

RECEIVER DESIGN FOR QUANTUM COMMUNICATION

Ph.D. Candidate : *Nicola Dalla Pozza*

Supervisor: Nicola Laurenti

Scuola di Dottorato in Ingegneria dell'Informazione

Indirizzo Scienza e Tecnologia dell'Informazione

Ciclo XXV°

Anno Accademico 2012-2013

Abstract

Born about a century ago, Quantum Mechanics has revolutionized the description and the interpretation of Physics at sub-microscopic level. In the last decades, due to the influence of mathematical and engineering research fields, Quantum Mechanics has given birth to related research areas like Quantum Computation, Quantum Information and Quantum Communication.

With the discovery of the laser, and later the development of fiber optics and satellite networks, Quantum Communication and Quantum Optics seems to have a natural field of application in Communication Systems. Despite this, the interest in this technology and studies for communication purpose has been overshadowed by the great results in communication networks achieved in the last decades with classical paradigms. However, due to the increasing demand of communication data rate, system designers are now looking at Quantum Mechanics for new and more performing solutions in communication purposes.

Early theoretical studies on Quantum Discrimination Theory and Quantum Information predict better performance for Communication Systems that take advantage of the quantum laws.

In addition, Quantum Mechanics provides the deepest description of the physical phenomena, and there are scenarios where a quantum model fits best, as in in deep space communications, where the received signal is really weak, or in a satellite networks, where we are interested in strongly reducing the power of transmitted signals, possibly without sacrificing performance significantly.

However, if on one side Quantum Communication Theory promises great gains in the performance of communication systems, on the other hand it fails to describe how to implement physical devices that reach these ultimate limits. So far, only a few architectures achieving these performances are known, and only for simple modulation formats. We are interested in the scenario of optical communications, where the message transmitted is encoded in a sequence of coherent states. Transmitter devices for coherent modulation are well known and consist in laser pulse generators. Instead, receiver implementations working at the quantum

limit performance limit are yet to be found.

In this Thesis I deal with different topics in the quantum transmission scenario.

First, I review existing classical (suboptimal) and quantum (suboptimal and optimal) receiver schemes for the binary coherent modulation. I present a new formulation of the optimal scheme known as *Dolinar Receiver* with the multiple copies problem, focusing on the information gained during the measurement.

Second, I analyze the binary communication in a noisy environment, studying the error probability and the capacity of the binary channel induced. Given the description of the quantum channel, I optimize both the transmitted quantum states and the measurement operators employed in the communication.

Third, I consider the Pulse Position Modulation, that is particularly suitable for space and satellite communication due to its simplicity of implementation and high capacity. I review some known suboptimal receivers, and I propose a receiver scheme which approaches the limit performance predicted with quantum theory outperforming the existing schemes.

To sum up the results of this Thesis, in order to approach the limit performance predicted by Quantum Mechanics, an *optimization* is always necessary to exceed the classical effects and trigger the quantum phenomena. In particular, the *information* gained during the measurement plays an important role, for example in the definition of adaptive receivers. In this Thesis both these aspects have been deeply investigated.

Sommario

Formalizzata più di un secolo fa, la Meccanica Quantistica ha rivoluzionato la descrizione e l'interpretazione della Fisica a livello microscopico. Negli ultimi decenni, grazie all'influenza di studi affini nei campi della matematica e dell'ingegneria, la Meccanica Quantistica ha portato allo sviluppo di aree di ricerca quali la Computazione Quantistica, la teoria dell'Informazione Quantistica e le Comunicazioni Quantistiche.

Con l'invenzione del laser, e i successivi sviluppi delle fibre ottiche e delle reti satellitari, la comunicazione quantistica e l'ottica quantistica hanno un naturale campo di applicazione nello studio nei sistemi di comunicazione. Nonostante ciò, l'interesse in questa tecnologia e gli studi quantistici sulle telecomunicazioni sono stati messi in ombra dai risultati nelle reti di comunicazione ottenuti negli ultimi decenni con paradigmi classici. Solo recentemente, a causa dell'aumento della richiesta di rate trasmissivo, i progettisti di sistemi di comunicazione guardano alla meccanica quantistica in cerca di soluzioni nuove e più efficienti.

I primi studi teorici nella teoria quantistica della discriminazione e dell'informazione prevedono un notevole vantaggio nelle prestazioni se i sistemi di comunicazione sono progettati secondo le leggi della meccanica quantistica.

Inoltre, la meccanica quantistica fornisce la più profonda descrizione dei fenomeni quantistici, e in alcuni scenario tale descrizione è più appropriata, come nel caso di comunicazioni dallo spazio profondo, dove il segnale ricevuto è estremamente debole, o nelle reti satellitari, dove siamo interessati a ridurre la potenza trasmessa con il segnale, senza sacrificare significativamente le prestazioni.

Se da un lato le comunicazioni quantistiche promettono grandi guadagni in termini di performance, dall'altro lato non spiegano esplicitamente come costruire dispositivi che raggiungono questi limiti. Finora, solo pochi schemi di comunicazione che raggiungono questo limite sono conosciuti, e solo per formati di modulazione semplici. Lo scenario di nostro interesse è quello delle trasmissioni ottiche, dove un messaggio trasmesso viene codificato in una sequenza di stati coerenti. Dispositivi di trasmissione per la modulazione coerente sono noti (gen-

eratori laser), mentre ricevitori che lavorano nel regime quantistico sono ancora da sviluppare.

In questo lavoro di Tesi sviluppo diversi temi nello scenario delle comunicazioni quantistiche.

Inizialmente, riassumo gli schemi di ricezione classici (subottimi) e quantistici (ottimi e subottimi) per la modulazione binaria coerente. Successivamente presento una riformulazione dello schema ottimo noto come il *ricevitore di Dolinar* come un problema di copie multiple, focalizzandomi sull'informazione guadagnata durante l'operazione di misura.

Successivamente, analizzo la comunicazione binaria in un ambiente rumoroso, studiando la probabilità di errore e la capacità del canale binario che si possono ottenere. Data una descrizione quantistica del canale, ottimizzo rispetto sia gli stati trasmessi che gli operatori di misura impiegati nella comunicazione.

In seguito considero una modulazione più complessa, la *Pulse Position Modulation*, particolarmente adatta per le comunicazioni dallo spazio e satellitari, grazie alla semplicità di implementazione e all'alta capacità. In primo luogo rivedo alcuni ricevitori subottimi, e successivamente propongo uno schema di ricezione che approssima le prestazioni limite predette con la teoria quantistica, superando gli schemi esistenti in letteratura.

Riassumendo i risultati della Tesi, per approssimare le prestazioni ottime predette dalla meccanica quantistica un procedimento di ottimizzazione è sempre necessario per superare gli effetti classici e innescare i fenomeni quantistici. In particolare, l'informazione guadagnata durante il procedimento di misura gioca un ruolo fondamentale, ad esempio nella definizione di ricevitori adattativi. In questo lavoro di Tesi entrambi questi aspetti sono stati investigati a fondo.

Table of contents

1	Introduction	11
1.1	Classical vs quantum communication paradigm	12
1.2	Deep Space scenario	14
1.3	Capacity improvements by Quantum Mechanics	16
1.4	Communication System Scenario	20
1.4.1	Classical Communication over Quantum Channel	22
1.5	Capacity of a Quantum Channel	27
1.6	Summary of the Results	31
2	Binary Quantum Receivers	33
2.1	Introduction	33
2.1.1	Binary modulation implementations	36
2.2	Kennedy Receiver	37
2.3	Improved Kennedy Receiver	39
2.4	Multiple Copy State Discrimination	41
2.5	Revisiting the Dolinar receiver	46
2.6	A simple suboptimal receiver	50
3	Optimal Encoding and Decoding over Noisy Qubit Channels	55
3.1	Introduction	55
3.1.1	Error Probability	56
3.1.2	Capacity	58
3.2	Partial orderings for classical binary channels	60

3.3	Coherence Vector Representation and Geometric Picture	65
3.3.1	Optimization of input states for given projectors	67
3.3.2	Region of achievable transition probabilities	69
3.4	Optimization and Numerical Methods	72
3.4.1	Probability of correct decision	72
3.4.2	Capacity	74
3.5	Examples	76
Appendices		81
3.A	Proof of the necessary condition for optimality	81
3.B	Quadratic Optimization Problems with Quadratic Constraints . .	82
4	Quantum Receivers for Pulse Position Modulation	87
4.1	Introduction to Pulse Position Modulation	87
4.2	Optimal Performance	90
4.2.1	Geometrically Uniform Symmetry	90
4.2.2	Least Square Measurements	92
4.2.3	Least Square Measurements and Geometric Uniform Symmetry	97
4.2.4	Theoretical Limit Performances of PPM	99
4.3	Existing Receiver Schemes for PPM	100
4.3.1	Classical Receivers	101
4.3.2	Unconditional Nulling Receiver	103
4.3.3	Conditional Nulling Receiver	106
4.3.4	Improved Conditional Nulling Receiver	110
4.4	Qubit Framework Representation	111
4.5	An Adaptive Receiver for PPM	113
4.5.1	State of the Algorithm	116
4.5.2	Useful Lemmas	119
4.5.3	Expected reward function	122
4.5.4	Dynamic Programming Algorithm	126
4.6	Results and Numerical Issues	129

<i>TABLE OF CONTENTS</i>	9
4.7 Conclusion	133
Appendices	137
4.A Review of Dynamic Programming	137
5 Conclusions	143

Chapter 1

Introduction

The extraordinary progress we have witnessed in the last decades in mobile communications, wireless technologies and internet networking is just the last fraction of a longer and increasing interest in a broader area of studies gathered together under the name of *Information and Communication Technology*.

These research fields are founded on the studies of electromagnetic field theory in the second half of the XIXth century, trying to give a description of the physical phenomena involved in the propagation of electromagnetic waves. More than fifty years later, the invention of the laser and the deployment of cable, fiber, and free space networks first, the advance in coding and decoding techniques later, the development of the internet infrastructure and protocols up to the current diffusion of mobile technology have provided new scenarios for the communication studies, giving birth to many related research areas such as signal processing, information, coding and system theory.

In the search for new and better-performing solutions to the challenges imposed by the communication scenarios, the scientific community is now looking for answers in research fields outside the *classical* physic, and Quantum Mechanics seems a promising field to be investigated in order to find new and such solutions.

