currently prepared for publication [44].

3.4 Multipartite super dense coding in the presence of noise

The notion of multipartite super dense coding has been introduced by Bose *et al.* [8] which generalizes the Bennett-Wiesner scheme [6] of super dense coding to multipartites. In this scheme, it was shown that the use of the multipartite entangled state can allow a single receiver to read messages from more than one source through a single measurement. This task has been done in the following way. $(N+1)$ parties share an $(N+1)$-particle maximally entangled state, i.e. a $|GHZ\rangle$ state,

$$|GHZ\rangle = \frac{|0\rangle^{\otimes N+1} + |1\rangle^{\otimes N+1}}{\sqrt{2}}, \tag{3.97}$$

possessing one particle each. N-parties, whom are referred as senders, are tent to send their messages to one receiver (Bob). The N senders decide in advance

to perform only certain unitary operations on the particles given to them. One of the senders have any of four possible unitary operations at her disposal, while each of the others have any of two possible unitary operations. That is, one of the senders encodes two bits on her particle, while the other senders encode one bit each. The unitary operations must be chosen so that for each possible combination of unitary transformations performed by senders, the state of $(N+1)$ particles changes to another member of the set of the maximally entangled states. A known set of such unitary operations is the set of the Pauli operators (3.70). The number of possible combination of unitary transformations by N senders is then $4 \times 2 \times 2... \times 2 = 2^{N+1}$. In the other hand, there are exactly 2^{N+1} states in the set of maximally entangled states of $(N+1)$ particles. After performing their unitary transformations, N senders send their particles to Bob which have now all $(N+1)$ particles at disposal. Bob performs an appropriate measurement on $(N+1)$ particles, which identifies the maximally entangled state in which particles are. Bob can then learn about the messages sent by each of the senders by the result of his single measurement. Comparing the rate of the information gain in the above scheme with the case that Bob performs a super dense coding [6] with each of N senders shows the efficiency of the above scheme (for more details about the rate of information gain see[8]). It is more efficient because it requires only $(N+1)$ particle as apposed to super dense coding communication with each sender, which requires $2N$ particles [8]. A generalization of this multipartite super dense coding to higher dimensions has also been discussed by Liu *et al.* [35].

Another generalization of distributed super dense coding for noiseless channels has been widely discussed by Bruß *et al.* [11]. In this paper, they follow a different goal than the above multipartite scheme. The problem is to find the optimal unitary encoding for multipartite super dense coding. Two scenarios of many senders with either one or two receiver(s) has been discussed. It has been shown that for a single receiver, Alices do not need to apply global unitaries to gain the optimal super dense coding capacity but it's enough that each Alice performs a local encoding on her side. It has also been shown that bound entangled states of bipartite cut between Alices and Bob are not "multi" dense-codeable. For the case of two receivers, some of the Alices send their information to the first Bob while the others send theirs to the second Bob. Then the super dense coding capacity

depends on the possibility of interaction between the receivers. If two receivers are not allowed to communicate, the corresponding capacities are additive while when they are allowed to use global measurements, they can obtain higher capacity. Furthermore, a general classification of multipartite quantum states according to their dense-codeability has been investigated.

Up to now, for noisy channels, we have only investigated the case of a single sender and a single receiver. In this section, we give an example for multipartite super dense coding considering the Pauli channel and unitary encoding. The multipartite problem that will be discussed here, is a generalization of the bipartite scenario discussed in section 3.2.2 for the case of a *one-sided* Pauli channel and a bipartite Bell state. We assume that there exist k-Alices $(A_1, ..., A_k)$ and k-Bobs $(B_1, ..., B_k)$ with the shared state $\rho_{00}^{a_1 b_1} \otimes ... \otimes \rho_{00}^{a_k b_k}$. The Bell state $\rho_{00}^{a_j b_j} = |\Phi_{00}\rangle\langle\Phi_{00}|^{a_j b_j}$ is the shared state between the jth Alice & Bob, where $|\Phi_{00}\rangle$ is defined in (3.36). The dimension of the jth Alice & Bob system is given by d_j^2. In this scenario, the Alices apply a global unitary operation $W_{i_1...i_k}^{a_1...a_k}$ on their side of the shared state $\rho_{00}^{a_1 b_1} \otimes ... \otimes \rho_{00}^{a_k b_k}$ and encode the state through

$$\rho_{i_1...i_k} = \left(W_{i_1...i_k}^{a_1...a_k} \otimes \mathbb{1}^{b_1...b_k}\right) \left(\rho_{00}^{a_1 b_1} \otimes ... \otimes \rho_{00}^{a_k b_k}\right) \left(W_{i_1...i_k}^{a_1...a_k \dagger} \otimes \mathbb{1}^{b_1...b_k}\right). \quad (3.98)$$

Then each of Alices send their d_j-dimensional subsystem through a *one-sided* Pauli channel (3.31) to the corresponding Bob. The probability of sending the encoded state $\rho_{i_1...i_k}$ is given by $p_{i_1...i_k}$ (see also Figure 3.6). Then the ensemble that the Bobs receive is $\{\Lambda_{a_1...a_k}^P(\rho_{i_1...i_k}), p_{i_1...i_k}\}$ where $\Lambda_{a_1...a_k}$ is a channel which globally acts on Alices' subsystem with the definition

$$\Lambda_{a_1...a_k}^P(\cdot) = \sum_{m_1 n_1 ... m_k n_k} q_{m_1 n_1 ... m_k n_k} \left(V_{m_1 n_1}^{a_1} \otimes ... \otimes V_{m_k n_k}^{a_k}\right)(\cdot)\left(V_{m_1 n_1}^{a_1 \dagger} \otimes ... \otimes V_{m_k n_k}^{a_k \dagger}\right), \quad (3.99)$$

where $V_{m_j n_j}^{a_j}$ are the *displacement* operators for the jth Alice which are defined in (3.32). The probabilities $q_{m_1 n_1 ... m_k n_k}$ add to *one* and can represent the correlation between the channels. Similar to the bipartite case, the amount of classical

information that the Bobs receive is given by the Holevo quantity

$$\begin{aligned}\chi\left(\{\rho_{i_1\ldots i_k}, p_{i_1\ldots i_k}\}\right) &= S\left[\sum_i p_{i_1\ldots i_k}\Lambda^P_{a_1\ldots a_k}\left(\rho_{i_1\ldots i_k}\right)\right] \\ &- \sum_i p_{i_1\ldots i_k} S\left[\Lambda^P_{a_1\ldots a_k}\left(\rho_{i_1\ldots i_k}\right)\right], \end{aligned} \tag{3.100}$$

and the super dense coding capacity is the optimum of this quantity with respect to the encoding $\{W^{a_1\ldots a_k}_{i_1\ldots i_k}, p_{i_1\ldots i_k}\}$, i.e.

$$C = \max_{\{W^{a_1\ldots a_k}_{i_1\ldots i_k}, p_{i_1\ldots i_k}\}} \chi\left(\{\rho_{i_1\ldots i_k}, p_{i_1\ldots i_k}\}\right). \tag{3.101}$$

Now, we focus on the optimization procedure. We prove that the second term of the Holevo quantity (3.100) is invariant under unitary rotation $U^{a_1\ldots a_k}$ of the state $\rho^{a_1b_1}_{00}\otimes\ldots\otimes\rho^{a_kb_k}_{00}$. Therefore, the optimization only runs over the first term on the RHS of the Holevo quantity (3.100). To show this claim, we first prove the following Lemma.

Lemma 7. Let

$$\rho^{a_1b_1}_{00}\otimes\ldots\otimes\rho^{a_kb_k}_{00} = |\Phi^{a_1b_1}_{00}\ldots\Phi^{a_kb_k}_{00}\rangle\langle\Phi^{a_1b_1}_{00}\ldots\Phi^{a_kb_k}_{00}|, \tag{3.102}$$

be $k-$copies of the Bell states with different dimensions. Let us define

$$\begin{aligned}\pi_{m_1n_1\ldots m_kn_k} := &\left(V^{a_1}_{m_1n_1}\otimes\ldots\otimes V^{a_k}_{m_kn_k}\otimes\mathbb{1}^{b_1\ldots b_k}\right)\left(U^{a_1\ldots a_k}\otimes\mathbb{1}^{b_1\ldots b_k}\right) \\ &\left(\rho^{a_1b_1}_{00}\otimes\ldots\otimes\rho^{a_kb_k}_{00}\right)\left(U^{a_1\ldots a_k\dagger}\otimes\mathbb{1}^{b_1\ldots b_k}\right)\left(V^{a_1\dagger}_{m_1n_1}\otimes\ldots\otimes V^{a_k\dagger}_{m_kn_k}\otimes\mathbb{1}^{b_1\ldots b_k}\right),\end{aligned} \tag{3.103}$$

where $U^{a_1\ldots a_k}$ is a unitary operator with the dimension of the system of all Alices and $V^{a_j}_{m_jn_j}$ being the *displacement* operators (3.32). For different states $\pi_{m_1n_1\ldots m_kn_k}$,

$$\pi_{m_1n_1\ldots m_kn_k}\pi_{\tilde{m_1}\tilde{n_1}\ldots\tilde{m_k}\tilde{n_k}} = 0 \tag{3.104}$$

holds.

Proof. To prove the Lemma, first we show that the statement

$$\begin{aligned}&\langle\Phi^{a_1b_1}_{00}\ldots\Phi^{a_kb_k}_{00}|(U^{a_1\ldots a_k\dagger})\left(V^{a_1\dagger}_{m_1n_1}V^{a_1}_{\tilde{m_1}\tilde{n_1}}\otimes\ldots\otimes V^{a_k\dagger}_{m_kn_k}V^{a_k}_{\tilde{m_k}\tilde{n_k}}\otimes\mathbb{1}^{b_1\ldots b_k}\right) \\ &\quad(U^{a_1\ldots a_k})\,|\Phi^{a_1b_1}_{00}\ldots\Phi^{a_kb_k}_{00}\rangle = 0,\end{aligned} \tag{3.105}$$

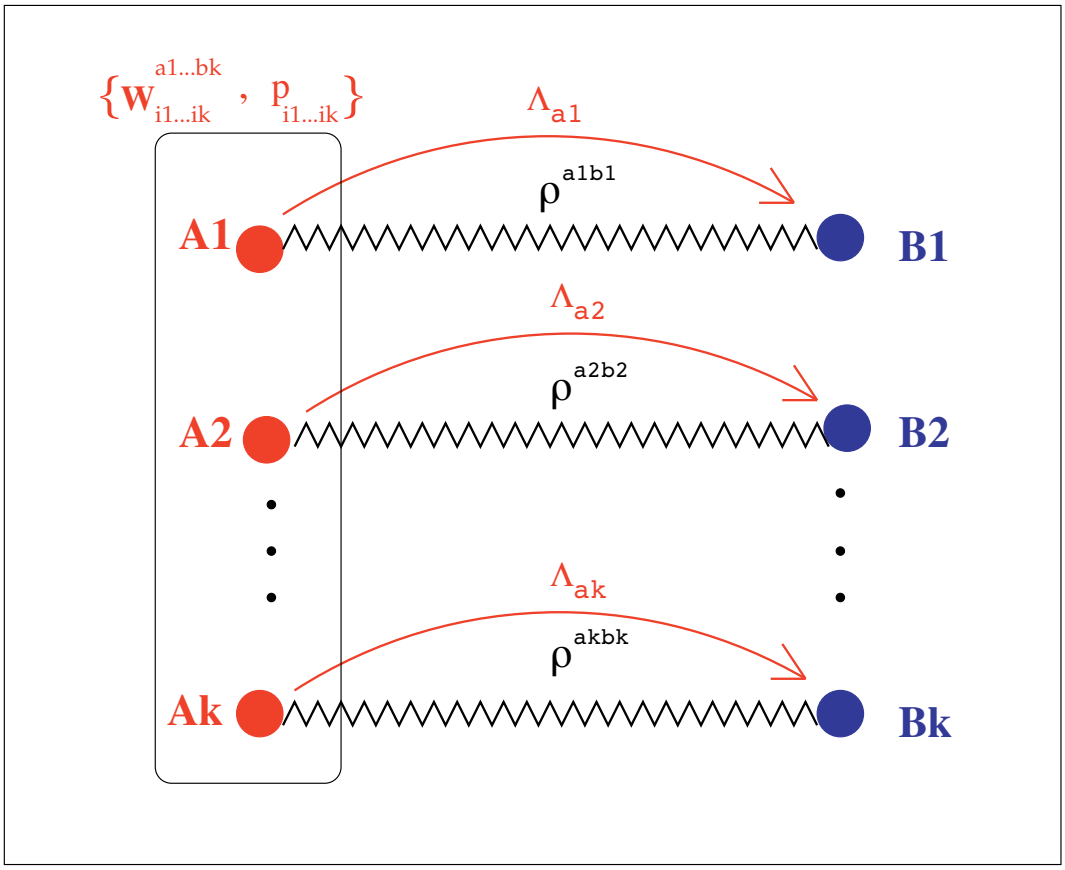

Figure 3.6: The multipartite super dense coding for k-copies of Bell states with different dimensions. Alices apply a global unitary operator $W_{i_1\ldots i_k}^{a_1\ldots a_k}$, taken from a set of $\{W_{i_1\ldots i_k}^{a_1\ldots a_k}\}$ with the probability $\{p_{i_1\ldots i_k}\}$, on their side of the shared state $\rho_{00}^{a_1 b_1} \otimes \ldots \otimes \rho_{00}^{a_k b_k}$. The Bell state $\rho_{00}^{a_j b_j}$ is shared between the jth Alice & Bob and the *one-sided* Pauli channel $\Lambda_{a_j}^P$ acts on the subsystem of the jth Alice after encoding. Since the channels can be correlated, the action of a global channel on Alices' system has been denoted by $\Lambda_{a_1,\ldots,a_k}^P$. We have considered that there is no noise on the Bobs' side.

