Classical Error-correcting Codes in Quantum Communications

Christopher E. Boyd

School of Electrical Engineering

Thesis submitted for examination for the degree of Master of Science in Technology.

Espoo 15.9.2014

Thesis supervisor:

Prof. Olav Tirkkonen

Thesis advisors:

D.Sc. Ülo Parts

D.Sc. Renaud-A. Pitaval

Aalto University
School of Electrical Engineering

AALTO UNIVERSITY
SCHOOL OF ELECTRICAL ENGINEERING

ABSTRACT OF THE
MASTER'S THESIS

Author: Christopher E. Boyd

Title: Classical Error-correcting Codes in Quantum Communications

Date: 15.9.2014 Language: English Number of pages: 7+80

Department of Communications and Networking

Professorship: Communications Engineering Code: S-72

Supervisor: Prof. Olav Tirkkonen

Advisors: D.Sc. Ülo Parts, D.Sc. Renaud-A. Pitaval

This thesis investigates the feasibility of utilising classical error-correcting codes in quantum communications, specifically in the entanglement-assisted communication of classical information over quantum depolarising channels. A classical-quantum communication system is presented, which relies on error-correction in the classical domain to achieve a classical information rate that approaches the entanglement-assisted capacity of a depolarising channel. Classical information is transmitted over the quantum channel by means of the superdense coding protocol. Different scenarios for the system are considered, including the noiseless and noisy distribution of initial entanglement resources and the use of higher-order entangled states. The overall transmission model corresponding to each scenario is shown to reduce to a classical, discrete and memoryless channel. The capacities of these equivalent classical channels are derived, and an inherent capacity loss when binary classical error correcting codes are employed is identified, motivating the use of non-binary codes. It is ultimately demonstrated that duo-binary turbo codes outperform binary turbo codes in all investigated scenarios, a result of the former's ability to exploit correlations between the pairs of classical bits communicated via the superdense coding protocol.

Keywords: Quantum communications, qubit, entanglement, superdense coding, error-correction coding, duo-binary turbo code

Preface

This thesis is the culmination of research work carried out in the Department of Communications and Networking at Aalto University School of Electrical Engineering, Helsinki, Finland.

The completion of the work detailed herein would not have been possible without the encouragement, guidance and support of my supervisors, colleagues and family.

I am immensely grateful to Prof. Olav Tirkkonen for setting me down such an interesting and challenging avenue of research, and to D.Sc. Ülo Parts for guiding my initial steps. Not having possessed any grounding in Quantum Mechanics or Quantum Information Theory, I would likely have stumbled without their patient tutelage.

I would also like to express my gratitude to my advisor and colleague D.Sc. Renaud-Alexandre Pitaval for his friendship, support, and critical eye, as well as for the countless hours of engaging conversation.

Finally, I'd like to thank my family for their unwavering love and emotional support from half the world away. I will never cease to appreciate it.

Christopher E. Boyd

Contents

Abbreviations

APP	A Posteriori Probability
BER	Bit Error Rate
BLER	Block Error Rate
BSC	Binary Symmetric Channel
CRSC	Circular Recursive Systematic Convolutional
CTC	Convolutional Turbo Code
DBTC	Duo-Binary Turbo Code
EPR	Einstein-Podolsky-Rosen
FEC	Forward Error Correction
FER	Frame Error Rate
IRCC	Irregular Convolution Code
LOCC	Local Operations and Classical Communications
LLR	Log Likelihood Ratio
MAP	Maximum A Poteriori
MDS	Maximum Distance Seperable
ML	Maximum Likelihood
NSC	Non-Systematic Convoluational
QECC	Quantum Error Correcting Code
RSC	Recursive Systematic Convolutional
SDC	Superdense Coding
SER	Symbol Error Rate
SISO	Soft Input Soft Output
SOVA	Soft Output Viterbi Algorithm
URC	Unitary Rate Code

1 Introduction

Recent technological advances have facilitated the development of ever faster and more efficient means of communicating and distributing information. Such technologies have become the cornerstone of modern society and the global economy, necessitating that they be both reliable and secure. The unrelenting demand for higher transmission rates and data security spurs the ongoing research and development of improved hardware, new encryption protocols and generally more efficient communication systems. However, the physical limitations of classical physics-based technologies could see progress slow, even cease, in the next decade, as the size of electronic components approaches the atomic scale. To break through this approaching barrier, it is necessary to venture into the quantum realm; to consider the implications of quantum physics and determine how the quantum properties of atomic and sub-atomic particles might be manipulated to facilitate computation and communication.

The relatively new field of quantum information theory, the convergence of classical information theory and quantum theory, has produced many exciting revelations and opportunities for improving upon classical communication by exploiting uniquely quantum phenomena. The primary of these being quantum entanglement, a means by which quantum information may be stored in the correlation between quantum states. The properties and consequences of entanglement cannot be described classically, necessitating and motivating the development of quantum information theory.

Quantum theory was well established by the 1930s, through the contributions of eminent physicists such as Plank, Einstein, Heisenberg and Schrödinger. Classical information theory was introduced by Shannon in 1948, with the paper "A Mathematical Theory of Communication". Bennet, Holevo, Schumacher and Shor were among those responsible for extending Shannon's work into the quantum realm throughout the remainder of the century, establishing the fundamentals of quantum information theory and developing multitudes of theoretical applications such as fault-tolerant quantum computation, quantum error-correcting codes and quantum communication protocols.

Quantum communications is the application of the fundamental results of quantum information theory for the purposes of communicating quantum or classical information. Quantum communication protocols, such as quantum teleportation and superdense coding, rely on the quantum correlation property induced by entanglement to perform seemingly miraculous communication tasks. The practical implementation of such protocols requires the consideration of the quantum noise effect of the environment on a quantum system, and the potential decoherence resulting in the loss of quantum information. Quantum channels have been defined which model the effect of discrete types of noise on a quantum state, which may induce errors. While quantum computation relies on strict error control, quantum communication can typically afford to employ error correction in order to achieve some acceptable level of transmission error.

The definition of the capacity of a noisy channel as the fastest rate at which information might be reliably transmitted over it, represents the most fundamental contribution of Shannon theory. Holevo was the first to extend this notion to quantum channels, and to define the single-use classical and quantum capacities of certain channels. Shannon's work proposed the existence of coding schemes which might enable transmission at rates near the capacity of a given channel. Since then, researchers have attempted to develop capacity-reaching, error-correcting codes, which protect information from transmission errors due to noise by the inclusion of appropriate redundancy. Turbo codes, in their various forms, represent one of the most successful schemes for error correction to date. Calberbank, Shor and Steane are among those who have pioneered quantum error-correcting codes, designed to preserve quantum information in the presence of noise.

Quantum information theory presents a wealth of new opportunities for exploiting the physical world for the purpose of communicating information. Research in the field of quantum communications is steadily moving beyond the theoretical, and protocols such as superdense coding have been successfully implemented experimentally using photons [2].

1.1 Motivation

Much like classical error-correcting codes, quantum error-correcting schemes employ a form of redundancy to protect information from corruption as a result of noise. Shor was the first to circumvent the no-cloning theorem of quantum mechanics, which had previously presented an obstacle to the formulation of quantum error-correcting codes. In 1995, he showed that quantum information could be stored in higher-order entangled states, and presented the first quantum error-correcting code. Since then, many such codes have been developed which play vital roles in the realisation of fault-tolerant quantum computing. However, it could be argued that, from a communications perspective, the use of quantum channel coding schemes would be inefficient under most circumstances. Not only is the generation and maintenance of higher-order entangled states increasingly difficult, but the coded information rate for even the most efficient quantum error correcting code to date would be unfeasibly low for communication purposes. For a communication system taking advantage of the inherent security of a quantum subsystem to transmit and receive classical information over a quantum channel, it may be preferable to transfer error-correction from the quantum to the classical domain.

There has been little research into the use of classical error-correcting codes in the entanglement-assisted communication of classical information over quantum channels; only one published paper could be found in the literature directly concerning the subject. A researcher at the University of South Hampton recently presented a custom, near-capacity code design for entanglement-assisted classical communication via the superdense coding protocol over a quantum depolarising channel [48]. The

1/2 rate code, a concatenation of a Unitary Rate Code (URC) and an Irregular Convolution Code (IRCC), was shown to provide an information rate within 0.6 dB of the capacity for a negligible bit-error-rate, and to outperform classical turbo codes. However, an information symbol-to-bit conversion inherent to the design of the URC-IRCC code limits the capacity of the overall transmission model. It is therefore proposed that non-binary or symbol-based classical error-correcting codes may be capable of providing increased performance.

1.2 Scope and Objectives

The scope of this thesis is classical error-correction coding in classical-quantum communication systems. The primary objective is to show that classical error-correcting codes are sufficient to provide acceptable error performance when transmitting classical information securely over noisy quantum channels, and that they represent a more efficient means of doing so than do quantum error-correcting codes. The secondary objective of the research is to determine whether or not, and to what degree, a non-binary classical channel code is capable of facilitating an error performance that exceeds that of an equivalent binary code. The intention is to achieve objectives by performing numerical simulations, in order to observe the performance of various classical error-correcting codes in a classical-quantum communication system employing practical variations of the superdense coding protocol to transmit classical information securely over a quantum depolarising channel. It is, in some regards, a cross-disciplinary venture, requiring an understanding of quantum theory, classical information theory and modern error-correcting methods.

1.3 Structure

This thesis is structured as follows. Chapter 2 provides a brief overview of the fundamental concepts in quantum information theory relevant to the research. Chapter 3 builds upon this theoretical foundation and introduces the field of quantum communications, which encompasses the various practical applications of the results of quantum information theory for the purposes of communicating quantum or classical information. Chapters 4 and 5 present the basic principles of quantum and classical error correction coding, respectively. Chapter 6 details a theoretical classical-quantum communication system, provides equivalent classical channels models for various operating scenarios, and derives the channel capacities. Chapter 7 evaluates the performance of various classical error correcting codes in the system via simulation, presenting and providing an analysis of the results obtained. Finally, conclusions regarding the research are drawn in Chapter 8.

2 Fundamentals of Quantum Information Theory

The following chapter provides a short overview of a number of fundamental topics in quantum information theory, presented in [1, 2, 3, 4]. Quantum information theory is the natural extension of the results of classical information theory into the quantum realm, and forms the theoretical background of quantum information processing and communication over quantum channels. A key result of quantum information theory is that the quantum nature of information and certain quantum phenomena can be exploited to expand beyond the boundaries of classical computation and communication systems.

2.1 The Quantum Bit and its Representations

The quantum bit or *qubit* is the quantum analogy of the classical bit, and forms the basic element or fundamental unit of quantum information. A qubit is the smallest possible two-state quantum system, whose physical implementation can be realised in a number of ways: in the spin states of electrons and atomic nuclei; the charge states of quantum dots; or the polarisation states of photons [2]. The first postulate of quantum mechanics asserts that a state of an isolated quantum system can be described by a unit vector in associated complex vector space, a finite-dimensional Hilbert space, called the *state vector* [4]. In quantum mechanics, Dirac or bra-ket notation is employed to represent state vectors. For a qubit, which has a two-dimensional state space, the state vector and the two orthonormal basis vectors in the state space are typically represented as *kets*, the latter written $|0\rangle$ and $|1\rangle$. The basis states form the *computational basis*, in which qubits are conventionally considered, manipulated and measured. In principle, it is possible to map a classical bit or *cbit* to a corresponding basis state as follows:

$$0 \rightarrow |0\rangle\,, \qquad 1 \rightarrow |1\rangle\,, \tag{2.1}$$

and if the state of a quantum system is limited to these orthogonal states, then the qubit is functionally equivalent to the cbit. What distinguishes the quantum bit from the classical is the qubit's ability to reside in a general superposition of the two basis states. A qubit may therefore be in any one of an infinite number of possible superposition states, i.e. linear combinations of states $|0\rangle$ and $|1\rangle$, while the cbit is limited to just two classical values. Such superposition states are considered *pure* states, and for a general noiseless qubit $|\psi\rangle$ have the form:

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle\,, \tag{2.2}$$

where α and β are complex numbers. These coefficients are probability amplitudes, whose squared magnitudes represent the probability that the qubit, when measured in the computational basis, will be found in the corresponding basis state. This is the measurement postulate of quantum information theory, known as the Born rule in quantum mechanics [3]. It follows that coefficients have the unit normalisation constraint:

$$|\alpha|^2 + |\beta|^2 = 1. \tag{2.3}$$

It is worthwhile to define a standard vector representation of the computational basis states and qubit superposition states. For the general qubit, an equivalent representation to the ket is a two-dimensional vector:

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle = \alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}. \tag{2.4}$$

In Dirac notation, the adjoint of a ket is its corresponding *bra*. The bra of the general qubit in a superposition state is written in vector form as:

$$\langle\psi| = \begin{bmatrix} \alpha^* & \beta^* \end{bmatrix}. \tag{2.5}$$

2.1.1 The Bloch Sphere

The Bloch sphere, named after the physicist Felix Bloch, provides a geometric representation of, and a useful way to visualise, the state space of a qubit. The Bloch sphere, depicted in Figure 2.1, has unit radius and its north and south poles are typically selected to represent the basis states $|0\rangle$ and $|1\rangle$, respectively. This choice is arbitrary, however, since all pairs of antipodal points on the surface correspond to mutually orthonormal states and could therefore be utilised as the basis. Any pure state can be mapped to a single point on the surface of the sphere, while points inside the sphere correspond to *mixed* states, which will be discussed later [2]. The centre of the sphere corresponds to the *maximally mixed state*. The unit vector from the origin to a point on the surface corresponding to a given pure state is called its Bloch vector.

The general idea is that the state of a qubit can be equivalently and completely described by unit spherical coordinates, namely the polar angle θ and the azimuth angle ϕ. Such a representation is possible because qubit states do not change under global phase shifts, i.e.,

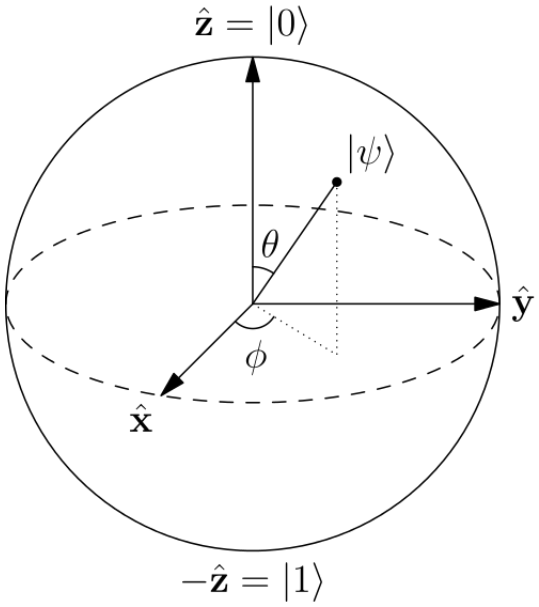


Figure 2.1: The Bloch Sphere [1].

$$|\psi\rangle \equiv e^{i\varphi} |\psi\rangle \,, \tag{2.6}$$

where $0 \leq \varphi < 2\pi$. As a consequence, the probability amplitude α can always be chosen to be a real number, and both coefficients α and β can be parameterised in terms of the polar angle θ and azimuthal angle ϕ, by utilising the constraint in equation 2.3 [1, 2]. This leads to the Bloch sphere representation of the general qubit:

$$|\psi\rangle = \cos\frac{\theta}{2} |0\rangle + e^{i\phi} \sin\frac{\theta}{2} |1\rangle \,, \tag{2.7}$$

where $0 \leq \theta < \pi$ and $0 \leq \phi < 2\pi$. The normalisation constraint is satisfied as:

$$\left|\cos\frac{\theta}{2}\right|^2 + \left|\sin\frac{\theta}{2}\right|^2 = 1, \tag{2.8}$$

and $|\psi\rangle = |0\rangle$ when $\theta = 0$, and $|\psi\rangle = |1\rangle$ when $\theta = \pi$, regardless of ϕ.

2.1.2 The Density Matrix and Density Operator

There exists another useful representation of the state of a qubit: the *density matrix*, a positive-semidefinite square matrix describing the projection operator onto the state. The density matrix of a general qubit in a pure state is the outer product of its bra and ket, as follows:

$$|\psi\rangle\langle\psi| = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \begin{bmatrix} \alpha^* & \beta^* \end{bmatrix} = \begin{bmatrix} \alpha\alpha^* & \alpha\beta^* \\ \beta\alpha^* & \beta\beta^* \end{bmatrix} = \begin{bmatrix} |\alpha|^2 & \alpha\beta^* \\ \beta\alpha^* & |\beta|^2 \end{bmatrix}. \tag{2.9}$$

It is worth noting that the rank of the density matrix for a pure state is equal to one, and that its diagonal entries are the respective probabilities of finding a qubit in the corresponding basis states upon measuring in the computational basis.

The density matrix provides a convenient approach to describing mixed states, where the state of a qubit is only partially known. Errors or imperfections in the preparation, evolution and measurement of pure quantum states can lead to uncertainty regarding the state of a qubit. The collective processes that produce a mixed state can be considered noise, and the uncertain state a noisy quantum state. While a pure state has a single state vector $|\psi\rangle$, a mixed state is a statistical *ensemble* or weighted sum of different pure states. This is described by the *density operator,* the quantum equivalent of a probability density function, which, for a qubit whose state is the ensemble $|\psi_n\rangle$ for some n, is defined as:

$$\rho = \sum_n p_n |\psi_n\rangle\langle\psi_n|, \tag{2.10}$$

where $p_n \geq 0$ are the classical probabilities that the pure state $|\psi_n\rangle$ occurs in the mixture, and $\sum p_n$=1. The density operator can replace the state vector as sufficient representation of a quantum state, mixed or otherwise, including classical uncertainty regarding the state.

Important properties of the density matrix and density operator may be revealed by the *trace* operation. The trace of a density matrix is the sum of the entries on the main diagonal, and therefore will always be equal to one because of the unit normalisation constraint [4]. In fact, for any density operator ρ, the trace is given by:

$$\mathrm{Tr}(\rho) = \mathrm{Tr}\left(\sum_n p_n |\psi_n\rangle\langle\psi_n|\right) = \sum_n p_n \mathrm{Tr}(|\psi_n\rangle\langle\psi_n|) = \sum_n p_n \langle\psi_n|\psi_n\rangle = 1. \tag{2.11}$$

2.2 Quantum Operations and Reversible Evolution

As in all physical systems, the state of a qubit evolves over time. The evolution of the state of a physical qubit can be directed, whether it be by applying a magnetic field to change the spin direction of an electron or elevating it to an excited state using a laser. State manipulation for the purposes of quantum information processing is achieved by applying such *control fields* in a known way, such that the background Hamiltonian which underlies the quantum system is modified and the qubit eventually evolves into the new state [1]. In quantum mechanics, the Hamiltonian represents the total energy of a system.

The evolution of a quantum state is described by the time-dependent Schrödinger equation [2]:

$$i\hbar\frac{\partial}{\partial t}\left|\psi\right\rangle = \mathcal{H}\left|\psi\right\rangle, \tag{2.12}$$

where $\hbar$ is the Dirac constant and $\mathcal{H}$ is the Hamiltonian operator. The partial differential equation has the solution:

$$\left|\psi\left(t\right)\right\rangle = U\left(t\right)\left|\psi\left(0\right)\right\rangle, \tag{2.13}$$

where

$$U\left(t\right) = e^{-i\frac{\mathcal{H}}{\hbar}t}. \tag{2.14}$$

From this result is derived the second postulate of quantum mechanics, which states that the evolution of a quantum system can be described by a unitary transformation U, and be represented by the following map:

$$\left|\psi\right\rangle \xrightarrow{\mathcal{H}} U\left|\psi\right\rangle.$$

The operator U is commonly referred to as the evolution operator or *propagator*, and is unitary because $\mathcal{H}$ is a Hermitian operator. The propagator alone is sufficient to describe the evolution, and every potential propagator U has an inverse, its conjugate transpose $U^{\dagger}$, where:

$$UU^{\dagger} = U^{\dagger}U = I, \tag{2.15}$$

implying that any evolution by such operators can be *reversed*. An exception to this is the measurement operator, which will be discussed later. An equivalent map describing unitary evolution in the density operator representation is:

$$\rho \xrightarrow{\mathcal{H}} U\rho U^{\dagger}. \tag{2.16}$$

Reversible evolution of quantum systems can be understood by considering the simple example of the quantum NOT or bit-flip gate, the quantum equivalent of the classical logic gate, operating on a single qubit. The quantum gate operates exactly like its classical counterpart on the basis states, and flips the latter for the general qubit in a superposition state. The propagator of the quantum NOT gate is denoted by X and its operation on the computational basis states and the general qubit $|\psi\rangle$ is described by:

$$|0\rangle \xrightarrow{X} |1\rangle\,, \qquad |1\rangle \xrightarrow{X} |0\rangle\,, \quad \text{and} \quad |\psi\rangle \xrightarrow{X} \alpha\,|1\rangle + \beta\,|0\rangle\,, \tag{2.17}$$

from which it is clear that the gate is its own inverse. Applying the operator twice will infallibly recover the original state, reversing the initial evolution.

A consequence of the unitary property of propagators is that transformations are inherently length preserving; they maintain the unit normalisation constraint from 2.3. The evolution of an isolated qubit is therefore described by a rotation of its corresponding unit Bloch vector around the origin of the Bloch sphere. For example, applying the quantum NOT gate is equivalent to rotating the Bloch vector 180° around the x axis.

2.3 Quantum Logic Gates

Quantum information processing requires logic gates and circuits, as does its classical counterpart, which transform information bits from one state to another in a useful way. In quantum logic circuits, gates perform quantum operations and cause qubits to evolve under specific, reversible unitary transformations[1].

A quantum gate that acts on a single qubit can be described by a two-by-two unitary matrix. In fact, any unitary matrix corresponds to a valid quantum gate [1]. The four most elementary and important quantum logic gates are based on the Pauli matrices:

[1]There also exists a number of special quantum logic gates which perform useful but irreversible evolutions, such as the SET, CLEAR and measurement gates.

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \; X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \; Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \; Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad (2.18)$$

all of which are Hermitian, unitary and square to identity. A complete list of the fundamental quantum operations that may be implemented by logic gates is included as Appendix A. I, X, Y and Z are the matrix representations (in the computational basis) of the identity, bit-flip (NOT), phase-flip and bit-phase-flip operators, respectively. The same notation is often used for the Pauli operators and their corresponding matrix representations. The Pauli matrices also define a matrix basis, by which any two-by-two matrix can be expanded as a weighted sum of the four matrices. This makes them additionally useful in describing single qubit states in density matrix form.

Perhaps even more important than the four basic quantum logic gates is the *Hadamard gate*. The Hadamard gate transforms the computational basis to a new basis, the Hadamard basis, according to:

$$|0\rangle \rightarrow |+\rangle \equiv \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \quad \text{and} \quad |1\rangle \rightarrow |-\rangle \equiv \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \quad (2.19)$$

The basis states $|+\rangle$ and $|-\rangle$ are equally weighted superpositions which lie on the equator of the Bloch sphere and have no classical interpretation. These states serve an important function in quantum information processing and communications, which will be revealed later. The matrix representation of the Hadamard operator H in the computational basis is:

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \quad (2.20)$$

2.4 Measurement

There exists no experiment or sequence of experiments by which the state of a single qubit might be determined, making it impossible to observe and accurately characterise the quantum system in a superposition state [2]. The reason for this is that the probability amplitudes α and β are unobservable quantities; a fact of nature which prevents a single qubit from being able to convey an infinite amount of classical information. It is possible, however, to retrieve some classical information from a qubit. The measurement or readout logic gate performs such an operation and is unique to quantum theory.

Observables in nature are those physical variables which may be represented as Hermitian operators, because their eigenvalues are real numbers which are valid outputs of the measurement gate. The Pauli operators are prime examples of single qubit observables: measuring the Z operator is equivalent to measuring in the computational basis. The eigenvalues of the Z operator are ± 1, which may be readily mapped to cbits.