Formalized and experimentally tested in the first decades of the previous century, Quantum Mechanics has established itself as the branch of Physics that provides the deepest description of the phenomena at sub-microscopic level. Later, in

conjunction with other research fields in mathematics and engineering, Quantum Mechanics has given birth to related research areas such as Quantum Computation, Quantum Information and Quantum Communication.

The initial investigations on the use of Quantum Mechanics for communication purposes have pointed out the possibility to achieve better performances if the communication system is designed exploiting quantum laws. Despite this, the interest in this studies and technology for communication purpose was overshadowed by the great results in communication networks achieved in the last decades with classical paradigms.

However, quantum effects are now being taken into account more and more often in the electronic hardware design and in deep space and satellite communication scenarios, so that communication engineers will soon have to face, but also to take advantage, of the laws of Quantum Physics. In my opinion, we can expect for the next decades a great technological development due to Quantum Mechanics as it has been in the previous decades for the study on the Electromagnetic Field Theory formulated more than a century ago.

In the next Sections I explain the need for a quantum approach to the design of communication systems, explaining some motivation for the shift from a classical to a quantum paradigm.

1.1 Main differences between the classical and quantum communication design paradigm

Quantum Physics provides the deepest description of the physical phenomena involved in Communication Systems. With the discovery of the laser, and later the development of fiber optics and satellite networks, Quantum Communication and Quantum Optics seem to have a natural field of application for transmission, propagation and detection tasks of the communication systems.

However, the deeper description provided by Quantum Physics introduces

profound differences with respect to the classical paradigm [1], which at first can appear as additional complications, but at the same time can be exploited at our own advantage.

Firstly, a measurement of a quantum system can be ambiguous even if there is no noise introduced by the environment. In the language of communication, the measurement at the receiver can be ambiguous even if the channel is noiseless. Given a quantum state, in general the outcome of a measurement is probabilistic, making it more difficult to discriminate among the transmitted quantum states. Although even in the classical case, sending non orthogonal signals through a noiseless channel prevents the receiver from perfectly discriminating among them, in the quantum paradigm randomness of the measurement outcomes is an intrinsic property that comes together with the definition of quantum state.

Secondly, the classical model adopted for the transmitted signal is not adequate to describe a whole set of effects well explained in Quantum Mechanics [2]. For example, the classical wave description for the electromagnetic field cannot describe accurately the uncertainty in the quadrature that we can observe and measure for low intensity fields. In this case, a better description is given by the model of a quantized field, leading to the notions of *annihilator* and *creation operators* to replace the description given by the classical phasor [3]. The non-commutativity of these operators explains the uncertainty in the quadrature measurements, bearing the possibility of errors in the discrimination between the coherent states, which are the quantum model for the laser pulse, even if the channel is noiseless.

Thirdly, one might suggest to perform multiple measurements on the quantum state to reduce the uncertainty in the discrimination. However, the act of measuring a quantum state fundamentally changes the properties of the system, such that successive measurement outcomes are biased by previous ones, and after a measurement the quantum state is different from the received one. In addition, simultaneous measurement has limitations in the uncertainty of the outcomes (Heisenberg uncertainty principle [4, 5]). Moreover, some measurements are dis-

tructive, and the quantum state is annihilated after the measurement.

Last, in order to overcome such limitation in the measurement, one could think to copy the receiver quantum state and then perform measurements in each copy. However, Quantum Mechanics prevents such copy of the quantum information, according to the so called No Cloning Theorem [6, 7].

Despite these complications, Quantum Mechanics predicts new and more efficient solutions for the communication scenario. In fact, in some cases these features of Quantum Physics are the base for new protocols and setups, as in the case of Quantum Key Distribution [8], where the impossibility to copy quantum states and the ambiguity that comes from the measurements are exploited to design secure communication links.

1.2 Deep Space scenario

In some communication scenario Quantum Mechanics is particularly suitable for the description of the signal transmission, propagation and detection. This is the case of Deep Space Communications.

Deep Space Communications are extremely challenging [9]. The spacecraft travels at billions of kilometers from the Earth, collecting data and sending them back in direction of the Earth. Due to the very long distance of propagation, the signal beam spreads due to the diffraction, and only a really attenuated signal reaches the receiver base station.

Due to its spreading, the beam intensity decreases as the square of the distance between transmitter and receiver, as the distance increases the problem becomes quadratically more difficult. For example, a geostationary satellite flies at an altitude of about 40,000 km. With the current technology, it is possible to achieve a communication link with a capacity of the order of Gigabits per second. If we consider the same system mounted on a spacecraft approaching Neptune or Pluto orbits, where the distance is of the order of 4,000,000,000 km, the beam spread would be 10 billions much higher. Assuming that we operate in the low SNR

regime, the consequent reduction of the beam power at the receiver reduces the capacity to 1 bit per second.

To overcome these difficulties, improvements at both transmitter and receiver side have been developed, such as more powerful transmitter devices and more sensitive receiver architectures. However, further increases are hard to accommodate. Bigger transmitting antennas are difficult to squeeze into launch shrouds, and increasing transmitted power is challenging due to the difficulties to generate electrical power from solar energy at such distance from the sun and to remove the excess heat from the device. At the receiver station, antennas are already enormous (34 m and 70 m in diameter for the NASA Deep Space Network) and operating at few degrees above the absolute zero to implement low-noise amplifier.

A turning point for deep space communication is the use of much higher frequency in the EM spectrum, such as those of optical signals. Currently, the carrier frequencies of the communication signal belongs to the X band (8 GHz). Although a great improvement has been obtained employing the Ka band (32 GHz), a greater one is expected considering the use of optical frequencies (about 300,000 GHz).

In Figure 1.2 is depicted the comparison of a radio and optical communication system sending data from the Saturn orbit. The left side of the Figure shows the transmitted beam spread from the Voyager spacecraft in direction of the Earth, where a communication system working in the X band is employed. The signal transmitted from a 3.7 m antennas, by the time it reaches the Earth it spreads out on an area 1000 Earth-diameter wide. On the contrary, the 10 cm optical telescope shown on the right side of the Figure, assuming it transmits an optical wavelength of $1\mu\text{m}$ (that corresponds to a frequency of 300,000 GHz) causes the beam to spread to a surface comparable with the Earth section. This represents an increase in the power density at the receiver of 10^6 times, achieved with a much smaller antenna, for a theoretically wavelength-squared advantage of about

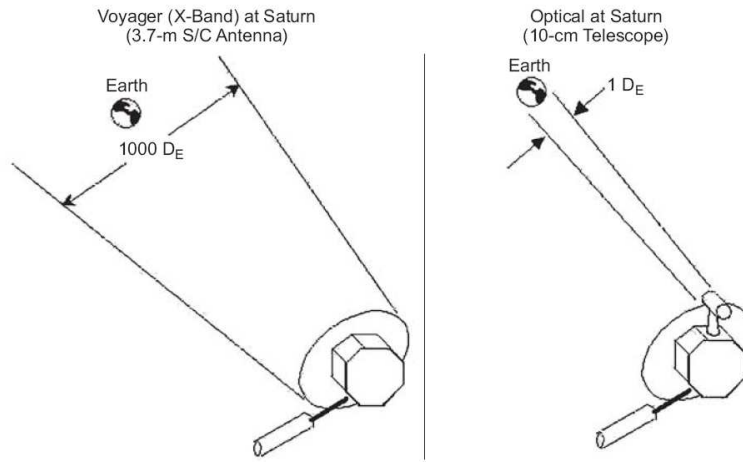


Figure 1.1: Comparison between radio and optical communication for Deep Space Scenario. On the left side, the radio frequencies cause the signal to spread more and require a bigger antenna, while in the case of optical frequencies on the right side the beam undergoes a lower spreading. Courtesy of “*Deep Space Optical Communications*”, Hamid Hemmati, Wiley, 2006.

90 dB.

In such a scenario the received signal is so weak and attenuated that can be considered as composed by a few photons. Therefore, a quantum description is necessary, in order to exploit the physical phenomena at the level of individual quanta. As we will see in the next sections, we can take advantage of this new paradigm to improve the performances of a communication system.

1.3 Capacity improvements by Quantum Mechanics

Information theory was originally formulated by Shannon [10] to seek the ultimate limit of the amount of information that can be reliably transmitted in a noisy communication link. This limit, which is referred to as the *capacity* of the channel,

was proved in Shannon's channel coding theorem to be equal to the maximum amount of mutual information between the input and the output of the channel.

The channel coding theorem has been applied to a variety of noisy communication links, where the disturbances were described at different levels of abstraction, ranging from the binary channel to the additive white gaussian noise. In all these descriptions, the underlying paradigm is classical. Indeed, when classical light source is employed at the transmitter and the receiver uses standard measurement apparatus such as homodyne, heterodyne or direct detection, a semiclassical description of the electromagnetic field suffices.

However, in optical communication it is the quantum noise that sets the fundamental limits, and the standard structural assumptions preclude the optimization on nonclassical light sources and nonstandard measurement.

As a study case consider the lossy channel [11]. The action of this channel is described in terms of input-output relation between the annihilator operator at the input and output mode, \hat{a} and \hat{a}' respectively. The Trace Preserving Completely Positive (TPCP) map associated with the lossy channel is

$$\hat{a}' = \sqrt{\eta} \hat{a} + \sqrt{1-\eta} \hat{b}, \quad (1.1)$$

with \hat{b} the annihilator operator associated with the noisy environment. In the case of $\eta = 1$, we have a lossless channel, otherwise for $\eta < 1$ we have a *pure-loss* channel if the environment mode is the vacuum.

It has been shown [12] that the capacity of a lossless channel, under a constraint on the transmitter's average number of sent photons N , is additive, and it can be achieved using *number states* with a Bose-Einstein statistic, leading to

$$C_{\text{lossless}}(N) = (N+1) \ln(N+1) - N \ln(N). \quad (1.2)$$

Later, it has been proved [11] that also the pure-loss channel is additive, with capacity

$$C_{\text{pure-loss}}(\eta N) = (\eta N + 1) \ln(\eta N + 1) - \eta N \ln(\eta N), \quad \text{for } 0 \leq \eta \leq 1 \quad (1.3)$$

that can be achieved with coherent states with a Gaussian prior distribution

$$p(\alpha) = \frac{e^{-|\alpha|^2/N}}{\pi N}. \quad (1.4)$$

Since (1.3) includes (1.2) as a particular case with $\eta = 1$, it is shown that the capacity of a lossless channel can be achieved both with nonclassical number state and classical coherent states. However, in the case of photon number states it suffices to use photon counting at the receiver¹ while in the case of random coding over coherent states a receiver structure achieving the capacity is still unknown.