holds. By using the definition (3.36) for a Bell state we have

$$\begin{aligned}
&\langle\Phi_{00}^{a_1b_1}...\Phi_{00}^{a_kb_k}|(U^{a_1...a_k\dagger})\left(V_{m_1n_1}^{a_1\dagger}V_{\tilde{m_1}\tilde{n_1}}^{a_1}\otimes...\otimes V_{m_kn_k}^{a_k\dagger}V_{\tilde{m_k}\tilde{n_k}}^{a_k}\otimes\mathbb{1}^{b_1...b_k}\right)(U^{a_1...a_k})\\
&|\Phi_{00}^{a_1b_1}...\Phi_{00}^{a_kb_k}\rangle\\
=&\sum_{j_1...j_k}\sum_{\tilde{j_1}...\tilde{j_k}}\langle j_1j_1...j_kj_k|(U^{a_1...a_k\dagger})\left(V_{m_1n_1}^{a_1\dagger}V_{\tilde{m_1}\tilde{n_1}}^{a_1}\otimes...\otimes V_{m_kn_k}^{a_k\dagger}V_{\tilde{m_k}\tilde{n_k}}^{a_k}\otimes\mathbb{1}^{b_1...b_k}\right)\\
&(U^{a_1...a_k})\,|\tilde{j_1}\tilde{j_1}...\tilde{j_k}\tilde{j_k}\rangle\\
=&\sum_{j_1...j_k}\langle j_1...j_k|U^{a_1...a_k\dagger}\left(V_{m_1n_1}^{a_1\dagger}V_{\tilde{m_1}\tilde{n_1}}^{a_1}\otimes...\otimes V_{m_kn_k}^{a_k\dagger}V_{\tilde{m_k}\tilde{n_k}}^{a_k}\right)U^{a_1...a_k}|j_1...j_k\rangle\\
=&\;tr_{a_1...a_k}\left[U^{a_1...a_k\dagger}\left(V_{m_1n_1}^{a_1\dagger}V_{\tilde{m_1}\tilde{n_1}}^{a_1}\otimes...\otimes V_{m_kn_k}^{a_k\dagger}V_{\tilde{m_k}\tilde{n_k}}^{a_k}\right)U^{a_1...a_k}\right]\\
=&\;tr_{a_1...a_k}\left[V_{m_1n_1}^{a_1\dagger}V_{\tilde{m_1}\tilde{n_1}}^{a_1}\otimes...\otimes V_{m_kn_k}^{a_k\dagger}V_{\tilde{m_k}\tilde{n_k}}^{a_k}\right]\\
=&\;\delta_{m_1\tilde{m_1}}\delta_{n_1\tilde{n_1}}...\delta_{m_k\tilde{m_k}}\delta_{n_k\tilde{n_k}},
\end{aligned}\tag{3.106}$$

where in the last line we have used $\mathrm{tr}\,V_{mn}V_{\tilde{m}\tilde{n}}^{\dagger}=d\delta_{m\tilde{m}}\delta_{n\tilde{n}}$. Different states $\pi_{m_1n_1...m_kn_k}$ have at lease one different indice for m_i or n_i. Then by using (3.106), the statement (3.105) is proved. Subsequently, we arrive at

$$\begin{aligned}
&\left(\rho_{00}^{a_1b_1}\otimes...\otimes\rho_{00}^{a_kb_k}\right)(U^{a_1...a_k\dagger})\left(V_{m_1n_1}^{a_1\dagger}V_{\tilde{m_1}\tilde{n_1}}^{a_1}\otimes...\otimes V_{m_kn_k}^{a_k\dagger}V_{\tilde{m_k}\tilde{n_k}}^{a_k}\otimes\mathbb{1}^{b_1...b_k}\right)\\
&(U^{a_1...a_k})\left(\rho_{00}^{a_1b_1}\otimes...\otimes\rho_{00}^{a_kb_k}\right)=0.
\end{aligned}\tag{3.107}$$

By using (3.107), for $\pi_{m_1n_1...m_kn_k}\pi_{\tilde{m_1}\tilde{n_1}...\tilde{m_k}\tilde{n_k}}$ we have

$$\begin{aligned}
\pi_{m_1n_1...m_kn_k}\pi_{\tilde{m_1}\tilde{n_1}...\tilde{m_k}\tilde{n_k}}=&\left(V_{m_1n_1}^{a_1}\otimes...\otimes V_{m_kn_k}^{a_k}\otimes\mathbb{1}^{b_1...b_k}\right)(U^{a_1...a_k})\\
&\underbrace{\left(\rho_{00}^{a_1b_1}\otimes...\otimes\rho_{00}^{a_kb_k}\right)(U^{a_1...a_k\dagger})\left(V_{m_1n_1}^{a_1\dagger}V_{\tilde{m_1}\tilde{n_1}}^{a_1}\otimes...\otimes V_{m_kn_k}^{a_k\dagger}V_{\tilde{m_k}\tilde{n_k}}^{a_k}\otimes\mathbb{1}^{b_1...b_k}\right)}\\
&\underbrace{(U^{a_1...a_k})\left(\rho_{00}^{a_1b_1}\otimes...\otimes\rho_{00}^{a_kb_k}\right)}_{=0}(U^{a_1...a_k\dagger})\left(V_{\tilde{m_1}\tilde{n_1}}^{a_1}\otimes...\otimes V_{\tilde{m_k}\tilde{n_k}}^{a_k}\otimes\mathbb{1}^{b_1...b_k}\right)=0,
\end{aligned}$$

which completes the proof. □

We now demonstrate that the second term of the Holevo quantity is independent of the chosen unitary encoding. Using the orthogonality property (3.104) and the purity of the density operator $\pi_{m_1n_1...m_kn_k}$, the channel output entropy

can be written as

$$
\begin{aligned}
& S\left[\Lambda_{a_1\ldots a_k}\left(\left(U^{a_1\ldots a_k}\otimes \mathbb{1}^{b_1\ldots b_k}\right)\left(\rho^{a_1b_1}\otimes\ldots\otimes\rho^{a_kb_k}\right)\left(U^{a_1\ldots a_k\dagger}\otimes \mathbb{1}^{b_1\ldots b_k}\right)\right)\right] \\
= \; & S\Big[\sum_{m_1n_1\ldots m_kn_k} q_{m_1n_1\ldots m_kn_k}\left(V^{a_1}_{m_1n_1}\otimes\ldots\otimes V^{a_k}_{m_kn_k}\otimes \mathbb{1}^{b_1\ldots b_k}\right)\left(U^{a_1\ldots a_k}\otimes \mathbb{1}^{b_1\ldots b_k}\right) \\
& \left(\rho_{00}^{a_1b_1}\otimes\ldots\otimes\rho_{00}^{a_kb_k}\right)\left(U^{a_1\ldots a_k\dagger}\otimes \mathbb{1}^{b_1\ldots b_k}\right)\left(V^{a_1\dagger}_{m_1n_1}\otimes\ldots\otimes V^{a_k\dagger}_{m_kn_k}\otimes \mathbb{1}^{b_1\ldots b_k}\right)\Big] \\
= \; & S\left[\sum_{m_1n_1\ldots m_kn_k} q_{m_1n_1\ldots m_kn_k}\pi_{m_1n_1\ldots m_kn_k}\right] = H\left(\{q_{m_1n_1\ldots m_kn_k}\}\right), \qquad (3.108)
\end{aligned}
$$

where $H\left(\{q_{m_1n_1\ldots m_kn_k}\}\right) = -\sum_{m_1n_1\ldots m_kn_k} q_{m_1n_1\ldots m_kn_k}\log q_{m_1n_1\ldots m_kn_k}$ is the Shannon entropy. Consequently, the channel output entropy, and thus the second term of the Holevo quantity, is just determined by the channel probabilities $q_{m_1n_1\ldots m_kn_k}$ and it is invariant under unitary encoding.

Here, by using the above result, we just focus on the first term of the Holevo quantity to find the optimal encoding. Since the von Neumann is subadditive, and since the maximum von Neumann entropy for a d-dimensional system is $\log d$, we find the upper bound

$$
\begin{aligned}
& S\left(\sum_{i_1\ldots i_k} p_{i_1\ldots i_k}\Lambda_{a_1\ldots a_k}(\rho_{i_1\ldots i_k})\right) \\
\leq \; & S\left(\mathrm{tr}_{b_1\ldots a_kb_k}\sum_{i_1\ldots i_k} p_{i_1\ldots i_k}\Lambda_{a_1\ldots a_k}(\rho_{i_1\ldots i_k})\right) \\
+ \; & S\left[\mathrm{tr}_{a_1\ldots a_kb_k}\sum_{i_1\ldots i_k} p_{i_1\ldots i_k}\Lambda_{a_1\ldots a_k}(\rho_{i_1\ldots i_k})\right] \\
+ \; & \ldots + S\left[\mathrm{tr}_{a_1b_1\ldots a_k}\sum_{i_1\ldots i_k} p_{i_1\ldots i_k}\Lambda_{a_1\ldots a_k}(\rho_{i_1\ldots i_k})\right] \\
\leq \; & \log d_1 + \log d_1 + \ldots + \log d_k. \qquad (3.109)
\end{aligned}
$$

The above bound can be reached by the set of the product encoding $V^{a_1\ldots a_k}_{i_1\ldots i_k} = V^{a_1}_{i_1}\otimes\ldots\otimes V^{a_k}_{i_k}$ which are chosen with equal probability $\tilde{p}_{i_1\ldots i_k} = \frac{1}{d_1^2d_2^2\ldots d_k^2}$. The local unitary operators $V^{a_j}_{i_j}$ are defined in (3.14).

Now we show that the above ensemble reaches the bound in (3.109). In the following the symbol $\tilde{\tau}_{i_1\ldots i_k}$ denotes the resource state after encoding with $V_{i_1\ldots i_k}^{a_1,\ldots,a_k}$. The ensemble average after the specific encoding with $\{V_{i_1\ldots i_k}^{a_1\ldots a_k}\}$, the probability distribution $\tilde{p}_{i_1\ldots i_k} = \frac{1}{d_1^2 d_2^2 \ldots d_k^2}$, and after action of the channel will be denoted as $\tilde{\rho}_{i_1\ldots i_k}$. Then, by using Lemma (1-a) and the linearity of the channel, the entropy of the average ensemble state $\tilde{\rho}_{i_1\ldots i_k}$ is given by

$$\begin{aligned}
S(\tilde{\rho}_{i_1\ldots i_k}) &= S\left(\sum_{i_1\ldots i_k} \tilde{p}_{i_1\ldots i_k} \Lambda_{a_1\ldots a_k}\left(\tilde{\tau}_{i_1\ldots i_k}\right)\right) \\
&= S\left(\Lambda_{a_1\ldots a_k}\left(\sum_{i_1\ldots i_k} \tilde{p}_{i_1\ldots i_k} \tilde{\tau}_{i_1\ldots i_k}\right)\right) \\
&= S\left(\Lambda_{a_1\ldots a_k}\left(\frac{\mathbb{1}}{d_1} \otimes \rho_{b_1} \otimes \ldots \frac{\mathbb{1}}{d_k} \otimes \rho_{b_k}\right)\right) \\
&= S\left(\Lambda_{a_1\ldots a_k}\left(\frac{\mathbb{1}}{d_1} \otimes \frac{\mathbb{1}}{d_1} \otimes \ldots \frac{\mathbb{1}}{d_k} \otimes \frac{\mathbb{1}}{d_k}\right)\right) \\
&= S\left(\frac{\mathbb{1}}{d_1} \otimes \frac{\mathbb{1}}{d_1} \otimes \ldots \frac{\mathbb{1}}{d_k} \otimes \frac{\mathbb{1}}{d_k}\right) \\
&= \log d_1^2 + \ldots + \log d_k^2, \qquad (3.110)
\end{aligned}$$

which is the bound in (3.109). The above optimal encoding illustrates that Alices cannot send more classical information by performing global unitaries and the optimal capacity is gained by local operations. By using (3.110) and (3.108) the multipartite super dense coding capacity is given by

$$C_{multi}^{P} = \log d_1^2 + \ldots + \log d_k^2 - H(\{q_{m_1 n_1 \ldots m_k n_k}\}). \qquad (3.111)$$

Therefore, for *one-sided* correlated Pauli channels, the super dense coding capacity is reduced be the amount of $H(\{q_{m_1 n_1 \ldots m_k n_k}\})$ with respect to the noiseless case. We also notice that the same capacity is achieved for any maximally entangled state, i.e. for any $(U_{a_1} \otimes U_{b_1} \otimes \ldots \otimes U_{a_k} \otimes U_{b_k}) \left|{\Phi_{00}}^{a_1 b_1} \ldots {\Phi_{00}}^{a_k b_k}\right\rangle$. In fact Lemma 7 still holds in this case and also the derivation of the capacity.