Quantum measurement is an irreversible evolution which is destructive, as the result of the measurement is the system "collapsing" from a superposition into a measurement basis state [1, 2, 3]. Measurement is the third postulate of quantum mechanics, which states that, if given a set of measurement (projection) operators $\{\Pi_n\}$ which covers the n subspaces of a quantum system, a measurement of the state of said system will produce classical outcome n with probability:

$$\Pr\left(n \,|\, |\psi\rangle\right) = \langle\psi|\, \Pi_n^\dagger \Pi_n \,|\psi\rangle\,, \tag{2.21}$$

or equivalently, in the density operator representation:

$$\Pr\left(n \,|\, \rho\right) = \mathrm{Tr}\left(\Pi_n^\dagger \rho \Pi_n\right). \tag{2.22}$$

The post-measurement or *posterior state* of the quantum system is given by:

$$\left|\psi'\right\rangle = \frac{\Pi_n \,|\psi\rangle}{\sqrt{\langle\psi|\, \Pi_n^\dagger \Pi_n \,|\psi\rangle}}, \tag{2.23}$$

or equivalently:

$$\rho' = \frac{\Pi_n^\dagger \rho \Pi_n}{\mathrm{Tr}\left(\Pi_n^\dagger \rho \Pi_n\right)}. \tag{2.24}$$

In the case of a measurement in the computational basis of a general qubit, there are two measurement operators:

$$\Pi_0 = |0\rangle\langle 0|\,, \quad \text{and} \quad \Pi_1 = |1\rangle\langle 1|\,. \tag{2.25}$$

There can only be one of two outcomes: a classical bit 0 if the measured state is $|0\rangle$; or a 1 if the state is $|1\rangle$, with probability $|\alpha|^2$ and $|\beta|^2$, respectively. A measurement may be performed in any orthonormal basis, but will only ever yield a binary result corresponding to a basis state. Post-measurement, the state of the qubit

is permanently changed to a basis state, which means that repeated measurements will not yield any additional information. This fact prevents α and β from being statistically determined.

To understand the consequences of measurement on the state of a qubit, it is useful to consider the case where the measurement outcome k is lost or inaccessible. The state of the qubit after such a *forgetful measurement,* rather than being definitively one basis state or the other, is a weighted sum of the basis states, that is, a mixed state. In the density matrix notation, the effect of a measurement in the computational basis on the general qubit in a superposition state is:

$$M(\rho) = \sum_n \Pi_n \rho \Pi_n^\dagger = \alpha\alpha^* |0\rangle \langle 0| + \beta\beta^* |1\rangle \langle 1| = \begin{bmatrix} |\alpha|^2 & 0 \\ 0 & |\beta|^2 \end{bmatrix}. \tag{2.26}$$

The off-diagonal elements in the density matrix contain information regarding the *coherent superposition* state of the qubit and are reduced to zero by the measurement, indicating that all such information has been completely lost. This is referred to as complete *quantum decoherence.*

Interactions between a quantum system and its classical environment are irreversible and equivalent to measurements in random bases [1]. A qubit that is not completely insulated from the environment will therefore be subject to decoherence, have its state become increasingly mixed, and eventually lose its quantum information to its surroundings.

2.5 Composite Quantum Systems

The power of quantum information processing becomes apparent as qubits are combined to form higher-dimensional quantum systems. A two qubit system, where the qubits are labeled a and b, has four basis states, to which ordered pairs of cbits can be mapped as might be expected:

$$00 \to |0_a 0_b\rangle, \quad 01 \to |0_a 1_b\rangle, \quad 10 \to |1_a 0_b\rangle, \quad \text{and} \quad 11 \to |1_a 1_b\rangle. \tag{2.27}$$

As with the single qubit, the composite system can reside in any possible linear combination or general superposition of these four basis states:

$$|\xi\rangle = \alpha |0_a 0_b\rangle + \beta |0_a 1_b\rangle + \gamma |1_a 0_b\rangle + \delta |1_a 1_b\rangle, \tag{2.28}$$

with the unit normalisation constraint on the probability amplitudes:

$$|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 = 1. \tag{2.29}$$

Evidently, the state of a two qubit system occupies a four-dimensional Hilbert space, while a three qubit system will occupy an eight-dimensional Hilbert space, and so forth. In this exponential increase in the state space with linear increase in the number of qubits lies the power of quantum computation.

The fourth and final postulate of quantum mechanics concerns composite systems. The *tensor* or *direct product* is employed to simplify the description of composite quantum system states in higher dimensional Hilbert spaces, when the states of the subsystems are independent and *separable*. Separable or product states are those in which the individual qubits can be considered separately, and represent only a small subset of the states accessible to a multi-qubit system. The basis state $|0_a 0_b\rangle$ of a two-qubit system, for example, is separable because it can be written as a direct product:

$$|0_a 0_b\rangle = |0_a\rangle \otimes |0_b\rangle \,. \tag{2.30}$$

Composite states that cannot be decomposed and represented as a direct product of the states of the individual qubits are considered to be *entangled*.

The standard vector representation of a two qubit system in the general superposition state $|\xi\rangle$ using the direct product is as follows:

$$\begin{aligned} |\xi\rangle &= \alpha\,|0\rangle \otimes |0\rangle + \beta\,|0\rangle \otimes |1\rangle + \gamma\,|1\rangle \otimes |0\rangle + \delta\,|1\rangle \otimes |1\rangle\,, \\ &= \alpha \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \gamma \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + \delta \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{bmatrix}. \end{aligned} \tag{2.31}$$

2.5.1 Evolution of Multi-qubit Systems

Composite quantum systems may also undergo reversible evolution. In a two qubit system, it is possible to perform a unitary operation on the states of either or both qubits simultaneously (bilateral operations). The direct product is used to generate the matrix representations of the unitary operators in composite systems.

For example, the matrix representation of the NOT operator in a two-qubit system, when it is to be applied only to the first qubit in the pair (the identity operator is applied to the second), can be found as follows:

$$X_a = X \otimes I = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \tag{2.32}$$

where X_a means that the NOT operator is applied to qubit a. Applying the above operator on the two qubit system in the general superposition state $|\xi\rangle$ yields:

$$|\xi\rangle \overset{X_a}{\rightarrow} \begin{bmatrix} \gamma \\ \beta \\ \alpha \\ \delta \end{bmatrix}. \tag{2.33}$$

An equivalent approach can be used for any operator and for composite systems with greater numbers of qubits.

2.5.2 Multi-qubit Logic Gates and Controlled Gates

There exists an important two-qubit logic gate and corresponding unitary evolution: the controlled-NOT (CNOT) gate. The gate is key in the generation of entangled states, as will be seen later. The matrix representation of the CNOT gate cannot be written as a direct product, but has the explicit form:

$$\text{CNOT} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \tag{2.34}$$

and has the effect, in the computational basis, of flipping the second qubit given that the state of the first is $|1\rangle$, i.e.,

$$|00\rangle \overset{\text{CNOT}}{\rightarrow} |00\rangle, \quad |01\rangle \overset{\text{CNOT}}{\rightarrow} |01\rangle, \quad |10\rangle \overset{\text{CNOT}}{\rightarrow} |11\rangle, \quad |11\rangle \overset{\text{CNOT}}{\rightarrow} |10\rangle, \quad \text{and}$$

$$|\xi\rangle \overset{\text{CNOT}}{\rightarrow} \alpha|00\rangle + \beta|01\rangle + \gamma|11\rangle + \delta|10\rangle. \tag{2.35}$$

The idea of the conditional control gate can be extended to apply to any unitary operation on a single or multiple qubits in a composite system, contingent on the state of any individual or number of the remaining qubits.

2.5.3 Measurement in Composite Quantum Systems

For a two-qubit system, there are four necessary measurement (projection) operators:

$$\Pi_{00} = |00\rangle\langle 00|, \;\; \Pi_{01} = |01\rangle\langle 01|, \;\; \Pi_{10} = |10\rangle\langle 10|, \;\; \text{and} \;\; \Pi_{11} = |11\rangle\langle 11|, \quad (2.36)$$

and four possible measurement outcomes corresponding to the four basis states. Similar to the single qubit case, a measurement in the computational basis of a two qubit system in the general superposition state will produce the result 00, 01, 10 or 11 with probability $|\alpha|^2$, $|\beta|^2$, $|\gamma|^2$, and $|\delta|^2$, respectively.

It is possible to measure only a subset of the qubits in a composite system, an operation referred to as a *partial measurement*. Measuring only the first qubit in a two-qubit system, for example, with the standard set of measurement operators $\{\Pi_0, \Pi_1\}$ requires the *global* measurement operators $\{\Pi_0 \otimes I, \Pi_1 \otimes I\}$, since the identity is to be applied to the second qubit. The probability of obtaining result 0 is $|\alpha|^2 + |\beta|^2$, such that the posterior state of the original quantum system becomes:

$$\left|\xi'\right\rangle = \frac{\alpha|00\rangle + \beta|01\rangle}{\sqrt{|\alpha|^2 + |\beta|^2}}, \quad (2.37)$$

where the state has been renormalised to maintain the unit norm condition for a valid quantum state. Similarly, a result of 1 will occur with probability $|\gamma|^2 + |\delta|^2$, and the posterior state of the system becomes:

$$\left|\xi'\right\rangle = \frac{\gamma|10\rangle + \delta|11\rangle}{\sqrt{|\gamma|^2 + |\delta|^2}}. \quad (2.38)$$

2.5.4 Quantum Entanglement

Quantum entanglement is a phenomenon unique to quantum mechanics, that plays a key role in many quantum information processing tasks and practical applications of quantum theory. In fact, entanglement is the major distinguishing feature between the quantum realm and the classical [4]. It occurs when two or more quantum systems interact in such a way that their properties become strongly correlated; a correlation which persists regardless of how far the individual systems might be separated in space. As previously discussed, entanglement in composite quantum systems is indicated by the fact that the combined state cannot be separated and represented as a direct product of the states of the individual qubits. In other words, each qubit can no longer be described independently and the composite system must be considered as a whole. Entangled quantum states are sometimes referred to as EPR states, after Einstein, Podolsky and Rosen [4]. The most important entangled states for two qubit systems are the four Bell states:

$$\left|\Phi^{\pm}\right\rangle = \frac{|00\rangle \pm |11\rangle}{\sqrt{2}}, \qquad \left|\Psi^{\pm}\right\rangle = \frac{|01\rangle \pm |10\rangle}{\sqrt{2}}, \tag{2.39}$$

which are all *maximally entangled*, that is they are as far removed from a direct product state as possible. The Bell states together form an orthonormal basis for two qubit states, and each can be generated from a corresponding computational basis composite state by a simple quantum circuit, depicted in Figure 2.2, consisting of a Hadamard gate followed by a CNOT gate, e.g.

$$|00\rangle \overset{H_a}{\rightarrow} \frac{|00\rangle + |10\rangle}{\sqrt{2}} \overset{\text{CNOT}}{\rightarrow} \frac{|00\rangle + |11\rangle}{\sqrt{2}} = \left|\Phi^{+}\right\rangle. \tag{2.40}$$

It is important to note that each Bell state can be transformed to any other by the application of a unitary operator (X, Y or Z) to a single qubit of the entangled pair.

The effect of the quantum entanglement phenomenon is most evident when a partial measurement of the composite system is performed. A maximally entangled state is

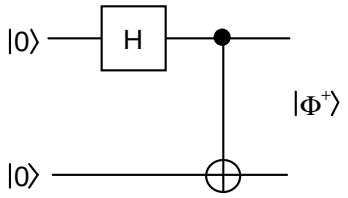


Figure 2.2: Quantum circuit for generating maximally-entangled Bell states. The Bell state which is generated is dependent on the input qubits.

effectively an equally weighted superposition of two basis states, and a measurement in the computational basis of one of the pair of qubits will return either a 0 (leaving a posterior state of $|00\rangle$), or 1 (leaving a posterior state of $|11\rangle$), with equal probability 1/2. The result of a partial measurement on a Bell state is the same regardless of which qubit of the pair is measured, indicative of the fact that the measurement outcomes are correlated [3]. In other words, the consequence of maximal entanglement is that finding one qubit in a particular state means that the second qubit must also be in that state. This means that it is possible to have immediate knowledge about the state of a potentially distant qubit by measuring its entangled counterpart. This consequence of entangling quantum systems has clear applications to communications.

2.5.5 Entanglement as a Shared Resource

A composite quantum system can be shared by multiple parties for the purposes of communicating information between them. Consider that two parties, one conventionally named Alice and the other Bob, share a two qubit system in the inseparable, entangled state:

$$\left|\Phi^+\right\rangle^{AB} = \frac{|0\rangle^A |0\rangle^B + |1\rangle^A |1\rangle^B}{\sqrt{2}}, \tag{2.41}$$

where the superscripts A and B indicate the part of the overall state which is possessed by Alice and Bob, respectively. The qubits must be in direct contact to be entangled, so the state is either prepared by a third party and distributed to Alice and Bob, or one party prepares the Bell state and transmits one of the qubits to the other party. It can be assumed that sharing the qubits is done in a perfect manner, such that the qubits remain maximally entangled once distributed. This shared entangled state is then a noiseless resource by which various communication tasks between the two parties can be performed.

While the qubits are physically separated, and the communication parties only have access to their *local* qubit, the entangled nature of the pair means that both Alice and Bob have equal control over the overall state of the shared system. For example, if Alice was to perform the *local operation* of applying a NOT gate to her qubit, the entangled state would become another Bell state, i.e.,

$$\left|\Phi^+\right\rangle^{AB} \xrightarrow{X_a} \frac{|1\rangle^A |0\rangle^B + |0\rangle^A |1\rangle^B}{\sqrt{2}} = \left|\Psi^+\right\rangle^{AB}. \tag{2.42}$$

If the shared qubits were not entangled and the composite system was in a separable state, a local operation performed by one party would have no consequence to the other.

2.5.6 Local Measurements and the Local Density Operator

Under certain circumstances, Alice or Bob may desire to measure their local qubit of an arbitrary shared, bipartite state ρ^{AB}. Such a *local measurement* is clearly equivalent to a partial measurement of the composite state. The states of the subsystems local to Alice and Bob can be represented by *local density operators*, which predict the results of local measurements and are expressed by the *partial trace*. For example, for the composite system:

$$\rho^{AB} = \left|\psi_n^A\right\rangle \left\langle\psi_n^A\right| \otimes \left|\psi_n^B\right\rangle \left\langle\psi_n^B\right|, \tag{2.43}$$

Alice's local density operator ρ_a^{AB} can be obtained as follows:

$$\rho_a^{AB} = \mathrm{Tr}_B\left(\rho^{AB}\right) \tag{2.44}$$

$$= \left|\psi_n^A\right\rangle \left\langle\psi_n^A\right| \mathrm{Tr}\left(\left|\psi_n^B\right\rangle \left\langle\psi_n^B\right|\right) \tag{2.45}$$

$$= \left|\psi_n^A\right\rangle \left\langle\psi_n^A\right| \left\langle\psi_n^B\middle|\psi_n^B\right\rangle \tag{2.46}$$

$$= \left|\psi_n^A\right\rangle \left\langle\psi_n^A\right|, \tag{2.47}$$

where Tr_B is the partial trace operator taken over the subspace of Bob's qubit, tracing out Bob's subsystem and leaving Alice's unchanged, and Tr is the partial trace operator taken over the entire space.

3 Quantum Communications

Quantum communications involves the transmission of a quantum state from one location to another. By communicating quantum information, certain processing tasks may be performed that would be inefficient, or even impossible, using classical information. Since the advent of quantum information theory, many quantum communication protocols have been developed which represent the direct application of the results of the theory to communications. These protocols typically involve a single sender, Alice, and receiver, Bob, and were initially devised as being ideal and noiseless. When the protocols are enacted in a practical environment, however, transmitted quantum states are subject to noise and the likelihood of undesirable evolutions.

3.1 Noisy Evolution and Quantum Channels

Imperfection in the evolution of a pure quantum state will inevitably result in a loss of information regarding the state, introducing noise and producing a mixed state [1]. Such noisy evolution is modeled by *quantum channels*, some of which are analogous to classical channel models. Quantum channels model communication at an abstract level, so it is sufficient to consider them from an information transmission perspective [4]. A noisy quantum channel essentially performs, with a given probability, some undesirable transformation on the input quantum system. Any given noisy quantum channel has the representation:

$$\rho \to \sum_j A_j \rho A_j^\dagger, \tag{3.1}$$

where A_j are called *Kraus operators*, which must satisfy:

$$\sum_j A_j A_j^\dagger \leq I, \tag{3.2}$$

and

$$\sum_j A_j^\dagger A_j \leq I. \tag{3.3}$$

Equality in the first condition on the Kraus operators guarantees trace preservation and in the second that the map is unitary. The perfect quantum channel has a single Kraus operator: the identity operator.

3.1.1 The Bit-flip Channel

The simplest example of a noisy quantum channel is the quantum bit-flip channel, which applies the X operator with probability p and the identity operator with probability $1 - p$. For the general qubit $|\psi\rangle$, the density operator describing the resulting mixed state is:

$$\rho = pX |\psi\rangle \langle\psi| X^\dagger + (1 - p) |\psi\rangle \langle\psi| . \tag{3.4}$$

The behaviour of the bit-flip channel can be generalised for any input quantum state ρ, by considering the channel to be a completely positive, trace-preserving and unital linear map which acts on the state as follows:

$$\rho \to pX\rho X^\dagger + (1 - p)\,\rho, \tag{3.5}$$

where the output is a "more mixed" version of the input, as a result of the channel.

3.1.2 The Dephasing Channel

The dephasing or phase-flip channel acts on a given density operator according to:

$$\rho \to pZ\rho Z^\dagger + (1 - p)\,\rho, \tag{3.6}$$

where p is the probability of the Z operator being applied. There is no classical equivalent to the dephasing channel, since phase is a property specific to quantum systems. Interestingly, the behaviour of the dephasing channel for $p = \frac{1}{2}$ is equivalent to that of a forgetful measurement, in which the input qubit is measured in the computational basis and the result is lost or never communicated to the receiver, ensuring complete decoherence.

3.1.3 The Pauli Channel

The Pauli channel is a generalisation of the bit-flip and phase-flip quantum channels. It applies a random Pauli operator according to a given probability distribution, resulting in the map:

$$\rho \to p_I\rho + p_X X\rho X^\dagger + p_Y Y\rho Y^\dagger + p_Z Z\rho Z^\dagger, \tag{3.7}$$

where p_I, p_X, p_Y and p_Z are the probabilities of the I, X, Y and Z operators being applied, respectively.

3.1.4 The Depolarising Channel

The depolarising channel is an important model for quantum noise, and represents the worst-case scenario of an input qubit being completely lost with some probability. In other words, an input qubit is replaced with the maximally mixed state π with probability p, where:

$$\pi = \frac{1}{4}\rho + \frac{1}{4}X\rho X^\dagger + \frac{1}{4}Y\rho Y^\dagger + \frac{1}{4}Z\rho Z^\dagger, \tag{3.8}$$

such that the channel map is:

$$\rho \to p\pi + (1-p)\rho = \frac{p}{4}X\rho X^\dagger + \frac{p}{4}Y\rho Y^\dagger + \frac{p}{4}Z\rho Z^\dagger + \frac{4-3p}{4}\rho. \tag{3.9}$$

The depolarising channel is useful when modeling physical channels about which no information is known or whose nature is yet to be estimated [1]. Literature on quantum error-correction exhibits a singular fondness for the depolarising channel, which can be attributed to the fact that the ability to error-correct the channel implies the ability to error-correct an arbitrary quantum operation on a qubit.

3.1.5 The Amplitude Damping Channel

The amplitude damping channel is a useful model for the noisy evolution that occurs in many physical systems, including the spontaneous emission of a photon from an atom. A qubit can be realised in the state of an atom with two energy levels: an arbitrary baseline energy state or ground state which can be considered $|0\rangle$; and an excited state that be can be considered $|1\rangle$. The process of spontaneous emission tends to decay such an atom from its excited state to its ground state, regardless of whether the atom resides in a superposition of the two states or not [1]. The Kraus operator that describes this behaviour is:

$$A_0 = \sqrt{\gamma}\,|0\rangle\langle 1|\,, \tag{3.10}$$

where γ is the probability of decay occurring. Applying the operator to the excited state decays it to the ground state with probability γ, i.e.,

$$A_0\,|1\rangle\langle 1|\,A_0^\dagger = \gamma\,|0\rangle\langle 0|\,, \tag{3.11}$$

and applying it to the ground state has no effect, i.e.,

$$A_0 \left|0\right\rangle \left\langle 0\right| A_0^\dagger = 0. \tag{3.12}$$

The operator A_0 alone is insufficient to specify a map for the channel, since it does not satisfy the condition on Kraus operators presented in equation 3.3. The following additional operator is required to met the condition:

$$A_1 = \left|0\right\rangle \left\langle 0\right| + \sqrt{1-\gamma}\left|1\right\rangle \left\langle 1\right|, \tag{3.13}$$

such that the map for the quantum amplitude damping channel becomes:

$$\rho \to A_0 \rho A_0^\dagger + A_1 \rho A_1^\dagger. \tag{3.14}$$

3.1.6 The Erasure Channel

The classical erasure channel has the function of replacing a transmitted bit of information with a known erasure symbol e with some probability ε. As a consequence, the output alphabet of the channel has an additional symbol compared with the input alphabet. The quantum erasure channel is simply the generalisation of the classical channel to the quantum realm, and can serve, for example, as a simplified model describing photon loss in optical systems. The quantum channel implements the following map:

$$\rho \to \varepsilon \left|e\right\rangle \left\langle e\right| + (1-\varepsilon)\,\rho, \tag{3.15}$$

where $\left|e\right\rangle$ is an erasure state that is orthogonal to the Hilbert space of the input state. The output space of the quantum erasure channel is one-dimension larger than the input space. Erasures may be detected at the receiving end of the channel via measurement, with the addition of the measurement operator $\Pi_e = \left|e\right\rangle \left\langle e\right|$ to the set of projectors on the input Hilbert space. Quantum information is preserved whenever erasures do not occur, since the erasure state and measurement operator are orthogonal to the input space.

3.2 Characterising Quantum Channels

From a communications perspective, a quantum channel represents the transfer of a quantum system from a transmitter to a receiver separated in space. The channel is considered noiseless if the system traversing it arrives at its destination intact, and noisy if it has interacted with other systems en route such that its state becomes mixed. In quantum information theory, there are several ways to characterise a quantum channel: fidelity, quantum discord and capacity.

3.2.1 Fidelity

In quantum information theory, fidelity is a measure of how close two quantum states are to one another [2]. In the context of quantum channels, fidelity measures how closely the output of a noisy channel resembles the input. In other words, it is a measure of how well a quantum channel preserves information; the *reliability* of the channel. If an arbitrary pure state $|\psi\rangle$ transmitted through a quantum channel emerges as the mixed state ρ, the fidelity of the channel is defined as:

$$F = \sqrt{\langle\psi|\rho|\psi\rangle}. \tag{3.16}$$

If the output state is equivalent to the input, i.e. the channel is noiseless, the fidelity is:

$$F^2 = \begin{bmatrix} \alpha^* & \beta^* \end{bmatrix} \begin{bmatrix} \alpha\alpha^* & \alpha\beta^* \\ \beta\alpha^* & \beta\beta^* \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = 1. \tag{3.17}$$

In the case of the depolarising channel with depolarising probability p, the fidelity between the output state and a pure input state $|\psi\rangle$ is:

$$F^2 = \langle\psi| \left(p\pi + (1-p)|\psi\rangle\langle\psi|\right) |\psi\rangle = 1 - \frac{p}{2}, \tag{3.18}$$

which concurs with the intuition that the noisier the channel, the lower the fidelity of the resultant state with the initial state.

3.2.2 Quantum Discord

Quantum discord is a measure of the "quantumness", or quantum excess, of correlations [16]. It can be used to describe the effect of a noisy quantum channel on the correlations between a pair of entangled states, when one of the pair is transmitted over the channel. This scenario is prevalent in quantum communication protocols, as will be shown later.