On the contrary, an homodyne and heterodyne detection scheme on a coherent state encoding gives the following capacity [13],

$$C_{heterodyne} = \frac{\ln(1 + 4\eta N)}{2}, \quad C_{homodyne} = \ln(1 + \eta N). \quad (1.5)$$

Figure 1.2 plots the capacity of pure-loss channel and the heterodyne and homodyne limit. It has been shown that heterodyne detection is asymptotically optimum,

$$\lim_{\eta N \rightarrow \infty} \frac{C_{heterodyne}}{C_{pure-loss}} = 1. \quad (1.6)$$

although in the free space and fiber propagation regime the typical situation is $\eta \ll 1$, such that to reach the asymptotic regime it is required a very high average number of transmitted photons.

The study case of the capacity for the lossy channel indicates that when the quantum paradigm is employed, better performances can be achieved with respect to the solution proposed by classical paradigm. Other examples can be found for the problem of discrimination among signals with minimal error probability, and Chapter 4 is dedicated to this issue when Pulse Position Modulation is employed.

¹Since the channel is lossless, a transmitted number state is not disturbed by the channel and perfectly detected by a photon counter at the receiver.

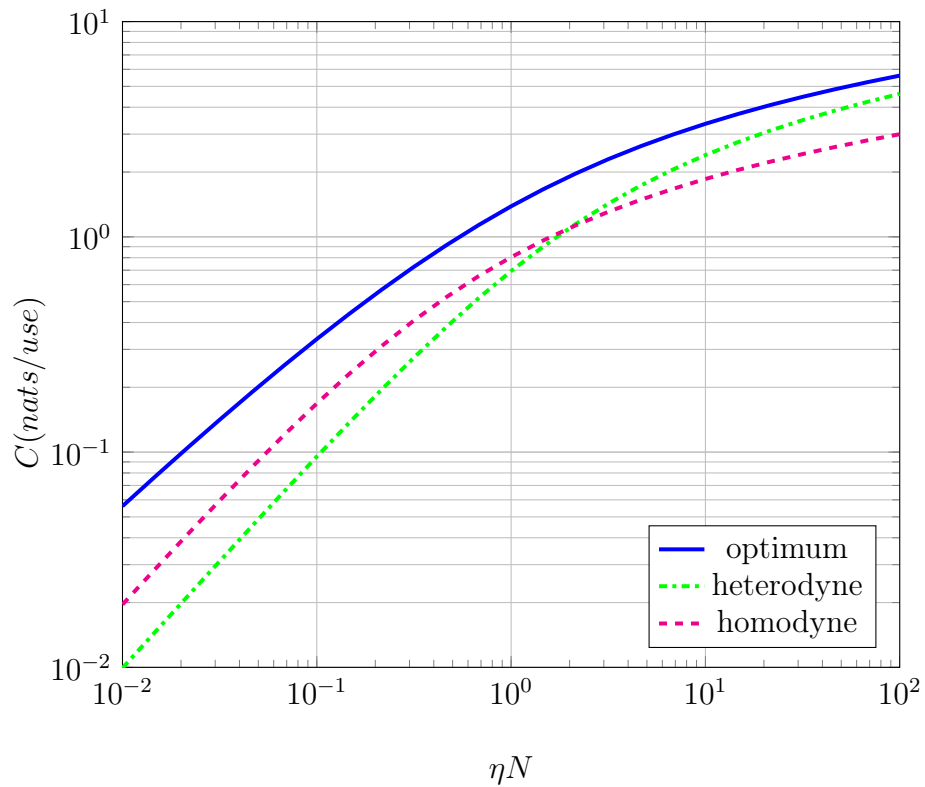


Figure 1.2: Capacity of the lossy bosonic channel as a function of the average received photons ηN . The solid blue line is the optimum capacity $C_{\text{pure-loss}}$ from eq. (1.3). The green dash-dotted line is the capacity $C_{\text{heterodyne}}$ achievable with coherent states and heterodyne detection, and the magenta dashed line is the capacity C_{homodyne} achievable employing coherent states and homodyne detection, evaluated from eq. (1.5).

1.4 Communication System Scenario

A Communication System is a collection of devices built together with the purpose of conveying information between its users, i.e., to move some kind of knowledge from a sender to a receiver.

In order to reach a user, the information needs a carrier, that is a physical observable that allows the sender to encode the message and the receiver to extract it. Depending on the physical phenomena involved in the process of transfer the observable from transmitter to receiver, we can discriminate *classical* and *quantum* channels. In the former case a classical description of the process suffices, while in the latter we need to use a quantum description.

This work is focused on *classical* communication over quantum channel, in the sense that *bits* of information are conveyed over a channel that has a quantum description. In classical communication system, the information is encoded in a message that can be regarded as a stream of symbols. It was Shannon that in 1948 [10] introduced formal notion of information. He was the first to understand that in order to quantify the information, we do not have to look at the meaning of the message. Instead, it is the *probability* of the realization of the message that is associated with an amount of information, and we can quantify the information produced by a stochastic process from its probability distribution.

The Communication System is usually represented as a sequence of devices that goes from one user, the transmitter Alice, i.e. the source of information, to the other user, the receiver, called Bob (see Figure 1.3).

On the transmitter side, a sequence of devices convert the information in a suitable way that can be sent through the media interposed between the two users.

The realization m of the random process that describe the message is converted into a sequence of bits b by an encoder. Subsequently, the sequence of bits is mapped in a sequence of signals $s(t)$ that are sent through the channel. Usually, in order to transfer the message over long distances, the signal is an

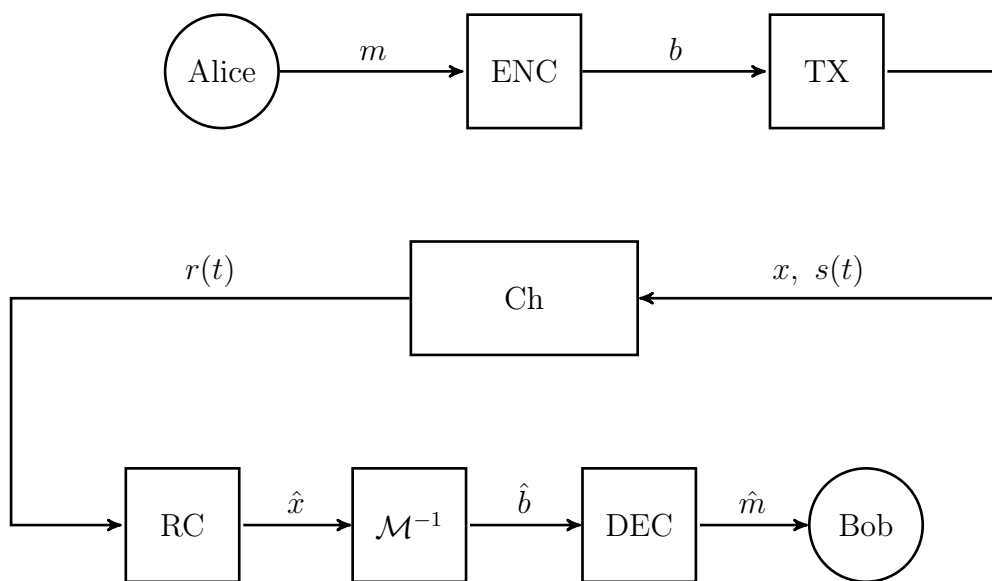


Figure 1.3: Block diagram of a communication system. The blocks corresponding to the user Alice, the encoder ENC, the transmitter map TX compose the transmitter side, while the blocks of the front-end receiver RC, the inverse map \mathcal{M}^{-1} , the decoder DEC and the user Bob compose the receiver side. Interposed between the two parts, the channel.

electromagnetic field that propagates in the media interposed between the sender and the receiver. The type of signals to be sent are chosen depending on the *modulation format*, i.e. a protocol that transmitter and receiver agree to follow. Usually, the possible signals can be chosen in a finite set of cardinality M called constellation, so that we can associate each signal with a numeric symbol x in the set $\{1, \dots, M\}$. Since there is a one-to-one correspondence between signals and symbols, we can refer indifferently to the former or the latter.

The signal travelling through the media is usually distorted and modified. The channel is a model of the interaction between the signal and the media, comprehensive of the disturbances introduced by the environment. In the case of an ideal channel, the travelling signal reaches the receiver side unmodified.

Out of the channel, at the receiver side, a device tries to estimate the transmitted sequence of symbols from the output $r(t)$ of the channel. Then, an inverse map converts the sequence of estimated symbols \hat{x} in the sequence of bits \hat{b} , and a decoder returns to the final user an information in the same form of the original process.

1.4.1 Classical Communication over Quantum Channel

In Figure 1.3, the communication system has been pictured as a cascade of logic blocks, each one representing a device. The decomposition in simple blocks is a common paradigm in engineering, that allows to focus only on the design and optimization of few parts of the communication system at a time.

In fact, in the description of a transmission system, we can ignore (at the transmitter side) the random process with values m , the encoder for the bit stream b and even the transmitter: a description of the possible signals $s(t)$, or the corresponding symbols x , with the probability distribution calculated from the chain source-encoder-transmitter suffices, and we can replace the chain with a single source, as in Figure 1.4.

For the same reason, at the receiver side, we can put together in a single block the chain \mathcal{M}^{-1} -decoder-Bob.

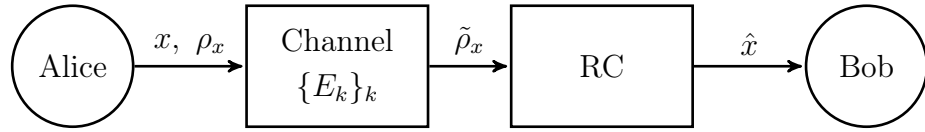


Figure 1.4: Simplified block diagram for a communication system with a quantum description of the channel, given by eq. (1.11). The transmitter Alice sends a symbol x associated with the density operator ρ_x through the channel, which modifies the density operator in $\tilde{\rho}_x$. From this quantum state, the receiver estimates the transmitted symbol, returning the estimation \hat{x} to the final user, Bob.