For $k-$copies of the $d-$dimensional Bell state ρ_{00}^{ab} with uncorrelated *one-sided* Pauli channels, i.e. $q_{m_1 n_1 \ldots m_k n_k} = q_{m_1 n_1} q_{m_2 n_2} \ldots q_{m_k n_k}$, the multipartite super dense capacity is k-times the one of the single copy capacity (3.40).

$$C_{multi}^{P,k-copy} = k\left(\log d^2 - H(\{q_{mn}\})\right). \qquad (3.112)$$

Note that the present section is being currently prepared for publication [43].

3.5 Conclusion

The first part of the present chapter covers the situation of uncorrelated noise. We investigated the bipartite super dense coding protocol in the presence of a unital noisy channel, which acts either on only Alice's subsystem after encoding (*one-sided* channel) or on both Alice's and Bob's subsystems (*two-sided* channel). For those cases where the von Neumann entropy fulfills a specific condition, the super dense coding capacity was derived. It was shown that a *one-sided* d-dimensional Pauli channel for the resource of Bell as well as Werner states fulfill the above-mentioned condition on the von Neumann entropy. The condition on the von Neumann entropy is also satisfied for a *two-sided* d-dimensional depolarizing channel. For these examples, we derived the explicit optimal super dense coding capacity, as a function of the initial resource state. When the initial state can be chosen, we found for the case of a *two-sided* 2D depolarizing channel that the optimal initial resource state is either a Bell state or a product state, depending on the value of the noise parameter. We also compared the classical capacity of the 2D depolarizing channel to the super dense coding capacities for an initial Bell state with a *one-sided* and *two-sided* 2D depolarizing channel. Our results showed that Alice and Bob may not win by sending classical information via a super dense coding protocol with unitary encoding if there is too much noise. Comparing the critical noise parameters for the quantum/classical boundary, we found that, in the scenario of the depolarizing channel, the protocols quantum teleportation and super dense coding are not equivalent, in the sense that they do not have the same critical noise parameter.

We then discussed the super dense coding protocol in the presence of a *correlated* d-dimensional Pauli channel considering both unitary and non-unitary encoding. Regarding the unitary encoding, it was shown that the problem of finding the super dense coding capacity reduces to the easier problem of finding a unitary operator which is applied to the initial state such that it minimizes the von Neumann entropy. It was proven that for the 2D quasiclassical channel and 2D fully correlated Pauli channel with Bell states and Werner states as

resources it is the identity operator which minimizes the von Neumann entropy. For those examples the super dense coding capacities were explicitly derived. It was also presented that by considering non-unitary encoding, the optimal strategy is to apply a pre-processing before unitary encoding. If the CPTP map that minimizes the von Neumann entropy is known, we found an expression for super dense coding capacity.

Finally, this chapter was concluded with an example for multipartite super dense coding. For k-copies of Bell states and *one-sided correlated* Pauli channels, the super dense coding capacity was explicitly calculated. Compared to the same scenario with noiseless channels, the capacity reduces by some amount of the Shannon entropy. It was illustrated that no global encoding by Alices can improve the capacity and that the local encodings by each of the Alices are sufficient to reach the capacity. We also showed that for *one sided uncorrelated* Pauli channels the multipartite capacity is simply additive, i.e. it is the sum of the single copy capacities.

As an outlook, it would be also interesting to study the following open problems. How can the super dense coding capacity be determined for other channels and states than the ones that fulfil the specific entropy condition in the case of the uncorrelated noise? What happens in the case of other correlated noisy channels than the Pauli channels? How does noise in general affect the multipartite super dense coding scenarios for the case of the many senders and either one or two receivers? Does global encoding in this case help?

Chapter 4

Optimal eavesdropping on noisy states in quantum key distribution

4.1 Introduction

Cryptography is the study of **secret** (Crypto-) **writing**(-graphy). The history of cryptography goes back to 400 B.C. It is the art or science of encompassing the principles and the methods of transforming an intelligible message into one that is unintelligible, and then re-transforming that message back to it's original form. In the late 1940s, mainly due to the paper of Shannon [47], cryptography became part of information theory and mathematics. He was one of the first modern cryptographers to attribute advanced mathematical techniques to the science of ciphers. Cryptography is nowadays defined as a mathematical system of transforming information so that it is unintelligible for any unauthorized party. In opposite, Cryptanalysis or codebreaking studies the principles and methods of transforming an unintelligible message back into an intelligible message without knowledge of the key. Both cryptography and cryptanalysis are called cryptology. Some basic concepts in cryptography are as follows.

Plaintext

The original intelligible message (information)

Cipher

An algorithm, a series of well-defined steps that can be followed as a procedure, for transforming an intelligible message into one that is unintelligible

Ciphertext

The transformed unintelligible message (information)

Key

A key is normally a string of bits used by a cryptographic algorithm to transform plain text into cipher text or vice versa and is only known by senders and receivers. The key should be the only part of the algorithm that is necessary to keep secret.

Encoding

The process of converting a plaintext to a ciphertext using a cipher and a key

Decoding

The process of converting a ciphertext to a plaintext using a cipher and a key

4.1.1 Quantum cryptography

Quantum cryptography stands at the crossroads between quantum mechanics and information theory. Some of the "negative" laws in quantum mechanics, due to its contrast to classical physics, has just recently been turned positive in the field of quantum information. Quantum cryptography is one of the best examples that reveals the positive side of these laws. Some of these laws read:

In general a measurement perturbs the quantum system and it is generally impossible to extract information about a quantum state without disturbing it. This principle enables *unconditionally secure key distribution* that is impossible classically. If a protocol is secure against the most powerful adversary limited only by laws of physics, its security is called *unconditional*.

No-cloning theorem

A known quantum state can be perfectly copied. One can prepare the system in a known well defined state, say one of its eigenstate, and then applies a sequence of unitary transformation that result in the desired state. But if the quantum system is not known, due to linearity and unitarity of quantum mechanics it is

impossible to create a perfectly cloned state [51]. This limitation on quantum information processing and communication has several important implications such as the possibility of realizing secure quantum communication channels and the impossibility of communicating faster than the speed of light.

Non-orthogonal states cannot be reliably distinguished

Classically, we are able to distinguish different items of information but there is no quantum measurement capable to distinguish perfectly, between two quantum states $|\psi_1\rangle$ and $|\psi_2\rangle$ with $\langle\psi_1|\psi_2\rangle \neq 0$. Indistinguishability of non-orthogonal states plays a key role in quantum cryptography.

4.1.2 Quantum key distribution (QKD)

In quantum cryptography or, more precisely, quantum key distribution a secret key is established between two trusted parties (Alice and Bob), by employing certain quantum states as signals, and suitable measurements . Two important quantum key distribution schemes are the BB84 and the Six state protocol.

BB84 protocol

BB84 is a QKD scheme developed by Charles Bennett and Gilles Brassard in 1984 [4]. In this protocol there exist two pairs of states with two orthogonal states within each pair. Pairs of orthogonal states are referred to as a basis. The two pairs of signal states of the protocol with pure state are $\{|0_x\rangle, |1_x\rangle\}$ and $\{|0_z\rangle, |1_z\rangle\}$, where $|0_\alpha\rangle$ and $|1_\alpha\rangle$ with $\alpha = x, z$ denote the eigenstates of Pauli operator σ_α. Here, the states $|0_\alpha\rangle$ symbolize the classical bit value 0, and $|1_\alpha\rangle$ represents the classical bit value 1. Alice randomly sends one of the four quantum states to Bob. No possible measurement can distinguish the 4 different states perfectly, as they are not all orthogonal. When Bob receives a state from Alice he chooses randomly one of the basis x or z for making a measurement. When Bob does the measurement in the same basis that Alice sends the state, there is a correlation between their results. After Alice sends the necessary number of the signals and Bob did the measurements, they communicate over a classical channel to tell when they used which basis. They discard all the signals in which

they used a different basis (sifting), which is half on average, leaving half the bits as a shared key. To check for the presence of eavesdropping Alice and Bob now compare a certain random subset of their remaining bit strings. If a third party, Eve, has gained any information about the signals, this introduced errors in Bob's measurements. If the level of eavesdropping is below a certain threshold, a key can be produced that is guaranteed to be secure (i.e. the eavesdropper has no information about), otherwise no secure key is possible and communication is aborted.

Six state protocol

In the six state protocol [2; 10] Alice enlarges the mentioned pairs of the BB84 protocol to three. The third pair is $\{|0_y\rangle, |1_y\rangle\}$. In this case Alice sends a state chosen randomly among the six states and Bob measures randomly in the x, y, or z basis. Here the a priori probability that Alice and Bob use the same basis is reduced to 1/3, which means that they have to discard 2/3 of the transmitted qubits before they can extract a cryptographic key. In the case of a symmetric protocol, i.e. $p_x = p_y = p_z = 1/3$. The advantage of this protocol compared to the BB84 protocol is its higher symmetry. There also exist asymmetric six state protocol with $p_x \neq p_y \neq p_z$.

4.1.3 Security

The usefulness of noisy channels for cryptographic purposes is now widely accepted by the scientific community. The first results dealing with noisebased cryptography were obtained by Wyner [52] for Discrete Memoryless Channels and later extended by Leung-Yan-Cheong and Hellman [31] for Gaussian channels as well as Csiszár and Körner [14], who proved that two parties connected by a noisy channel can communicate securely, even in the presence of an eavesdropper receiving the messages. The general idea discussed in [14] is that three parties Alice, Bob and Eve do measurements on their quantum systems where the outcomes are the classical random variables A, B and E with the joint probability distribution $p(A, B, E)$. It provides a necessary and sufficient condition on the possibility that Alice and Bob extract a secret key from $p(A, B, E)$. Here

we mention this idea in the following theorem without going through the proof. However, it is rather intuitive.

Theorem 4.1 (Csiszár and Körner [14]): *For a given $p(A, B, E)$, Alice and Bob can establish a secret key if and only if the mutual information between Alice & Bob is bigger than the mutual information between Alice & Eve or Bob & Eve, i.e. $I(A:B) \geqslant I(A:E)$ or $I(A:B) \geqslant I(B:E)$, where the mutual information $I(A:B)$ is defined in equation (1.5).*

Key rate

The key rate K is defined by the ratio of the length of the final secure key to the length of the raw key.

Theorem 4.2 (Csiszár and Körner [14]): *If Alice, Bob and Eve share independent realizations of the classical probability distribution that results from the protocol and eavesdropping strategy, the key rate K is lower bounded by $K \geq I(A:B) - I(A:E)$.*

In this chapter, we present an intuitive understanding for the counter-intuitive fact that additional noise on the quantum level may help the trusted parties to improve the performance of a quantum cryptographic protocol. A related question has been studied in [41]: There, it has been shown that if one of the parties (Alice or Bob) adds some noise to their classical measurement data before error correction, then the BB84 [4], B92 [3] and six state protocol [2; 10] are more robust with respect to quantum noise, i.e. the secret key rate is non-zero up to higher values of the quantum bit error rate. Adding noise at the quantum level, i.e. before the measurement, will also lead to additional noise at the classical level. In this sense our scenario should lead, at least qualitatively, to a similar result as shown in [41].

In the following in the asymptotic limit of many signals we will discuss the six state protocol with additional (equal) noise on all signal states. We also consider that Alice's and Bob's apparatuses are perfect while they do not have total control over the channel.

4.2 Eavesdropping on noisy states

In the six state protocol with mixed states Alice sends instead of pure states one of the following six mixed states (either deliberately, or due to unavoidable noise in the transmission channel):

$$\rho_i = (1-p)|i\rangle\langle i| + \frac{p}{2}\mathbb{1}, \quad i \in \{0_\alpha, 1_\alpha\}, \tag{4.1}$$

with $\alpha = x, y, z$. The parameter p describes the amount of noise, with $0 \leq p \leq 1$. Here, we assume the noise to be equal in all bases, i.e. we study the depolarizing channel. (In a more general model, polarization dependent noise could be treated in an analogous way, by letting p_α depend on $\alpha = x, y, z$.)

For the eavesdropping strategy we assume that Eve is restricted to interfering separately with each of the single systems sent by Alice (i.e. individual attack). In this class of attacks she attaches to each qubit an independent probe which is initially in the state $|X\rangle$ and applies some unitary transformation. The dimension of the probes and the interaction are in principle arbitrary, but in [19] it has been shown that the most general unitary eavesdropping attack on a d-dimensional signal state needs only d^2 linearly independent ancilla states of Eve. (This argument also holds when the signal states are mixed, as the unitary transformation of the basis states already uniquely defines the transformation of any superposition, due to linearity, and thus also of a mixture of projectors onto superpositions of basis states.) Thus, it is enough for Eve to use two qubits for her probe states.