In classical information theory, the correlation between two random variables $\mathcal{X}$ and $\mathcal{Y}$ is given by the following two equivalent equations for the mutual information:

$$I(\mathcal{X};\mathcal{Y}) = H(\mathcal{X}) + H(\mathcal{Y}) - H(\mathcal{X},\mathcal{Y}), \tag{3.19}$$

and

$$J(\mathcal{X};\mathcal{Y}) = H(\mathcal{X}) - H(\mathcal{X}|\mathcal{Y}), \tag{3.20}$$

where H is the classical entropy function, $H(\mathcal{X})$ is the entropy of $\mathcal{X}$, $H(\mathcal{X}|\mathcal{Y})$ is the conditional entropy of $\mathcal{X}$ given $\mathcal{Y}$, and $H(\mathcal{X},\mathcal{Y})$ is the joint entropy of $\mathcal{X}$ and $\mathcal{Y}$. While $I(\mathcal{X};\mathcal{Y}) = J(\mathcal{X};\mathcal{Y})$ in classical information theory, it has been shown that, when the concept of mutual information is generalised to quantum systems, the equations are no longer equivalent [16]. To define I for two entangled quantum systems $\mathcal{A}$ and $\mathcal{B}$, it is sufficient to replace the classical probability distributions with the combined density operator $\rho_{\mathcal{AB}}$ and reduced density operators $\rho_{\mathcal{A}}$, $\rho_{\mathcal{B}}$, and the Shannon entropy H with the von Neumann entropy S, where:

$$S(\mathcal{A}) = S(\rho_{\mathcal{A}}) = -\mathrm{Tr}(\rho_{\mathcal{A}}\ln\rho_{\mathcal{A}}), \tag{3.21}$$

such that:

$$I(\mathcal{A};\mathcal{B}) = S(\mathcal{A}) + S(\mathcal{B}) - S(\mathcal{A},\mathcal{B}) \tag{3.22}$$

$$= S(\rho_{\mathcal{A}}) + S(\rho_{\mathcal{B}}) - S(\rho_{\mathcal{AB}}). \tag{3.23}$$

The generalisation of J is not as straight forward, as there is no directly equivalent expression for the conditional entropy of $\mathcal{A}$ given $\mathcal{B}$ in quantum theory. Determining the state of $\mathcal{A}$ given the state of $\mathcal{B}$ requires first the selection of a measurement basis. Typically, perfect measurements of $\mathcal{B}$ are defined by a set of projectors:

$$\left\{\Pi_i^{\mathcal{B}}\right\} = \{I\otimes\Pi_0,\ I\otimes\Pi_1\}. \tag{3.24}$$

The state of $\mathcal{A}$ after measuring the outcome corresponding to each projector is given by:

$$\rho_{\mathcal{A}|\Pi_i^{\mathcal{B}}} = \frac{\Pi_i^{\mathcal{B}} \rho_{\mathcal{AB}} \Pi_i^{\mathcal{B}}}{\mathrm{Tr}\left(\Pi_i^{\mathcal{B}} \rho_{\mathcal{AB}}\right)}, \tag{3.25}$$

with probability $p_i = \mathrm{Tr}\left(\Pi_i^{\mathcal{B}} \rho_{\mathcal{AB}}\right)$. From this, it is possible to define a quantum generalisation of J, as follows:

$$J\left(\mathcal{A};\mathcal{B}\right)_{\left\{\Pi_i^{\mathcal{B}}\right\}} = S\left(\mathcal{A}\right) - S\left(\mathcal{A}|\left\{\Pi_i^{\mathcal{B}}\right\}\right) \tag{3.26}$$

$$= S\left(\rho_{\mathcal{A}}\right) - S\left(\rho_{\mathcal{A}|\Pi_i^{\mathcal{B}}}\right), \tag{3.27}$$

which is the amount of information obtained regarding the state of $\mathcal{A}$, as a result of the measurement $\left\{\Pi_i^{\mathcal{B}}\right\}$.

Finally, the quantum discord δ is defined as the difference between the quantum generalisations of the classical expressions for the mutual information, i.e.,

$$\delta\left(\mathcal{A};\mathcal{B}\right)_{\left\{\Pi_i^{\mathcal{B}}\right\}} = I\left(\mathcal{A};\mathcal{B}\right) - J\left(\mathcal{A};\mathcal{B}\right)_{\left\{\Pi_i^{\mathcal{B}}\right\}} \tag{3.28}$$

$$= S\left(\rho_{\mathcal{B}}\right) - S\left(\rho_{\mathcal{AB}}\right) + S\left(\rho_{\mathcal{A}|\Pi_i^{\mathcal{B}}}\right). \tag{3.29}$$

3.2.3 Capacity

As will be shown in the following section, quantum channels may be used to communicate both quantum and classical information, or even share entangled pairs, between remote parties. As a consequence, quantum channels have several distinct capacities, which define the limit on the reliable transmission of each type of information over the channel. This is in contrast to classical channels, which may be fully characterised by a single capacity. The limits on the transmission rate for classical and quantum information over a quantum channel are predictably referred to as the classical capacity C and quantum capacity Q, respectively. In addition, each quantum channel has both a *classically-assisted* quantum capacity Q_C (its capacity for transmitting intact quantum states with the aid of a duplex classical side-channel),

and an *entanglement-assisted* classical capacity C_E (for transmitting classical information by utilising an unlimited amount of noiseless entanglement between sender and receiver). These various capacities may be further defined as *single-use* or *asymptotic* capacities. The former referring to the limit on the information that might be transmitted in a single use of the channel, and the latter if arbitrarily many uses of the channel are permitted.

Holevo was the first to investigate the classical capacity of quantum channels [18]. His initial discoveries showed that the classical capacity of a channel was related to the *accessible information* of a quantum system; the amount of classical information that might be extracted from a system via an optimal measurement when said information is encoded using a particular ensemble of states. In general, the accessible information of an ensemble is difficult to calculate, but Holevo was able to define a readily-determined upper bound. This bound, referred to as the *Holevo information* χ, for the ensemble pf pure states ρ_i, where:

$$\rho = \sum_i p_i \rho_i \tag{3.30}$$

is given by:

$$\chi = S(\rho) - \sum_i p_i S(\rho_i). \tag{3.31}$$

The information rate χ was later shown to be asymptotically achievable over a quantum channel, and to therefore represent its classical capacity [43].

3.3 Quantum Communication Protocols

Quantum communication protocols represent practical applications of the postulates of quantum mechanics and the results of quantum information theory for the purposes of communicating classical information, quantum information or entanglement.

3.3.1 Elementary Coding

The most trivial of all the quantum communication protocols, elementary coding, is simply the bitwise encoding of cbits to qubits, in order to transmit classical information over quantum channels. A classical bit is mapped to one of the two orthogonal basis states of a qubit (e.g. as per equation (2.1)), such that the receiver can reliably recover the classical information sent over a noiseless quantum channel via a measurement. The steps of the protocol are:

1. Alice prepares a qubit in the computational basis state corresponding to the classical information bit to be transmitted.

2. Alice transmits the qubit to Bob over a noiseless quantum channel.

3. Bob performs a measurement in the computational basis on the qubit and, based on the result, determines what was sent.

The classical capacity or bound on encoding efficiency for this protocol is one classical bit of information per use of a noiseless qubit channel. This suggests that nothing is gained by using qubits instead of cbits to transmit information, however, there is a certain degree of data security to be gained in the use of quantum states to convey information. While the information is vulnerable to eavesdropping, observation (measurement) by an uninvited party destroys the quantum information contained in a transmitted state, such that the presence of the eavesdropper may be immediately detected by the receiver. Fortunately, the limit on the information rate can be exceeded, and the level of security enhanced, by exploiting the properties and consequences of quantum entanglement.

3.3.2 Entanglement Distribution

The entanglement distribution (ED) protocol is concerned with the generation and sharing of a maximally entangled Bell state between Alice and Bob, to be used as a communication resource. Quantum communication protocols that make use of entanglement typically assume that the involved parties initially possess half of an entangled bipartite state, sometimes referred to as one bit of entanglement or *ebit*. The resource can be created and distributed by a third party, conventionally called Eve, or by either communicating party, and may be noiseless or noisy depending on the quantum channels involved. The resulting shared state will not be a pure Bell state in the case of noisy distribution, but rather a mixed state. Figure 3.1 depicts two variations of the protocol. In the case where Eve prepares and distributes the bipartite state, the steps of the protocol are:

1. Eve prepares and combines two qubits $|0\rangle^{E}$ and $|0\rangle^{E'}$ into the composite state $|0\rangle^{E} \otimes |0\rangle^{E'}$.

2. Eve entangles the qubits by means of two local operations, a Hadamard gate followed by a controlled-NOT gate, producing the Bell state $|\Phi^{+}\rangle^{EE'}$.

3. Eve sends qubit E to Alice and E' to Bob over noiseless quantum channels, such that the two parties share one bit of entanglement $|\Phi^{+}\rangle^{AB}$.

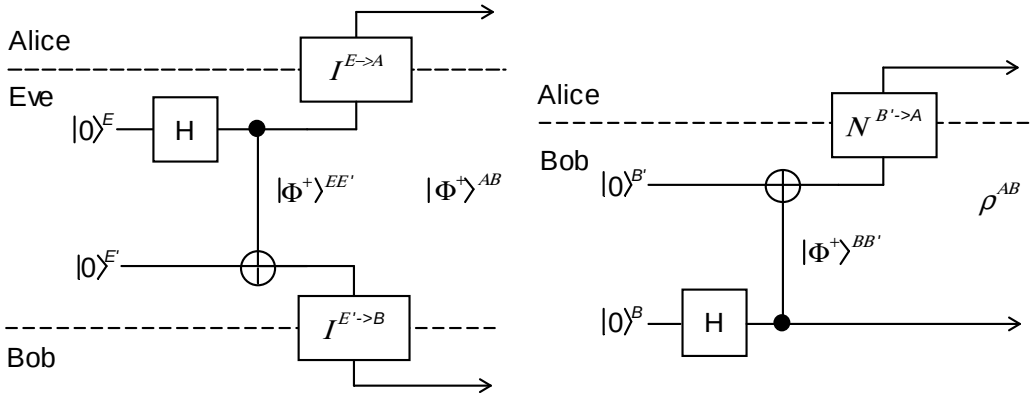


Figure 3.1: Quantum circuits depicting two arrangements of the entanglement distribution protocol.

3.3.3 Entanglement Distillation

The entanglement distillation protocol is concerned with the recovery of a pure, maximally-entangled state from a number of arbitrarily-entangled (mixed) states, such as Bell states that have been subject to noise during distribution between communication parties. Such mixed states are undesirable as an initial state for many quantum communication protocols, and so it is often preferable to sacrifice entanglement resources to generate a superior resource with which to enact the protocols. Entanglement distillation is also known as entanglement purification, since the protocol essentially generates an entangled pair that is purer than the input pairs. The protocol could play a singularly important role in combating decoherence, overcoming the degenerative effects of noisy quantum channels, and ultimately realising practical quantum communication. *Quantum repeaters*, theoretical devices which will potentially facilitate long-distance transmission of quantum information, are based on the principle of quantum distillation [4]. There are various methods of distilling entanglement, but the fundamental protocol, depicted in Figure 3.2, consists of the following steps:

1. Alice and Bob share two previously prepared and noisily distributed ebits ρ_1^{AB} and ρ_2^{AB}. Alice and Bob assign their local qubit of ρ_1^{AB} to be the *control* qubit, and their local qubit of ρ_2^{AB} to be the *target* qubit.

2. The parties combine the states of their local qubits by means of a local CNOT operation, generating a new entangled pair.

3. Alice and Bob measure their target qubits (destroying them), and communicate their respective results to one another over classical channels. If the results are equivalent, the entangled state of the unmeasured control qubits is known to have been purified.

4. The parties may choose to iterate the protocol and sacrifice additional noisy ebits to enable further purification, and eventually obtain a maximally-entangled state.

After sufficient iterations, the entanglement distillation protocol effectively simulates noiseless entanglement distribution over a noisy quantum channel, with the aid of local operations and classical communication (LOCC).

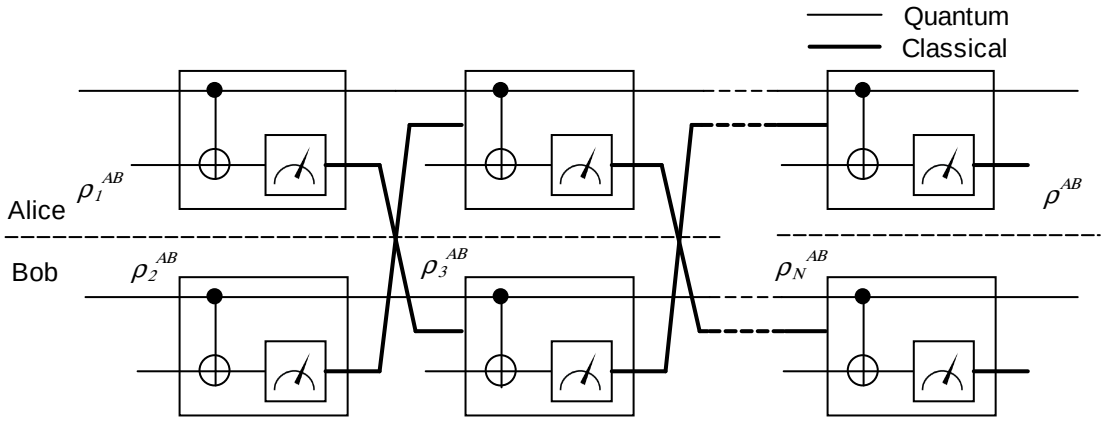


Figure 3.2: Quantum circuit depicting the entanglement distillation/purification protocol, consuming $N-1$ noisy ebits to purify a single noisy resource.

3.3.4 Superdense Coding

Perhaps the most important application of quantum information theory in terms of communications is known as the superdense coding (SDC) protocol, introduced by Bennet and Wieser [9]. The protocol is considered *dense* because it exploits the non-local properties of entangled states in order to allow two classical bits to be communicated in the transmission of a single qubit of an entangled pair. The classical capacity of the protocol is therefore double that of elementary coding; two classical bits of information per use of a noiseless qubit channel. The classical capacity is reduced for a noisy quantum channel, to a degree dependent on the channel properties. Figure 3.3 depicts the superdense protocol, whose steps are:

1. Alice and Bob share a previously prepared and distributed noiseless ebit $|\Phi^+\rangle^{AB}$. Alice applies a unitary operation on her local qubit, conditional on the two-bit classical message she desires to transmit. A potential encoding arrangement is shown in Table 3.1.

2. Alice transmits her qubit to Bob over a quantum channel.

3. Bob performs a Bell measurement (measurement in the Bell basis) to distinguish between the four orthonormal states and so determines the two message bits; perfectly if transmission is made over a noiseless channel or with certain error probability if over a noisy channel.

Message bits	Unitary Operation	Resulting Bell State
00	I	$\left\|\Phi^+\right\rangle^{AB}$
01	X	$\left\|\Psi^+\right\rangle^{AB}$
10	Z	$\left\|\Phi^-\right\rangle^{AB}$
11	$Y \equiv XZ$	$\left\|\Psi^-\right\rangle^{AB}$

Table 3.1: Encoding table for the superdense coding protocol.

A benefit of the superdense coding protocol, whose importance is not to be understated, is an inherent information security. Interception of the qubit passing through the channel between Alice and Bob will not enable an eavesdropper to determine which message was sent, since possession of both qubits and a measurement of the overall Bell state is required to determine it. In fact, a partial measurement of the intercepted qubit will produce the same result regardless of the message sent by Alice and offer no information whatsoever to the eavesdropper [1].

The entanglement-assisted classical capacity of the superdense coding protocol is the maximum rate a sender may reliably transmit classical information to a receiver, for a given initial state ρ^{AB}. If the entanglement distribution between communication parties, Alice and Bob, is made via a noiseless channel, the initial state for the SDC is a Bell state, as above. Over a noiseless channel and for unitary encoding, the superdense coding capacity is given by:

$$C_E = 1 + S\left(\rho_b^{AB}\right) - S\left(\rho^{AB}\right), \tag{3.32}$$

where ρ_b^{AB} is Bob's reduced or local density operator and $S\left(\rho^{AB}\right)$ is the Von Neumann entropy of the quantum state, as defined in subsection 3.2. If the bipartite state is not entangled, then:

$$S\left(\rho_b^{AB}\right) - S\left(\rho^{AB}\right) = 0, \tag{3.33}$$

and the capacity reduces to that of elementary coding: one cbit per channel use. Inversely, if the state is maximally entangled, then:

$$S\left(\rho_b^{AB}\right) - S\left(\rho^{AB}\right) = 1, \tag{3.34}$$

and the capacity is doubled to two cbits per channel use. Therefore, initial states with some degree of entanglement such that:

$$S\left(\rho_b^{AB}\right) - S\left(\rho^{AB}\right) > 0, \tag{3.35}$$

are those useful for superdense coding [44].

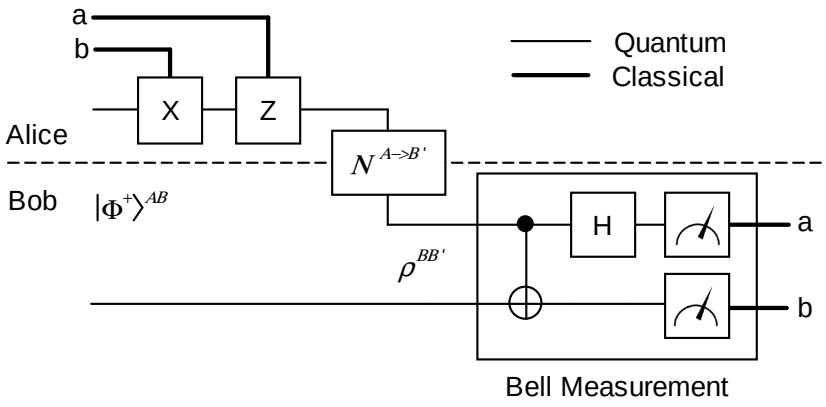


Figure 3.3: Quantum circuit depicting the superdense coding protocol.

3.3.5 Teleportation

The *quantum teleportation* protocol utilises classical communications to transport the state of a qubit from one party to another. The qubit state is destroyed at the origin, and, after some time (dependent on the limitations of the classical communication system), is recreated at the destination. The process of "teleporting" the qubit consumes one entanglement resource. The teleportation protocol prevents the need for the qubit to traverse a potentially noisy quantum channel between Alice and Bob, as in the case of superdense coding, and is, in some sense, the inverse of SDC. The protocol is depicted in Figure 3.4, and has the following procedure:

1. Alice and Bob share a previously prepared and distributed ebit $|\Phi^+\rangle^{AB}$. Alice possesses a qubit in the arbitrary state $|\psi\rangle^{A'} = \alpha\,|0\rangle^{A'} + \beta\,|1\rangle^{A'}$, which she would like to teleport to Bob. The state of the composite system is $|\psi\rangle^{A'}\,|\Phi^+\rangle^{AB}$, which can be written in terms of the outcome vectors of a measurement on the two qubits local to Alice as follows:

$$\begin{aligned} |\psi\rangle^{A'}\left|\Phi^+\right\rangle^{AB} &= \frac{1}{\sqrt{2}}\Big(\alpha\,|000\rangle^{A'AB} + \alpha\,|011\rangle^{A'AB} \\ &\quad +\beta\,|100\rangle^{A'AB} + \beta\,|111\rangle^{A'AB}\Big), \qquad (3.36) \\ &= \frac{1}{2}\Big(\left|\Phi^+\right\rangle^{A'A}\left[\alpha\,|0\rangle^{B} + \beta\,|1\rangle^{B}\right] \\ &\quad +\left|\Phi^-\right\rangle^{A'A}\left[\alpha\,|0\rangle^{B} - \beta\,|1\rangle^{B}\right] \\ &\quad +\left|\Psi^+\right\rangle^{A'A}\left[\beta\,|0\rangle^{B} + \alpha\,|1\rangle^{B}\right] \\ &\quad +\left|\Psi^-\right\rangle^{A'A}\left[-\beta\,|0\rangle^{B} + \alpha\,|1\rangle^{B}\right]\Big), \qquad (3.37) \end{aligned}$$

$$= \frac{1}{2}\left(\left|\Phi^+\right\rangle^{A'A}|\psi\rangle^B + \left|\Phi^-\right\rangle^{A'A} Z|\psi\rangle^B + \left|\Psi^+\right\rangle^{A'A} X|\psi\rangle^B + \left|\Psi^-\right\rangle^{A'A} XZ|\psi\rangle^B\right). \quad (3.38)$$

2. Alice performs a Bell measurement on the qubit to be teleported and her local qubit of the entangled pair, such that state of the system collapses to one of the four superposition terms in equation 3.38 with uniform probability. The outcome of the measurement instantly informs Alice about the state of Bob's qubit, which is in the state $U|\psi\rangle^B$, where U is a Pauli operator. Alice encodes her measurement result in a 2-bit classical message, according to a scheme such as the one shown in Table 3.2.

3. Alice transmits the message corresponding to her measurement result to Bob over a classical channel.

4. Bob performs a unitary transformation (that is, applies the appropriate Pauli operator) on his local qubit, according to the message received, to recover the original state of the teleported qubit. After the recovery operation, Bob's qubit is guaranteed to be in the state $|\psi\rangle^B$, regardless of the outcome of Alice's measurement.

Outcome	State of Bob's qubit	Message bits
$\|\Phi^+\rangle^{AB}$	$I\|\psi\rangle^B$	00
$\|\Psi^+\rangle^{AB}$	$X\|\psi\rangle^B$	01
$\|\Phi^-\rangle^{AB}$	$Z\|\psi\rangle^B$	10
$\|\Psi^-\rangle^{AB}$	$XZ\|\psi\rangle^B$	11

Table 3.2: Encoding table for the teleportation protocol.

There are a few aspects of the quantum teleportation protocol important to note. Firstly, teleportation is an *oblivious* protocol, in that neither Alice nor Bob require any knowledge of the state of a qubit in order to successfully teleport it between them. The protocol can therefore be considered universal, since it works for any input state.

Secondly, the protocol does not violate the non-cloning theorem, in that it never copies the state of the input qubit. Rather, Alice's local qubits become entangled as a result of the Bell measurement, and the information regarding the state $|\psi\rangle^{A'}$ is destroyed, to be recreated by Bob.

Thirdly, as with superdense coding, quantum teleportation is an inherently secure communication protocol. The classical message transmitted from Alice to Bob

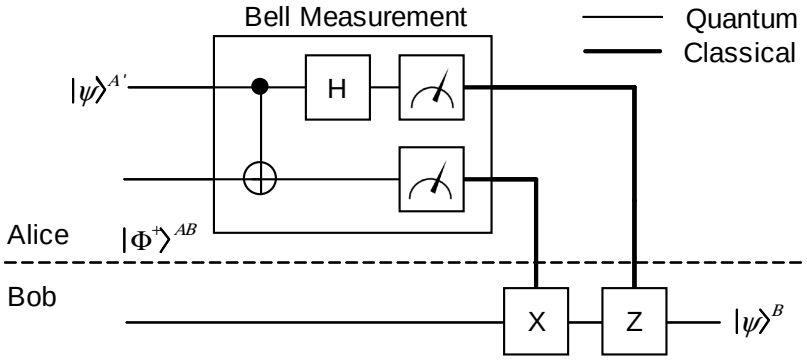


Figure 3.4: Quantum circuit depicting the teleportation protocol.

contains no information regarding the state of the qubit to be teleported. In fact, the result of the Bell measurement is entirely independent of the probability amplitudes α and β corresponding to the input qubit, and the four measurement outcomes are equiprobable, such that each of the corresponding messages occur with probability $\frac{1}{4}$. An eavesdropper may not therefore learn anything about the state of the input qubit by intercepting the message, and only Bob may recreate the input state, since he possesses half of the original entanglement resource.

Finally, quantum teleportation does not present a means for the instantaneous transmission of quantum information, as the name might suggest. There exists no action that might be performed by Alice on her half of the entanglement resource that would have an effect that Bob may instantly observe [1]. If Bob was to perform a partial measurement of his subsystem, even the instant after Alice's Bell measurement, he would observe that his qubit is in the maximally mixed state and obtain no useful information. Only after Bob receives the classical message regarding the Bell measurement outcome, and performs the necessary recovery operation, is the input state teleported to him. Teleportation is not therefore a superliminal communication protocol, but is limited by the speed of the classical communication system. The protocol's usefulness lies in its ability to transfer quantum states from one place to another, while avoiding intervening sources of quantum noise.

4 Quantum Error Correction Coding

To make quantum communication viable, it is necessary to protect quantum information from errors arising as a result of transmission over noisy channels and decoherence from interactions with the environment. Just as in classical communications, efficient transmission of quantum information over noisy channels requires error detection and correction at the receiver. Quantum error correcting codes (QECC) have been developed to this end, which encode quantum states in such a way as to make them resilient to noise. QECCs utilise essentially the same principles as classical error correcting codes. Encoding involves the addition of redundant information to a message at the transmitter such that, if the encoded message is only partially corrupted by the transmission channel, there might remain sufficient information to decode and recover the original message at the receiver. The amount of redundancy necessary to be able to successfully recover the message depends on the severity of the noise in the channel.

It is necessary to circumnavigate certain difficulties unique to quantum information processing in order to develop feasible quantum error-correcting codes. First and most important of these is the *no-cloning theorem*, which states that it is impossible to create identical copies of an arbitrary, unknown quantum state [9]. The theorem implies that there exists no such two-qubit unitary operator U that might act as a *universal copier* of quantum information, copying the state of an arbitrary qubit $|\psi\rangle$ to another, ancillary qubit (initially in state $|0\rangle$), i.e.,

$$\nexists U\colon\ U\,|\psi\rangle \otimes |0\rangle = |\psi\rangle \otimes |\psi\rangle\,. \tag{4.1}$$

This seems to suggest that quantum error correction is precluded from utilising classic redundancy and repetition codes to protect information from transmission errors. Fortunately, however, it has been shown that it is possible to embed or spread the state of a single qubit over the highly-entangled composite state of multiple qubits. More generally, a QECC can be considered to be a mapping of k *logical* qubits, (occupying a 2^k dimensional Hilbert space) to n *encoded* qubits (occupying a Hilbert space of dimension 2^n), where $n > k$ [31]. A $(n,\, k)$ quantum code can be defined as a k-dimensional subspace of the n-dimensional Hilbert space, called the *coding space*. A unitary encoding circuit is typically responsible for rotating the global state into the coding space.