In Figure 1.4 we resume the scheme of Figure 1.3 with the simplification mentioned above.

The model that represents a logic block can have more or less fine details, and in general different assumptions lead to use different models. For example, in the case of the channel, great development has been achieved with a classical model of physical phenomena, but as previously seen in Sections 1.2,1.3 in some regimes a quantum model fits best and allows to improve performance.

In this case, it seems appropriate to talk about *Classical Communication* over a *Quantum Channel*. As a consequence, the transmitted signal $s(t)$ and the received signal $r(t)$ must be described according to the quantum model of the channel.

Denoted with \mathcal{H} the quantum system that describes the transmitted and received signals, and with $\mathcal{L}(\mathcal{H})$ the set of linear operators on \mathcal{H} , the transmitter associates to each symbol $x \in \{1, \dots, i, \dots, M\}$ a quantum state described by a density matrix in the set

$$\rho_x \in \{\rho_1, \rho_2, \dots, \rho_M\} \subset \mathcal{D}(\mathcal{H}), \quad (1.7)$$

where $\mathcal{D}(\mathcal{H})$ denotes the set of all the density matrices of \mathcal{H} ,

$$\mathcal{D}(\mathcal{H}) = \{\rho \in \mathcal{L}(\mathcal{H}) | \rho = \rho^\dagger \geq 0, \text{tr}(\rho) = 1\}. \quad (1.8)$$

Sometimes, e.g. in the binary case, it is customary to enumerate the possible values for the symbols x starting from 0.

In the case of pure quantum states, the density matrix ρ_x reduces to

$$\rho_x = |\gamma_x\rangle\langle\gamma_x|. \quad (1.9)$$

The symbols, and hence the quantum states, are drawn with a priori probability

$$\{p_1, p_2, \dots, p_M\}. \quad (1.10)$$

The channel, usually an optical fiber or the free space in telecommunication system, modifies the transmitted density operator ρ_x , and the output density operator $\tilde{\rho}_x \in \mathcal{D}(\mathcal{H})$ can be described by a linear map $\mathcal{E} : \mathcal{D}(\mathcal{H}) \rightarrow \mathcal{D}(\mathcal{H})$ that acts as in

$$\tilde{\rho}_x = \sum_k E_k^\dagger \rho_x E_k, \quad (1.11)$$

where the operators $\{E_k\}$ are called *Kraus operators* and (1.11) is called *Kraus representation*. In order for the map represented by the Kraus operator $\{E_k\}$ to be stochastic, that is, Completely Positive and Trace Preserving (CPTP), it is required that

$$\sum_k E_k E_k^\dagger = I_{\mathcal{H}}. \quad (1.12)$$

In some cases, the map is also unital, that is it verifies $\Phi(I) = I$, if it holds

$$\sum_k E_k^\dagger E_k = I_{\mathcal{H}}. \quad (1.13)$$

The receiver, given the density operator $\tilde{\rho}_x$, performs a measurement on the Hilbert space. The measurement is defined by a set of operators $\{P_k\}$ which in general are *Positive Operator Valued Measurement* (POVM), i.e. a set of positive Hermitian operator

$$P_k^\dagger = P_k, \quad P_k \geq 0 \quad (1.14)$$

that sum up to the identity

$$\sum_k P_k = I_{\mathcal{H}}. \quad (1.15)$$

The constraint (1.15) is called *relation of completeness*.

As a particular case, the POVM $\{P_k\}$ can be projectors $\{\Pi_k\}$ on some subspace of the Hilbert space \mathcal{H} , such that they also verify

$$\Pi_j \Pi_k = \delta_{j,k} \Pi_j, \quad \forall j, k. \quad (1.16)$$

The measurement operators $\{P_k\}$ are defined on the Hilbert Space \mathcal{H} . Each P_k is associated with an outcome of the measurement. As already anticipated, the receiver tries to estimate the transmitted quantum state ρ_x and hence the symbol x . In general the cardinality of the outcome set can be greater than the cardinality of the symbol set. The receiver sets a rule for the association between the outcome and the estimate quantum state, with the meaning

“When I measure the outcome of P_k , I estimate that ρ_k has been transmitted.” (1.17)

From the mathematical point of view, the estimate is a new random variable, \hat{x} , with values

$$\hat{x} \in \{1, \dots, M\}. \quad (1.18)$$

that is with the same alphabet as x (see Figure 1.4). Sometimes, we will use the symbols $\{\hat{1}, \dots, \hat{M}\}$ as a shorthand notation to indicate the realization of the measurement.

Following the block-paradigm described above, we can describe the sequence channel-receiver in a new block that describes with transition probabilities the transformation between the random variables x and \hat{x} . From the point of view of the measurement design, this means that we can group all the $\{P_k\}_k$ associated with the same estimate $\hat{\rho}_{\hat{x}}$, and define a new POVM \tilde{P}_k . For this reason, I will consider without loss of generality a set of $\{P_k\}_k$ with cardinality M .

From the definition of the POVM, we can calculate the probabilities of the outcomes given that the state ρ_i has been sent (Born Rule)

$$\begin{aligned} P[\hat{x} = j | x = k] &= P[\text{outcome is } j \text{ of } P_j | \rho_k \text{ has been transmitted}] \\ &= \text{tr}(P_j \rho_k) \end{aligned} \quad (1.19)$$

(if ρ_k is a pure state) $= \langle \gamma_k | P_j | \gamma_k \rangle$.

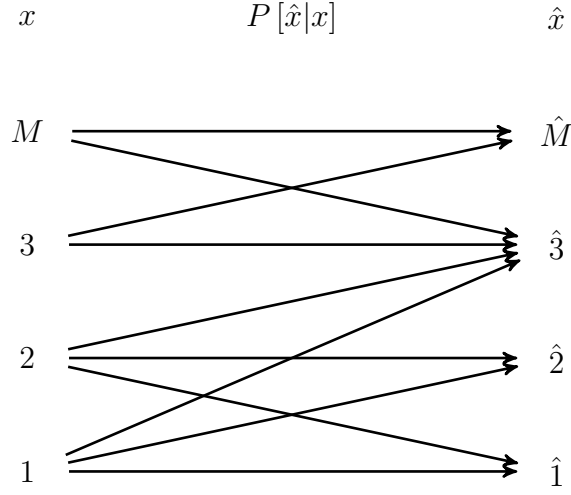


Figure 1.5: Representation of the channel between symbols x and outcomes \hat{x} , described through arrows that indicate the positive transition probabilities $P[\hat{x}|x]$.

In order to simplify the equation, sometimes we will use the following shorthand notation

$$\begin{aligned} p_{\hat{j}|k} &= P[\hat{x} = \hat{j} | x = k], & p_{j,k} &= P[\hat{x} = \hat{j}, x = k], \\ p_k &= P[x = k], & p_{\hat{j}} &= P[\hat{x} = \hat{j}], \end{aligned} \quad (1.20)$$

where the vertical bar or the comma in the subscript indicates if we are referring to a conditional or joint probability, respectively. These transition probabilities define the channel between the values of the random variable x and \hat{x} , as in Figure 1.5.

In order to evaluate the system, a figure of merit that can be considered is the *probability of correct decision*

$$P_c = P[\hat{x} = x] = \sum_{k=1}^M P[\hat{x} = k, x = k] = \sum_{k=1}^M P[\hat{x} = k | x = k] p_k, \quad (1.21)$$

that can be rewritten as

$$P_c = \sum_{k=1}^M \text{tr}(P_k \rho_k) p_k = \sum_{k=1}^M \text{tr}(P_k \rho_k p_k). \quad (1.22)$$

Sometimes it is more useful to define the complementary probability, the *error probability*

$$P_e = 1 - P_c = \sum_i \sum_{k \neq i} P[\hat{x} = k | x = i] p_i = \sum_i \sum_{k \neq i} \text{tr}(P_k \rho_i) p_i \quad (1.23)$$

When the correct decision probability is assumed as a performance figure to evaluate the Communication System, the problem statement is to find the set of POVM, satisfying the constraints of Hermiticity and semi-definite positiveness (1.14) and the completeness relation (1.15), that maximizes P_c .

1.5 Capacity of a Quantum Channel

In the design of a Communication System, particular attention must be paid to the type of information that we want to convey from the sender to the receiver. The information transmitted could be of two kind,

1. classical information, when we aim at transmitting *bits*,
2. quantum information, when we aim at transmitting *qubit*.

Of course, the choice of which information is delivered depends on the application for the two users. In both cases, the communication system must be devised to counteract the noise and the effect of the quantum channel interposed between sender and receiver.

A channel absolute parameter for the limit performances for the realible information delivered is given by the concept of *capacity*. As we recall from Section 1.3, the capacity of a classical channel is obtained from the mutual information between Alice's transmitted symbols x and Bob's estimated symbols \hat{x} ,

$$H(x; \hat{x}) = \sum_{x, \hat{x}} p_{\hat{x}|x} p_x \left(\log p_{\hat{x}|x} - \log \sum_{x'} p_{\hat{x}|x'} p_{x'} \right) \quad (1.24)$$

which depends on both the transition probabilities $p_{\hat{x}|x}$ induced by the channel and the a priori probabilities p_x . The Capacity of the channel described with

a given transition probabilities is defined as the maximum of (1.24) over the a priori distribution of the symbols,

$$\mathcal{C}_{Shan} = \sup_{\{p_x\}} H(x; \hat{x}). \quad (1.25)$$

The operational meaning of the capacity comes from the Shannon's noisy channel coding theorem [10]. The theorem says that in the presence of a channel with capacity \mathcal{C}_{Shan} , and with many (actually, in the limit of infinite) uses of the channel, there exist a coding and decoding scheme that allows to convey up to $R < \mathcal{C}_{Shan}$ bits of information per channel use, with arbitrarily vanishing probability of error.

The proof of the theorem is beyond the purpose of this work, but we want to highlight that it considers the transmission of long sequences of n symbols $\bar{x} = [x_1, x_2, \dots, x_n]$, with each $x_i \in \{1, \dots, M\}$ drawn with its a priori probability $p_{x_i} = p_x$, while at the receiver each symbols is measured individually and at the end an estimation on the global sequence is performed taking advantage of the concept of *typicality*.

In the case of quantum channels, the classical model is not rich enough to include quantum effects. *Quantum Information Theory* extends the classical counterpart to take into account the peculiarity of quantum mechanics. As a consequence, a quantum channel can be characterized with different capacities depending of the communication task, such as to convey classical or quantum information.