The most general unitary transformation U that Eve can design is defined via its action on the basis states (where we use for the computational basis the notation $|0\rangle = |0_z\rangle$ and $|1\rangle = |1_z\rangle$),

$$U|0\rangle|X\rangle = \sqrt{1-D}|0\rangle|A\rangle + \sqrt{D}|1\rangle|B\rangle, \tag{4.2a}$$

$$U|1\rangle|X\rangle = \sqrt{1-D}|1\rangle|C\rangle + \sqrt{D}|0\rangle|D\rangle, \tag{4.2b}$$

where D is called the disturbance, with $0 \leq D \leq \frac{1}{2}$. Eve's normalized probes after interaction are $|A\rangle$, $|B\rangle$, $|C\rangle$, and $|D\rangle$. They have to be chosen such that U is a unitary operator,

$$\langle X|\langle 0|U^\dagger U|1\rangle|X\rangle = \left(\sqrt{1-D}\langle A|\langle 0| + \sqrt{D}\langle B|\langle 1|\right)\left(\sqrt{1-D}|1\rangle|C\rangle + \sqrt{D}|0\rangle|D\rangle\right),$$

where $\langle 0|1\rangle = 0$. The unitarity of U leads to

$$\langle A|D\rangle + \langle B|C\rangle = 0 \tag{4.3}$$

We notice that the same unitary transformation U should be applied to the other initial states $|0_{x,y}\rangle$ and $|1_{x,y}\rangle$. After Eve's interaction ρ_i changes to a shared state between Alice and Eve ρ_i^{AE}. In the following we explicitly calculate ρ_i^{AE} for $i = 0_\alpha, 1_\alpha$:

$$\begin{aligned}
\rho_0^{AE} &= (1-\frac{p}{2})\Big[(1-D)(|0\rangle\langle 0|)(|A\rangle\langle A|) + D(|1\rangle\langle 1|)(|B\rangle\langle B|) \\
&+ \sqrt{D(1-D)}(|0\rangle\langle 1|)(|A\rangle\langle B|) + \sqrt{D(1-D)}(|1\rangle\langle 0|)(|B\rangle\langle A|)\Big] \\
&+ \frac{p}{2}\Big[(1-D)(|1\rangle\langle 1|)(|C\rangle\langle C|) + D(|0\rangle\langle 0|)(|D\rangle\langle D|) \\
&+ \sqrt{D(1-D)}(|1\rangle\langle 0|)(|C\rangle\langle D|) + \sqrt{D(1-D)}(|0\rangle\langle 1|(|D\rangle\langle C|)\Big], \quad (4.4)
\end{aligned}$$

$$\begin{aligned}
\rho_1^{AE} &= (1-\frac{p}{2})\Big[(1-D)(|1\rangle\langle 1|)(|C\rangle\langle C|) + D(|0\rangle\langle 0|)(|D\rangle\langle D|) \\
&+ \sqrt{D(1-D)}(|1\rangle\langle 0|)(|C\rangle\langle D|) + \sqrt{D(1-D)}(|0\rangle\langle 1|)(|D\rangle\langle C|)\Big] \\
&+ \frac{p}{2}\Big[(1-D)(|0\rangle\langle 0|)(|A\rangle\langle A|) + D(|1\rangle\langle 1|)(|B\rangle\langle B|) \\
&+ \sqrt{D(1-D)}(|0\rangle\langle 1|)(|A\rangle\langle B|) + \sqrt{D(1-D)}(|1\rangle\langle 0|(|B\rangle\langle A|)\Big], \quad (4.5)
\end{aligned}$$

$$\begin{aligned}
\rho_{0_x}^{AE} &= \frac{1}{2}\Big[|0\rangle\langle 0|[(1-D)|A\rangle\langle A| + D|D\rangle\langle D|] + |1\rangle\langle 1|[(1-D)|C\rangle\langle C| \\
&+ D|B\rangle\langle B|] + \sqrt{D(1-D)}|0\rangle\langle 1|[|A\rangle\langle B| + |D\rangle\langle C|] \\
&+ \sqrt{D(1-D)}|1\rangle\langle 0|[|B\rangle\langle A| + |C\rangle\langle D|]\Big] \\
&+ (\frac{1-p}{2})\Big[|0\rangle\langle 1|[(1-D)|A\rangle\langle C| + D|D\rangle\langle B|] + |1\rangle\langle 0|[(1-D)|C\rangle\langle A| \\
&+ D|B\rangle\langle D|] + \sqrt{D(1-D)}|0\rangle\langle 0|[|A\rangle\langle D| + |D\rangle\langle A|] \\
&+ \sqrt{D(1-D)}|1\rangle\langle 1|[|C\rangle\langle B| + |B\rangle\langle C|]\Big], \quad (4.6)
\end{aligned}$$

$$\begin{aligned}
\rho_{1_x}^{AE} &= \frac{1}{2}\Big[|0\rangle\langle 0|[(1-D)|A\rangle\langle A| + D|D\rangle\langle D|] + |1\rangle\langle 1|[(1-D)|C\rangle\langle C| \\
&+ D|B\rangle\langle B|] + \sqrt{D(1-D)}|0\rangle\langle 1|[|A\rangle\langle B| + |D\rangle\langle C|] \\
&+ \sqrt{D(1-D)}|1\rangle\langle 0|[|B\rangle\langle A| + |C\rangle\langle D|]\Big] \\
&- (\frac{1-p}{2})\Big[|0\rangle\langle 1|[(1-D)|A\rangle\langle C| + D|D\rangle\langle B|] + |1\rangle\langle 0|[(1-D)|C\rangle\langle A| \\
&+ D|B\rangle\langle D|] + \sqrt{D(1-D)}|0\rangle\langle 0|[|A\rangle\langle D| + |D\rangle\langle A|] \\
&+ \sqrt{D(1-D)}|1\rangle\langle 1|[|C\rangle\langle B| + |B\rangle\langle C|]\Big], \qquad (4.7)
\end{aligned}$$

$$\begin{aligned}
\rho_{0_y}^{AE} &= \frac{1}{2}\Big[|0\rangle\langle 0|[(1-D)|A\rangle\langle A| + D|D\rangle\langle D|] + |1\rangle\langle 1|[(1-D)|C\rangle\langle C| \\
&+ D|B\rangle\langle B|] + \sqrt{D(1-D)}|0\rangle\langle 1|[|A\rangle\langle B| + |D\rangle\langle C|] \\
&+ \sqrt{D(1-D)}|1\rangle\langle 0|[|B\rangle\langle A| + |C\rangle\langle D|]\Big] \\
&+i\ (\frac{1-p}{2})\Big[|0\rangle\langle 1|[(1-D)|A\rangle\langle C| + D|D\rangle\langle B|] + |1\rangle\langle 0|[(1-D)|C\rangle\langle A| \\
&+ D|B\rangle\langle D|] + \sqrt{D(1-D)}|0\rangle\langle 0|[|A\rangle\langle D| + |D\rangle\langle A|] \\
&+ \sqrt{D(1-D)}|1\rangle\langle 1|[|C\rangle\langle B| + |B\rangle\langle C|]\Big], \qquad (4.8)
\end{aligned}$$

$$\begin{aligned}
\rho_{1_y}^{AE} &= \frac{1}{2}\Big[|0\rangle\langle 0|[(1-D)|A\rangle\langle A| + D|D\rangle\langle D|] + |1\rangle\langle 1|[(1-D)|C\rangle\langle C| \\
&+ D|B\rangle\langle B|] + \sqrt{D(1-D)}|0\rangle\langle 1|[|A\rangle\langle B| + |D\rangle\langle C|] \\
&+ \sqrt{D(1-D)}|1\rangle\langle 0|[|B\rangle\langle A| + |C\rangle\langle D|]\Big] \\
&-i\ (\frac{1-p}{2})\Big[|0\rangle\langle 1|[(1-D)|A\rangle\langle C| + D|D\rangle\langle B|] + |1\rangle\langle 0|[(1-D)|C\rangle\langle A| \\
&+ D|B\rangle\langle D|] + \sqrt{D(1-D)}|0\rangle\langle 0|[|A\rangle\langle D| + |D\rangle\langle A|] \\
&+ \sqrt{D(1-D)}|1\rangle\langle 1|[|C\rangle\langle B| + |B\rangle\langle C|]\Big]. \qquad (4.9)
\end{aligned}$$

$\rho_i^B = \mathrm{tr}_E(\rho_i^{AE})$ corresponds to Bob's state when Alice sends the signal state

$|i\rangle$. The partial trace over Eve's system in (4.4),...,(4.9) read :

$$\begin{aligned}\rho_0^B &= \mathrm{tr}_E(\rho_0^{AE}) = (1-\frac{p}{2})\Big[(1-D)(|0\rangle\langle 0|) + D(|1\rangle\langle 1|) \\ &+ \sqrt{D(1-D)}(|0\rangle\langle 1|)\langle B|A\rangle + \sqrt{D(1-D)}(|1\rangle\langle 0|)\langle A|B\rangle\Big] \\ &+ \frac{p}{2}\Big[(1-D)(|1\rangle\langle 1|) + D(|0\rangle\langle 0|) + \sqrt{D(1-D)}(|1\rangle\langle 0|)\langle D|C\rangle \\ &+ \sqrt{D(1-D)}(|0\rangle\langle 1|\langle C|D\rangle\Big]. \end{aligned} \tag{4.10}$$

$$\begin{aligned}\rho_1^B &= \mathrm{tr}_E(\rho_1^{AE}) = (1-\frac{p}{2})\Big[(1-D)(|1\rangle\langle 1|) + D(|0\rangle\langle 0|) \\ &+ \sqrt{D(1-D)}(|1\rangle\langle 0|)\langle D|C\rangle + \sqrt{D(1-D)}(|0\rangle\langle 1|\langle C|D\rangle\Big] \\ &+ \frac{p}{2}\Big[(1-D)(|0\rangle\langle 0|) + D(|1\rangle\langle 1|) + \sqrt{D(1-D)}(|0\rangle\langle 1|)\langle B|A\rangle \\ &+ \sqrt{D(1-D)}(|1\rangle\langle 0|)\langle A|B\rangle\Big], \end{aligned} \tag{4.11}$$

$$\begin{aligned}\rho_{0_x}^B &= \mathrm{tr}_E(\rho_{0_x}^{AE}) = \frac{1}{2}\Big[\sqrt{D(1-D)}|0\rangle\langle 1|\left(\langle B|A\rangle + \langle C|D\rangle\right) \\ &+ \sqrt{D(1-D)}|1\rangle\langle 0|\left(\langle A|B\rangle + \langle D|C\rangle\right) + |0\rangle\langle 0| + |1\rangle\langle 1|\Big] \\ &+ (\frac{1-p}{2})\Big[|0\rangle\langle 1|\left[(1-D)\langle C|A\rangle + D\langle B|D\rangle\right] \\ &+ |1\rangle\langle 0|\left[(1-D)\langle A|C\rangle + D\langle D|B\rangle\right] + \sqrt{D(1-D)}|0\rangle\langle 0|[\langle D|A\rangle + \langle A|D\rangle] \\ &+ \sqrt{D(1-D)}|1\rangle\langle 1|(\langle B|C\rangle + \langle C|B\rangle)\Big], \end{aligned} \tag{4.12}$$

$$\begin{aligned}\rho_{1_x}^B &= \mathrm{tr}_E(\rho_{1_x}^{AE}) = \frac{1}{2}\Big[\sqrt{D(1-D)}|0\rangle\langle 1|\left(\langle B|A\rangle + \langle C|D\rangle\right) \\ &+ \sqrt{D(1-D)}|1\rangle\langle 0|\left(\langle A|B\rangle + \langle D|C\rangle\right) + |0\rangle\langle 0| + |1\rangle\langle 1|\Big] \\ &- (\frac{1-p}{2})\Big[|0\rangle\langle 1|\left[(1-D)\langle C|A\rangle + D\langle B|D\rangle\right] \\ &+ |1\rangle\langle 0|\left[(1-D)\langle A|C\rangle + D\langle D|B\rangle\right] + \sqrt{D(1-D)}|0\rangle\langle 0|[\langle D|A\rangle + \langle A|D\rangle] \\ &+ \sqrt{D(1-D)}|1\rangle\langle 1|(\langle B|C\rangle + \langle C|B\rangle)\Big], \end{aligned} \tag{4.13}$$

$$\begin{aligned}
\rho_{0_y}^B &= \mathrm{tr}_E(\rho_{0_y}^{AE}) = \frac{1}{2}\Big[\sqrt{D(1-D)}|0\rangle\langle 1|\left(\langle B|A\rangle + \langle C|D\rangle\right) \\
&+ \sqrt{D(1-D)}|1\rangle\langle 0|\left(\langle A|B\rangle + \langle D|C\rangle\right) + |0\rangle\langle 0| + |1\rangle\langle 1|\Big] \\
&+i \ (\frac{1-p}{2})\Big[|0\rangle\langle 1|\left[(1-D)\langle C|A\rangle + D\langle B|D\rangle\right] \\
&+ |1\rangle\langle 0|\left[(1-D)\langle A|C\rangle + D\langle D|B\rangle\right] + \sqrt{D(1-D)}|0\rangle\langle 0|[\langle D|A\rangle + \langle A|D\rangle] \\
&+ \sqrt{D(1-D)}|1\rangle\langle 1|(\langle B|C\rangle + \langle C|B\rangle)\Big], \qquad (4.14)
\end{aligned}$$

$$\begin{aligned}
\rho_{1_y}^B &= \mathrm{tr}_E(\rho_{1_y}^{AE}) = \frac{1}{2}\Big[\sqrt{D(1-D)}|0\rangle\langle 1|\left(\langle B|A\rangle + \langle C|D\rangle\right) \\
&+ \sqrt{D(1-D)}|1\rangle\langle 0|\left(\langle A|B\rangle + \langle D|C\rangle\right) + |0\rangle\langle 0| + |1\rangle\langle 1|\Big] \\
&-i \ (\frac{1-p}{2})\Big[|0\rangle\langle 1|\left[(1-D)\langle C|A\rangle + D\langle B|D\rangle\right] \\
&+ |1\rangle\langle 0|\left[(1-D)\langle A|C\rangle + D\langle D|B\rangle\right] + \sqrt{D(1-D)}|0\rangle\langle 0|[\langle D|A\rangle + \langle A|D\rangle] \\
&+ \sqrt{D(1-D)}|1\rangle\langle 1|(\langle B|C\rangle + \langle C|B\rangle)\Big]. \qquad (4.15)
\end{aligned}$$

The quantum bit error rate in the z-basis is denoted as Q_z, and is given as the fraction of original signals $|0\rangle(|1\rangle)$ sent by Alice, but interpreted as $|1\rangle(|0\rangle)$ by Bob, namely

$$Q_z = \tfrac{1}{2}\langle 0|\rho_1^B|0\rangle + \tfrac{1}{2}\langle 1|\rho_0^B|1\rangle, \qquad (4.16)$$

where ρ_0^B and ρ_1^B are the states that Bob receives when Alice sends $|0\rangle$ and $|1\rangle$, respectively. We define $Q_{x,y}$ in an analogous way for the x and y-basis.