Secondly, as there are an infinite number of possible unitary transformations of a quantum state, so there are a *continuum* of different errors that may occur for a single qubit. To detect exactly which error occurred in order to correct it would require infinite precision and resources. It is practical therefore to limit the consideration to the conventional quantum errors: bit-flips and phase-flips. Even though these are very artificial types of quantum noise, developing means to correct them leads to useful results for correcting more realistic types of noise [46].

Lastly, classical error-correction relies on the ability to observe the output of a channel in order to determine which decoding procedure should be employed. Since observation destroys quantum states, this is not a possibility for quantum error-correction coding.

4.1 Fundamental Quantum Error-correcting Codes

If sender Alice wishes to transmit quantum information to receiver Bob via a noisy quantum channel, she may protect the information from errors by employing a quantum error correcting code. Most existing quantum codes are based upon, or derived from, a number of fundamental error-correcting codes, devised to protect quantum information from the most basic and discrete forms of quantum noise: bit-flips and phase-flips.

4.1.1 The Bit-flip Code

In order to ensure that a particular quantum coding scheme will be effective and efficient at protecting quantum information during transmission over a noisy quantum channel, Alice must obtain some knowledge regarding the type of noise in the channel. Assuming that the noise is limited to bit-flips, that is, the channel is a bit-flip channel which flips qubits with probability p, the bit-flip code represents the most practical QECC that Alice might employ. The bit-flip QECC encodes the state of a single qubit into the composite state (tensor product) of three qubits, the information qubit plus two ancillary qubits, according to:

$$|0\rangle \to |000\rangle\,, \qquad |1\rangle \to |111\rangle\,, \quad \text{and} \quad |\psi\rangle \to \alpha\,|000\rangle + \beta\,|111\rangle\,. \tag{4.2}$$

The quantum circuit performing this encoding is shown in Figure 4.1, and consists of two CNOT gates acting upon the ancillary qubits according to the state of the information qubit. If the three qubits serially traverse independent copies of the bit-flip channel, the bit-flip code is able to provide such error protection as to perfectly recover the potentially corrupted quantum state at the receiver (without destroying the superposition), assuming at most one of the encoded qubits have been flipped. This means that, if the noisy channel between Alice and Bob in the SDC protocol is a bit-flip channel, the bit-flip code will provide error tolerance while reducing the maximum classical information rate to 2 cbits per 3 uses of the channel.

Error detection may be realised through s*yndrome diagnosis*. A measurement of the composite quantum state produces a result called the *error syndrome*, which can be used to determine which qubit (if any), experienced a bit-flip error, as well as facilitate correction. Importantly, syndrome measurements do not perturb the

state $\left|\psi'\right\rangle$, since the syndrome contains no information regarding the state being protected by the quantum error-correcting code [3]. There are four error syndromes, each corresponding to a specific measurement (projection) operator:

$$\Pi_0 = |000\rangle\langle 000| + |111\rangle\langle 111| \qquad \text{no error,} \tag{4.3}$$

$$\Pi_1 = |100\rangle\langle 100| + |011\rangle\langle 011| \qquad \text{bit flip on qubit 1,} \tag{4.4}$$

$$\Pi_2 = |010\rangle\langle 010| + |101\rangle\langle 101| \qquad \text{bit flip on qubit 2, and} \tag{4.5}$$

$$\Pi_3 = |001\rangle\langle 001| + |110\rangle\langle 110| \qquad \text{bit flip on qubit 3.} \tag{4.6}$$

For example, if a bit-flip has occurred on the first qubit such that the received state is $\left|\psi'\right\rangle = \alpha\,|100\rangle + \beta\,|011\rangle$, the error syndromes (outcomes of the four measurements) are given by:

$$\left\langle\psi'\right|\Pi_0\left|\psi'\right\rangle = 0, \tag{4.7}$$

$$\left\langle\psi'\right|\Pi_1\left|\psi'\right\rangle = 1, \tag{4.8}$$

$$\left\langle\psi'\right|\Pi_2\left|\psi'\right\rangle = 0, \text{ and} \tag{4.9}$$

$$\left\langle\psi'\right|\Pi_3\left|\psi'\right\rangle = 0. \tag{4.10}$$

The receiver is thus informed of the error having occurred and its position, and may correct it by simply applying the bit-flip operation to the qubit in error. The above procedure recovers the original state with perfect accuracy, given that bit flip occurs on at most one of the three qubits transmitted. The probability that an error cannot be corrected is obviously then equivalent to the probability that two or more bit flips occur, and is given by:

$$P_e = 1 - \left[(1-p)^3 + 3p\,(1-p)^2\right] \tag{4.11}$$

$$= 3p^2 - 2p^2. \tag{4.12}$$

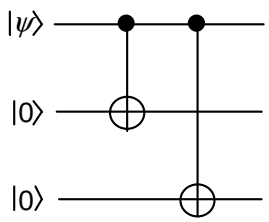


Figure 4.1: Encoding circuit for the 3-qubit bit-flip quantum error correction code, capable of correcting a bit-flip error on a single qubit. The state $|\psi\rangle$ is encoded into the composite state of three qubits using two "blank", ancillary qubits.

From this, it is clear that the quantum error correcting code improves upon the reliability of transmission provided that $p < 1/2$.

There is an alternate and more efficient procedure for the detection of bit-flips, which is especially useful when generalising the three qubit bit-flip code. The presence and position of bit-flip errors can be revealed by performing only two measurements: projecting out the eigenvalues of the observables $Z_1Z_2 = Z \otimes Z \otimes I$ and $Z_2Z_3 = I \otimes Z \otimes Z$. The measurements each return a single bit of information, indicating whether or not the corresponding qubit pairs are the same, and together provide a total of four possible error syndromes. These syndromes can be interpreted as binary numbers which directly indicate the location of a potential error within the three qubit codeword. For example, if measuring Z_1Z_2 returns $+1$ and Z_2Z_3 returns -1, the first and second qubits are the same while the second and third qubits differ, indicating that the third qubit has been flipped with high probability, and the associated error syndrome is (1, 0). From this, the receiver of the quantum information is informed that the third qubit has experienced a bit-flip and requires correction through the application of a bit-flip operation.

4.1.2 The Phase-flip Code

The bit-flip code can be modified to alternatively correct phase-flip errors which may occur as a result of dephasing noise. Encoding, error-detection and recovery operations are performed just as for the bit-flip code, but with respect to the Hadamard basis instead of the computational basis. By encoding the state of a qubit to be protected according to:

$$|0\rangle \to |+++\rangle\,, \quad |1\rangle \to |---\rangle\,, \quad |\psi\rangle \to \alpha\,|+++\rangle + \beta\,|---\rangle\,, \tag{4.13}$$

phase-flip errors act upon the state exactly like bit-flip errors. The change of basis is facilitated by a Hadamard gate acting on all three qubits (the information and ancillary qubits), once they are encoded as for the bit-flip channel, as depicted in Figure 4.2. Error detection and correction is performed using the same measurement

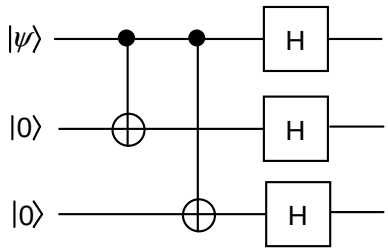


Figure 4.2: Encoding circuit for the 3-qubit phase-flip quantum error correction code, capable of correcting a phase-flip error on a single qubit.

operators as for the bit-flip code, but conjugated by Hadamard gates. The observables of interest are $X_1X_2 = H^{\otimes 3}Z_1Z_2H^{\otimes 3}$ and $X_2X_3 = H^{\otimes 3}Z_2Z_3H^{\otimes 3}$, where the notation $H^{\otimes 3}$ means that the Hadamard operator is applied to each of the three qubits. Measurement of the observables compares the sign of the corresponding qubit pairs, generates a binary error syndrome, and allows for the presence and position of phase-flips to be detected and corrected in exactly the same way as the bit-flip code does flipped bits.

4.1.3 The Shor Code

The first quantum error-correcting code capable of correcting *arbitrary* single-qubit errors was introduced by Shor in 1995 [26]. The Shor code is eminently suited to correcting errors caused by depolarising noise. The code is a combination, or, more specifically, a concatenation of the bit- and phase-flip codes. The information qubit is first encoded by the phase-flip code, and each of the resulting three qubits is further encoded using three copies of the bit-flip encoding circuit. The overall encoding circuit for the Shor code is shown in Figure 4.3. The code protects the state of an information qubit in the composite state of nine qubits, according to:

$$|0\rangle \rightarrow \frac{(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)}{2\sqrt{2}}, \text{ and} \tag{4.14}$$

$$|1\rangle \rightarrow \frac{(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)}{2\sqrt{2}}. \tag{4.15}$$

Evidently, the overall encoded states consist of three *clusters*, each of which contains three qubits and is prepared in the same composite quantum state. Solitary bit flips in any cluster can be detected and corrected by performing the same procedure as for the bit-flip code on the corresponding subset of qubits. Phase errors, however, are detected by comparing the relative phases of the clusters. A phase-flip on any of the qubits in a cluster has the effect of flipping the sign of that cluster.

Comparing the phases of pairs of clusters can be done by measuring the observables $X_1X_2X_3X_4X_5X_6$ and $X_4X_5X_6X_7X_8X_9$, which have eigenvalues ± 1, generating a two bit error syndrome which indicates if a phase-flip has occurred and in which cluster. The original quantum state may be recovered by simply applying a unitary phase transformation (the Z operator) to one of the qubits in the corrupted cluster.

By performing the two syndrome measurements in succession, the Shor code facilitates the correction of combined bit- and phase-flip errors on a single qubit. More interestingly, however, is the code's ability to protect quantum information from arbitrary, single qubit errors. This facility can be attributed to the surprising fact that the continuum of potential errors that may occur for a single qubit can be corrected by considering only a discrete subset of those errors [3].

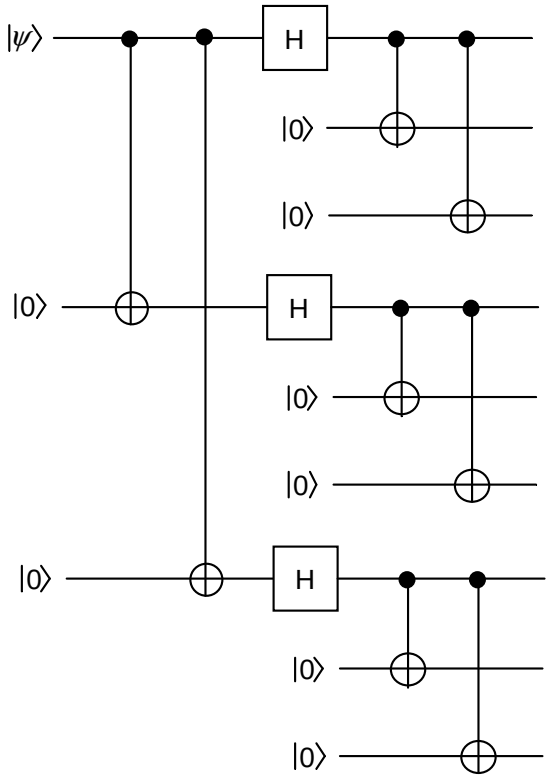


Figure 4.3: Encoding circuit for the 9-qubit Shor quantum error correction code, capable of correcting an arbitrary error on a single qubit.

4.2 CSS Codes

Quantum error-correcting codes based on the principles of, and, in some cases, constructed directly from, classical linear codes are collectively referred to as Calderbank-Shor-Steane (CSS) codes. The Steane code, formulated by Andrew Steane in 1996, is an important example of a CSS code, constructed using the self-dual [7, 3, 4]

Hamming code. The Hamming code and its dual are capable of correcting bit- and phase-flip errors, respectively. The Steane code ultimately performs the same function as the Shor code: correcting for arbitrary errors on a single qubit. However, it is able to do so by encoding the quantum information in the highly-entangled composite state of only seven qubits. The quantum Hamming bound provides a constraint on the number of encoded qubits n required to protect one logical qubit ($k = 1$) from a single qubit error, and is given by:

$$1 + 3n \leq 2^{n-1}, \tag{4.16}$$

which is satisfied for $n \geq 5$. In 2012, a class of 5-qubit codes which meet this bound and provide tolerance to all possible errors on a single qubit was discovered [32].

5 Classical Error Correction Coding

This chapter introduces the basics of classical error correction coding and its role in digital communications. The concept of error correcting codes was pioneered by Claude Shannon in 1948, when he published a ground-breaking paper presenting the *noisy channel coding theorem*. The theorem proved the existence of coding schemes enabling near error-free transmission of information over channels inhibited by noise, for all code rates up to the channel capacity. While Shannon's proof was non-constructive, many capacity-approaching coding schemes have since been discovered that provide receivers with the ability to recover corrupted data without requiring retransmission. The inclusion of redundancy at the transmitter to provide protection from transmission errors is indicative of a *forward* error-correcting (FEC) code. FEC codes require only simplex communication, making them especially attractive for wireless communication systems, since they improve spectral efficiency. There are two important classes of FEC codes: block codes and convolutional codes.

5.1 Block Codes

Block codes are a family of error-correcting codes which have the function of dividing a stream of input data into discrete blocks of fixed length and performing encoding on each block. They are often referred to as (n, k) codes, since a binary block code will encode an input block or *message* comprised of k bits into a larger, n-bit *codeword*. Each codeword contains $n - k$ parity or redundancy bits, which facilitate error detection and message recovery at the receiver. The efficiency of the block code and the number of errors it is able to correct per block is typically dependent on the number of parity bits added. The ratio between the length of the message k and the length of the codeword n is the *code rate*, i.e.,

$$R = \frac{k}{n}. \tag{5.1}$$

For a binary block code operating on input blocks of length k, there are 2^k possible, distinct messages which may be transmitted. Each message is mapped by a block encoder to a corresponding codeword from a set of 2^n available codewords, each of which can be decoded independently. An important result of the noisy channel coding theorem is that the probability of error in the decoded output can be reduced by increasing the block length n, for the same code rate R. However, for larger block lengths, attainment of the codewords and decoding becomes increasingly more difficult. The process can quickly become hampered by hardware limitations, necessitating large amounts of storage and more computational power. *Linear block codes* overcome this and reduce decoding complexity to some degree by restricting the set of possible codewords to those cases where the sum of any two codewords produces

a third, valid codeword. *Cyclic codes* are a subclass of linear block codes which further increase the ease by which codewords may be obtained and decoded for large block lengths, by defining a set of codewords wherein a cyclic shift on any codeword results in another codeword. Bose-Chaudhuri-Hocquenghem (BCH) codes are a popular class of cyclic codes, which are constructed using the mathematical concept of *finite fields* and are capable of detecting and correcting multiple transmission errors.

5.1.1 Reed-Solomon Codes

Reed-Solomon (RS) codes are an important, non-binary subclass of BCH codes, and are arguably some of the most successful FEC codes in use today due to their exceptional performance in the presence of *burst errors* over erasure channels. RS codes are non-binary in that they operate on a message sequence of m-bit data symbols, where $m \geq 2$. Reed-Solomon codes are optimal because they achieve the largest minimum distance d_{min} possible for a $(n,\, k)$ linear code (the Singleton bound), namely:

$$d_{min} = n - k + 1. \tag{5.2}$$

Such a code is referred to as being maximum distance separable (MDS). The distance between codewords is defined for non-binary codes as the number of symbols by which two encoded sequences differ, as it is by the Hamming distance for binary codes.

A $(n,\, k)$ RS code exists for all n and k such that:

$$0 < k < n < 2^m + 2, \tag{5.3}$$

where k is the message length and n is the number of code symbols in an encoded block. The most conventional RS code has parameters:

$$(n,\, k) = (2^m - 1,\, 2^m - 1 - 2t) \tag{5.4}$$

where t is the number of symbol errors the code is capable of correcting. Without prior knowledge of the locations of the error symbols, RS codes can correct up to:

$$t = \frac{n - k}{2}, \tag{5.5}$$

erroneous symbols in a received codeword. That is, they are able to correct half as many symbol errors are there are redundant, parity symbols in the codeword. A RS code is also capable of correcting erasures, where corrupted symbols are replaced by error symbols such that the error locations become known to the receiver. Any combination of errors and erasures may be corrected, provided that the number of errors ϵ and the number of erasures ε satisfies:

$$2\epsilon + \varepsilon \leq n - k. \tag{5.6}$$

5.1.1.1 Encoding Reed-Solomon Codes

In Reed-Solomon encoding, a message $x = (x_1, x_2, \ldots, x_k)$ is first linearly mapped to a polynomial p_x of degree less than k over the finite (Galois) field GF (2^m), i.e.,

$$p_x(X) = \sum_{i=1}^{k} x_i X^{i-1}. \tag{5.7}$$

In the original encoding procedure devised by Reed and Solomon, the message polynomial p_x is then evaluated at n distinct points of F, and the sequence of resulting values $a_1, a_2, ..., a_n$ is the codeword corresponding to the message. In the now conventional method, the message polynomial is instead multiplied by a *generator polynomial* g of degree $n - k$, known to both transmitter and receiver, whose roots are exactly $\alpha^1, \alpha^2, ..., \alpha^{n-k}$ and which has the form:

$$g(X) = \prod_{i=1}^{n-k} \left(X - \alpha^i\right), \tag{5.8}$$

$$= g_0 + g_1 X + \cdots + g_{n-k-1} X^{n-k-1} + X^{n-k}. \tag{5.9}$$

Codewords are then defined as the sequence of n coefficients of the polynomial $c(X) = p_x(X) \cdot g(X)$. Consequently, the receiver can identify valid codewords as those polynomials that are exactly divisible by the generator polynomial.

5.1.1.2 Decoding Reed-Solomon Codes

The Berlekamp-Massey algorithm implements the efficient decoding of both binary BCH codes and non-binary Reed-Solomon codes. Once a codeword has been transmitted, having been subjected to noise in the channel and potentially corrupted, the receiver interprets it as the coefficients of a polynomial $r(X)$. To decode the original message polynomial, the receiver performs the polynomial division:

$$p_x(X) = \frac{r(X)}{g(X)}. \tag{5.10}$$

If the received polynomial is perfectly divisible by the generator polynomial, then it is equivalent to what was transmitted and without error. Conversely, a remainder indicates that at least one error has occurred and constitutes an error-pattern polynomial $e(X)$, where:

$$r(X) = c(X) + e(X), \tag{5.11}$$

and:

$$e(X) = e_0 + e_1 X + \cdots + e_{n-1} X^{n-1}. \tag{5.12}$$

Each of the n coefficients of the error-pattern polynomial is an error value in GF(2^m), with the error's position in the codeword indicated by the degree of the associated term of the polynomial. The receiver may then use this information to modify the received codeword and determine the codeword that was most likely transmitted. If more than t coefficients of e are non-zero, than the error correction capability of the RS code is exceeded and the received codeword is uncorrectable.

5.2 Convolutional Codes

Convolutional codes differ from block codes in that encoding is performed on a continuous basis rather than on discrete portions of the input data, thereby generating a stream of encoded output data. Message length k and codeword length n are typically small when compared with block codes.

5.2.1 Viterbi-decoded Convolutional Codes

One of the most prolific techniques for forward error correction involves decoding a convolutionally encoded bitstream according to the Viterbi algorithm [34].

5.2.1.1 Encoding Convolutional Codes

In convolutional encoding, mapping of input data to a codeword is not only dependent on the current message, but on a certain amount of preceding messages. More specifically, the output of a convolutional encoder is a linear combination of the current input and a certain number of *delayed* inputs, such that each message influences N codewords, where $N > 1$. This necessitates that the convolutional encoder possess *memory*, such that it may retain n past messages, where $n = N - 1$. As a consequence, a convolutional encoder is typically implemented using N-stage shift registers, where N is the *constraint length* of the code. Alternatively, a code can be described as a memory-n code.

The two most important types of convolutional code are non-systematic (NSC) and recursive systematic convolutional codes (RSC). The latter employ feedback to enable recursive encoding and are systematic because there is a direct path between the input and the output, such that the output contains the input and only the redundancy is dependent on previous input. In other words, for a binary convolutional code, the output codeword consists of both *systematic bits* and *parity bits*. Figure 5.1 provides simple examples of the two types of convolutional encoders.

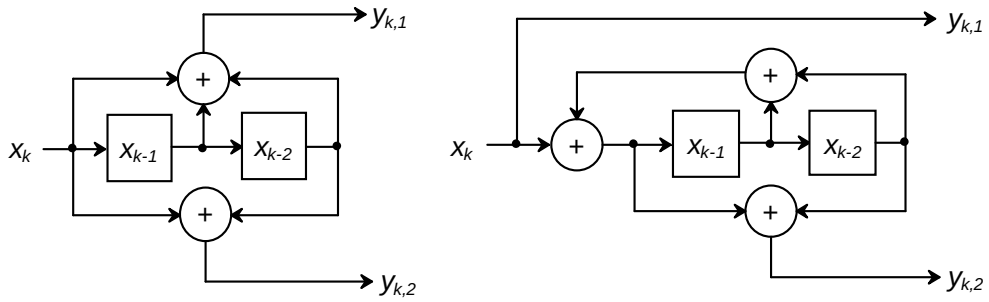


Figure 5.1: Memory-2, 1/2 rate NSC (left) and RSC (right) convolutional encoders.

The behaviour of a convolutional encoder can be described as a finite state machine, where the current state of the encoder is defined by the input values contained in the memory. For a binary convolutional encoder with constraint length N, the number of possible states is 2^{N-1}. The half-rate convolutional encoders in Figure 5.1 can reside in four different states: 00, 01, 10 or 11, and the output is dependent on both this state and the input bit, according to its design. Upon each encoding, the next input bit is fed into the shift register and the state of the encoder transitions to a new state. There are two equivalent ways to represent the transitions between states: the state diagram and the trellis diagram. Figure 5.2 shows both the state and trellis diagrams for the NSC encoder in Figure 5.1, from which the input-output relation of the encoder can be clearly observed.

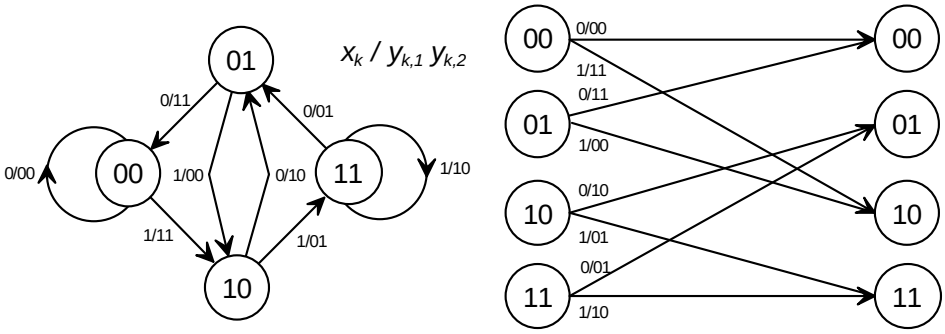


Figure 5.2: State diagram (left) and trellis diagram for the NSC encoder in Figure 5.1, showing the state transitions and the output bits.

5.2.1.2 Decoding Convolutional Codes

In digital communications, where computational complexity is a major concern, Viterbi's maximum likelihood algorithm represents arguably the best technique for decoding convolutional codes. Since its introduction in 1967, Viterbi's algorithm has become the predominant technique for decoding convolutionally-encoded data due to its fixed decoding time [34, 37]. Maximum likelihood (ML) decoding ultimately involves the selection of the codeword $\hat{x}$ closest to the received data y (the codeword most likely to have been transmitted given the received sequence), such that:

$$P\left(y \,|\, \hat{x}\right) = \max_{x} P\left(y \,|\, x\right), \tag{5.13}$$

where x is a valid codeword and $P\left(y \,|\, x\right)$ is referred to as the likelihood. By following this decoding rule, the average error probability is minimised, when all codewords are equiprobable. Direct implementation of a binary ML decoder operating on received frames of k bits requires the storage of 2^k codewords, and comparison of the frame to each codeword, which for large k is computationally demanding. In 1967, Andrew

Viterbi proposed a simplification of the ML decoding algorithm, which reduced hardware requirements and complexity while maintaining decoding performance, thereby making it feasible for communication at high data rates.

The Viterbi algorithm (VA) involves the traversal of a multistage trellis, where each possible path between states represents a valid codeword and the optimum path produces the most likely codeword to match the received data at each stage. Only the best path at a given stage needs to be stored, since all other paths to any given stage will only increase the path distance metric to the most likely codeword. As a result, the complexity of the algorithm is proportional only to the number of states in the trellis. Viterbi's algorithm performs a single recursion through the received data and outputs the most likely codeword as a *hard* output (that is, the bits are quantised to a 0 or 1), using the Hamming distance as a metric. The algorithm can be adapted to accept *soft* input and produce soft output bits for improved decoding performance. Soft bits include likelihood information regarding the value of the bit. The algorithm which performs soft-input soft-output (SISO) decoding, with the square Euclidean distance as a metric, is referred to as the soft-output Viterbi algorithm (SOVA).