Here we focus in the transmission of classical information on a quantum channel. A reader interested in the transmission of quantum information can find a brief introduction in [14] and references therein.

Even when considering the delivery of classical information, several definitions of capacities has been proposed, depending on the coding and decoding strategy allowed to the sender and to the receiver [14, 15]. Bennett and Shor [16] have identified four possibilities, \mathcal{C}_{PP} , \mathcal{C}_{PE} , \mathcal{C}_{EP} and \mathcal{C}_{EE} , depending whether product states only or entangled states (first digit P or E) are allowed in the transmission

and product measurement or entangled measurement (second digit P or E) are allowed in the detection. Leaving aside \mathcal{C}_{EP} , is it clear that since transmission and detection strategies employing entangled states and measurement are more general, it holds

$$\mathcal{C}_{PP} \leq \mathcal{C}_{PE} \leq \mathcal{C}_{EE}. \quad (1.26)$$

When no entanglement is allowed at either end, the channel is equivalent to a classical noisy channel with the transition probabilities defined from the states and the measurement operators as in (1.19). Following the strategy of the classical coding, we can consider the n -th extension of the Hilbert space $\mathcal{H}^{\otimes n} = \mathcal{H} \otimes \cdots \otimes \mathcal{H}$ where are well defined the sequences of n states transmitted successively,

$$\rho^{(n)} = \rho_1 \otimes \cdots \otimes \rho_n. \quad (1.27)$$

In the assumption of *memoryless* channel, the corresponding channel for the sequence is just the juxtaposition of the action of the channel to each symbol,

$$\mathcal{E}^{(n)}(\rho^{(n)}) = \mathcal{E}(\rho_1) \otimes \mathcal{E}(\rho_2) \otimes \cdots \otimes \mathcal{E}(\rho_n), \quad (1.28)$$

The capacity \mathcal{C}_{PP} can be evaluated as

$$\mathcal{C}_{PP} = \sup_{p_x, \rho_x, P_{\hat{x}}} H(x; \hat{x}). \quad (1.29)$$

and it really resemble Shannon capacity \mathcal{C}_{Shan} except that it requires an optimization with respect the a priori probabilities $\{p_x\}$, the transmitted states $\{\rho_x\}$ and the measurement operators $\{P_{\hat{x}}\}$. Of course, the capacity \mathcal{C}_{PP} with the optimal states and measurements operators equal \mathcal{C}_{Shan} with the same states and POVM.

In the case of product states and entanglement measurements, the transmitter can encode the message in product sequences as (1.27) and the receiver can employ global measurement on the whole sequence. In this case the classical capacity is indicated with \mathcal{C}_{PE} or with $\mathcal{C}^{(1)}$, where the superscript indicates that for the coding are allowed only sequences of quantum states defined in a *single* Hilbert space.

The possibility of entanglement measurements gives rise to the first differences in terms of results between the classical and the quantum case. An important result that constitutes the counterpart of the classical noisy channel coding theorem has been formulated first in [17] for the restricted case of pure states, and then in [18, 19] in the more general case of mixed states.

Theorem 1.1. *Holevo–Schumacher–Westmoreland, Noisy Channel Coding for Classical Information*

Let \mathcal{E} be a quantum channel, define the Holevo quantity or Holevo capacity

$$\chi(\mathcal{E}) = \max_{p_i, \rho_i} S \left(\mathcal{E} \left(\sum_i p_i \rho_i \right) \right) - \sum_i p_i \mathcal{E}(\rho_i) \quad (1.30)$$

where the maximization is made with respect to the input quantum states $\{\rho_i\}$ and their a priori distribution $\{p_i\}$ employed in the communication. The quantity (1.30) is the classical capacity of \mathcal{E} when considering the transmission of product states and allowing for entangled measurements,

$$\chi(\mathcal{E}) = \mathcal{C}_{PE}. \quad (1.31)$$

Therefore, in the case of \mathcal{C}_{PE} , there is an expression that allow to evaluate the capacity. As already stated, it is clear that

$$\mathcal{C}_{Shan} \leq \mathcal{C}_{PE} \quad (1.32)$$

and examples in which this inequality is strict have been provided by Holevo [20]. In addition, a necessary and sufficient condition for the strict inequality results to be the commutativity of the input quantum states $\{\rho_i\}$ [21]

A third possibility is to employ quantum states entangled in n uses of the channel for the encoding as well as entangled measurement operators. In this case, the capacity of the channel is indicated with \mathcal{C}_{EE} , or with $\mathcal{C}^{(n)}$, where the superscript indicates that the quantum states are entangled over n uses of the channel. Finding the capacity of the channel \mathcal{E} when the states are encoded over n uses of the channel and entangled measurements are allowed, is equivalent to

find the capacity $\mathcal{C}^{(1)}$ of the channel $\mathcal{E}^{(n)}$. Despite this, question whether the use of entangled quantum states can actually improve the capacity of the channel, that is wheter \mathcal{C}_{EE} can striclty exceed \mathcal{C}_{PE} , is still an open question.

1.6 Summary of the Results

In this Thesis we consider the communication scenario and in particular the discrimination problem associated with the detection of the transmitted quantum states. We focus on the design of a receiver that can outperform the classical schemes and approach the theoretical quantum limit. Since at the moment no clear instructions are known to implement such a receiver, an optimization process is always necessary in order to overcome the classical phenomena and trigger the quantum effects.

In the next Chapter we set up the discrimination problem between two coherent states. This issue applies to the detection of signals in an optical binary coherent communication scenario, where an optimal receiver has been theoretically described and experimentally tested. Due to its properties, a coherent state of duration T can be viewed as a sequence of shorter and weaker modes of duration T/N . The discrimination of the coherent states can be hence interpreted as a discrimination among multiple copies of the same state. With this interpretation, we were able to infer the Dolinar receiver, the only example of implementation of a quantum receiver reaching the limit performance, as a multiple copy adaptive measurement on the segmented states. In addition, using the multiple copies theory, we propose a suboptimal simplified version of the Dolinar's scheme.

In the third Chapter we consider the problem of transmitting classical information over a noisy quantum channel. Given a description in terms of a Completely Positive Trace Preserving qubit map, we optimize the input states and the output measurements with respect to both error probability and classical capacity figures of merit. Via the coherence vector representation we are able to geometrically characterize the action of the channel, and find necessary condition for the

optimality of the quantum states and measurements, allowing us to reformulate the optimization in a form that can be solved by standard numerical algorithms.

In the fourth Chapter we study the discrimination problem among the Pulse Position Modulation signals. This modulation is particularly suitable for satellite and deep space communication scenario due to its high energy efficiency. Although a quantum receiver reaching the limit performance still has to be discovered, several suboptimal schemes has been proposed. We review these schemes and reformulate them in a qubit framework in order to highlight their limitations. Then, we propose an adaptive scheme to overcome these limitations. By applying a dynamic programming optimization, we are able to optimize the measurements operator, obtaining a receiver scheme which outperform all the existing architectures.

The last Chapter summarizes the results and concludes the Thesis.

Chapter 2

Binary Quantum Receivers

Binary modulation is the simplest modulation format by many points of view. In terms of encoding, it provides a direct mapping between the bit stream and the signals sequence to be sent through the channel, without additional encodings. In terms of hypothesis discrimination the problem is simplified since discarding one hypothesis implies the assumption of the other. In terms of signal generation and transmission, it requires a simple two-valued modulation for the physical quantity involved. In this section, I review the main receiver schemes for quantum binary modulation, leading the attention of the reader into the concept of adaptive measurement.

2.1 Introduction

In Quantum Detection Theory, the binary modulation of quantum states has been the most studied modulation, used since the first developments in quantum mechanics to test theoretical predictions. This effort has led to great results in both theoretical and experimental analysis.

Referring to the communication scenario in Figure 1.3, in the case of binary modulation the bit stream $\{b\}_i$ encoding the sender message is directly mapped into the symbol sequence $\{x\}_i$. Each symbol $x \in \{0, 1\}$ is then associated to a quantum state $\rho \in \{\rho_0, \rho_1\}$, $\rho \in \mathcal{D}(\mathcal{H})$ and sent through the channel. As usual,

symbol x is described as a random variable with a priori probability p_0, p_1 .

In the scenario of binary hypothesis testing, the problem of discrimination between ρ_0 and ρ_1 has been completely addressed by Helstrom [22], who found the optimal performance and the operators to achieve it in terms of POVM. Due to the simplicity of the binary modulation, optimal POVM can be easily found.

Consider in fact the pair of POVM (P_0, P_1) employed by the receiver to discriminate between the transmitted quantum states. The receiver estimates the transmitted quantum state $x = 0$ if the outcome $\hat{x} = \hat{0}$ is observed, otherwise the hypothesis $x = 1$ is taken.

The probability of correct decision reads

$$P_c = P[\hat{x} = x] = P[\hat{x} = \hat{0}|x = 0] p_0 + P[\hat{x} = \hat{1}|x = 1] p_1 \quad (2.1)$$

$$= \text{tr}(P_0 \rho_0) p_0 + \text{tr}(P_1 \rho_1) p_1 \quad (2.2)$$

The completeness relation in the case of binary modulation becomes $P_0 + P_1 = I_{\mathcal{H}_0}$, thus

$$P_c = \text{tr}(P_0 \rho_0 p_0 + P_1 \rho_1 p_1) = \text{tr}((I - P_1) \rho_0 p_0 + P_1 \rho_1 p_1) \quad (2.3)$$

$$= p_0 + \text{tr}(P_1(\rho_1 p_1 - \rho_0 p_0)) \quad (2.4)$$

The solution to the maximization of (2.4) comes from the eigenvalues decomposition of $\Delta = \rho_1 p_1 - \rho_0 p_0$. Define the eigenvector $|\nu\rangle$

$$(\rho_1 p_1 - \rho_0 p_0) |\nu\rangle = \nu |\nu\rangle \quad (2.5)$$

In order to maximize the probability of correct decision (2.4), the POVM P_1 must be the projector on the subspace of $\rho_1 p_1 - \rho_0 p_0$ associated with positive eigenvalues¹,

$$P_1 = \sum_{\nu > 0} |\nu\rangle\langle\nu|, \quad P_0 = I - P_1 = \sum_{\nu < 0} |\nu\rangle\langle\nu| \quad (2.6)$$

and the probability of correct decision becomes

$$P_c = p_0 + \sum_{\nu > 0} \nu \quad (2.7)$$

¹The projector associated to possible null eigenvalues can be indifferently included in P_0 or in P_1 .