As we assume the noise to be uniform, a quantum bit error rate that is basis-dependent indicates the presence of an eavesdropper. We therefore assume that Eve uses a strategy that produces the same quantum bit error rate in the three different bases, i.e.

$$Q = Q_z = Q_x = Q_y. \qquad (4.17)$$

which can be tested by Alice and Bob, by comparing a part of their bit string for the z-basis. Again, an analogous requirement has to hold in the other two bases,

too. Additionally, we restrict Eve to attack in such a way that the two terms of $Q_{x,y,z}$ are identical, i.e.

$$Q = \langle 0|\rho_1^B|0\rangle = \langle 1|\rho_0^B|1\rangle \,, \tag{4.18a}$$
$$Q = \langle 0_x|\rho_{1_x}^B|0_x\rangle = \langle 1_x|\rho_{0_x}^B|1_x\rangle \,, \tag{4.18b}$$
$$Q = \langle 0_y|\rho_{1_y}^B|0_y\rangle = \langle 1_y|\rho_{0_y}^B|1_y\rangle \,, \tag{4.18c}$$

By substituting from (4.10), (4.11) in (4.18a), it can be easily verified that the relationship between the quantum bit error rate Q and D is

$$Q = D(1-p) + \frac{p}{2}. \tag{4.19}$$

For $\langle 0_x|\rho_{1_x}^B|0_x\rangle$ and $\langle 1_x|\rho_{0_x}^B|1_x\rangle$, using (4.12), (4.13) and (4.3),we have

$$\begin{aligned}\langle 0_x|\rho_{1_x}^B|0_x\rangle &= \frac{1}{2}\Big[1+\sqrt{D(1-D)}\mathrm{Re}\left(\langle \mathrm{A}|\mathrm{B}\rangle + \langle \mathrm{D}|\mathrm{C}\rangle\right)\Big] \\ &- (\frac{1-p}{2})\Big[(1-D)\mathrm{Re}\langle \mathrm{A}|\mathrm{C}\rangle + \mathrm{DRe}\langle \mathrm{B}|\mathrm{D}\rangle\Big], \end{aligned}\tag{4.20a}$$
$$\begin{aligned}\langle 1_x|\rho_{0_x}^B|1_x\rangle &= \frac{1}{2}\Big[1-\sqrt{D(1-D)}\mathrm{Re}\left(\langle \mathrm{A}|\mathrm{B}\rangle + \langle \mathrm{D}|\mathrm{C}\rangle\right)\Big] \\ &- (\frac{1-p}{2})\Big[(1-D)\mathrm{Re}\langle \mathrm{A}|\mathrm{C}\rangle + \mathrm{DRe}\langle \mathrm{B}|\mathrm{D}\rangle\Big]. \end{aligned}\tag{4.20b}$$

Substituting (4.20a) and (4.20b) in (4.18b) results a restriction over Eve's states,

$$\mathrm{Re}\left(\langle \mathrm{A}|\mathrm{B}\rangle + \langle \mathrm{D}|\mathrm{C}\rangle\right) = 0. \tag{4.21}$$

The expressions Q_x is then,

$$Q_x = \frac{1}{2} - (\frac{1-p}{2})\Big[(1-D)Re\langle A|C\rangle + DRe\langle B|D\rangle\Big] \tag{4.22}$$

Similar to the x direction we consider the y direction and arrive at

$$\begin{aligned}\langle 0_y|\rho_{1_y}^B|0_y\rangle &= \frac{1}{2}\Big[1+i\sqrt{D(1-D)}\mathrm{Im}\left(\langle \mathrm{A}|\mathrm{B}\rangle + \langle \mathrm{D}|\mathrm{C}\rangle\right)\Big] \\ &- (\frac{1-p}{2})\Big[(1-D)\mathrm{Re}\langle \mathrm{A}|\mathrm{C}\rangle - \mathrm{DRe}\langle \mathrm{B}|\mathrm{D}\rangle\Big], \end{aligned}\tag{4.23a}$$
$$\begin{aligned}\langle 1_y|\rho_{0_y}^B|1_y\rangle &= \frac{1}{2}\Big[1-i\sqrt{D(1-D)}\mathrm{Im}\left(\langle \mathrm{A}|\mathrm{B}\rangle + \langle \mathrm{D}|\mathrm{C}\rangle\right)\Big] \\ &- (\frac{1-p}{2})\Big[(1-D)\mathrm{Re}\langle \mathrm{A}|\mathrm{C}\rangle - \mathrm{DRe}\langle \mathrm{B}|\mathrm{D}\rangle\Big]. \end{aligned}\tag{4.23b}$$

From the equality $\langle 0_y|\rho^B_{1_y}|0_y\rangle = \langle 1_y|\rho^B_{0_y}|1_y\rangle$ we find another restriction on Eve's states,

$$\mathrm{Im}\left(\langle \mathrm{A}|\mathrm{B}\rangle + \langle \mathrm{D}|\mathrm{C}\rangle\right) = 0. \tag{4.24}$$

Then the quantum bit error rate in the $y-$ direction is

$$Q_y = \frac{1}{2} - \left(\frac{1-p}{2}\right)\Big[(1-D)\mathrm{Re}\langle \mathrm{A}|\mathrm{C}\rangle - \mathrm{DRe}\langle \mathrm{B}|\mathrm{D}\rangle\Big] \tag{4.25}$$

Since it has been considered that the quantum bit error rate Q is equal in all directions, i.e. using $Q_x = Q_y$ and (4.22) and (4.25), we find a new restriction on Eve's states that reads,

$$\mathrm{Re}\left(\langle \mathrm{B}|\mathrm{D}\rangle\right) = 0. \tag{4.26}$$

By using (4.26) in $2Q = Q_y$ we obtain for the scalar product of $|A\rangle$ and $|C\rangle$

$$\mathrm{Re}\langle \mathrm{A}|\mathrm{C}\rangle = 2 - \frac{1}{1-\mathrm{D}} = \frac{2(1-2\mathrm{Q})}{2-\mathrm{p}-2\mathrm{Q}}. \tag{4.27}$$

For convenience, we summarizes the four conditions between Eve's probes mentioned in (4.3),(4.21),(4.24),(4.26) and (4.27):

$$\begin{aligned}
\langle B|D\rangle &= 0, && \text{(4.28a)}\\
\mathrm{Re}\langle \mathrm{A}|\mathrm{C}\rangle &= \frac{2(1-2Q)}{2-p-2Q}, && \text{(4.28b)}\\
\langle A|B\rangle + \langle D|C\rangle &= 0, && \text{(4.28c)}\\
\langle A|D\rangle + \langle B|C\rangle &= 0. && \text{(4.28d)}
\end{aligned}$$

Note that instead of $\mathrm{Re}\langle \mathrm{B}|\mathrm{D}\rangle = 0$ we used the stronger condition $\langle B|D\rangle = 0$. We also note that the quantum bit error rate Q only depends on the scalar product between $|A\rangle$ and $|C\rangle$. Eve's two-qubit states can be written as an expansion of four basis vectors with complex coefficients. As explained above, Eve's states only need to be four-dimensional. We have the freedom to choose $|B\rangle = |00\rangle$. Equation (4.28a) allows to assign $|D\rangle$ one of the other three basis vectors, e.g., $|D\rangle = |11\rangle$. The general expansion for the normalized vectors $|A\rangle$ and $|C\rangle$ are

$$|A\rangle = \alpha_A|00\rangle + \beta_A|10\rangle + \gamma_A|01\rangle + \delta_A|11\rangle, \tag{4.29}$$

with

$$|\alpha_A|^2 + |\beta_A|^2 + |\gamma_A|^2 + |\delta_A|^2 = 1 \ , \tag{4.30}$$

and

$$|C\rangle = \alpha_C|00\rangle + \beta_C|10\rangle + \gamma_C|01\rangle + \delta_C|11\rangle, \tag{4.31}$$

with

$$|\alpha_C|^2 + |\beta_C|^2 + |\gamma_C|^2 + |\delta_C|^2 = 1. \tag{4.32}$$

We have to determine the free parameters $\alpha_A, ..., \delta_A$ and $\alpha_C, ..., \delta_C$ such that Eve's transformation is optimized. As a figure of merit we will calculate the mutual information between Eve and Alice, and optimize Eve's transformation such that she acquires the maximal mutual information.

4.3 Optimal eavesdropping in terms of mutual information

The mutual information measures the information that two parties share. Here the parties have variables X, Y that can take values x, y, respectively. The mutual information is defined [40] as

$$I^{XY} := I(X:Y) \ = \ \sum_{x,y} p(x,y) \log p(y|x) - \sum_{y} p(y) \log p(y), \tag{4.33}$$

where $p(x,y) = p(x)p(y|x)$, is the joint probability to find x and y, and $p(y|x)$ is the conditional probability of y, given x. All logarithms are taken to base 2. We determine the mutual information between Alice and Eve as a measure of the amount of information that Eve extract from the original state. Arbitrarily, we choose the $z-$basis for the rest of our calculations. We need to find Eve's state after interaction with the original state ρ_i. Accordingly, we trace out Alice's

subsystem form the shared states ρ_0^{AE} and ρ_1^{AE},

$$\begin{aligned}\rho_0^E &= \mathrm{tr}_A\,\rho_0^{AE} = (1-\frac{p}{2})\left((1-D)|A\rangle\langle A| + D|B\rangle\langle B|\right) \\ &+ \frac{p}{2}\left((1-D)|C\rangle\langle C| + D|D\rangle\langle D|\right), \qquad (4.34a)\\ \rho_1^E &= \mathrm{tr}_A\,\rho_1^{AE} = (1-\frac{p}{2})\left((1-D)|C\rangle\langle C| + D|D\rangle\langle D|\right) \\ &+ \frac{p}{2}\left((1-D)|A\rangle\langle A| + D|B\rangle\langle B|\right). \qquad (4.34b)\end{aligned}$$

Let $M_1, ..., M_4$ and $M_5, ..., M_8$ to be the conditional probabilities of measuring $|00\rangle, |10\rangle, |01\rangle$ and $|11\rangle$ by Eve, given ρ_0^E and ρ_1^E, respectively. These conditional probabilities M_i are

$$\begin{aligned}M_1 := p(00|\rho_0^E) &= (1-\frac{p}{2})\left((1-D)|\alpha_A|^2 + D\right) + \frac{p}{2}(1-D)|\alpha_C|^2, \qquad (4.35a)\\ M_2 := p(10|\rho_0^E) &= (1-\frac{p}{2})(1-D)|\beta_A|^2 + \frac{p}{2}(1-D)|\beta_C|^2, \qquad (4.35b)\\ M_3 := p(01|\rho_0^E) &= (1-\frac{p}{2})(1-D)|\gamma_A|^2 + \frac{p}{2}(1-D)|\gamma_C|^2, \qquad (4.35c)\\ M_4 := p(11|\rho_0^E) &= \frac{p}{2}\left((1-D)|\delta_C|^2 + D\right) + (1-\frac{p}{2})(1-D)|\delta_A|^2, \qquad (4.35d)\\ M_5 := p(00|\rho_1^E) &= \frac{p}{2}\left((1-D)|\alpha_A|^2 + D\right) + (1-\frac{p}{2})(1-D)|\alpha_C|^2, \qquad (4.35e)\\ M_6 := p(10|\rho_1^E) &= (1-\frac{p}{2})(1-D)|\beta_C|^2 + \frac{p}{2}(1-D)|\beta_A|^2, \qquad (4.35f)\\ M_7 := p(01|\rho_1^E) &= (1-\frac{p}{2})(1-D)|\gamma_C|^2 + \frac{p}{2}(1-D)|\gamma_A|^2, \qquad (4.35g)\\ M_8 := p(11|\rho_1^E) &= \frac{p}{2}(1-D)|\delta_A|^2 + (1-\frac{p}{2})\left((1-D)|\delta_C|^2 + D\right). \qquad (4.35h)\end{aligned}$$