Another important method for decoding convolutional codes is the SISO Maximum-A-Posteriori (MAP) algorithm, which estimates the sequence of input bits and provides the probability that each received bit matches what was transmitted: the *a-posteriori probability* (APP). In contrast to the ML decoding rule, MAP decoding chooses a codeword $\hat{x}$ after receiving y such that:

$$P\left(\hat{x} \,|\, y\right) = \max_{x} P\left(x \,|\, y\right), \tag{5.14}$$

where $P\left(x \,|\, y\right)$ is the a-posteriori probability of codeword x given that y was received, in order to minimise the average probability of error. The respective decoding rules for ML and MAP decoding are related by Baye's Theorem, in that:

$$P\left(x \,|\, y\right) = \frac{P\left(x\,,\, y\right)}{P\left(y\right)} \tag{5.15}$$

$$= \frac{P\left(x\right) P\left(y \,|\, x\right)}{P(y)}, \tag{5.16}$$

where $P(y)$ is constant if all codewords are equiprobable, and $P\left(x\right)$ is referred to as the *a-priori probability* of codeword x. The MAP algorithm performs two recursions of the received data, one forward and one backwards, producing metrics

from both directions in the associated trellis and soft outputs. The complexity of the MAP algorithm is approximately double that of the Viterbi algorithm, but the former is eminently suited to the decoding of modern, high-performance, forward error-correcting codes.

5.2.2 Turbo Codes

Convolutional turbo codes (CTC) are a class of FEC codes first introduced by Berrou, Glavieux, and Thitimajshima in 1993, which exhibit what was then unprecedented and unmatched error-correcting performance [38]. Turbo codes were shown to provide bit error rates (BER) within 1dB of the Shannon limit with reasonable computational complexity; a result which astounded the research community, as conventional codes had previous struggled to breach 3dB. Today, turbo codes and their various evolved forms are widely utilised in digital communication systems for their outstanding ability to facilitate reliable information transmission over noise limited channels.

5.2.2.1 Encoding Turbo Codes

A turbo encoder consists of at least two identical RSC encoders, referred to as *constituent encoders*. In the subclass of parallel concatenated convolutional codes (PCCC), the constituent encoders are in a parallel arrangement and are separated by an interleaver. The purpose of the interleaver is to scramble the input data stream in a pseudo-random, but predetermined, manner. The first encoder is fed the input data directly and the subsequent encoders are fed interleaved versions of the data, such that the output of the constituent encoders is different and uncorrelated. The action of the interleavers is known to both the transmitter and the receiver, and serves to improve decoding performance. Typically, the number of constituent codes is kept to a minimum, since the performance gained by increasing the number of parallel codes does not justify the additional complexity and overhead [34].

The decoders for each of the constituent codes have been shown to perform best when their respective encoders begin and end at a known state, such as the all-zero state [38]. This can be achieved by independently terminating the trellis of each encoder with an input sequence which returns the encoder to the all-zero state, known as a *tail*. Including a tail in every encoded frame has the drawback of reducing the code rate, although for large frame lengths the reduction is negligible.

Figure 5.3 illustrates the design of an example turbo encoder. It consists of two 1/2 rate RSC constituent encoders, each generating a single parity bit y_k for every input bit x_k, such that the overall rate of the turbo encoder is 1/3 including the systematic bit.

To increase the rate of the turbo code, at the cost of error correcting performance, the outputs of the constituent encoders may be selectively and systematically discarded

in a process called *puncturing*. For example, by disregarding half of the total parity bits generated for a given input frame, the rate of the turbo code might be increased to 1/2. The inverse procedure of *depuncturing* is performed prior to decoding the turbo code, which simply replaces the punctured bits with undefined data (typically by padding with zeros). Puncturing is traditionally not performed on the systematic bits.

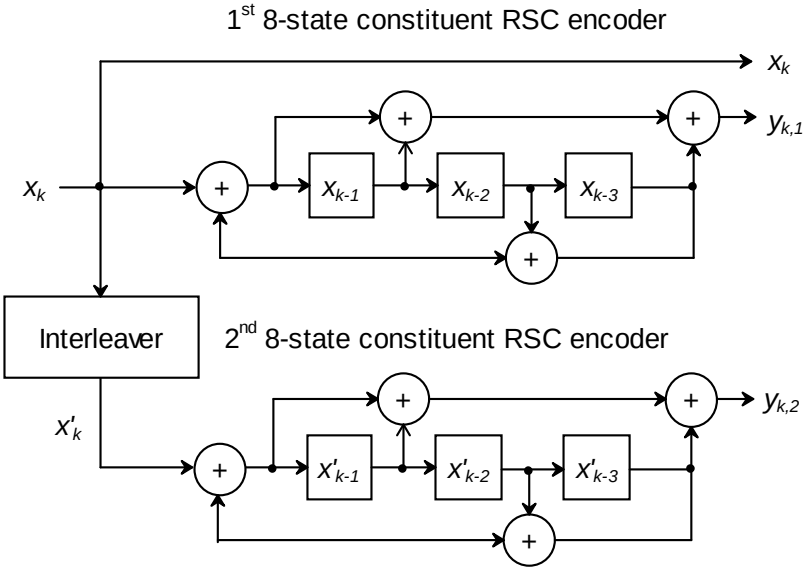


Figure 5.3: Memory-3, 1/3 rate turbo encoder.

5.2.2.2 Decoding Turbo Codes

A turbo code consists of a combination of component RSC codes, which have relatively few states (2^n for a memory-n RSC encoder) and might therefore be routinely decoded by the SOVA or MAP algorithms. However, decoding the overall code using such traditional algorithms directly is considerably more challenging, due to the large number of states ($2^{m \times n}$ for a parallel concatenation of m memory-n RSC constituent encoders), and the presence of the interleaver. Fortunately, this can be overcome using a method called *iterative decoding*, where the component codes are individually decoded by low-complexity decoders which iteratively exchange information between them in order to converge to a result. Turbo codes typically employ variations of the MAP algorithm for decoding, such as the optimal log-MAP or simplified max-log-MAP algorithms, which operate in the log-domain rather than the probability domain.

A turbo decoder corresponding to the encoder in Figure 5.3 consists of two decoders, which ultimately cooperate to estimate the a-posteriori probabilities for the message bits, i.e.,

$$\Pr\left(x_k = x \mid x_k,\, y_{k,1},\, y_{k,2}\right) \quad \text{for } x \in \{0,\, 1\} \text{ and } 1 \leq k \leq m, \tag{5.17}$$

where m is the length of the input message frame. The exchange of information necessary for iterative decoding is only beneficial in the presence of soft bits, so the decoders typically employ either the SOVA or MAP decoding algorithms. The latter has been shown to result in the best error correction performance in the majority of cases [34]. Figure 5.4 illustrates the turbo decoding process. Initially, the first decoder is fed the systematic bits and the parity bits output by the first encoder, and generates soft decisions regarding the message bits derived from the parity bit $y_{k,1}$. For the MAP-based algorithms, this is the log-likelihood ratio (LLR), the logarithm of the ratio of the a-posteriori probabilities for x_k, i.e.,

$$L\left(x_k\right) = \ln\left[\frac{\Pr\left(x_k = 1\right)}{\Pr\left(x_k = 0\right)}\right]. \tag{5.18}$$

The sign and magnitude of the LLR together indicate the likelihood of the bit towards 1 or 0. The first decoder then passes this *extrinsic information* $\Delta_{1\to 2}$ to the second decoder, which performs the same process on the parity bits output by the second encoder $y_{k,2}$, and returns its extrinsic information $\Delta_{2\to 1}$ to the first decoder as a-priori information (initially unavailable), that the latter might update its decisions accordingly during the subsequent iteration. The passing of soft information between the first and second decoders, and back to the first, constitutes a single iteration of the decoder. Multiple iterations may greatly improve the estimation of the message bits, although typically with diminishing returns [34]. After a predetermined number of iterations, once satisfactory convergence has been achieved, a hard decision on the message bit $\hat{m}_k$ is made using the extrinsic information from both decoders and the systematic bits.

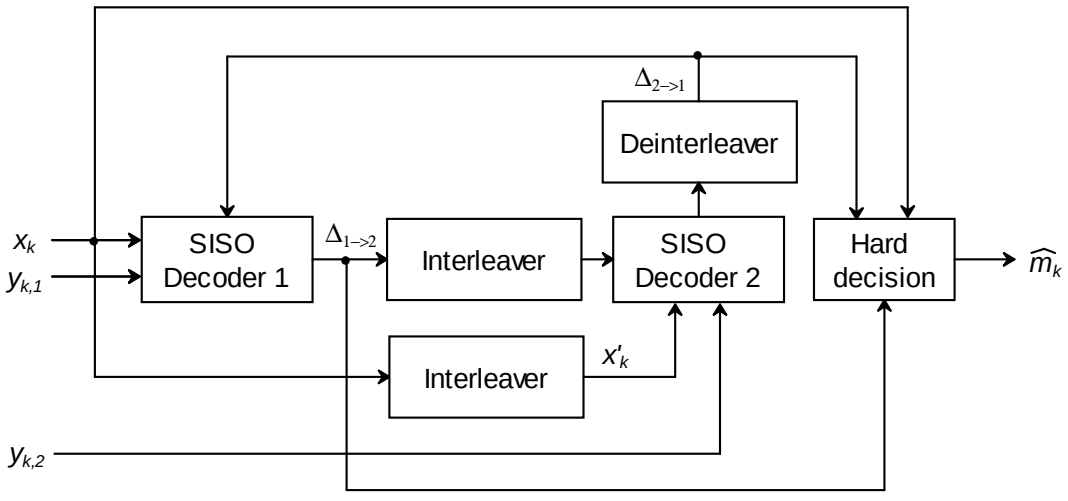


Figure 5.4: Turbo decoder.

5.2.3 Duo-binary Turbo Codes

Replacing the 1/2 rate constituent codes of a binary turbo code with m-input, $m/(m+1)$ rate binary RSC codes, produces a non-binary or m-ary turbo code which, under certain circumstances, has advantages over the binary code [40, 41]. *Duo-binary* turbo codes (DBTC) with $m = 2$, encode information bits *pair wise*, and have already been adopted into many digital communication standards because they exhibit improved performance over classical turbo codes and provide natural coding rates for higher order modulations, while maintaining reasonable decoding complexity.

The advantages of duo-binary CTCs are a consequence of the bi-dimensionality of the code, and include larger minimum distances and better convergence of the iterative decoding process due to a decrease in the correlation effect between the constituent decoders [41]. The trellis of duo-binary encoder has half the number of states as a binary encoder with the same constraint length, therefore requiring half as much memory to implement. Duo-binary turbo codes require less puncturing to achieve higher code rates and throughput is doubled, and latency halved, per decoding cycle. The decoder is more robust, and the relative performance of log-MAP and the sub-optimal max-log-MAP decoding are more comparable than for a binary turbo code.

5.2.3.1 Encoding Duo-binary Turbo Codes

Figure 5.5 illustrates the general structure of a memory-3 duo-binary turbo encoder. The parallel, constituent encoders are fed frames of k message bits grouped into $N = k/2$ *couples*, with the second operating on an interleaved version of each frame. A major contributing factor to the enhanced performance of duo-binary turbo codes is an improved interleaver design. Interleaving is performed on two levels: within the couples and between the couples, leading to further reduced correlations between the outputs of the encoders. Both constituent encoders produce a pair of parity bits for each input couple $(x_{k,1}, x_{k,2})$, resulting in a total of three output couples per input couple, including the systematic bits, and an overall code rate of 2/6. As with binary turbo codes, puncturing can be utilised to increase the code rate. To achieve a code rate of 2/4, the encoder simply discards parity outputs $y_{k,2}$ and $y_{k,4}$ and transmits the couples $(x_{k,1}, x_{k,2})$ and $(y_{k,1}, y_{k,3})$.

Duo-binary turbo codes have been optimised for relatively short frame sizes and high data rates, meaning that the inclusion of a tail, for the purposes of forcing the encoder to end in a known state, has a non-negligible impact on the code rate. For this reason, DBTCs implement circular recursive systematic convolutional (CRSC) encoding instead. CRSC codes are based on the concept of *tailbiting*, and operate in such a way as to ensure that the ending state of the encoder always matches the starting state. The process involves input frames being encoded twice by the

constituent encoders. Each encoder is initialised to the all-zero state $\mathbf{S}_0$, and the first encoding (precoding) pass is used to derive its own, independent *circulation state* $\mathbf{S}_c$. The frame is then encoded with each encoder initialised to $\mathbf{S}_c$, such that their final state is guaranteed to be $\mathbf{S}_c$ also.

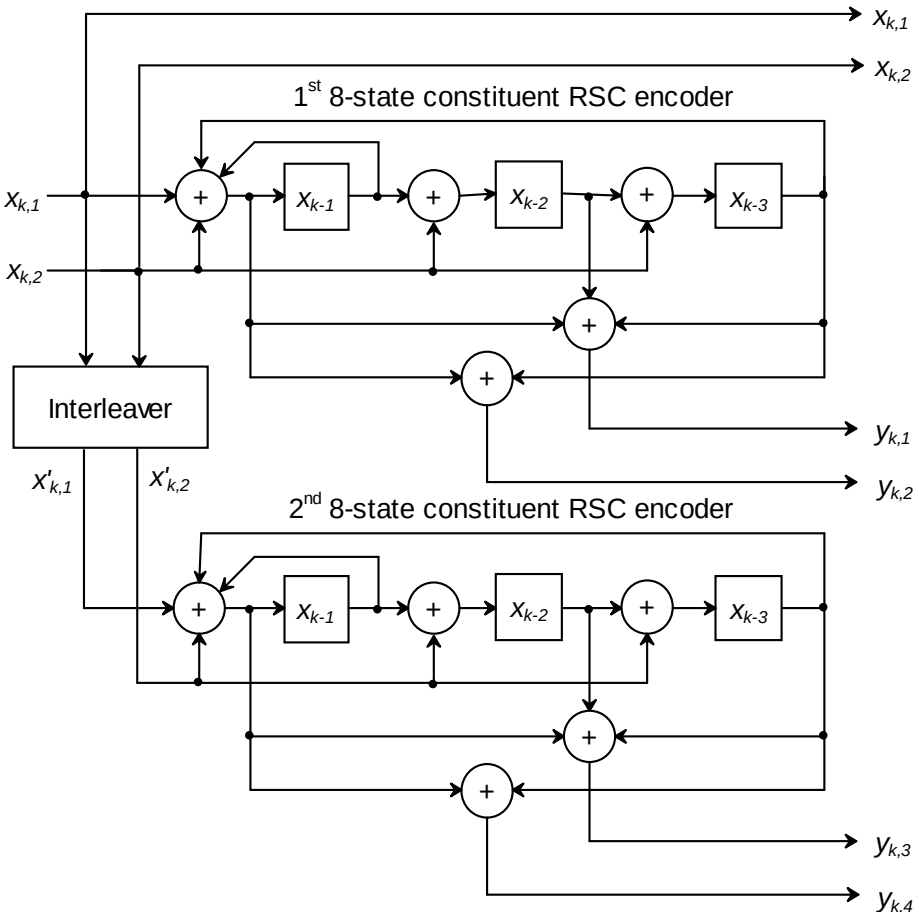


Figure 5.5: Memory-3, 2/6 rate duo-binary turbo encoder.

5.2.3.2 Decoding Duo-binary Turbo Codes

The use of duo-binary constituent encoders defined over GF (4) rather than GF (2) in duo-binary turbo codes increases decoding complexity, but ultimately facilitates faster decoding in hardware and increases performance over their binary counterparts, for the same code rate. The structure of the duo-binary decoder is shown in Figure 5.6. As with binary turbo codes, decoding duo-binary turbo codes is typically performed according to the log-MAP or max-log-MAP algorithms, and involves the iterative exchange of extrinsic information between two component decoders. Unlike binary turbo codes, however, duo-binary codes require three log-likelihood ratios to represent the current soft decision regarding the message couple. For example,

the set of LLRs that comprise the extrinsic information $\Delta_{1\to 2}$ between the first and second constituent decoders are given by:

$$L\left(x_{k,1}, x_{k,2}\right) = \ln\left[\frac{\Pr\left(x_{k,1}=a,\, x_{k,2}=b\right)}{\Pr\left(x_{k,1}=0,\, x_{k,2}=0\right)}\right], \tag{5.19}$$

where $(a,\, b) \in \{(0,\, 1), (1,\, 0), (1,\, 1)\}$. Once the decoder has completed a given number of iterations (or met some other convergence criterion), it makes and outputs a hard decision regarding each bit of the message couple.

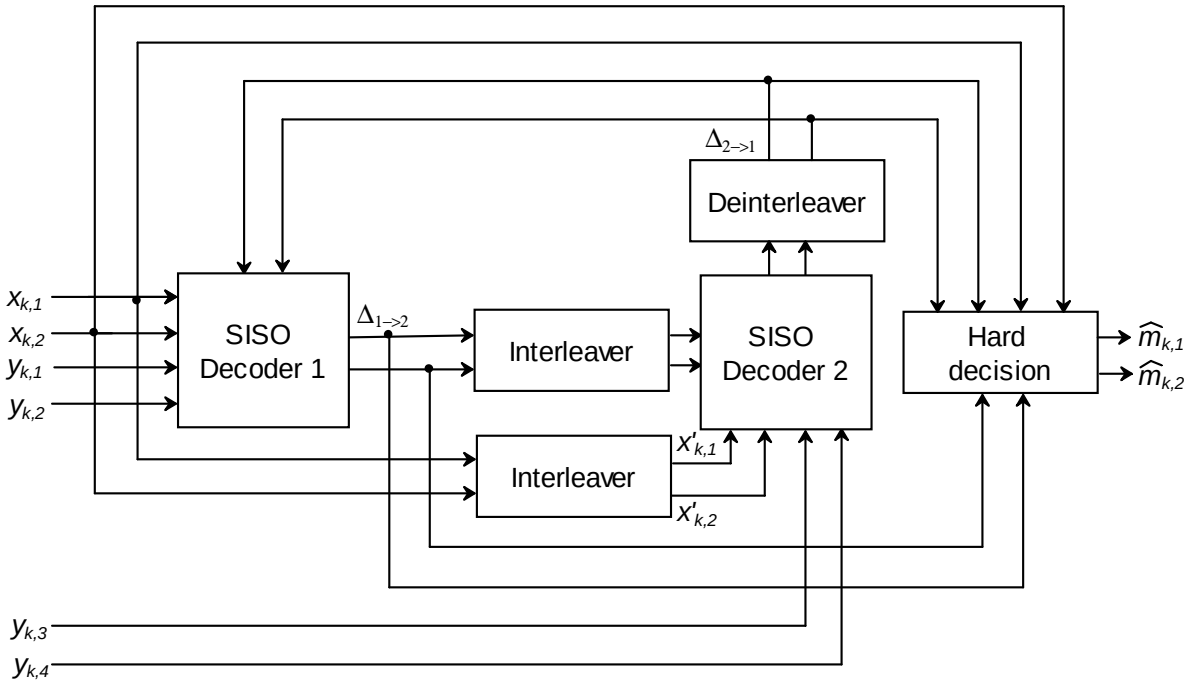


Figure 5.6: Duo-binary turbo decoder.

6 Channel Models and Capacities for Classical-Quantum Systems

To assess whether classical error-correcting codes have the potential to provide quantum communication systems with the ability to reliably communicate classical information over noisy quantum channels, at higher rates than might be achieved using QECCs, it is necessary to first devise a suitable classical-quantum communication system. The architecture of a proposed system is depicted in Figure 6.1.

The devised system implements the superdense coding protocol, with noiseless entanglement distribution, in order to communicate classical information from one party to another over a quantum depolarising channel. Error protection is provided by a classical FEC code, which encodes information at the transmitter at a given code rate. The encoder is assumed to include adaptive rate functionality, whether it be by puncturing or some other means, such that it might improve error-correcting performance at the cost of the information rate. Codewords are fed into the superdense encoder, which performs a conversion from the classical to the quantum domain by mapping groups of n coded classical bits (or symbols consisting of n bits) to 2^n orthogonal, maximally-entangled Bell states (which may also be considered symbols). In this way, the superdense encoder operates in a similar fashion to a classical M-ary modulation scheme, mapping multiple bits to M complex-valued phasors.

The qubits of each output state that are local to the transmitter are then sent over the quantum depolarising channel serially. The subsequent mixed or noisy quantum state is decoded at the receiver, by means of a symbol-by-symbol Bell measurement, to produce codewords that are potentially in error. Conventionally, the outputs of the superdense decoder are hard bits rather than soft bits as they might be for a classical communication channel. In order to access the error performance gain associated with iterative decoding, it is necessary to conceive of a soft-decision superdense decoder which outputs the extrinsic probabilities for the bits or symbols that comprise a transmitted classical codeword. Finally, this soft output is fed into the decoder corresponding to the implemented FEC code, which recovers the original information bits or symbols with some degree of reliability.

The success of the classical error-correcting coding scheme is measured in how closely the achievable information rate is to the classical capacity of the quantum subsystem, for a vanishingly low bit error rate. The entanglement-assisted classical capacity of the superdense coding protocol over a noiseless channel and for unitary encoding, was shown in Section 3.3 to be contingent on the initial, shared entangled state ρ^{AB}, i.e.,

$$C_E = 1 + S\left(\rho_b^{AB}\right) - S\left(\rho^{AB}\right). \tag{6.1}$$

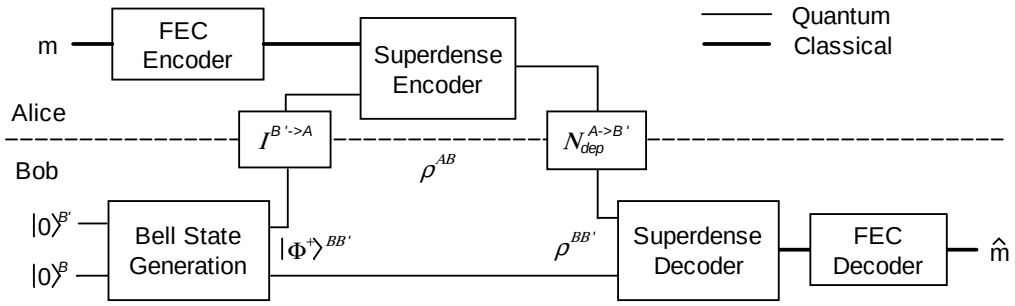


Figure 6.1: An example of a classical-quantum communication system, employing the superdense coding protocol to communicate classical information over a quantum depolarising channel. Entanglement distribution may be noisy or noiseless.

In practice, the detrimental effect of quantum noise on the capacity must also be taken into consideration. Two different scenarios of superdense coding over noisy quantum channels are considered in this thesis. In the first scenario, communicating parties Alice and Bob are assumed to share a bipartite, maximally-entangled, pure state prior to enacting the protocol. This may be achieved via noiseless entanglement distribution, where a third party, Eve, generates the entangled pair and transmits a qubit to each of the communicating parties over noiseless channels. Alternatively, entanglement distillation might be enacted to purify a noisy initial state at the cost of ebits. Alice then performs the encoding operation on her local qubit and transmits it to Bob via a noisy quantum channel. This arrangement can be referred to as superdense coding with noiseless entanglement distribution, or SDC over a *one-sided* channel. The second scenario to be considered involves multiple uses of the noisy channel between Alice and Bob to establish the shared resource state and enact the SDC protocol. In this case, Bob prepares the Bell state and sends one qubit of the entangled pair to Alice over the noisy channel. The initial state is therefore mixed and the capacity is reduced as a result. Alice performs the local encoding and returns her portion of the shared state to Bob via the same noisy channel. This arrangement can be referred to as superdense coding with noisy entanglement distribution, or SDC over a *two-sided* channel.