In the case of pure states $\rho_0 = |\gamma_0\rangle\langle\gamma_0|, \rho_1 = |\gamma_1\rangle\langle\gamma_1|$ the analysis may be simplified. Define \mathcal{H}_0 as the subspace of \mathcal{H} spanned by $\{|\gamma_0\rangle, |\gamma_1\rangle\}$. Since \mathcal{H}_0 has dimension two, i.e. $\mathcal{H}_0 \sim \mathbb{C}^2$, we can define an orthonormal basis $|x\rangle, |y\rangle \in \mathcal{H}_0$ such that without loss of generality we can write

$$\begin{aligned} |\gamma_0\rangle &= \cos\theta |x\rangle + \sin\theta |y\rangle \\ |\gamma_1\rangle &= \cos\theta |x\rangle - \sin\theta |y\rangle \end{aligned}, \quad \theta \in [0, \pi/4] \quad (2.8)$$

with inner product $\chi = \langle\gamma_0|\gamma_1\rangle = \cos(2\theta)$. As we have seen, the optimal POVM are projectors, and it has been proved elsewhere [23] that in the case of pure states, these projectors have rank 1. Define the measurement operator in the basis $\{|x\rangle, |y\rangle\}$ as

$$|\mu_0\rangle = \cos\phi |x\rangle + \sin\phi |y\rangle \quad (2.9)$$

$$|\mu_1\rangle = \sin\phi |x\rangle - \cos\phi |y\rangle \quad (2.10)$$

The probability of correct decision can be written as

$$P_c = |\langle\mu_0|\gamma_0\rangle|^2 p_0 + |\langle\mu_1|\gamma_1\rangle|^2 p_1 \quad (2.11)$$

$$\begin{aligned} &= \cos^2(\theta - \phi)p_0 + \sin^2(\theta + \phi)p_1 \\ &= \frac{1 + \cos(2\theta - 2\phi)}{2} p_0 + \frac{1 - \cos(2\theta + 2\phi)}{2} p_1 \end{aligned} \quad (2.12)$$

The search for stationary points of (2.12) with respect to the angle ϕ reads

$$\tan 2\phi = \frac{1}{p_0 - p_1} \tan 2\theta \quad (2.13)$$

that is solved for

$$\phi = \frac{1}{2} \arctan \left[\frac{1}{p_0 - p_1} \tan 2\theta \right] + l \frac{\pi}{2}, \quad l \in \mathbb{Z}. \quad (2.14)$$

An analysis of the second derivative reveals that the point of maximum is obtained for even values of l , such as

$$\phi = \frac{1}{2} \arctan \left[\frac{1}{p_0 - p_1} \tan 2\theta \right], \quad (2.15)$$

verifying

$$\begin{aligned}\sin 2\phi &= \frac{1}{R} \sin 2\theta \\ \cos 2\phi &= \frac{p_0 - p_1}{R} \cos 2\theta\end{aligned}\tag{2.16}$$

with

$$R = \sqrt{\sin^2 2\theta + (p_0 - p_1)^2 \cos^2 2\theta} = \sqrt{1 - 4p_0p_1 \cos^2 2\theta},\tag{2.17}$$

while the point of minimum, obtained for odd values of l , has the opposite sign in the right-hand side of both (2.16).

The maximized probability of correct decision is called the Helstrom Bound,

$$P_c = \frac{1}{2} \left[1 + \sqrt{1 - 4p_0p_1\chi^2} \right].\tag{2.18}$$

2.1.1 Binary modulation implementations

The binary modulation can be easily implemented with coherent states. Both intensity and phase modulations can be considered in order to encode the symbol x .

In the case of intensity modulation, the modulation is called *On Off Keying* and makes use of the following associations

$$\begin{aligned}x = 0 &\longrightarrow |\gamma_0\rangle = |0\rangle \\ x = 1 &\longrightarrow |\gamma_1\rangle = |2\alpha\rangle\end{aligned}\tag{2.19}$$

with inner product $\chi = |\langle 0|2\alpha\rangle| = e^{-2|\alpha|^2}$. This modulation can be easily implemented with a laser source and an intensity modulator that shuts down the laser pulse in correspondence of the time interval of the symbol $x = 0$.

In the case of phase modulation, the association between symbol x and transmitted quantum state reads

$$\begin{aligned}x = 0 &\longrightarrow |\gamma_0\rangle = |\alpha\rangle \\ x = 1 &\longrightarrow |\gamma_1\rangle = |-\alpha\rangle\end{aligned}\tag{2.20}$$

with inner product $\chi = |\langle -\alpha|\alpha\rangle| = e^{-2|\alpha|^2}$. The modulation is called *Binary Phase Shift Keyed*, BPSK. In this chapter, we will focus on the latter binary coherent modulation.

Note that following the definition (2.19) and (2.20), the inner product of the transmitted quantum states is the same and so is the optimal performance by (2.18). This is also motivated by the fact that, in the absence of noise, at the receiver side we can go from one modulation to the other applying a displacement operation on the incoming quantum state, as used in the Kennedy receiver described in the next section.

2.2 Kennedy Receiver

Classical coherent communication relies on homodyne and heterodyne detection to discriminate between the two signals in BPSK modulation. These receiver schemes implement a phase measurement of the incoming unknown coherent state: the received optical mode is mixed with a strong local oscillator through a balanced beam splitter, and an intensity measurement is performed on the two output modes. In the case of homodyne detection, the local oscillator is in phase with the signal, that is at the same optical frequency, while in the case of heterodyne detection the oscillator is detuned, i.e. with a much higher frequency.

In the absence of thermal noise, with BPSK detection both homodyne and heterodyne receiver have the same performance. For signal with equal a priori probability, the probability of correct decision is given by²

$$P_c^{(homodyne)} = 1 - \frac{\operatorname{erfc}(\sqrt{2}|\alpha|^2)}{2} = 1 - Q\left(\sqrt{4|\alpha|^2}\right) \quad (2.23)$$

The first receiver scheme that shows the possibility to beat the classical ho-

²Different textbooks and papers [24, 9, 25] use different notations. The Complementary Error Function erfc is defined by

$$\operatorname{erfc}(x) = \frac{1}{\sqrt{\pi} \int_x^\infty e^{-t^2} dt} \quad (2.21)$$

while $Q(\cdot)$ is the tail probability of the standard normal distribution,

$$Q(x) = \frac{1}{\sqrt{2\pi} \int_x^\infty e^{-\frac{t^2}{2}} dt} \quad (2.22)$$

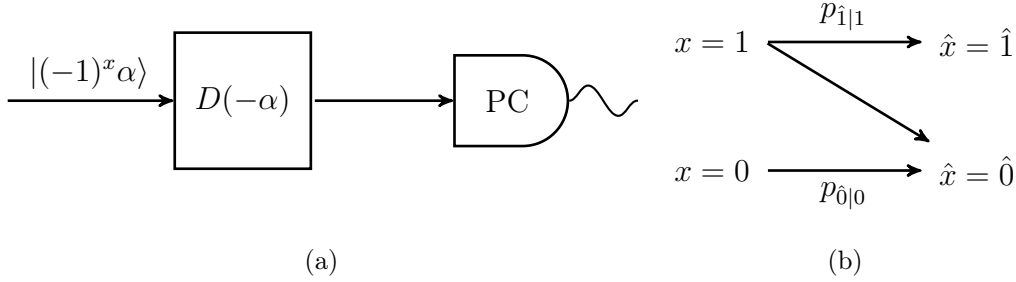


Figure 2.1: (a) Scheme of the Kennedy receiver. The incoming coherent state $|(-1)^x \alpha\rangle$ is first displaced by $D(-\alpha)$ and then a photon counter PC seeks for photons in the symbol time interval. (b) Transition probabilities.

modyne limit was proposed by Kennedy [23]. A diagram of the Kennedy receiver is depicted in Figure 2.1. The incoming unknown coherent state is displaced with $D(-\alpha)$, such that in the case of coherent state $|\alpha\rangle$ it is displaced to the vacuum state $|0\rangle$, while in the case of the coherent state $|\alpha\rangle$ it is displaced by $|-2\alpha\rangle$. After the displacement operation, a photon counting is performed in the time symbol interval.

Let us assume the photon counter has unit efficiency and no dark counts. Define as $z = \hat{0}, \hat{1}$ the outcomes from the photon counter, where

$$P_{\hat{0}} = |0\rangle\langle 0|, \quad P_{\hat{1}} = I - P_{\hat{0}} = \sum_{n \in \mathcal{N}} |n\rangle\langle n| \quad (2.24)$$

Let us also define the following rule for the estimation of the symbol x ,

$$\begin{aligned} z = \hat{0} &\longrightarrow \hat{x} = 0 \\ z = \hat{1} &\longrightarrow \hat{x} = 1 \end{aligned} \quad (2.25)$$

The scheme is designed such that, in the absence of noise, the symbol $x = 0$ corresponding to the quantum state $|\alpha\rangle$ is displaced to the vacuum and perfectly detected by $P_{\hat{0}}$, that is

$$p_{\hat{0}|0} = \text{tr}(P_{\hat{0}} D(-\alpha) |\alpha\rangle\langle \alpha| D^\dagger(-\alpha)) = 1. \quad (2.26)$$

Therefore, the only error in this receiver scheme arises when the symbol $x = 1$, displaced in $|-2\alpha\rangle$, is misdetected to an outcome $z = \hat{0}$. Since this event can occur

with probability

$$\text{tr} (P_{\hat{0}} D(-\alpha) |-\alpha\rangle \langle -\alpha| D^\dagger(-\alpha)) = \text{tr} (P_{\hat{0}} | -2\alpha\rangle \langle -2\alpha|) = |\langle 0| -2\alpha\rangle|^2 = e^{-4|\alpha|^2} \quad (2.27)$$

the global probability of correct decision of the Kennedy receiver is given by

$$P_c^{(Kennedy)} = p_0 + p_1(1 - e^{-4|\alpha|^2}) \quad (2.28)$$

or in general

$$P_c^{(Kennedy)} = 1 - \min\{p_0, p_1\} e^{-4|\alpha|^2}. \quad (2.29)$$

In Figure 2.3, the error probability of the homodyne detector and the Kennedy receiver are compared with other receiver schemes presented in the following sections with respect to the average photon number. The Kennedy receiver performs better than the homodyne limit for an average photon per symbol greater than about 0.5, and the gap between the performances spread out as the photon number increases. However, in practice for high average photon numbers the performance of the Kennedy receiver is limited by dark counts, that with a probability of 10^{-6} prevents to go underneath this threshold.