$p(y) = \sum_x p(x)p(y|x)$ is the probability that Eve detects y. Therefore, the probabilities of measuring $|00\rangle, |10\rangle, |01\rangle$ and $|11\rangle$ by Eve are,

$$\begin{aligned}p(00) &= p(\rho_0^E)p(00|\rho_0^E) + p(\rho_1^E)p(00|\rho_1^E) = \frac{1}{2}(M_1 + M_5), \qquad (4.36a)\\ p(10) &= p(\rho_0^E)p(10|\rho_0^E) + p(\rho_1^E)p(10|\rho_1^E) = \frac{1}{2}(M_2 + M_6), \qquad (4.36b)\\ p(01) &= p(\rho_0^E)p(01|\rho_0^E) + p(\rho_1^E)p(01|\rho_1^E) = \frac{1}{2}(M_3 + M_7), \qquad (4.36c)\\ p(11) &= p(\rho_0^E)p(11|\rho_0^E) + p(\rho_1^E)p(11|\rho_1^E) = \frac{1}{2}(M_4 + M_8). \qquad (4.36d)\end{aligned}$$

Using (4.35a-4.35h) and (4.36a-4.36d), the two terms of the mutual information in (4.33) read,

$$\sum_y p(y)\log p(y) = \frac{1}{2}\Big[(M_1+M_5)\log(M_1+M_5)+(M_2+M_6)\log(M_2+M_6)$$
$$+(M_3+M_7)\log(M_3+M_7)+(M_4+M_8)\log(M_4+M_8)\Big]-1, \tag{4.37a}$$
$$\sum_{x,y} p(x)p(y|x)\log p(y|x) = \frac{1}{2}\Big[M_1\log M_1+M_2\log M_2+M_3\log M_3$$
$$+M_4\log M_4+M_5\log M_5+M_6\log M_6+M_7\log M_7+M_8\log M_8\Big]. \tag{4.37b}$$

Then the mutual information between Alice and Eve is,

$$\begin{aligned} I^{AE} &= \sum_{x,y} p(x,y)\log p(y|x) - \sum_y p(y)\log p(y) \\ &= \frac{1}{2}\Big[M_1\log M_1+M_2\log M_2+M_3\log M_3+M_4\log M_4+M_5\log M_5 \\ &+ M_6\log M_6+M_7\log M_7+M_8\log M_8\Big]-\frac{1}{2}\Big[(M_1+M_5)\log(M_1+M_5) \\ &+ (M_2+M_6)\log(M_2+M_6)+(M_3+M_7)\log(M_3+M_7) \\ &+ (M_4+M_8)\log(M_4+M_8)\Big]+1 \\ &= 1+\frac{1}{2}\left(\tau[M_1,M_5]+\tau[M_2,M_6]+\tau[M_3,M_7]+\tau[M_4,M_8]\right), \end{aligned} \tag{4.38}$$

where we used the definition

$$\tau[x,y] = x\log x + y\log y - (x+y)\log(x+y). \tag{4.39}$$

Eve wishes to retrieve the maximal information, i.e. she has to choose the optimal coefficients $\alpha_{A,C},\beta_{A,C},\gamma_{A,C},\delta_{A,C}$, for fixed p and Q. The full problem can be simplified with the following argument: As mentioned above, for a fixed noise parameter p the quantum bit error rate Q only depends on the real part of the overlap between $|A\rangle$ and $|C\rangle$, see equation (4.28b). Therefore, Eve is free to choose those states on which Q does not depend in such a way that her information is maximal, as long as the constraints given in equations (4.28a)-(4.28d) are fulfilled. Thus, Eve will choose her ancilla states orthogonal (whenever possible), i.e. $\langle A|B\rangle = \langle B|C\rangle = \langle A|D\rangle = \langle D|C\rangle = 0$, which corresponds to $\alpha_A =$

$\delta_A = \alpha_C = \delta_C = 0$. One realizes by looking at equation (4.2a) and (4.2b) that in this way Eve's probe states are made as distinguishable as possible (for given $\langle A|C\rangle \neq 0$). The best measurement for the two remaining non-orthogonal states $|A\rangle$ and $|C\rangle$ is rank one and orthogonal [16; 40]. With $\alpha_A = \delta_A = \alpha_C = \delta_C = 0$ and $D = (Q - \frac{p}{2})/(1-p)$, the conditional probabilities M_i in (4.35a-4.35h) reduce to

$$
\begin{aligned}
M_1 = M_8 &= (1-\frac{p}{2}) \cdot \frac{Q-\frac{p}{2}}{1-p}, && (4.40a)\\
M_2 &= (\frac{1-\frac{p}{2}-Q}{1-p})\{(1-\frac{p}{2})|\beta_A|^2 + \frac{p}{2}|\beta_C|^2\}, && (4.40b)\\
M_3 &= (\frac{1-\frac{p}{2}-Q}{1-p})\{(1-\frac{p}{2})|\gamma_A|^2 + \frac{p}{2}|\gamma_C|^2\}, && (4.40c)\\
M_4 = M_5 &= \frac{p}{2} \cdot \frac{Q-\frac{p}{2}}{1-p}, && (4.40d)\\
M_6 &= (\frac{1-\frac{p}{2}-Q}{1-p})\{(1-\frac{p}{2})|\beta_C|^2 + \frac{p}{2}|\beta_A|^2\}, && (4.40e)\\
M_7 &= (\frac{1-\frac{p}{2}-Q}{1-p})\{(1-\frac{p}{2})|\gamma_C|^2 + \frac{p}{2}|\gamma_A|^2\}. && (4.40f)
\end{aligned}
$$

By substituting these M_i's in (4.38), the mutual information between Alice and Eve I^{AE} is,

$$
\begin{aligned}
I^{AE} &= 1 + \frac{1}{2}(\frac{1-\frac{p}{2}-Q}{1-p}) \cdot \Big\{\tau\Big[(1-\frac{p}{2})|\beta_A|^2 + \frac{p}{2}|\beta_C|^2 , \frac{p}{2}|\beta_A|^2 + (1-\frac{p}{2})|\beta_C|^2\Big]\\
&+ \tau\Big[(1-\frac{p}{2})|\gamma_A|^2 + \frac{p}{2}|\gamma_C|^2, \frac{p}{2}|\gamma_A|^2 + (1-\frac{p}{2})|\gamma_C|^2\Big]\Big\}\\
&+ (\frac{Q-\frac{p}{2}}{1-p}).\tau\Big[1-\frac{p}{2} \ , \ \frac{p}{2}\Big]. \qquad (4.41)
\end{aligned}
$$

We continue the problem of optimizing I^{AE} by using the Lagrange multiplier method. To do this, first, we redefine the complex coefficients $\beta_{A,C}, \gamma_{A,C}$ in polar coordinates:

$$
\begin{aligned}
\beta_A &= r_{\beta_A}, && (4.42)\\
\beta_C &= r_{\beta_C}, && (4.43)\\
\gamma_A &= r_{\gamma_A} \exp(i\Phi_{\gamma_A}), && (4.44)\\
\gamma_C &= r_{\gamma_C} \exp(i\Phi_{\gamma_C}). && (4.45)
\end{aligned}
$$

We note that because of the unphysical global phase we have the freedom to choose β_A and β_C to be real. Using the Lagrange multiplier method we then write the Lagrangian L as

$$L = I^{AE} + \lambda_1 g_1 + \lambda_2 g_2 + \lambda_3 g_3, \tag{4.46}$$

where g_1 , g_2 and g_3 are the constraints (4.28b),(4.30) and (4.32). The derivative $dL = 0$ yields the following system of equations:

$$g_1 = r_{\beta_A} r_{\beta_C} + r_{\gamma_A} r_{\gamma_C} \cos(\Phi_{\gamma_A} - \Phi_{\gamma_C}) - \frac{2(1-2Q)}{2-p-2Q} = 0, \tag{4.47a}$$

$$g_2 = {r_{\beta_A}}^2 + {r_{\gamma_A}}^2 - 1 = 0, \tag{4.47b}$$

$$g_3 = {r_{\beta_C}}^2 + {r_{\gamma_C}}^2 - 1 = 0, \tag{4.47c}$$

$$r_{\gamma_A} r_{\gamma_C} \sin(\Phi_{\gamma_A} - \Phi_{\gamma_C}) = 0, \tag{4.47d}$$

$$r_{\beta_A}\{(\frac{1-\frac{p}{2}-Q}{1-p})((1-\frac{p}{2})\log M_2 + \frac{p}{2}\log M_6 - \log(M_2+M_6)) + 2\lambda_2\}$$
$$+\lambda_1 r_{\beta_C} = 0, \tag{4.47e}$$

$$r_{\beta_C}\{(\frac{1-\frac{p}{2}-Q}{1-p})(\frac{p}{2}\log M_2 + (1-\frac{p}{2})\log M_6 - \log(M_2+M_6)) + 2\lambda_3\}$$
$$+\lambda_1 r_{\beta_A} = 0, \tag{4.47f}$$

$$r_{\gamma_A}\{(\frac{1-\frac{p}{2}-Q}{1-p})((1-\frac{p}{2})\log M_3 + \frac{p}{2}\log M_7 - \log(M_3+M_7)) + 2\lambda_2\}$$
$$+\lambda_1 r_{\gamma_C}\cos(\Phi_{\gamma_A} - \Phi_{\gamma_C}) = 0, \tag{4.47g}$$

$$r_{\gamma_C}\{(\frac{1-\frac{p}{2}-Q}{1-p})(\frac{p}{2}\log M_3 + (1-\frac{p}{2})\log M_7 - \log(M_3+M_7)) + 2\lambda_3\}$$
$$+\lambda_1 r_{\gamma_A}\cos(\Phi_{\gamma_A} - \Phi_{\gamma_C}) = 0. \tag{4.47h}$$

with M_i given in (4.40a- 4.40f). It is not straightforward to extract the solution from this set of equations. We will follow a strategy based on analytical and numerical methods. Due to equation (4.47d) there are two possible solutions, which are $\cos(\Phi_{\gamma_A} - \Phi_{\gamma_C}) = 1$ and $\cos(\Phi_{\gamma_A} - \Phi_{\gamma_C}) = -1$ (as r_{γ_A} and r_{γ_C} cannot be zero).

Let us first assume the option $\cos(\Phi_{\gamma_A} - \Phi_{\gamma_C}) = 1$. Note that in this case the set of equations (4.47a) - (4.47h) is invariant under the simultaneous exchange $r_{\beta_A} \leftrightarrow r_{\gamma_A}$ and $r_{\beta_C} \leftrightarrow r_{\gamma_C}$. As we can see from (4.41), the mutual information function is also symmetric under this exchange. Now, we combine the set of the

equations (4.47a)-(4.47h) to one joint equation in terms of p, Q and ${r_{\beta_A}}^2$. The task is to find all roots for ${r_{\beta_A}}^2$. From equations (4.47b) and (4.47c) we have ${r_{\gamma_A}}^2 = 1 - {r_{\beta_A}}^2$ and ${r_{\gamma_C}}^2 = 1 - {r_{\beta_C}}^2$. This fact, together with the symmetry mentioned above, means that there has to be an even number of roots for ${r_{\beta_A}}^2$ (if one finds a solution for $r^2_{\beta_A}$, then $1 - r^2_{\beta_A}$ is also a solution). Numerically (by plotting the joint equation in terms of ${r_{\beta_A}}^2$) we show that for different p and Q there are always exactly two roots. Analytically, ${r_{\beta_C}}^2 = 1 - {r_{\beta_A}}^2$ is a possible solution for the equations (4.47b)-(4.47h). By inserting this expression for ${r_{\beta_C}}^2$ as well as ${r_{\gamma_A}}^2$ and ${r_{\gamma_C}}^2$ (see above) into (4.47a) we find two solutions for ${r_{\beta_A}}^2$ which are parametrized in terms of p and Q. One of them is

$$ {r_{\beta_A}}^2 = |\beta_A|^2 = \tfrac{1}{2}\Big(1 + \tfrac{1}{1-\frac{p}{2}-Q}\sqrt{(Q-\tfrac{p}{2})(2-3Q-\tfrac{p}{2})}\,\Big) \tag{4.48} $$

and the other one is

$$ {r_{\beta_A}}^2 = |\beta_A|^2 = \tfrac{1}{2}\Big(1 - \tfrac{1}{1-\frac{p}{2}-Q}\sqrt{(Q-\tfrac{p}{2})(2-3Q-\tfrac{p}{2})}\,\Big). \tag{4.49} $$

Both of them lead to the same mutual information (this is clear from the symmetry, as explained above). Thus, we find that the maximal mutual information between Alice & Eve is

$$ \begin{aligned} I^{AE} &= 1 + (\frac{1-\frac{p}{2}-Q}{1-p})\Big\{\big((1-p)|\beta_A|^2 + \frac{p}{2}\big)\log\big((1-p)|\beta_A|^2 + \frac{p}{2}\big) \\ &+ \big(1 - \frac{p}{2} - (1-p)|\beta_A|^2\big)\log\big(1 - \frac{p}{2} - (1-p)|\beta_A|^2\big)\Big\} \\ &+ (\frac{Q-\frac{p}{2}}{1-p})\Big\{\frac{p}{2}\log\frac{p}{2} + (1-\frac{p}{2})\log(1-\frac{p}{2})\Big\}, \end{aligned} \tag{4.50} $$

Using these analytical and numerical considerations ensured that $|\beta_C|^2 = 1 - |\beta_A|^2$ is the unique relation between ${r_{\beta_A}}^2$ and ${r_{\beta_C}}^2$.