6.1 Characterising the Quantum Depolarising Channel

As previously stated, a noisy quantum transmission channel can be described mathematically as a positive trace-preserving linear map acting on the quantum state. The quantum depolarising channel acting on Alice's side is denoted by the mapping of quantum state ρ to a linear combination of itself and the maximally mixed state, i.e.,

$$\Lambda_a^{\text{dep}}(\rho) = \sum_{j=0}^{3} q_j A_j \rho A_j^\dagger \tag{6.2}$$

$$= \frac{p}{4} X\rho X^\dagger + \frac{p}{4} Y\rho Y^\dagger + \frac{p}{4} Z\rho Z^\dagger + \frac{4-3p}{4}\rho, \tag{6.3}$$

where: the Kraus operators A_j for the channel are the Pauli matrices; q_j are the probabilities of the respective unitary operations being applied; and p is the depolarising probability. The depolarising channel above is favoured in literature related to quantum computation and communication because of its particular characteristics [20, 44, 48]. The classical capacity of the depolarising channel has been defined and shown to be equivalent to that of a binary symmetric channel (BSC) with crossover probability $p/2$ [44]. For such a channel, with input x and output y, the capacity may be easily derived using the classical entropy function H, as follows:

$$C^{\text{dep}} \equiv 1 - H\left(y \,|\, x\right) \tag{6.4}$$

$$= 1 + \frac{2-p}{2}\log_2\frac{2-p}{2} + \frac{p}{2}\log_2\frac{p}{2} \qquad \text{cbits/channel use.} \tag{6.5}$$

Figure 6.2 depicts the various characterising measures for the channel, all of which degrade consistently with an increase in depolarising noise and are minimised when the noise parameter is maximised. This is not the case for other types of quantum noise. For example, if one qubit of an EPR pair in a Bell state was to traverse a fully dephasing channel (that is, the probability of dephasing is 1), the process would be equivalent to applying the Z operator to the state of the transmitted qubit. The result would therefore be another maximally-entangled, maximally-correlated Bell state. For this reason, the discord of the dephasing channel has a parabolic shape: the channel deteriorates the correlation between the states of the two qubits up to a dephasing probability of 1/2, where complete decoherence of the input state occurs. Beyond that, the correlations increase and are maximised again at a dephasing probability of 1. Dephasing noise does not effect the fidelity between the input and output states of the channel, therefore the measure is maximum and constant over the dephasing probability p and does not explicitly characterise the channel. The converse is true for the quantum bit-flip channel, where maximum noise is equivalent to applying the X operator to the input state instead. The bit-flip channel does not effect the correlations between an EPR pair when one of the pair traverses it, such that the discord is maximum and constant over the bit-flip probability p and the fidelity better characterises the channel. The classical capacity for the bit-flip and dephasing channels are identical and equivalent to that of a BSC with a crossover probability of p.

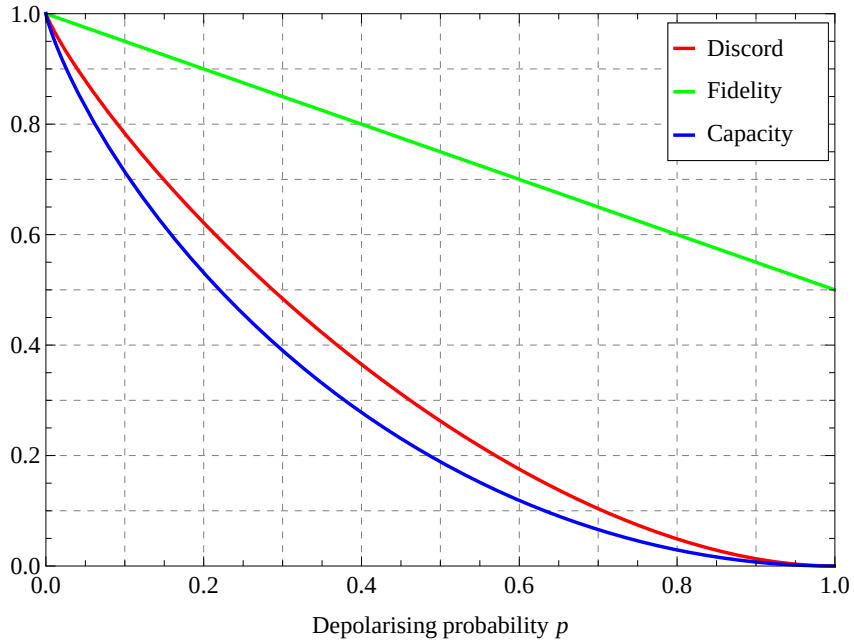


Figure 6.2: Characterising measures of the quantum depolarising channel, as functions of the depolarising probability.

6.2 Equivalent Classical Channel Models

In order to simulate the performance of classical error-correcting codes in the proposed classical-quantum communication system, it is necessary to define an equivalent classical channel model for each variation of the quantum superdense coding protocol over the depolarising channel to be studied. There are three such variations considered herein.

6.2.1 Noiseless Entanglement Distribution

The entanglement-assisted classical capacity of the superdense coding protocol over the depolarising channel, with a Bell state for the initial resource, has been derived based on the equivalence of the overall transmission model to a M-ary symmetric discrete classical channel with $M = 4$ [44]. This comparability is a direct consequence of the fact that the superdense decoder performs measurements on a symbol-by-symbol basis. The transition probabilities of the equivalent classical channel model are given by:

$$P(y \mid x) = \begin{cases} 1 - \frac{3p}{4} & \text{if } x = y \\ \frac{p}{4} & \text{otherwise} \end{cases}, \tag{6.6}$$

where p is the depolarising probability. The channel is graphically depicted in Figure 6.3. The entanglement-assisted classical capacity of SDC over a quantum depolarising channel, with noiseless entanglement distribution, is therefore given by:

$$C_{E,\,2\mathrm{SDC}}^{\text{one-sided dep}} = 1 + S\left(\rho_b^{AB}\right) - S\left(\Lambda_a^{\text{dep}}\left(\rho^{AB}\right)\right) \tag{6.7}$$

$$\equiv \log_2 M - \mathrm{E}\left[\sum_{n=0}^{M-1} P\left(y_n \,|\, x\right) \log_2 P\left(y_n \,|\, x\right)\right] \tag{6.8}$$

$$= 2 + \frac{4-3p}{4}\log_2\frac{4-3p}{4} + \frac{3p}{4}\log_2\frac{p}{4} \quad \text{cbits/channel use.} \tag{6.9}$$

This does not represent the achievable capacity of the system when a binary FEC code is employed, however. The symbol-to-bit conversion that must inevitably take place at the output of the superdense decoder incurs an inherent capacity loss, which bit-based error-correcting codes are unable to recover [48]. The reduced capacity is a result of the loss of information regarding the correlations between X and Y errors in the depolarising channel, when the output bits of the superdense coding protocol are treated as independent and uncorrelated by the decoder. In this case, the classical channel model, from the perspective of the FEC decoder, is decomposed from a 4-ary symmetric discrete classical channel into two, parallel binary-symmetric channels. The FEC decoder observes that the pair of constituent bits output by the superdense decoder have each independently traversed a corresponding BSC and been subject to the possibility of a bit-flip. The cross-over probability of the BSCs can be determined by marginalising the symbol-based probabilities $P\left(y \,|\, x\right)$ of the above 4-ary channel model to bit-based probabilities $P\left(y_i \,|\, x_i\right)$, where i is the index of the output bit and $i \in \{1,\, 2\}$, i.e.,

$$P\left(y_i \,|\, x_i\right) = \begin{cases} 1 - \frac{3p}{4} + \frac{p}{4} = 1 - \frac{p}{2} & \text{if } x_i = y_i \\ \frac{p}{4} + \frac{p}{4} = \frac{p}{2} & \text{otherwise} \end{cases}. \tag{6.10}$$

The BSC pair are identical and have a crossover probability of $p/2$, such that the bit-based system capacity (as opposed to the symbol-based capacity) is twice the classical capacity of the channel, i.e.,

$$C_{E,\,2\mathrm{SDC},\,\text{bit-based}}^{\text{one-sided dep}} = 2C^{\text{dep}} \quad \text{cbits/channel use.} \tag{6.11}$$

This reduced capacity motivates the use of symbol-based classical error correcting codes, such as duo-binary turbo codes or Reed Solomon codes, instead of bit-based FECs.

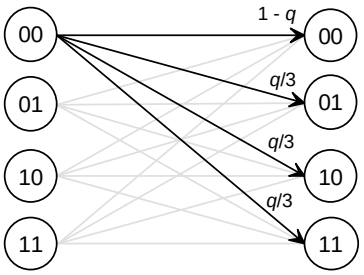


Figure 6.3: 4-ary symmetric channel representation for the superdense coding protocol over the one-sided depolarising channel. Note that $q = \frac{3p}{4}$, where p is the depolarising probability.

6.2.2 Noisy Entanglement Distribution

The symbol-based and bit-based capacities of superdense coding with noisy entanglement distribution over the quantum depolarising channel might be similarly found as above by determining the associated classical transmission models [44]. The two-sided depolarising channel is denoted by the mapping:

$$\Lambda_{ab}^{\text{dep}}(\rho) = \sum_{i,j=0}^{3} q_i q_j \left(A_i \otimes A_j\right) \rho \left(A_i^{\dagger} \otimes A_j^{\dagger}\right). \tag{6.12}$$

The equivalent classical channel for SDC over the two-sided channel is the serial concatenation of two 4-ary symmetric channels, as shown in Figure 6.4. This corresponds to the fact that the qubit representing half of the entangled bipartite state, initially generated and distributed by receiver Bob to sender Alice, traverses the same depolarising channel twice in the process of enacting the SDC protocol. The transition probabilities of the overall classical channel are given by:

$$P(y \,|\, x) = \begin{cases} \left(1 - \frac{3p}{4}\right)^2 + \frac{3p^2}{16} & \text{if } x = y \\ 2\left(1 - \frac{3p}{4}\right)\left(\frac{p}{4}\right) + \frac{p^2}{8} & \text{otherwise} \end{cases}, \tag{6.13}$$

such that the entanglement-assisted classical capacity of the channel is:

$$C_{E,\,2\text{SDC}}^{\text{two-sided dep}} = 1 + S\left(\Lambda_b^{\text{dep}}\left(\rho_b^{AB}\right)\right) - S\left(\Lambda_{ab}^{\text{dep}}\left(\rho^{AB}\right)\right) \tag{6.14}$$

$$\equiv 2 + \frac{1 + 3\left(1-p\right)^2}{4} \log_2 \frac{1 + 3\left(1-p\right)^2}{4}$$

$$+ 3\left(\frac{1 - \left(1-p\right)^2}{4} \log_2 \frac{1 - \left(1-p\right)^2}{4}\right) \quad \text{cbits/channel use.} \tag{6.15}$$

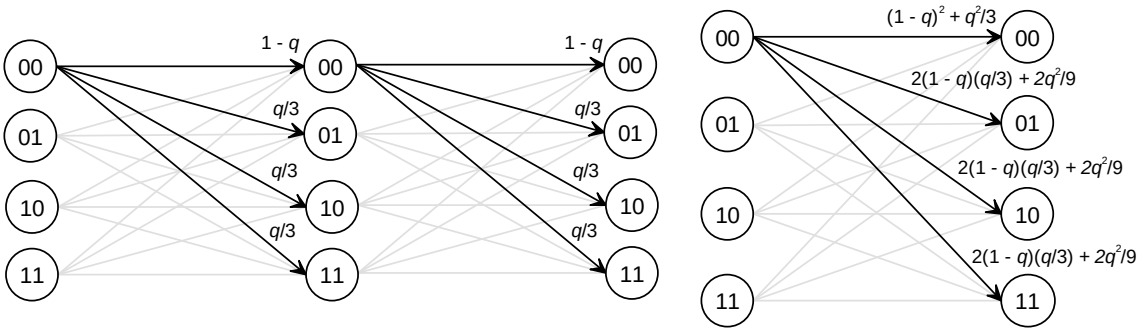


Figure 6.4: Equivalent 4-ary symmetric channel representations for the superdense coding protocol over the two-sided depolarising channel. Note that $q = \frac{3p}{4}$, where p is the depolarising probability.

As before, this capacity represents the symbol-based capacity of the channel, and employing a bit-based error correcting code results in the system capacity being reduced to that of a parallel pair of independent binary-symmetric channels, whose transition probabilities are given by:

$$P\left(y_i \,|\, x_i\right) = \begin{cases} \left(1 - \frac{3p}{4}\right)^2 + \frac{3p^2}{16} + 2\left(1 - \frac{3p}{4}\right)\left(\frac{p}{4}\right) + \frac{p^2}{8} = 1 + p\left(\frac{p}{2} - 1\right) & \text{if } x_i = y_i \\ 4\left(1 - \frac{3p}{4}\right)\left(\frac{p}{4}\right) + \frac{p^2}{4} = p\left(1 - \frac{p}{2}\right) & \text{otherwise} \end{cases}, \tag{6.16}$$

for $i \in \{1, 2\}$. From this, the bit-based system capacity for this channel model can be readily determined as follows:

$$\begin{aligned} C_{E,\mathrm{2SDC,bit-based}}^{\text{two-sided dep}} &= 2 + 2p\left(1 - \frac{p}{2}\right)\log_2 p\left(1 - \frac{p}{2}\right) \\ &\quad + 2\left[1 + p\left(\frac{p}{2} - 1\right)\right]\log_2\left[1 + p\left(\frac{p}{2} - 1\right)\right] \quad \text{cbits/channel use.} \end{aligned} \tag{6.17}$$

6.2.3 Higher-order Entanglement

A similar approach might also be employed to determine the classical capacity of superdense coding exploiting higher-order entanglement. The 2-qubit superdense coding protocol (2SDC) can be generalised to make use of a N-qubit maximally-entangled initial state. There are a number of potential scenarios which arise with the use of more than two entangled qubits.

In the conventional scenario, with a single sender Alice and receiver Bob, the receiver may distribute from one to $N - 1$ of the available, entangled qubits to the sender.

Since Alice may only perform unitary operations on her local qubits in order to influence the global state, the number of qubits distributed to her determines the number of unambiguously distinguishable, maximally-entangled states which she may produce, and therefore the amount of classical information that she may encode. In the case where Bob distributes the maximum number of qubits to Alice, she performs the encoding and the $N-1$ qubits are returned to the receiver, who measures the global state, extracting $\log_2 2^N = N$ cbits. In an alternative scenario, a N-qubit maximally-entangled initial state may be shared amongst N parties, such that a receiver might detect messages from $N-1$ senders simultaneously via a single measurement of the multipartite state [51]. This behaviour could potentially be useful in implementing a distributed quantum network. For each transmission, one sender is granted permission to utilise any of four possible unitary operations such that they might encode two cbits onto their qubit. The remaining senders are then able to encode only a single cbit onto their qubit, via any two unitary operations at their disposal. The condition on the unitary operations is that, for each possible combination of operations performed by the senders, the resulting state of the N-qubit system is a member of the set of maximally-entangled states for N-qubits. This is possible because the $N-1$ senders may collectively perform $4 \times 2 \times ... \times 2 = 2N$ combinations of unitary operations on the initial state, and there are $2N$ maximally entangled states for a N-qubit system. All $N-1$ qubits are returned over individual channels to the receiver, who extracts N cbits, simultaneously obtaining $N-1$ messages. The entanglement-assisted classical capacity between two users for both N-qubit SDC scenarios over noiseless channels is the same: $N/(N-1)$ cbits/channel use. Clearly, and perhaps disappointingly, this capacity is reduced as the order of the entanglement utilised is increased.

For the 3-qubit superdense coding protocol (3SDC), the overall transmission model reduces to an 8-ary symmetric classical channel. The transition probabilities corresponding to one of the possible outputs is given by:

$$P(y = 000 \,|\, x = c) = \begin{cases} \left(1 - \frac{3p}{4}\right)^2 + \frac{p^2}{16} & \text{if } c = 000 \\ 2\left(1 - \frac{3p}{4}\right)\left(\frac{p}{4}\right) & \text{if } c = 001 \\ \left(1 - \frac{3p}{4}\right)\left(\frac{p}{4}\right) + \frac{p^2}{16} & \text{if } c \in \{010, 011, 110, 111\} \\ \frac{p^2}{8} & \text{if } c \in \{100, 101\} \end{cases}, \tag{6.18}$$

where c is the 3-bit codeword corresponding to each of the eight transmission symbols. The channel is graphically depicted in Figure 6.5. The entanglement-assisted classical capacity of 3SDC over a quantum depolarising channel is given by:

$$C_{E,\mathrm{3SDC}}^{\mathrm{one-sided\,dep}} = \frac{3}{2} - \frac{1}{2}\sum_{n=0}^{7} P(y_n \,|\, x) \log_2 P(y_n \,|\, x) \quad \text{cbits/channel use.} \tag{6.19}$$

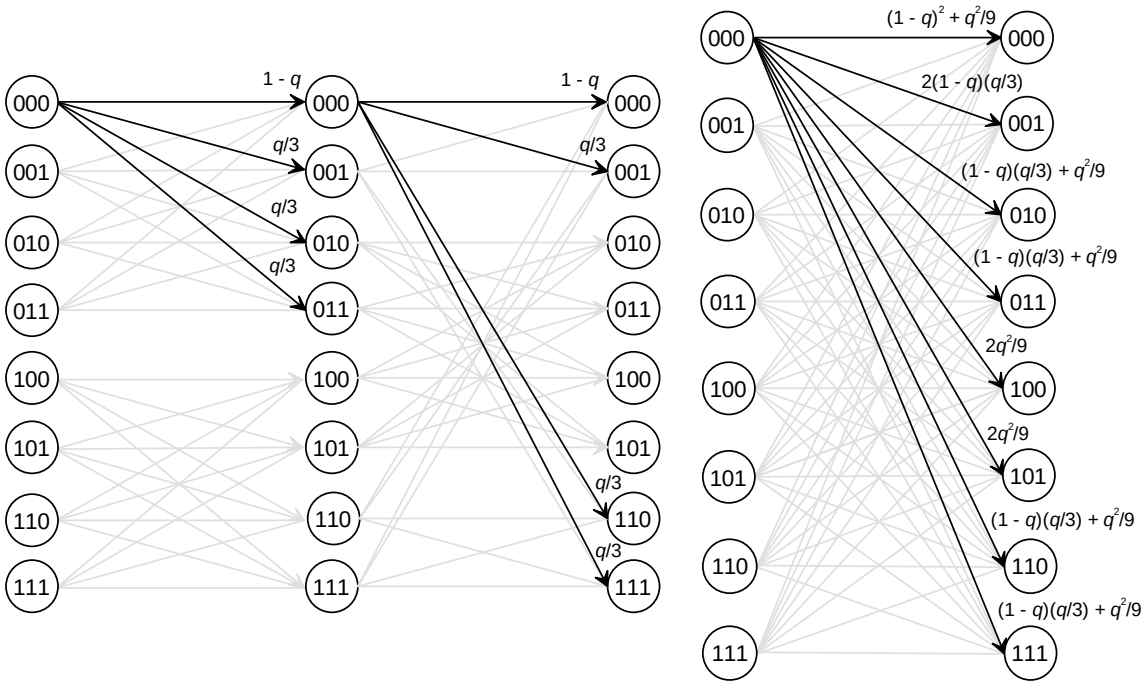


Figure 6.5: Equivalent 8-ary symmetric channel representations for the 3-qubit superdense coding protocol over the one-sided depolarising channel. Note that $q = \frac{3p}{4}$, where p is the depolarising probability.

The 8-ary classical channel model can be decomposed into three parallel BSCs to determine the bit-based capacity. By marginalising the symbol-based probabilities $P(y \,|\, x)$ to bit-based probabilities $P(y_i \,|\, x_i)$, where i is the index of the output bit and $i \in \{1, 2, 3\}$, the crossover probabilities of the respective BSCs can be found to be:

$$P(y_i \,|\, x_i) = \begin{cases} 1 + p\left(\frac{p}{2} - 1\right) & \text{if } x_i = y_i \\ p\left(1 - \frac{p}{2}\right) & \text{otherwise} \end{cases}, \quad \text{for } i \in \{1, 2\}, \tag{6.20}$$

and

$$P(y_i \,|\, x_i) = \begin{cases} 1 - \frac{p}{2} & \text{if } x_i = y_i \\ \frac{p}{2} & \text{otherwise} \end{cases}, \quad \text{for } i = 3. \tag{6.21}$$

Interestingly, the first two bits experience the same respective channels as in the case of noisy entanglement distribution, and the third is simply a depolarising channel. The bit-based system capacity for 3SDC over a one-sided depolarising channel is therefore given by:

$$C^{\text{one-sided dep}}_{E,\text{3SDC,bit-based}} = C^{\text{two-sided dep}}_{E,\text{2SDC,bit-based}} + C^{\text{dep}} \quad \text{cbits/channel use.} \tag{6.22}$$

6.2.4 Channel Capacities

Figure 6.6 depicts the classical capacity of the quantum depolarising channel and entanglement-assisted classical capacities of the aforementioned variations of the superdense coding protocol over said channel, as functions of the depolarising probability or noise parameter p. As expected, superdense coding with noiseless entanglement distribution represents the best case scenario, facilitating the transmission of classical information at the maximum rate in the presence of noise. Noisy entanglement distribution clearly has a dramatic impact on the superdense coding capacity, to such a degree that in the regime $0.345 \leq p \leq 1$, the use of an entanglement resource no longer increases the maximum transmission rate for classical information over elementary encoding. This reinforces the necessity of the entanglement distillation protocol for entanglement-assisted communication over especially noisy quantum channels. The capacity loss inherent when utilising a bit-based over a non-binary or symbol-based classical error-correcting code to provide error protection is also evident in the figure, and can be seen to be most severe for 3SDC.

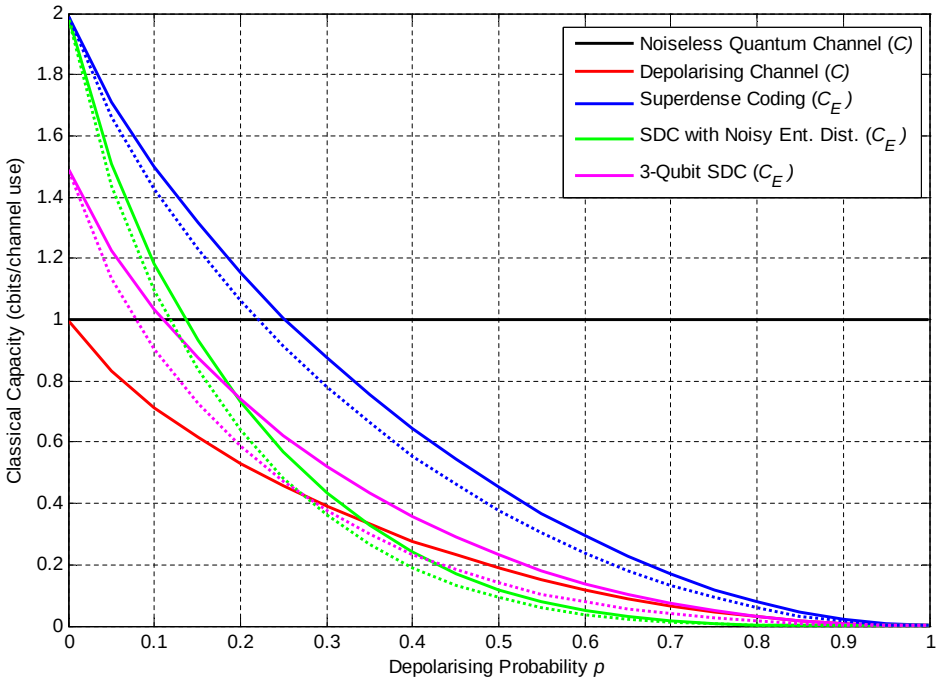


Figure 6.6: The classical capacity C^{dep} of the depolarising channel, the superdense coding capacities over a one-sided and two-sided depolarising channel, $C_{E,\text{2SDC}}^{\text{one-sided dep}}$ and $C_{E,\text{2SDC}}^{\text{two-sided dep}}$, respectively, and the capacity $C_{E,\text{3SDC}}^{\text{one-sided dep}}$ of the 3-qubit SDC protocol over a one-sided depolarising channel. The symbol-based and bit-based capacities in each case are shown by the solid and dashed lines, respectively.

7 Performance of Classical Error-correcting Codes

This chapter presents Monte Carlo simulations performed to assess the error performance of various classical error-correcting codes in the classical-quantum communication system described in Chapter 6. Figure 7.1 depicts the general block diagram of the simulation process, designed to collect relevant classical error statistics such as bit error rate, symbol error rate and frame (block) error rate. The error-correcting codes considered include non-binary Reed-Solomon and duo-binary turbo codes, as well as Viterbi-decoded convolutional codes and binary turbo codes. The latter are implemented for context and comparison, as the performance of the symbol-based FEC codes is of primary interest, since they are not subject to the inherent capacity loss discussed in the previous chapter. Collectively, the results are expected to show that, by employing error-correction in the classical rather than the quantum domain, the system is able to transmit classical information over a quantum depolarising channel with an acceptable tolerance to noise. The superdense encoder and decoder can be considered to be analogous to a M-ary modulator and demodulator, respectively. The puncturing functionality is only relevant to binary and duo-binary turbo codes.