2.3 Improved Kennedy Receiver

The Kennedy receiver provides a near optimal scheme for the discrimination of the BPSK signal that approach the Helstrom bound exponentially as $|\alpha|^2 \rightarrow \infty$,

$$P_c^{Helstrom} - P_c^{Kennedy} \sim e^{-4\alpha^2}. \quad (2.30)$$

Despite that, the scheme is not robust against dark counts and thermal noise [26], and a mode mismatch between the BPSK signal and the added local oscillator causes additional dark counts. Furthermore, the *complete* nulling of the coherent state $|-\alpha\rangle$ is not proven to be the optimal choice in order to maximize the probability of correct decision with this architecture.

Takeoka and Sasaki [24] have proposed some improvements for the Kennedy receiver by means of a Gaussian measurement. An operator is said to be Gaussian if it maps Gaussian states into Gaussian states, that is an operation that

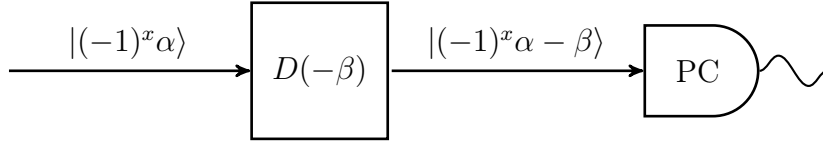


Figure 2.2: Scheme of the Generalized Kennedy receiver. A displacement operation $D(-\beta)$ is performed on the incoming unknown coherent state, and subsequently a photon counting is performed.

generalizes the phase rotation, displacement and squeezing operation on coherent states. A Gaussian measurement usually consists of adding some ancillary Gaussian states, apply Gaussian operations and perform homodyne measurements. In [24] they show that a receiver employing only Gaussian measurement cannot achieve the Helstrom bound, and can at most reach the homodyne error probability.

In order to approach the Helstrom limit, we need to consider non-Gaussian operation devices, such as photon counters. Consider the receiver scheme of Figure 2.2, where a Gaussian operation is followed by a photon counting. The scheme reconduces to the Kennedy receiver when the gaussian operation is the displacement $D(-\alpha)$. Since a Gaussian operation generalize the displacement operation, we expect to improve the error probability by the optimization of the Gaussian operation.

In [24] it is shown that phase rotations are unnecessary, and the optimal conditions for displacement and squeezing are evaluated. Here we consider only the optimization of the displacement value, for a receiver architecture such as the one in Figure 2.2, called *Generalized* or *Improved Kennedy receiver*.

Consider a displacement $D(-\beta)$ of the incoming unknown BPSK signal, with β to be optimized. Using the same estimation rule (2.25) of the Kennedy receiver, the correct detection probability becomes

$$P_c^{(Gen.Kennedy)} = p_0 \langle \alpha - \beta | P_0 | \alpha - \beta \rangle + p_1 \langle -\alpha - \beta | P_1 | -\alpha - \beta \rangle \quad (2.31)$$

$$= p_0 e^{-(\alpha-\beta)^2} + p_1 (1 - e^{-(\alpha+\beta)^2}) \quad (2.32)$$

By nulling the derivative with respect to β , we find that the displacement value β^* maximizing (2.32) satisfies the transcendental equation

$$\frac{p_0}{p_1} = \frac{\beta^* + \alpha}{\beta^* - \alpha} e^{-4\beta^*\alpha} \quad (2.33)$$

which can be numerically solved and the corresponding value of $P_c^{(Gen.Kennedy)}$ evaluated.

In Figure 2.3 the performance of the generalized Kennedy receiver is compared with the Kennedy receiver, the Helstrom bound and the homodyne limit, assuming equal a priori probability $p_0 = p_1 = 1/2$. For large values of α the improvement obtained by optimizing the displacement β is negligible. On the other hand, as α goes to 0, the performance of the Generalized Kennedy approximates the Helstrom bound.

2.4 Multiple Copy State Discrimination

The scenario of the multiple copies problem is slightly different from the one presented in Section (1.4). In this case, Alice chooses a binary symbol $x \in \{0, 1\}$ with a priori probability $\{p_0, p_1\}$, and accordingly to the chosen symbol she sends to the receiver Bob N copies of the quantum state $|\gamma_0\rangle$ or $|\gamma_1\rangle$. Bob has to guess which symbol has been chosen, possibly taking advantage of the multiple resources.

The ensemble sent to Bob can be appropriately described in the tensorial product Hilbert Space $\mathcal{H} = \mathcal{H}_0^{\otimes N}$, namely

$$\begin{aligned} |\bar{\gamma}_0\rangle &= |\gamma_0\rangle \otimes \dots \otimes |\gamma_0\rangle \\ |\bar{\gamma}_1\rangle &= |\gamma_1\rangle \otimes \dots \otimes |\gamma_1\rangle \end{aligned} \quad (2.34)$$

The multiple copies problem is still a binary discrimination problem, but formulated on quantum states $|\bar{\gamma}_0\rangle$ and $|\bar{\gamma}_1\rangle$. The Helstrom theory introduced in Section 2.1 still applies, and the maximum probability of correct decision is given by the Helstrom bound (2.18)

$$P_c = \frac{1}{2} \left[1 + \sqrt{1 - 4p_0p_1X^2} \right] = \frac{1}{2} \left[1 + \sqrt{1 - 4p_0p_1\chi^{2N}} \right] \quad (2.35)$$

where $X = \langle \bar{\gamma}_0 | \bar{\gamma}_1 \rangle$ is the inner product between $|\bar{\gamma}_0\rangle$ and $|\bar{\gamma}_1\rangle$, and by definition (2.34) is equal to $X = \chi^N = \langle \gamma_0 | \gamma_1 \rangle^N$, with χ the inner product of $|\gamma_0\rangle$ and $|\gamma_1\rangle$.

The interest for this problem comes from the different measurement scheme that Bob can use to approach (2.35). By Helstrom discrimination theory, Bob may achieve this bound by a global measurement using suitable von Neumann projectors $\Pi_{\hat{0}} = |\bar{\mu}_{\hat{0}}\rangle\langle\bar{\mu}_{\hat{0}}|$, $\Pi_{\hat{1}} = |\bar{\mu}_{\hat{1}}\rangle\langle\bar{\mu}_{\hat{1}}|$ over the product space \mathcal{H} . Unfortunately, the optimum measurement vectors $|\bar{\mu}_{\hat{0}}\rangle, |\bar{\mu}_{\hat{1}}\rangle$ turn out to be a linear superposition of the pure states $|\bar{\gamma}_0\rangle, |\bar{\gamma}_1\rangle$, entangled over the multiple copies in the spaces \mathcal{H}_0 .

On the other hand, Bob may employ local measurement schemes on single spaces \mathcal{H}_0 . Interestingly, the same performance (2.35) can be achieved using an optimized adaptive local scheme [27, 28, 29].

Assume that the local measurement orthonormal vectors are described as in (2.10), namely

$$|\mu_{\hat{0}}(\bar{z}_{k-1})\rangle = \cos \phi_{\bar{z}_{k-1}} |x\rangle + \sin \phi_{\bar{z}_{k-1}} |y\rangle \quad (2.36)$$

$$|\mu_{\hat{1}}(\bar{z}_{k-1})\rangle = \sin \phi_{\bar{z}_{k-1}} |x\rangle - \cos \phi_{\bar{z}_{k-1}} |y\rangle \quad (2.37)$$

where the angle $\phi_{\bar{z}_{k-1}}$ that defines the k -th measurement operators depends on the list of the previous individual measurements $\bar{z}_{k-1} = [z_1, z_2, \dots, z_{k-1}]$. The adaptive optimization problem consists in finding a starting measurement angle ϕ_{\emptyset} (the symbol \emptyset indicates that does not depend on any previous outcomes) and a recursive rule that defines the next measurement angle,

$$\phi_{\bar{z}_{k-1}} = f_k(z_1, z_2, \dots, z_{k-1}) . \quad (2.38)$$

The global measurement operators are given by

$$|\bar{\mu}_{\bar{z}_N}\rangle = |\mu_{z_1}\rangle \otimes |\mu_{z_2}(z_1)\rangle \otimes \dots \otimes |\mu_{z_N}([z_1 \dots z_{N-1}])\rangle \quad (2.39)$$

and we can readily see that all these operator sum up to the identity in \mathcal{H} ,

$$\sum_{\bar{z}_N \in \mathcal{Z}_N} |\bar{\mu}_{\bar{z}_N}\rangle\langle\bar{\mu}_{\bar{z}_N}| = I_{\mathcal{H}} \quad (2.40)$$

with \mathcal{Z}_N the set of all the binary sequence of outcomes of length N .