For the case $\cos(\Phi_{\gamma_A} - \Phi_{\gamma_C}) = -1$, we repeat the above process. However, equation (4.47a) is now *not* symmetric under the exchange $r_{\beta_A} \leftrightarrow r_{\gamma_A}$ and $r_{\beta_C} \leftrightarrow r_{\gamma_C}$. If we plot the joint function for equations (4.47a)-(4.47h)

in terms of ${r_{\beta_A}}^2$, we just find one root, and thus just expect one solution of the set of equations. Analytically we obtain $r_{\beta_C} = r_{\beta_A}$ as a possible solution. This leads to the following mutual information between Alice and Eve:

$$ I^{AE} = (\frac{Q-\frac{p}{2}}{1-p})\Big\{1 + \frac{p}{2}\log\frac{p}{2} + (1-\frac{p}{2})\log(1-\frac{p}{2})\Big\}. \tag{4.51} $$

Comparing the equations (4.50) and (4.51) analytically, we see that for all p and Q the function in (4.50) is bigger than (4.51). Therefore, equation (4.50) is the optimal mutual information.

$$\begin{aligned} I^{AE} &= 1 + (\frac{1-\frac{p}{2}-Q}{1-p}).\Big\{\big((1-p)|\beta_A|^2+\frac{p}{2}\big)\log\big((1-p)|\beta_A|^2+\frac{p}{2}\big) \\ &+ \big(1-\frac{p}{2}-(1-p)|\beta_A|^2\big)\log\big(1-\frac{p}{2}-(1-p)|\beta_A|^2\big)\Big\} \\ &+ (\frac{Q-\frac{p}{2}}{1-p})\Big\{\frac{p}{2}\log\frac{p}{2}+(1-\frac{p}{2})\log(1-\frac{p}{2})\Big\}, \end{aligned} \tag{4.52}$$

where, $|\beta_A|^2$ is given by equation (4.49). In Figure

4.4 Mutual information between Alice and Bob

Obtaining the mutual information between Alice and Bob is an easy task. Bob does a measurement in $z-$direction. The probabilities that he detects $|0\rangle$ and $|1\rangle$, given ρ_0^B and ρ_1^B are

$$p(0|\rho_0^B) = \langle 0|\rho_0^B|0\rangle = (1-\frac{p}{2})(1-D)+\frac{p}{2}D, \tag{4.53a}$$

$$p(0|\rho_1^B) = \langle 0|\rho_1^B|0\rangle = (1-\frac{p}{2})D+\frac{p}{2}(1-D), \tag{4.53b}$$

$$p(1|\rho_0^B) = \langle 0|\rho_0^B|0\rangle = (1-\frac{p}{2})(1-D)+\frac{p}{2}D, \tag{4.53c}$$

$$p(1|\rho_1^B) = \langle 0|\rho_0^B|0\rangle = (1-\frac{p}{2})D+\frac{p}{2}(1-D). \tag{4.53d}$$

Therefore, the probabilities of detecting $|0\rangle$ and $|1\rangle$ by Bob, are

$$p(0) = \frac{1}{2}\Big[(1-\frac{p}{2})(1-D)+\frac{p}{2}D+(1-\frac{p}{2})D+\frac{p}{2}(1-D)\Big] = \frac{1}{2}, \tag{4.54a}$$

$$p(1) = \frac{1}{2}\Big[(1-\frac{p}{2})(1-D)+\frac{p}{2}D+(1-\frac{p}{2})D+\frac{p}{2}(1-D)\Big] = \frac{1}{2}. \tag{4.54b}$$

From the definition (4.33), the mutual information between Alice and Bob is

then,

$$\begin{aligned} I^{AB} &= \sum_{x,y} p(x)p(y|x)\log p(y|x) - \sum_{y} p(y)\log p(y) \\ &= \left[(1-\frac{p}{2})(1-D)+\frac{p}{2}D\right]\log\left[(1-\frac{p}{2})(1-D)+\frac{p}{2}D\right] \\ &+ \left[(1-\frac{p}{2})D+\frac{p}{2}(1-D)\right]\log\left[(1-\frac{p}{2})D+\frac{p}{2}(1-D)\right]+1. \end{aligned}$$

Using(4.19), we rewrite the above mutual information (4.55) in terms of quantum bit error rate Q,

$$I^{AB} = 1 + Q\log Q + (1-Q)\log(1-Q). \tag{4.55}$$

One easily confirms that in the absence of white noise, i.e. for $p = 0$, the mutual information functions (Alice&Bob and Alice&Eve) reduce to the noiseless case described in [10].

4.5 Security proof

In this part we conclude that how the added quantum noise can make quantum key distribution more robust against eavesdropping. To do so, we start for an example with the noise parameter $p = 0.05$ which is introduced due to equation (4.1). For this amount of noise we plot the mutual information curves I^{AE} (equation 4.52) and I^{AB} (equation 4.55) as a function of the qubit error rate Q and we compare them with the noiseless case ($p = 0$). The plots are shown in **??**. As we observe for $p = 0.05$ both I^{AB} and I^{AE} start at a non-zero value for Q. Compare to the pure state case ($p = 0$), the mutual information between Alice and Bob I^{AB} remains invariant while the mutual information between Alice and Eve I^{AE} decreases. We also observe that for $p = 0.05$, the crossing-point between two mutual information curves (I^{AB} and I^{AE}) has moved towards a larger Q.

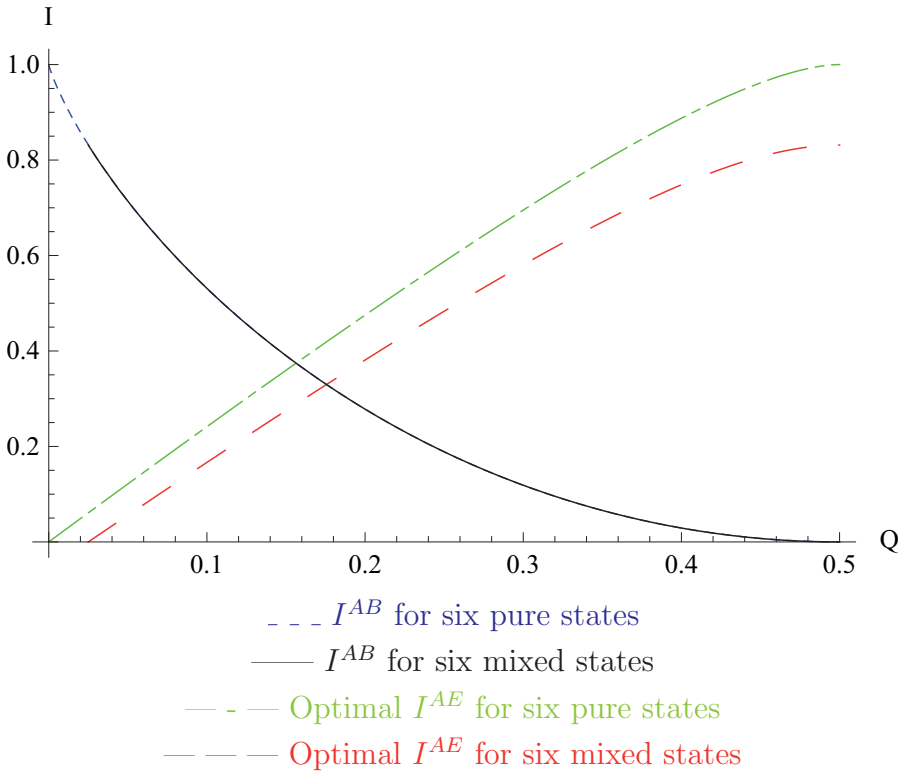

Figure 4.1: Mutual information between Alice & Bob and Alice & Eve for six pure ($p = 0$) and six mixed state ($p = 0.05$) cases, as a function of qubit error rate (Q).

Now, we pay attention on arbitrary p. For any $p \neq 0$ both information curves start at a non-zero value for Q, as we have $Q \geq p/2$ from equation (4.19)(like the mentioned example for $p = 0.05$). In Figure 4.2 for an example with quantum bit error rate $Q = 0.11$, the mutual information I^{AE} as a function of p has been depicted. It confirms again that with increasing the noise parameter p Eve loses more and more information on Alice's qubit.

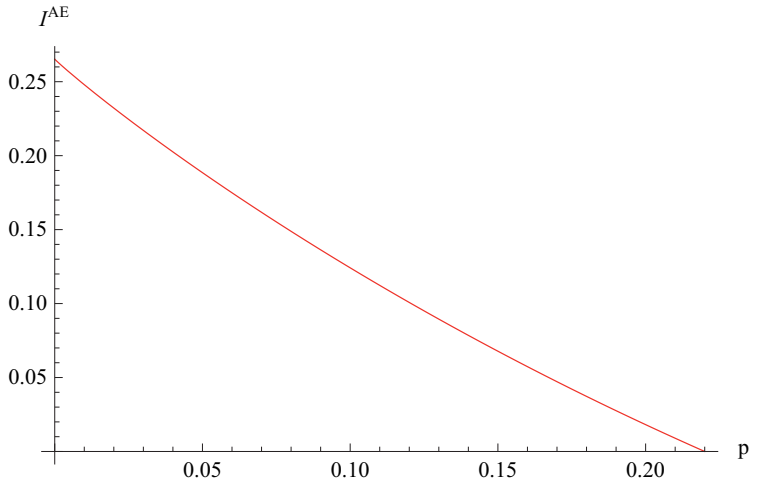

Figure 4.2: Mutual information I^{AE} as a function of noise parameter p for an example with quantum bit error rate $Q = 0.11$.

On the other hand, from equation (4.55), it is evident that I^{AB} does only depend on the quantum bit error rate Q and is "independent" on p. Intuitively speaking, both the trusted and untrusted parties undergo some degradation of their information due to the noise. The crossing point $Q(p) := Q$ as a function of noise p is give by

$$Q(p) = \texttt{Root}\left[I^{AB}(Q) - I^{AE}(Q,p) = 0\right], \tag{4.56}$$

where $I^{AE}(Q,p)$ and $I^{AB}(Q)$ are the same mutual information functions given in equation (4.52) and (4.55), respectively. We plot this Q-value for the crossing-

point (equation 4.56) as a function of noise p. The result is shown in Figure 4.3. For $p = 0$ the crossing point is $Q(p = 0) = 0.15637$. In equation (4.19) the relation between qubit error rate Q, disturbance D and noise p is given. One might expect that the crossing point of the two information curves obeys this linear dependence, i.e. $Q = D(1 - p) + p/2$ with $D = 0.15637$; this is the dashed line in Figure 4.3. However, the true value for the crossing point is the solid line and lies above that straight line. From Figure 4.3 (solid line) it is easy to see that as p increases the crossing point between two mutual information curves also moves to higher values.

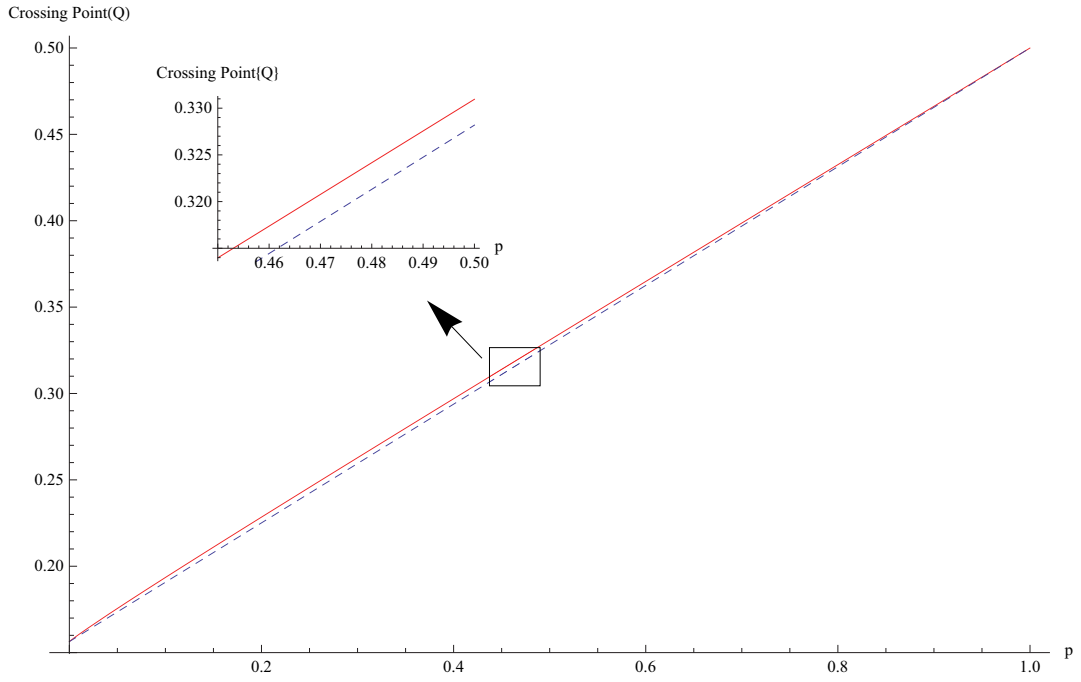

Figure 4.3: The solid line is the value of Q for the crossing point of I^{AB} and I^{AE}, as a function of the noise parameter p. The dashed straight line corresponds to equation (4.19) when D=0.15637. For details see text.