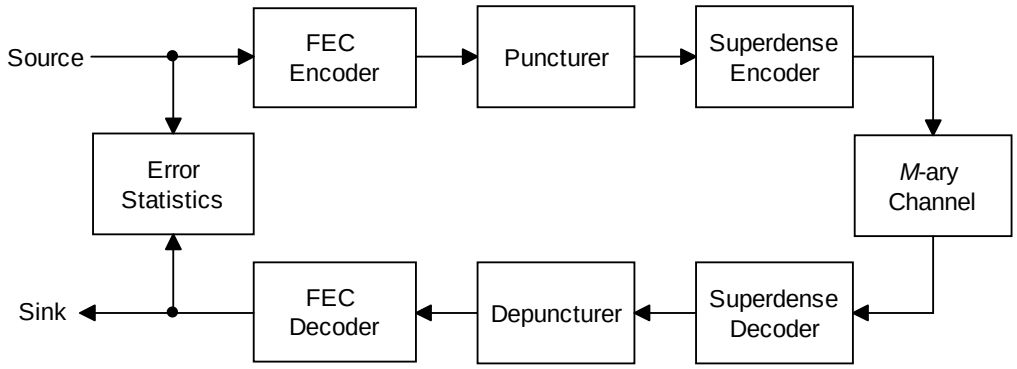


Figure 7.1: Block diagram describing the function of the simulator.

7.1 Reed-Solomon Codes

Reed-Solomon codes present a natural choice for a non-binary classical error correcting scheme capable of providing a one-to-one correspondence between classical transmission symbols and SDC symbols (Bell states), so as to avoid the capacity loss inherent when using bit-based FEC codes. The error performance of the non-binary FEC code was evaluated over the various M-ary classical channel models corresponding to the three variations of SDC over the quantum depolarising channel detailed in Chapter 6.

For 2SDC, where the number of cbits associated with each transmission symbol is 2, the available codes are those defined over GF (4), which includes only the single-

error-correcting $(3, 1)$ RS code with rate 1/3. While employing this code, the overall coding rate of the classical-quantum communication system becomes $1/3 \times 2 = 2/3$ cbits/channel use. From the previous chapter's Figure 6.6, the maximum tolerable depolarising probability or noise associated with this rate can be found to be 0.39 for 2SDC with noiseless entanglement distribution and 0.22 with noisy entanglement distribution. These values represent the symbol-based capacity limits in each case. Similarly, for 3SDC, the code set includes those codes defined over GF (8), i.e. the double error-correcting$(7, 1)$ and the single error-correcting $(7, 3)$ RS codes. When utilising the code with rate 1/7, the system code rate becomes $1/7 \times 3/2 = 3/14$ cbits/channel use and the maximum tolerable depolarising probability is 0.45, for 3SDC with noiseless entanglement distribution. For the $(7, 3)$ RS code, the system code rate increases to 9/14 cbits/channel use and the capacity limit occurs at a depolarising probability of 0.17.

The average classical BER over thirty thousand simulations for the $(3, 1)$ RS code and the $(7, 1)$ and $(7, 3)$ RS codes was found for 2SDC with noisy and noiseless entanglement distribution and 3SDC with noiseless distribution, respectively. The results of the simulations are presented in Figures 7.2 and 7.3. It can be observed immediately that the error performance in all cases is far from capacity-approaching, and the coding gains are minimal at low BERs. For 2SDC, the coding gain (in terms of the depolarising probability) at a BER of 10^{-2} is reduced from approximately $10\log_{10}\left(\frac{0.105}{0.22}\right) = 6.79$ dB to 5.64 dB when the entanglement is distributed such that the initial resource is noisy.

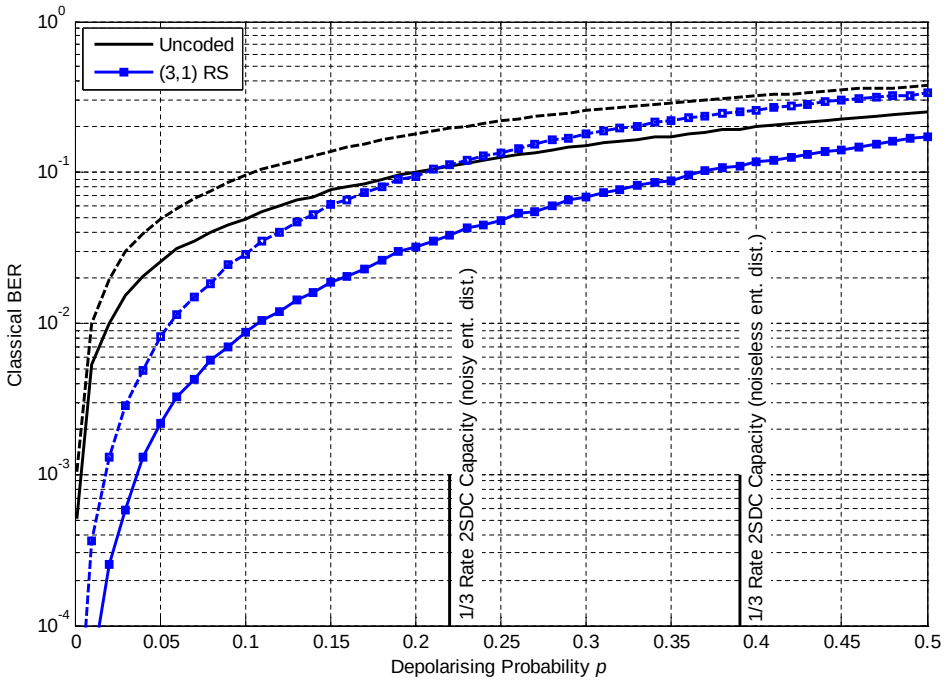


Figure 7.2: BER vs. quantum depolarising probability for 2SDC with noiseless (solid) and noisy (dashed) entanglement distribution employing a single error-correcting (3, 1) RS code.

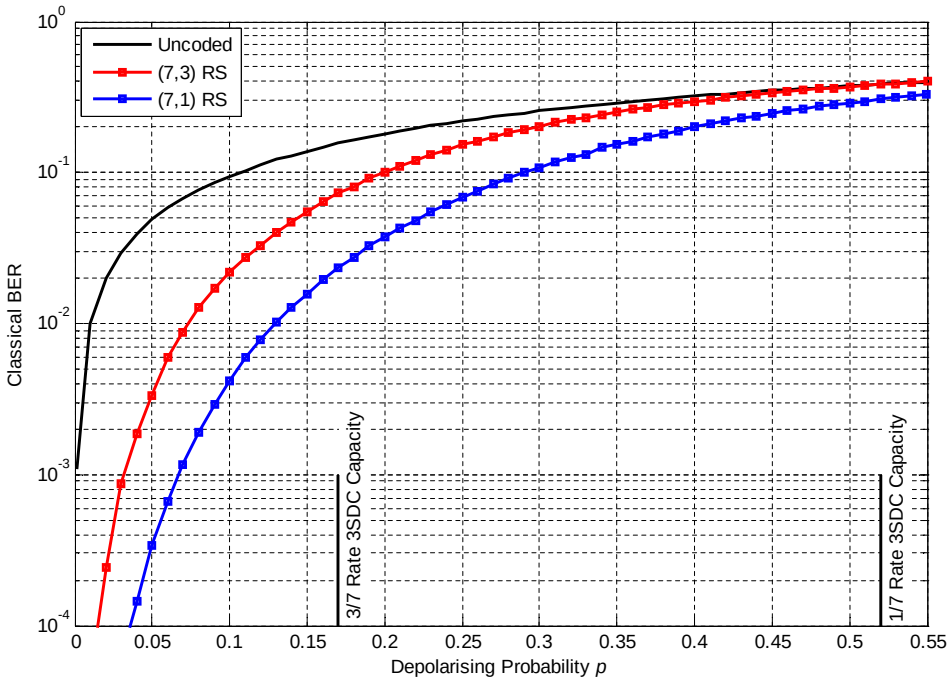


Figure 7.3: BER vs. quantum depolarising probability for 3SDC with noiseless entanglement distribution employing the double error-correcting (7, 1) RS code and the single error-correcting (7, 3) RS code.

The relative coding performance between 2SDC and 3SDC is not directly comparable, since the coding rates are different. In the latter case, the gain from reducing the code rate from 3/7 to 1/7 is only 3.68 dB at a BER 10^{-4}, despite the greatly enhanced capacity are the latter rate.

The limited error performance observed can be attributed to the fact that the RS code used is designed primarily for *erasure* channels. There are also limitations due to the discrete nature of the M-ary channels involved and the limited number of transmission symbols available. In an analogous classical scenario, with M-ary modulation over an Additive While Gaussian Noise (AWGN) channel, the error-correcting capability or BER performance of the Reed-Solomon code can be readily enhanced by increasing the block length, effectively increasing M. The code rate can be kept constant by simply maintaining the same ratio of information to redundancy, or lowered to further improve performance at the cost of information rate. The order of the polynomials involved in the encoding and decoding of RS codes with larger block sizes are such that the time taken, and processing power required, increases exponentially. In other words, there is an inherent trade-off between the performance gained by increasing the block size and the complexity imposed. Unfortunately, this holds no bearing for 2SDC and 3SDC, for which the one-to-one mapping of symbols between the classical and quantum domains means that the RS codes are limited to block sizes corresponding to four and eight transmission symbols, respectively. The set of available RS codes in each case are those of low complexity and relatively poor performance. Convolutional turbo codes and other iteratively-

decoded codes can therefore be expected to outperform Reed-Solomon codes in the described classical-quantum communication system.

7.2 Convolutional and Turbo Codes

Convolutional turbo codes represent the most likely candidates for classical error-correcting codes capable of enabling the classical-quantum communication system, detailed in the previous chapter, to operate at a classical information rate approaching the various capacity limits of the given scenarios. While it is theoretically possible for a block code with a sufficiently large block length (or a convolutional code with a large enough constraint length), to achieve the same rate, the discrete nature of the quantum channel renders the use of such codes nonviable. The superdense coding protocol, generalised to a N-particle system, would need to operate on an entangled quantum state whose dimension N coincides with the large block length of the code; a resource that would potentially be severely difficult to generate and maintain. Even if this was not the case, the complexity of, and processing power required for, the decoding of such codes would make their use impractical. Binary and duo-binary turbo codes, however, overcome this limitation through the use of recursive encoder and iterative decoders. These elements ultimately serve to make a convolutional code with a shorter constraint length appear to be, and exhibit the performance of, a block code with large block length.

Figure 7.4 depicts and compares the simulated BER versus quantum depolarising probability for an "off-the-shelf" Viterbi-decoded convolutional code, conventional turbo code (TC) and duo-binary turbo code, over the 4-ary classical channel corresponding to 2SDC with noiseless entanglement distribution over a quantum depolarising channel. Both the TC and DBTC are memory-3 codes with code rate 1/3, which can be readily increased to 1/2 by puncturing the encoder output. The rate of the convolutional code is adapted between 1/2 and 1/3 by adjusting the generator polynomial appropriately. With these code rates available, the classical-quantum communication system is capable of reducing its overall rate from 1 to 2/3 cbits/channel use in order to remain operable in the presence of noise which exceeds a given threshold.

For the iteratively-decoded TC and DBTC, eight iterations of the max-log-MAP algorithm (a simplified version of the MAP algorithm, operating in the log domain) was chosen as a suitable compromise between error performance and decoding efficiency. Each simulation point in Figure 7.4 represents the average BER for 200 frames of 1728 bits for the TC, 1728 bits (864 couples) for the DBTC (the largest supported frame size in the DVB-RCS standard), and 10000 bits for the Viterbi-decoded convolutional code. It is immediately evident that the DBTC exhibits the best error performance of the three convolutional codes, thereby facilitating a negligible BER up to a depolarising probability closest to the maximum tolerable probability for each code rate, as expected.

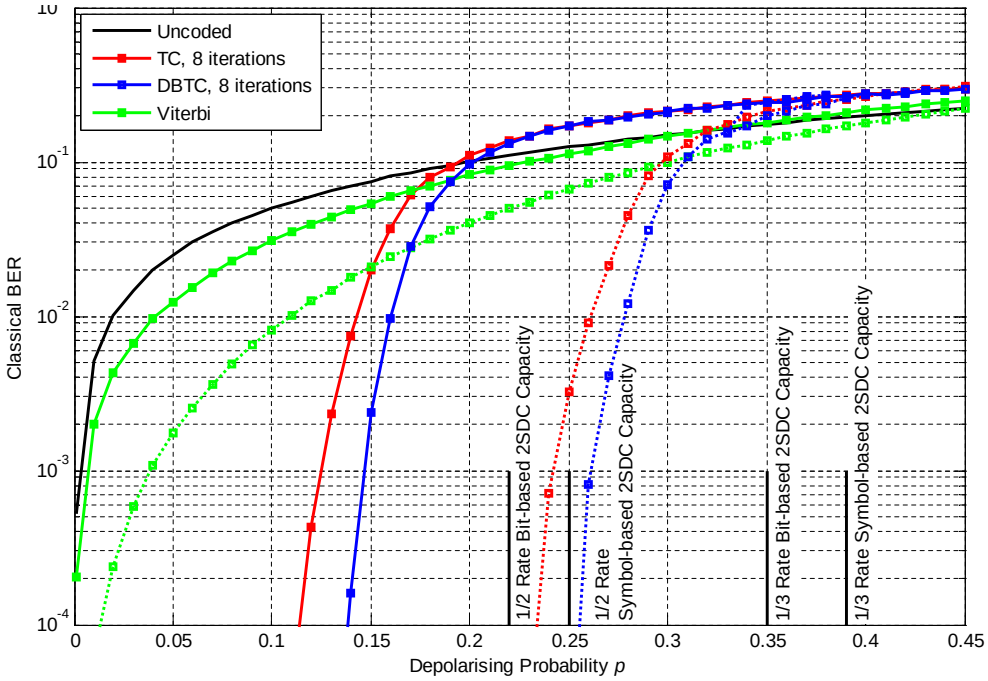


Figure 7.4: BER vs. quantum depolarising probability for 2SDC with noiseless entanglement distribution employing three different classical, convolutional error-correcting codes. BER curves for a code rate of 1/2 and 1/3 are shown by solid and dashed lines, respectively. Turbo codes are punctured to achieve a code rate of 1/2 and utilise iterative max-log-MAP decoding.

The low-complexity, Viterbi-decoded convolutional code provides a limited error tolerability that is comparable to that of the single-error correcting (3, 1) RS code, at a code rate of 1/3: a coding gain of approximately 6.99 dB at a BER of 10^{-2}. The binary and duo-binary turbo codes improve greatly upon this error performance, with BER curves which converge much more rapidly, a characteristic that can be attributed to their use of iterative decoding. The DBTC outperforms its binary counterpart by as much as 1.05 dB at the negligibly low BER of 10^{-4} for a code rate of 1/2, and 0.45 dB for a rate of 1/3.

The uncoded and coded BER curves in Figure 7.4 all exhibit a clear crossover point, beyond which the error performance of the associated classical error-correcting code becomes worse than in the case where the transmitted information is uncoded. For example, for $p > 0.2$ the employment of the punctured DBTC is no longer of any benefit to the classical-quantum communication system, and actually serves to increase the probability of bit error. This undesirable behaviour can be attributed to the fact that, as the probability of depolarisation occurring in the channel increases, the probability that the number of resulting bit errors exceeds the error-correcting capabilities of the codes also increases. The crossover points indicate the average depolarising probability at which this occurs.

Figure 7.5 provides a complementary depiction of the results in Figure 7.4, describing the classical information rate or throughput as a function of the depolarising probability, based on the block error rate given by the formula:

$$P_B = 1 - [1 - P_b(p)]^n , \tag{7.1}$$

where P_b is the bit error rate, p is the depolarising probability and n is the block or frame length in bits. The classical throughput for 2SDC is then given by:

$$T = 2 \times R \times [1 - P_B(p)] \quad \text{cbits/channel use}, \tag{7.2}$$

where R is the FEC rate. The throughput curves in particular provide a clear picture of how close the classical information rate of the system is to the capacity, and more readily identifies the depolarising probability thresholds at which the error performance curves converge.

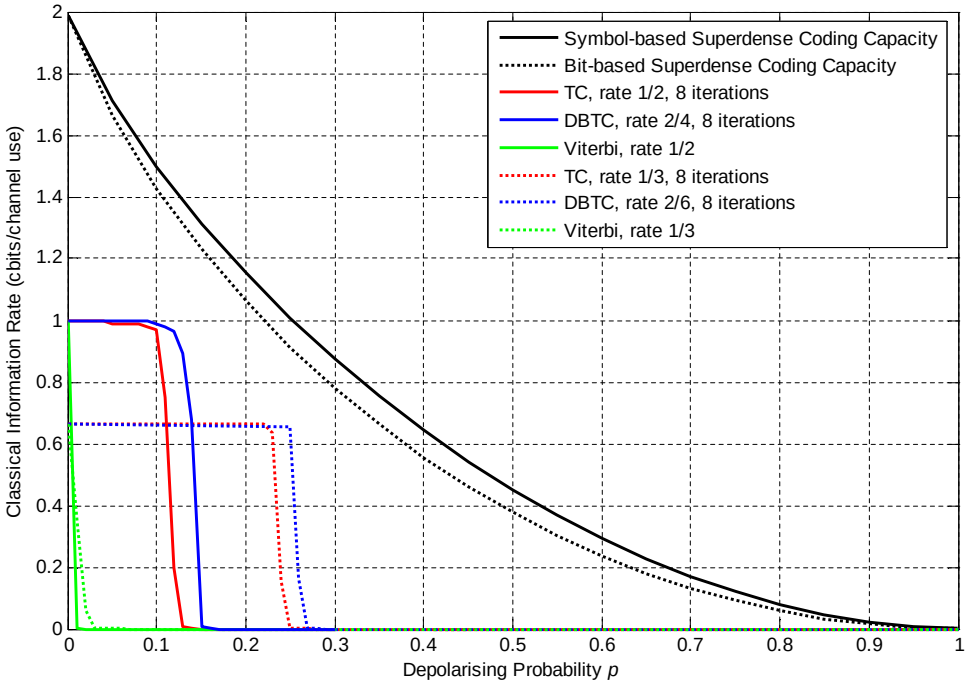


Figure 7.5: Achievable classical information rates (based on the BLER) for the classical-quantum communication system employing 2SDC with noiseless entanglement distribution with three different classical, convolutional error-correcting codes.

The relative performances of the three classical error-correcting codes under consideration for 2SDC with noisy entanglement distribution and 3SDC with noiseless entanglement distribution are similarly shown in Figures 7.6 through 7.9. Interestingly, noisy 2SDC can be observed in all cases to exhibit comparable error performance to noiseless 3SDC at a BER of 10^{-4}, although the curves corresponding to the former converge earlier and the throughput in the latter case is reduced.

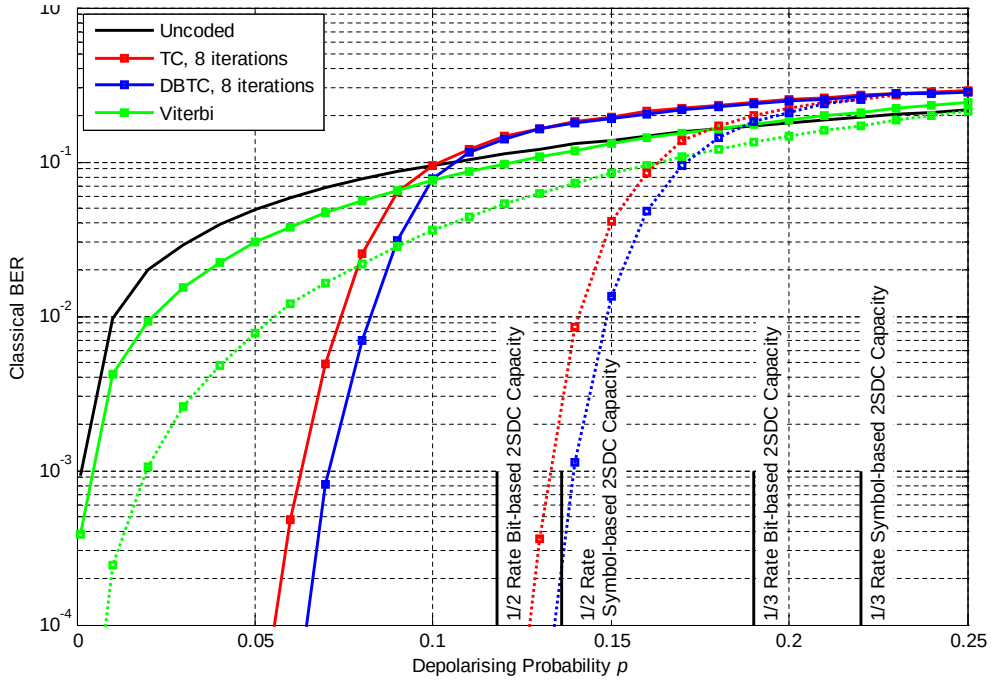


Figure 7.6: BER vs. quantum depolarising probability for 2SDC with noisy entanglement distribution employing three different classical, convolutional error-correcting codes. BER curves for a code rate of 1/2 and 1/3 are shown by solid and dashed lines, respectively. Turbo codes are punctured to achieve a code rate of 1/2 and utilise iterative max-log-MAP decoding.

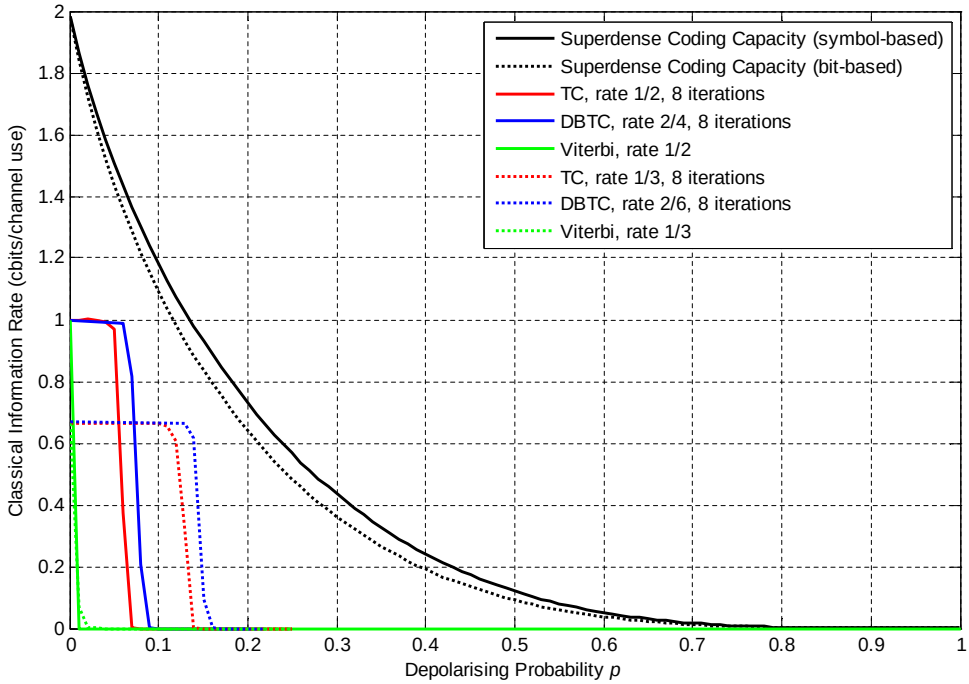


Figure 7.7: Achievable classical information rates (based on the BLER) for the classical-quantum communication system employing 2SDC with noisy entanglement distribution with three different classical error-correcting codes.

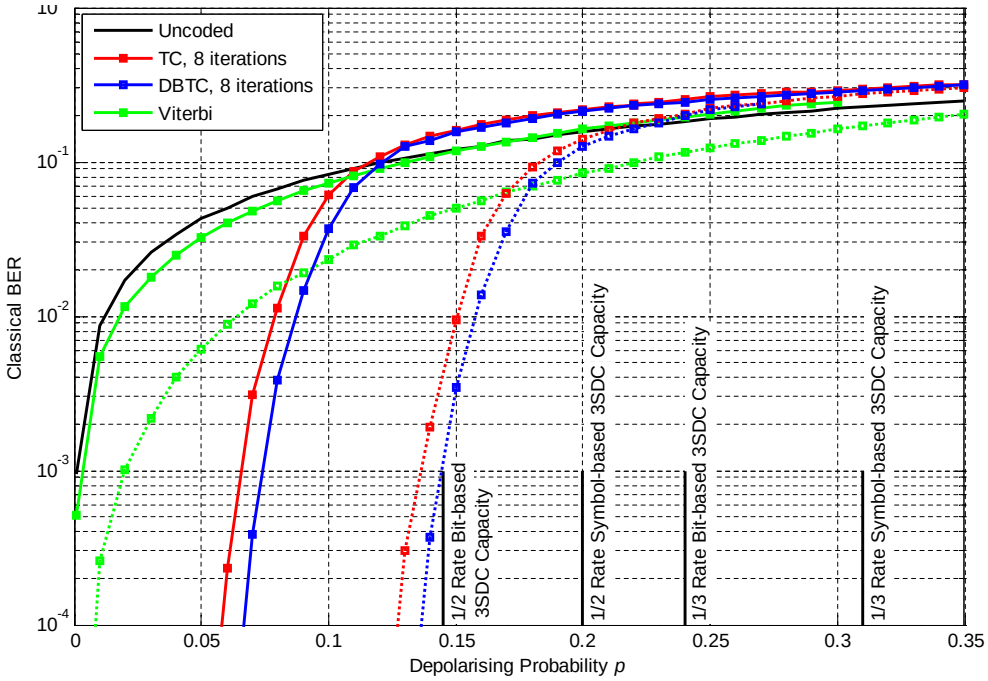


Figure 7.8: BER vs. quantum depolarising probability for 3SDC with noiseless entanglement distribution employing three different classical, convolutional error-correcting codes. BER curves for a code rate of 1/2 and 1/3 are shown by solid and dashed lines, respectively. Turbo codes are punctured to achieve a code rate of 1/2 and utilise iterative max-log-MAP decoding.

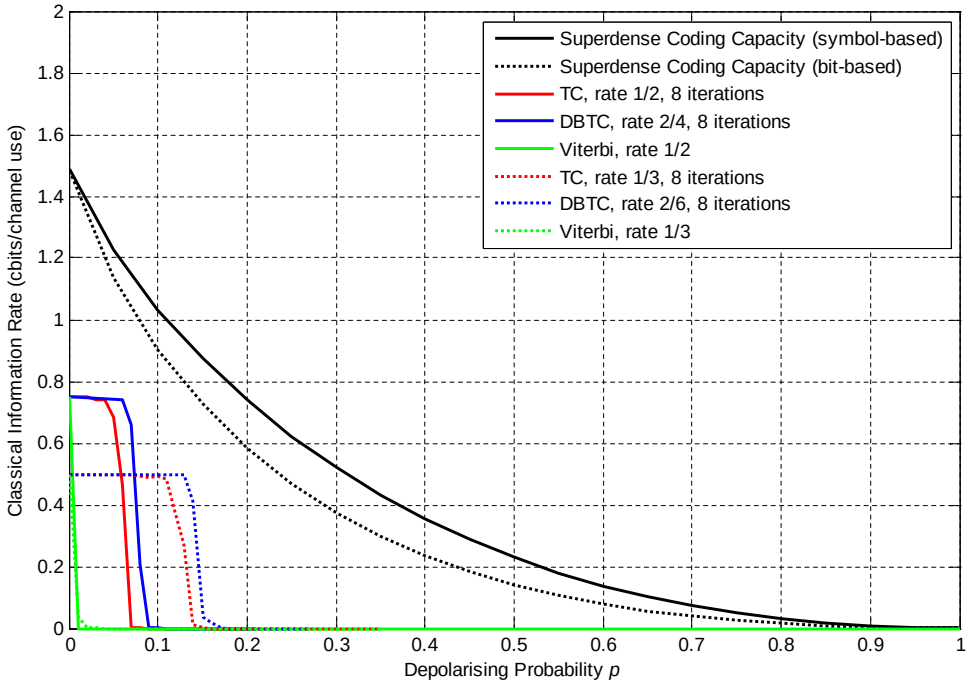


Figure 7.9: Achievable classical information rates (based on the BLER) for the classical-quantum communication system employing 3SDC with noiseless entanglement distribution with three different classical error-correcting codes.

The results collectively allow the maximum quantum depolarising noise up to which the classical-quantum communication system may operate with an acceptable rate of transmission error to be identified in each of the given scenarios. For example, the system employing 2SDC with noiseless entanglement distribution and a DBTC for error-correction can tolerate a depolarising probability of up to 0.255 before experiencing a consequential BER. Another particularly useful observation is that, given that the properties of the noisy channel between the communicating parties are such that the probability of depolarisation is less than 0.14, the system may forgo the use of entanglement distillation to purify a nosily distributed initial state (and the associated entanglement resource cost), and still communicate classical information with a negligibly low BER.

Table 7.1 details the deviations from the respective capacities for each of the classical error-correcting codes at a BER of 10^{-4}. Unfortunately, in none of the cases is the duo-binary code able to approach or exceed the maximum tolerable depolarising probability corresponding to the bit-based capacity at a sufficiently low BER, or indeed outperform the URC-IRCC code in [48]. The latter was shown to enable an error performance, for a classical-quantum communication system functionally equivalent to the one detailed herein, within 0.6 dB and 0.75 dB of the bit-based capacity for 2SD and 3SDC (with noiseless entanglement distribution) at a BER of 10^{-4}, respectively. Comparatively, the duo-binary turbo code facilitates performance within 1.96 dB and 3.29 dB of the corresponding capacity limit. There remains a possibility that compounding the decoding complexity of the DBTC by increasing the number of iterations may yield improved error performance.

Table 7.1: Deviation from the symbol-based (bit-based) capacity at a BER of 10^{-4}.

	Ent. Dist.	Rate	Viterbi	TC	DBTC
2SDC	noiseless	1/2	-	3.37 (2.82) dB	2.52 (1.96) dB
		1/3	12.22 (11.66) dB	2.26 (1.73) dB	1.95 (1.37) dB
	noisy	1/2	-	3.93 (3.32) dB	3.21 (2.96) dB
		1/3	13.52 (12.90) dB	2.38 (1.76) dB	2.22 (1.59) dB
3SDC	noiseless	1/2	-	5.23 (3.91) dB	4.69 (3.29) dB
		1/3	14.91 (13.80) dB	3.77 (2.66) dB	3.61 (2.68) dB

7.2.1 Limits on the Error Performance of DBTCs

The error performance of certain classical error-correcting codes can be improved by increasing the capability and complexity of the decoding process. While this may not always be efficient or feasible in communications, where a balance between error rate and coding delay must typically be struck, it can provide an indication of the limit of a code's ability to protect from errors. Since the frame length for duo-binary turbo codes is limited to 1728 bits by the interleaver design, the error performance

of the DBTC featured herein may only be enhanced by increasing the number of decoding iterations [41].

Figure 7.10 presents the simulated bit errors rates for the classical-quantum communication system employing 2SDC with noiseless entanglement distribution and a punctured DBTC for error protection, after a number of decoding iterations, as functions of the depolarising probability. The convergence threshold for the code may be readily observed, beyond which the BER curves converge towards the asymptotic performance of the system. The curves reveal the performance of the "off-the-shelf" duo-binary turbo code to be non-ideal, in that convergence occurs late and the asymptotic performance is relatively poor. Ideally, through optimisation and tuning, the convergence threshold would be reduced and the curves trend to a negligible BER more rapidly. It can also be observed that, as expected, the error performance of the code improves as the number of iterations increases, although there is clear evidence of a diminishing return on the performance at the cost of compounding complexity. The gain from increasing the number of iterations beyond 16 is negligible, which provides an indication of the limits of the code's performance for a frame length of 1728 bits. For 32 iterations of the max-log-MAP algorithm, the duo-binary turbo codes error performance comes within 1.81 dB of the bit-based capacity at the negligible BER of 10^{-4}.

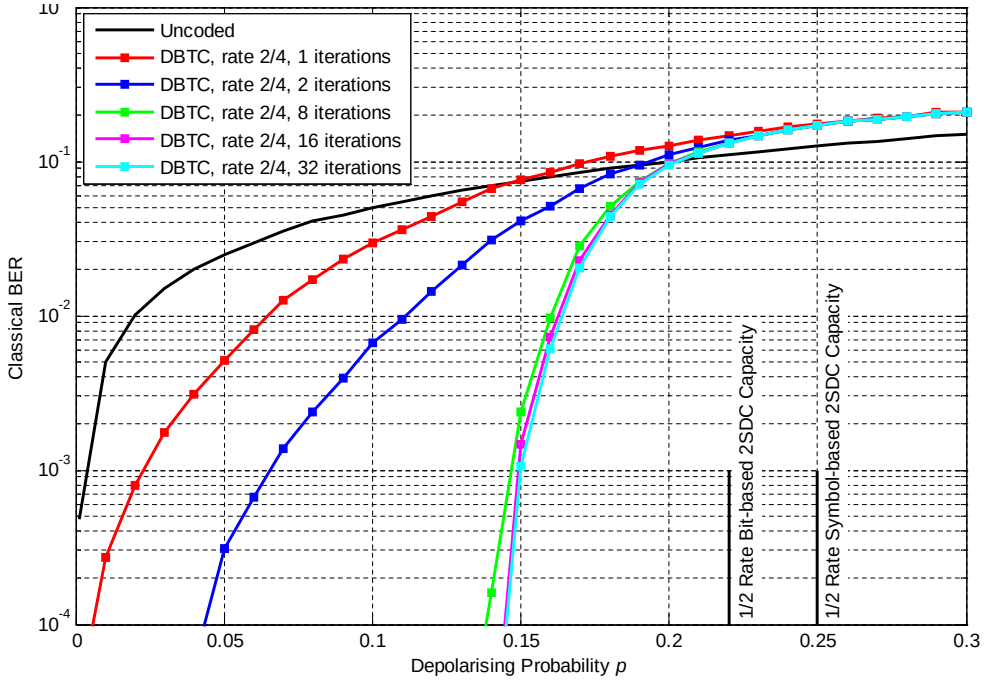


Figure 7.10: BER vs. quantum depolarising probability for 2SDC with noiseless entanglement distribution employing a duo-binary turbo code, punctured to achieve a code rate of 1/2, utilising iterative max-log-MAP decoding.

8 Conclusion

Quantum communication involves the exploitation of phenomena specific to quantum theory for the purposes of communicating information with perfect security. Information is carried in a quantum property of a physical system, such as the polarisation of a photon or spin state of an atom. In accordance with the postulates of quantum theory, interception of a transmitted qubit by an unauthorised party destroys the quantum information and immediately alerts the receiver to the presence of the eavesdropper. Quantum entanglement may be exploited to facilitate the transmission of multiple classical bits with that of a single qubit, or to circumvent a noisy quantum channel and transmit quantum states perfectly with the assistance of classical communications. Such capabilities obviously make quantum technologies a viable and attractive prospect for improving information security in future communication networks.

As the transmission of quantum states becomes a physical reality, it is necessary to consider methods of protecting quantum information and controlling errors occurring as a result of interactions with the environment and quantum noise. Quantum error-correcting codes have been developed which protect a single qubit from arbitrary errors resulting from the traversal of noisy quantum channels, by spreading its state over the highly-entangled composite state of multiple qubits. The most efficient quantum error-correcting codes to date encode quantum information in the state of at least five entangled qubits. Since maintaining even a single pair of qubits in a maximally entangled state presents enough of a challenge to researchers, until such time as entanglement becomes an economical resource, it is necessary to consider if more efficient and established methods of error protection from the classical realm might be sufficient. Another major concern with the use of quantum error-correcting codes for communication lies in the relatively low information rates achievable when compared with classical codes. For a theoretical classical-quantum communication system which transmits classical information from one party to another over a noisy quantum channel via the superdense coding protocol, the most efficient QECC could facilitate an information rate of two classical bits per five uses of the channel. In comparison, a half-rate classical FEC would provide for the substantially superior information rate of one classical bit per channel use. This thesis has shown that moving the error correction from the quantum to the classical domain allows for the reliable communication of classical information over a noisy quantum channel at higher rates and without necessitating the generation and maintenance of higher-order entangled states.

For the classical-quantum communication system described herein, multiple operating scenarios were considered. These included the noiseless and noisy distribution of entanglement resources and the use of higher-order entangled states in the enactment of the superdense coding protocol. For each scenario, the overall transmission model was shown to reduce to a classical, discrete and memoryless channel, such that the channel's capacity might be derived. The use of binary classical error-correcting codes

was shown to result in an inexorable capacity loss, attributed to the fact that binary codes disregard correlations between the pairs of classical output bits corresponding to the transmission symbols in the quantum domain. This motivated the use of non-binary FECs such as duo-binary turbo codes, which were speculated to potentially be capable of breaching the reduced, bit-based capacity through a one-to-one mapping of symbols between the classical and quantum domains. Through simulation, DBTCs were shown to provide superior error performance when compared with binary turbo codes and other FEC codes, allowing the system with to operate relatively close to the bit-based capacity at a negligible BER, in a number of practical scenarios. A primary contribution of this thesis has been to show and measure via simulation the gain to be obtained by employing non-binary classical error correction codes over binary codes in the entanglement-assisted communication of classical information over a quantum depolarising channel. A punctured duo-binary code was shown to outperform a binary counterpart by as much as 1.05 dB at a BER of 10^{-4}, as result of the former's ability to exploit the correlations between the pairs of bits output by the superdense coding protocol.

The potential avenues for future work are many, as the amount of existing research into the use of classical error-correcting codes in quantum communications is meagre. There exists many more non-binary classical error-correcting codes that may provide near-capacity error performance than those investigated herein. Candidates include non-binary Low Density Parity Check (LDPC) codes and polar codes. Investigation into the performance of classical FEC codes for quantum communication over different quantum channels may also be warranted, especially quantum erasure channels. The classical error-correcting codes used to generate the results presented herein were "off-the-shelf" codes, having undergone no optimisation for the purpose at hand. It is speculated that a degree of "handcrafting" or tuning could potentially see a DBTC facilitate an error performance enabling the classical-quantum communication system to operate at a rate in closer proximity to the capacity at a negligible BER.

References

[1] M. Wilde. “From Classical to Quantum Shannon Theory.” arXiv:1106.1445v4, 2012.

[2] J. A. Jones, and D. Jaksch. *Quantum Information, Computation and Communication.* New York, Cambridge University Press, 2012.

[3] M. A. Nielson, and I. L. Chuang. *Quantum Computation and Quantum Information.* New York, Cambridge University Press, 2010.

[4] S. Imre, and L. Gyongyosi. *Advanced Quantum Communications: An Engineering Approach.* New Jersey, Wiley-IEEE Press, 2012.

[5] A. Einstein, B. Podolsky, and N. Rosen. "Can quantum-mechanical description of physical reality be considered complete?" *Physical Review*, vol. 47, no. 10, pp. 777-780, 1935.

[6] C. E. Shannon. "A mathematical theory of communication." *The Bell System Technical Journal*, vol. 27, pp. 379-423, July 1948.

[7] R. W. Hamming. "Error detecting and error correcting codes." *Bell System Technical Journal*, vol. 29, no. 2, pp. 147-160, 1950.

[8] J. S. Bell. "On the einstein-podolsky-rosen paradox." *Physics*, vol. 1, pp. 195-200, 1964.

[9] C. H. Bennett, and D. P. DiVincenzo. "Quantum information and computation." *Physics Today*, vol. 48, no. 10, pp. 24-30, 1995.

[10] C. H. Bennett, and S. J. Wiesner. "Communication via one-and two-particle operators on Einstein-Podolsky-Rosen states." *Physical Review Letters,* vol. 69, no. 20, pp. 2881-2884, Nov. 1992.

[11] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters. "Purification of noisy entanglement and faithful teleportation via noisy channels." *Physical Review Letters*, vol. 76, no. 5, pp. 722-725, 1996.

[12] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wooters. "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels." *Physical Review Letters*, vol. 70, no. 13, pp. 1895-1902, 1993.

[13] B. Schumacher, and M. A. Nielsen. "Quantum data processing and error correction." *Physical Review A*, vol. 54, no. 4, pp. 2629-2635, Oct. 1996.

[14] K. Wen, and G. L. Long. "One-party Quantum Error Correcting Codes for Unbalanced Errors: Principles and Application to Quantum Dense Coding and Quantum Secure Direct Communications." *International Journal of Quantum Information*, vol. 8, pp. 697-719, 2010.

[15] Jun. W. J. “Superdense coding in noisy environments: A quantum trajectory approach.” *Physical Review A*, vol. 73, no. 6, pp. 64301, 2006.

[16] H. Ollivier, and W. H. Zurek. "Quantum discord: a measure of the quantumness of correlations." *Physical Review Letters*, vol. 88, no. 1, 017901, 2001.

[17] S. M. Barnett, S. J. D. Phoenix. "Entropy as a measure of quantum optical correlation." *Physical Review A*, vol. 40, no. 5, pp. 2404-2409, 1989.

[18] A. S. Holevo. "Bounds for the quantity of information transmitted by a quantum communication channel." *Problemy Peredachi Informatsii,* vol. 9, no. 3, pp. 3-11, 1973.

[19] S. Lloyd. "Capacity of the noisy quantum channel." *Physical Review A,* vol. 55, no. 3, pp. 1613-1622, 1997.

[20] C. King. "The capacity of the quantum depolarizing channel." *IEEE Transactions on Information Theory*, vol. 49, no. 1, pp. 221-229, 2003.

[21] C. H. Bennett, D. P. DiVincenzo, and J. A. Smolin. "Capacities of quantum erasure channels." *Physical Review Letters*, vol. 78, no. 16, pp. 3217-3220, 1997.

[22] A. S. Holevo. “The capacity of the quantum channel with general signal states,” *IEEE Transactions on Information Theory*, vol. 44, no. 1, pp. 269-273, Jan. 1998.

[23] C. H. Bennett, G. Brassard, S. Popescu, et. al. "Purification of noisy entanglement and faithful teleportation via noisy channels." *Physical Review Letters*, vol. 76, no. 5, pp. 722-725, 1996.

[24] H-J. Briegel, W. Dür, J. I. Cirac, et. al. "Quantum repeaters: The role of imperfect local operations in quantum communication." *Physical Review Letters*, vol. 81, no. 26, pp. 5932- 5939.

[25] B. Schumacher. "Quantum coding." *Physical Review A*, vol. 51, no.4, pp. 2738-2747, 1995.

[26] P. W. Shor. "Scheme for reducing decoherence in quantum computer memory." *Physical review A*, vol. 52, no. 4, pp. 2493-2496, 1995

[27] R. A. Calderbank, and P. W. Shor. "Good quantum error-correcting codes exist." *Physical Review A*, vol. 54, no. 2, pp. 1098-1105, 1996.

[28] D. P. DiVincenzo, P. W. Shor. "Fault-tolerant error correction with efficient quantum codes." *Physical Review Letters*, vol. 77, no. 15, pp. 3260-3263, 1996.

[29] A. M. Steane. “Error correcting codes in quantum theory”, *Physical Review Letters*, vol 77, no. 5, pp. 793-797, 1996.

[30] A. M. Steane. "Simple quantum error-correcting codes." *Physical Review A*, vol. 54, no. 6, pp. 4741-4751, 1996.

[31] E. Knill, and R. Laflamme. "Theory of quantum error-correcting codes." *Physical Review A*, vol. 55, no. 2, pp. 900-943, 1997.

[32] B. Criger, O. Moussa, and R. Laflamme. "Quantum error correction with mixed ancilla qubits." *Physical Review A*, vol. 85, no. 4, 2012.

[33] S. Loepp, and W. Wootters. *Protecting Information: From Classical Error Correction to Quantum Cryptography.* New York, Cambridge University Press, 2006.

[34] B. P. Lathi, and Z. Ding. *Modern Digital and Analog Communication Systems.* New York, Oxford University Press, 2010.

[35] W. C. Huffman, and V. Pless. *Fundamentals of Error Correcting Codes.* New York, Cambridge University Press, 2003.

[36] I. S. Reed, and G. Solomon. "Polynomial codes over certain finite fields." *Journal of the Society for Industrial & Applied Mathematics*, vol. 8, no. 2, pp. 300-304, 1960.

[37] A. J. Viterbi. "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm." *IEEE Transactions on Information Theory*, vol. 13, no. 2, pp. 260-269, 1967.

[38] G. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error correcting coding: Turbo codes," in *Proceedings of the IEEE International Conference of Communications*, pp. 1064-1070, May 1993.

[39] J. Hagenauer, E Offer, and L. Papke. "Iterative decoding of binary block and convolutional codes." *IEEE Transactions on Information Theory*, vol. 42, no. 2, pp. 429-445, Mar. 1996.

[40] C. Berrou, M. Jézéquel, C. Douillard, and S. Kerouédan, "The advantages of non-binary turbo codes." in *Proceedings of the IEEE Information Theory Workshop*, pp. 61 – 63, Sep. 2001.

[41] C. Douillard, and C. Berrou. "Turbo codes with rate-$m/(m+1)$ constituent convolutional codes." *IEEE Transactions on Communications*, vol. 53, no. 10, pp. 1630-1638, Oct. 2005.

[42] C. Bennett, P. Shor, J. Smolin, and A. Thapliyal. "Entanglement-assisted classical capacity of noisy quantum channels." *Physical Review Letters,* vol. 83, no. 15, pp. 3081-3084, Feb. 2008.

[43] B. Schumacher, M. D. Westmoreland. "Sending classical information via noisy quantum channels." *Physical Review A*, vol. 56, no. 1, pp. 131-139, July 1997.

[44] Z. Shadman, H. Kampermann, C. Macchiavello, and D. Bruß. "Optimal super dense coding over noisy quantum channels." *New Journal of Physics,* vol. 12, no. 7, 2010.

[45] M. Wilde, and G. Saikat. "Polar codes for classical-quantum channels." Information Theory, *IEEE Transactions on Information Theory*, vol. 59, no. 2, pp. 1175-1187, 2013.

[46] A. Steane. "A tutorial on quantum error correction." in *Proceedings of the International School of Physics "Enrico Fermi" course CLXII*, vol. 162, pp. 1-32, 2006.

[47] D. J. C. MacKay, and P. L. McFadden. "Sparse-graph codes for quantum error correction." *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2315-2330, 2004.

[48] Z. Babar, S. Ng, and L. Hanzo. "Near-capacity code design for entanglement-assisted classical communication over quantum depolarizing channels." *IEEE Transactions on Communications*, vol. 61, no. 12, pp. 4801-4807, Dec. 2013.

[49] Z. Shadman, H. Kampermann, and D. Bruß. "Distributed superdense coding over noisy channels." *Physical Review A*, vol. 85, no. 5, 2012.

[50] Z. Shadman, H. Kampermann, C. Macchiavello, and D. Bruß. "A review on super dense coding over covariant noisy channels." *Quantum Measurements and Quantum Metrology*, vol. 1, pp. 21-33, 2013.

[51] S. Bose, V. Vedral, and P. L. Knight. "Multiparticle generalization of entanglement swapping." *Physical Review A*, vol. 57, no. 2, pp. 822-830, 1998.

[52] J. I. Cirac, A. K. Ekert, S. F. Huelga, and C. Macchiavello. (1999). "Distributed quantum computation over noisy channels." *Physical Review A*, vol. 59, no. 6, pp. 4249-4255, 1999

Appendix A Fundamental Quantum Operations

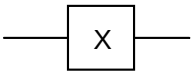
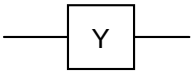
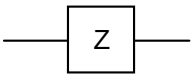
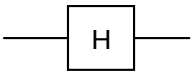
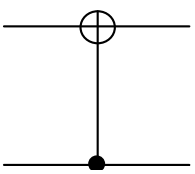
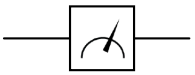
Gate	Operation	Rotation	Matrix Form	Quantum Circuit
I	$\|0\rangle \xrightarrow{I} \|0\rangle$ $\|1\rangle \xrightarrow{I} \|1\rangle$ $\|\psi\rangle \xrightarrow{I} \alpha\|0\rangle + \beta\|1\rangle$	$0°$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	
X	$\|0\rangle \xrightarrow{X} \|1\rangle$ $\|1\rangle \xrightarrow{X} \|0\rangle$ $\|\psi\rangle \xrightarrow{X} \alpha\|1\rangle + \beta\|0\rangle$	$180°_x$	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$	
Y	$Y\|0\rangle \xrightarrow{Y} i\|1\rangle$ $\|1\rangle \xrightarrow{Y} -i\|0\rangle$ $\|\psi\rangle \xrightarrow{Y} i\alpha\|1\rangle - i\beta\|0\rangle$	$180°_y$	$\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$	
Z	$\|0\rangle \xrightarrow{Z} \|0\rangle$ $\|1\rangle \xrightarrow{Z} -\|1\rangle$ $\|\psi\rangle \xrightarrow{Z} \alpha\|0\rangle - \beta\|1\rangle$	$180°_z$	$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	
H	$\|0\rangle \xrightarrow{H} \frac{\|0\rangle+\|1\rangle}{\sqrt{2}}$ $\|1\rangle \xrightarrow{H} \frac{\|0\rangle-\|1\rangle}{\sqrt{2}}$ $\|\psi\rangle \xrightarrow{H} \frac{\alpha+\beta}{\sqrt{2}}\|0\rangle + \frac{\alpha-\beta}{\sqrt{2}}\|1\rangle$		$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$	
CNOT	$\|00\rangle \xrightarrow{\text{CNOT}} \|00\rangle$ $\|01\rangle \xrightarrow{\text{CNOT}} \|01\rangle$ $\|10\rangle \xrightarrow{\text{CNOT}} \|11\rangle$ $\|11\rangle \xrightarrow{\text{CNOT}} \|10\rangle$ $\|\xi\rangle \xrightarrow{\text{CNOT}} \alpha\|00\rangle + \beta\|01\rangle$ $+\gamma\|11\rangle + \delta\|10\rangle$		$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$	
M				

Table A1: The fundamental quantum operations and their Bloch, matrix and quantum circuit representations.