A natural question then arises that: Which consequence does this have for the creation of a secret key? Following the mentioned theorem by Csiszár and Körner (Theorem 4.1), a secret key can be established if $I^{AB} - I^{AE} \geq 0$, i.e. for values of Q which are smaller than the value for the crossing point. Therefore, the area that Alice and Bob can establish a secret after error correction and privacy amplification has been increased. In fact it shows that additional noise on the quantum level helps the trusted parties to improve the performance of a quantum cryptographic protocol. In other words, the six state protocol with mixed states is more robust against eavesdropping than the six state protocol with pure states. Following Theorem 4.2 we also find a lower bound on the key rate ($K \geq I^{AB} - I^{AE}$). The lower bound is illustrated in Figure 4.4 for the pure state case ($p = 0$) and for the mixed state case with $p = 0.05$. It shows that adding the noise also improves the lower bound on the key rate.

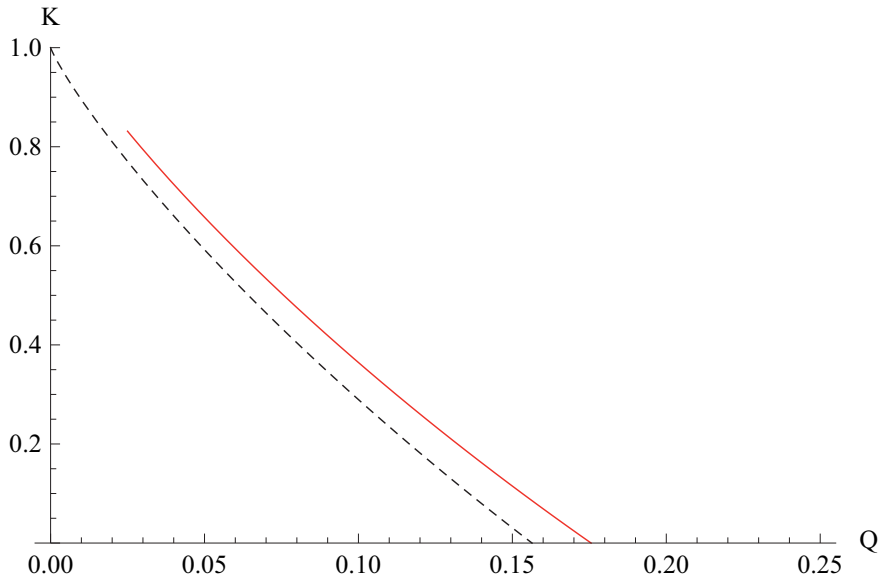

Figure 4.4: Lower bound on the key rate K as function of Q, for the individual eavesdropping strategy as described in the text. Dashed line: $p = 0$, solid line: $p = 0.05$.

4.6 Conclusion

In this chapter, a positive aspect of the quantum noise in an eavesdropping problem was provided. We investigated the six state protocol with additional noise on all signal states in the presence of an eavesdropper. We derived the optimal mutual information that Eve can obtain, when using individual attacks on noisy quantum signals, and compared it to the mutual information achievable by eavesdropping on pure states. As we expected, Alice and Bob, but also Eve, lose some information due to the additional noise. It was also shown that the threshold value of the quantum bit error rate, below which the mutual information between Alice & Bob is bigger than Alice & Eve, moves towards higher values (depending on the noise parameter p). This lead, by following a theorem by Csiszár and Körner, to an enlargement in the eligible area for which the trusted parties can produce a secret key. It was also depicted that due to the additional noise the lower bound on the key rate upgrades. In fact the extra noise improves the performance of the six state protocol and it becomes more robust against eavesdropping.

The present chapter has been mainly published in [46]. As an outlook, it would be interesting to study other protocols and/or other types of noise. It would be also interesting to look at the present eavesdropping problem from another point of view. That is to calculate the key rate as a function of the noise parameter p, number of the signals n, the quantum bit error rate Q, and a security parameter which is called ε. By adding extra noise we predict an improvement in the key rate. However, this claim should be mathematically verified.

References

[1] Julio T. Barreiro, Tzu-Chieh Wei, and Paul G. Kwiat. *Nature Phys.*, **4**:282, 2008. 28

[2] H. Bechmann-Pasquinucci and N. Gisin. *Phys. Rev. A*, **59**:4238, 1999. 82, 83

[3] C. H. Bennett. *Phys. Rev. Lett.*, **68**:3121, 1992. 83

[4] C. H. Bennett and G. Brassard. *Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore, India*, pages 175–179, 1984. 81, 83

[5] C. H. Bennett and et. al. *Phys. Rev. Lett.*, **83**:3081, 1999. 54

[6] C. H. Bennett and S. J. Wiesner. *Phys. Rev. Lett.*, **69**:2881, 1992. 27, 68, 69

[7] S. Bose. *Phys. Rev. Lett.*, **91**:207901, 2003. 20, 21

[8] S. Bose, V. Vedral, and P. L. Knight. *Phys. Rev. A*, **57**:822, 1998. 68, 69

[9] G. Bowen and S. Mancini. *Phys. Rev. A*, **69**:012306, 2004. 21

[10] D. Bruss. *Phys. Rev. Lett.*, **81**:3018, 1998. 82, 83, 98

[11] D. Bruss, G. M. D'Ariano, M. Lewenstein, C. Macchiavello, A. Sen(De), and U. Sen. *Phys. Rev. Lett.*, **93**:210501, 2004. 29, 30, 31, 69

[12] D. Bruss, G. M. D'Ariano, M. Lewenstein, C. Macchiavello, A. Sen(De), and U. Sen. *Int. J. Quant. Inform.*, **4**:415, 2006. 11, 29, 30, 31, 32

[13] D. Bruss, L. Faoro, C. Macchiavello, and G. M. Palma. *J. Mod. Optics*, **47**:325, 2000. 48

[14] I. Csiszár and J.Körner. *IEEE Trans. Inf. Theory*, **IT-24**:339, 1978. 82, 83

[15] S. Daffer, K. Wódkiewicz, and J. K. McIver. *Phys. Rev. A*, **67**:062312, 2003. 21

[16] E. B. Davies. *IEEE Inf. Theory*, **IT-24**:596, 1978. 94

[17] M. J. Donald. *Math. Proc. Camb. Phil. Soc.*, **101**:363, 1987. 40

[18] D. I. Fivel. *Phys. Rev. Lett.*, **74**:835, 1995. 41

[19] C. A. Fuchs and A. Peres. *Phys. Rev. A*, **53**:2038, 1996. 84

[20] J. P. Gordon. *in Proc. Int. School. Phys. "Enrico Fermi, Course XXXI", ed. P.A. Miles*, page 156, 1964. 29

[21] T. Hiroshima. *J. Phys. A Math. Gen.*, **34**:6907, 2001. 29, 30, 37, 39

[22] A. S. Holevo. *Prol. Inf. Transm.*, **9**:110, 1973. 28, 29

[23] A. S. Holevo. The capacity of the quantum channel with general signal states. *IEEE Trans. Inf. Theory*, **44**:269–273, 1998. 21, 29

[24] M. Horodecki, P. Horodecki, and R. Horodecki. *Phys. Rev. Lett.*, **80**:5239, 1998. 30

[25] M. Horodecki, P. Horodecki, and R. Horodecki. *Phys. Rev. A*, **60**:1888, 1999. 54

[26] M. Horodecki, P. Horodecki, R. Horodecki, D. W. Leung, and B. Terhal. *Quantum Inf. Comput.*, **70**, 2001. 30, 66

[27] M. Horodecki and M. Piani. *quant-ph/0701134v2.* 31, 66, 67

[28] R. Horodecki, P. Horodecki, and M. Horodecki. *Phys. Lett. A*, **210**:377, 1996. 30

[29] E. Karpov, D. Daems, and N. J. Cerf. *Phys. Rev. A*, **74**:032320, 2006. 21

[30] C. King. The capacity of the quantum depolarizing channel. *IEEE Transactions on Information Theory*, **49**:221–229, 2003. 51, 53

[31] S. K. Leung-Yan-Cheong and M. E. Hellman. The gaussian wire-tap channel. *IEEE Trans. Inf. Theory*, **24**:451–456, 1978. 82

[32] L. B. Levitin. *Inf. Theory, Tashkent*, page pp. 111, 1969. 29

[33] X. Li, Q. Pan, J. Jing, J. Zhang, C. Xie, and K. Peng. 31

[34] S. Lioyd. The capacity of a noisy quantum channel. *Phys. Rev. A*, **55**:1613–1622, 1997. 21

[35] X. S. Liu, G. L. Long, D. M. Tong, and Feng Li. *Phys. Rev. A*, **65**:022304, 2002. 69

[36] C. Macchiavello and G. M. Palma. *Phys. Rev. A*, **65**:050301(R), 2002. 53

[37] C. Macchiavello, G. M. Palma, and S. Virmani. *Phys. Rev. A*, **69**:010303(R), 2004. 21, 60

[38] S. Mancini. *J. Phys.: Conf. Ser.*, **36**:121, 2006. 21

[39] K. Mattle, H. Weinfurter, P. G. Kwiat, and A. Zeilinger. *Phys. Rev. Lett.*, **76**:4656–4659, 1996. 28

[40] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information.* Cambridge University Press, Cambridge, United Kingdom, 2000. 52, 91, 94

[41] R. Renner, N. Gisin, and B. Kraus. *Phys. Rev. A*, **72**:012332, 2005. 83

[42] B. Schumacher and M. D. Westmoreland. Sending classical information via noisy quantum channels. *Phys. Rev. A*, **56**:131–138, 1997. 21, 29

[43] Z. Shadman, H. Kampermann, C. Macchiavello, and D. Bruss. Multipartite super dense coding with noisy quantum channels. *In preparation*. 76

[44] Z. Shadman, H. Kampermann, C. Macchiavello, and D. Bruss. Optimal super dense coding with correlated noisy channels. *In preparation*. 68

[45] Z. Shadman, H. Kampermann, C. Macchiavello, and D. Bruss. Optimal super dense coding over noisy quantum channels. *New J. Phys.*, **12**:073042, 2010. 54

[46] Z. Shadman, H. Kampermann, T. Meyer, and D. Bruss. Optimal eavesdropping on noisy states in quantum key distribution. *Int. J. Quant. Inform.*, **7**:297, 2009. 103

[47] C. E. Shannon. Mathematical theory of communication. *Bell Syst. Tech. J*, **27**:379–423, 623–656, 1948. 13, 79

[48] C. E. Shannon and W. Weaver. *Mathematical Theory of Communication*. University of Illinois Press, 1963. 13

[49] R. F. Werner. *J. Phys. A Math. Gen.*, **34**:7081–7094, 2001. 54

[50] A. Winter. *J. Math. Phys.*, **43**:4341, 2002. 66

[51] W. K. Wootters and W. H. Zurek. A single quantum cannot be cloned. *Nature*, **299**:802, 1982. 81

[52] A. D. Wyner. The wire-tap channel. *The Bell System Technical Journal*, **54**:1355–1367, 1975. 82

[53] M. Ziman and V. Bužek. *Phys. Rev. A*, **67**:042321, 2003. 30

Acknowledgements

First and foremost, I offer my sincerest gratitude to my supervisor Prof. Dr. Dagmar Bruß who has supported me throughout this thesis with her patience, encouragement and knowledge whilst allowing me the room to work in my own way.

I would especially like to thank Dr. Hermann Kampermann for his open door policy to discuss the problems and providing me with intersting comments.

In my daily work, I have been blessed with a friendly and cheerful group colleagues. To Dr. Razmik Unanyan, Dr. Tim Meyer, Dr. Matthias Kleinmann, Dr. Colin Wilmott, Markus Mertz, Alexander Streltsov, Sylvia Bratzik, Silvestre Abruzzo and Christian Koch, I offer my thanks for the nice atmosphere and interesting discussions. I wish also to thank the institute system administrator Jens Bremer and the group secretary Cornelia Glowacki for their generouse help on the technical and official details.

Furthermore, I would like to express my gratefulness to Chiara Macchiavello and Barbara Kraus for helpful discussions.

I am grateful to all of my friends who supported me in any respect during the completing of the project, especially Marjan Bazrafshan, Bahar Rezaei and Daniela Marzi.

Finally, to my loving parents and family for their endless love, understanding and protection through the duration of my studies.

Erklärung

Hiermit erkläre ich vorliegende Dissertation selbstständig verfasst, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt, sowie Zitate kenntlich gemacht zu haben.

Düsseldorf, den 31.01.2011

(Zahra Shadman)