The estimation rule is defined in such a way that the last outcome (the rightmost binary digit of \bar{z}_N) gives the final hypothesis for $|\gamma_0\rangle$ or $|\gamma_1\rangle$. Thus, we have

$$P_c = \sum_{\bar{z} \in \mathcal{Z}_{N-1}} (P[z_N = \hat{0} | \bar{z}, x = 0] p_0 + P[z_N = \hat{1} | \bar{z}, x = 1] p_1) \quad (2.41)$$

$$\begin{aligned} &= \sum_{\bar{z} \in \mathcal{Z}_{N-1}} (P[z_N = \hat{0} | \bar{z}, x = 0] P[\bar{z}, x = 0] + P[z_N = \hat{1} | \bar{z}, x = 1] P[\bar{z}, x = 1]) \\ &= \sum_{\bar{z} \in \mathcal{Z}_{N-1}} (|\langle \mu_{\hat{0}}(\bar{z}) | \gamma_0 \rangle|^2 P[\bar{z}, x = 0] + |\langle \mu_{\hat{1}}(\bar{z}) | \gamma_1 \rangle|^2 P[\bar{z}, x = 1]) \end{aligned} \quad (2.42)$$

The maximization with respect to $\phi_{\bar{z}}$, $\bar{z} \in \mathcal{Z}_{N-1}$, has a solution similar to (2.15), leading to the relation³

$$\tan 2\phi_{\bar{z}} = \frac{p_{\bar{z},0} + p_{\bar{z},1}}{p_{\bar{z},0} - p_{\bar{z},1}} \tan 2\theta, \quad (2.44)$$

that is

$$\sin 2\phi_{\bar{z}} = \frac{p_{\bar{z},0} + p_{\bar{z},1}}{R(\bar{z})} \sin 2\theta, \quad (2.45)$$

$$\cos 2\phi_{\bar{z}} = \frac{p_{\bar{z},0} - p_{\bar{z},1}}{R(\bar{z})} \cos 2\theta, \quad (2.46)$$

with $R(\bar{z})$ the normalization factor

$$R(\bar{z}) = \sqrt{[p_{\bar{z},0} + p_{\bar{z},1}]^2 \sin^2 2\theta + (p_{\bar{z},0} - p_{\bar{z},1})^2 \cos^2 2\theta} \quad (2.47)$$

$$= \sqrt{[p_{\bar{z},0} + p_{\bar{z},1}]^2 - 4p_{\bar{z},0}p_{\bar{z},1} \cos(2\theta)^2}. \quad (2.48)$$

Substituting back in (2.42) we obtain

$$P_c = \frac{1}{2} + \frac{1}{2} \sum_{\bar{z} \in \mathcal{Z}_{N-1}} R(\bar{z}) \quad (2.49)$$

Although this appears to be a dynamic programming problem [30], and as such has been dealt with and solved in [28], it turns out that there is a much simpler

³We use the shorthand notation of equations (1.20) for the joint probabilities

$$p_{\bar{z},i} = P[\bar{z}, x = i]. \quad (2.43)$$

solution. The problem reduces to a Bayesian updating problem with recursive relation

$$\phi_{[z_1 \dots z_{k-1}]} = f_k(z_{k-1}) \quad (2.50)$$

so that the optimal angle at the k -th measurement depends only on the previous outcome. It is convenient to change the notation on the dependence of the angle, and simply define as $\phi_k(z_{k-1}) = \phi_{[z_1 \dots z_{k-1}]}$ the angle in (2.50).

By the Bayesian update, the optimal measurement angle ϕ_k is the solution of the binary discrimination problem obtained replacing the a priori probabilities $P[x = 0]$ and $P[x = 1]$ with the a posteriori probabilities given the last outcome, that are $P[x = 0|z_{k-1} = i]$ and $P[x = 1|z_{k-1} = i]$. In particular, the following result holds.

Proposition 2.1. *The a posteriori probabilities after the measurement on the k -th copy are given by the Helstrom bound*

$$P[x = i|z_k = i] = \frac{1}{2} \left[1 + \sqrt{1 - 4p_0p_1\chi^{2(k)}} \right] \quad i = 0, 1, \quad (2.51)$$

and the subsequent optimal measurement angle is given by

$$\begin{aligned} \phi_k(z_{k-1}) &= \frac{1}{2} \arctan \left[\frac{1}{P[x = 0|z_{k-1} = i] - P[x = 1|z_{k-1} = i]} \tan 2\theta \right] \quad (2.52) \\ &= \frac{1}{2} \arctan \left[(-1)^{z_{k-1}} \frac{1}{\sqrt{1 - 4p_0p_1\chi^{2(k-1)}}} \tan 2\theta \right], \quad k = 1, \dots, N. \end{aligned} \quad (2.53)$$

Proof. This result can be proven by induction. Given the a priori probability p_0 and p_1 , the optimization of the first angles leads to (2.15), and the probability of correct decision $P_c^{(1)}$ after the first measurement is given by the Helstrom bound

$$P_c^{(1)} = \frac{1}{2} \left[1 + \sqrt{1 - 4p_0p_1\chi^2} \right] \quad (2.54)$$

Then, the result is proven for $k = 1$. The a posteriori probabilities read

$$P[x = 0|z = \hat{0}] = \frac{P[z = \hat{0}|x = 0] p_0}{P[z = \hat{0}]} = \frac{\cos^2(\theta - \phi_1)p_0}{\cos^2(\theta - \phi_1)p_0 + \cos^2(\theta + \phi_1)p_1} \quad (2.55)$$

$$= \frac{\left[1 + \frac{\cos^2 2\theta}{R}(p_0 - p_1) + \frac{\sin^2 2\theta}{R}\right] p_0}{1 + \frac{\cos^2 2\theta}{R}(p_0 - p_1) + \frac{\sin^2 2\theta}{R}(p_0 - p_1)} \quad (2.56)$$

$$= \frac{[R + \cos^2 2\theta(p_0 - p_1) + \sin^2 2\theta] p_0}{R + p_0 - p_1} \cdot \frac{R - (p_0 - p_1)}{R - (p_0 - p_1)} \quad (2.57)$$

$$= \frac{1}{2} [1 + R] \quad (2.58)$$

where in (2.56) we substitute the optimal angle ϕ_1 , and R has the usual definition (2.17). Similarly, one can see that it holds for $P[x = 1|z = \hat{1}]$. Now suppose that the result holds true for k . From the inductive hypothesis, the provisional correct detection coincides with the Helstrom bound, the adaptive measurement up to the k -th copy coincides with the optimal global measurement and the a posteriori probabilities are

$$P[x = i|z_k = i] = P_c^{(k)}. \quad (2.59)$$

The optimization of ϕ_{k+1} gives the solution obtained by replacing p_0 and p_1 in the expression (2.15) with the a posteriori probability, leading to

$$P_c^{(k+1)} = \frac{1}{2} \left[1 + \sqrt{1 - 4P_c^{(k)} [1 - P_c^{(k)}] \chi^2} \right] = \frac{1}{2} \left[1 + \sqrt{1 - 4p_0p_1\chi^{2(k+1)}} \right]$$

□

We can evaluate the a posteriori probability and the angles sequence offline. The optimum local adaptive measurement can be summarized by the following step-by-step procedure.

1. From the overlap coefficient χ and the input probabilities p_0 and p_1 evaluate

$$\phi_k = \frac{1}{2} \arctan \left[(-1)^{z_{k-1}} \frac{1}{\sqrt{1 - 4p_0p_1\chi^{2(k-1)}}} \tan 2\theta \right] \quad (2.60)$$

and compute the double sequence of measurement angles

$$\begin{array}{cccc} \phi_1 & \phi_2 & \dots & \phi_N \\ -\phi_1 & -\phi_2 & \dots & -\phi_N \end{array} \quad (2.61)$$

2. Start with angle ϕ_1 if $P[x = 0] \geq 1/2$ and $-\phi_1$ otherwise.
3. Use the angles of the first sequence until the measurement result is 0.
4. Change angle sequence every time the result changes and accept z_N as the global result.

In conclusion, the problem of multiple copies can be solved with an adaptive scheme, achieving the best performance in terms of probability of correct decision predicted by the Helstrom bound. The adaptive scheme requires local measurement in each copy and classical communication of the outcome, in order to optimize the next copy measurement. The scheme can be interpreted by the means of Bayesian updating: the outcome of the measurements allows to calculate the a posteriori probability of the symbol $x = 0$ or $x = 1$, that becomes the a priori probability for the next measurement. In addition, the next measurement can be evaluated as the optimal solution for the binary discrimination of a single copy with these new a priori probabilities.

2.5 Revisiting the Dolinar receiver through the multiple copy discrimination theory

The discrimination between two coherent states $|\alpha\rangle$ and $|\!-\alpha\rangle$ of a travelling single mode harmonic oscillator presents a difficulty similar to that of collective measurements on multiple copies. Namely, the optimal POVMs predicted by Helstrom theory are linear combinations of $|\alpha\rangle$ and $|\!-\alpha\rangle$, but do not correspond to any measurable observable. On the other hand, the coherent states $|\alpha\rangle$ and $|\!-\alpha\rangle$ of duration T can be thought as sequences of shorter and weaker modes of

duration T/N , namely,

$$\begin{aligned} |\alpha\rangle &= \left| \frac{\alpha}{\sqrt{N}} \right\rangle \otimes \cdots \otimes \left| \frac{\alpha}{\sqrt{N}} \right\rangle \\ |-\alpha\rangle &= \left| -\frac{\alpha}{\sqrt{N}} \right\rangle \otimes \cdots \otimes \left| -\frac{\alpha}{\sqrt{N}} \right\rangle \end{aligned} \quad (2.62)$$

This interpretation of coherent states suggests that the theory of multiple copies discrimination could be applied for the binary coherent state discrimination.

As previously seen in Sections 2.2 and 2.3, in the presence of weak coherent states the optimal measurements can be well approximated by a displacement operation followed by a photon counting. Then, in principle, the sequence of measurement angles (2.61) can be reinterpreted in this context as a sequence of displacements. Consequently, the Dolinar receiver that makes use of a continuous displacement operation appears as an adaptive scheme employing projective measurements on multiple copies of a coherent state.

This interpretation has already been noticed in [31], although a complete proof of the Dolinar receiver in such terms was not given. In the paper, as N goes to infinity, each copy can be considered as a qubit on the two dimensional space spanned by the number state $|0\rangle$ and $|1\rangle$, that are exactly the eigentates of the photon counter. The displacement operation is hence interpreted as a rotation in the space of qubits, to improved the measurement in the basis of $|0\rangle, |1\rangle$.

Consider the input field $\psi(t)$, $0 < t < T$, corresponding to the coherent state $|\pm\alpha\rangle$, represented by

$$\psi(t) = \pm\psi e^{i2\pi f_0 t}, \quad (2.63)$$

where f_0 is the optical frequency and T is the pulse duration. The mean number of photons arriving at the detector is given by

$$\alpha^2 = \int_0^T |\psi(t)|^2 dt = \psi^2 T. \quad (2.64)$$

As the number of the copies N goes to infinity, the measurement in each copy (2.62) of the coherent state becomes infinitesimal and the sequence of displacements gives birth to a continuous time feedback scheme. The Dolinar receiver

subtracts from the unknown input field (2.63) a time-varying field generated by a local oscillator, producing a time variant displacement operation. The coherent state of the displacing operation is chosen between either $u_0(t)$ or $u_1(t)$, accordingly to the value of $z(t)$, a binary signal with possible values $\hat{0}$ and $\hat{1}$, giving the provisional decision at time t . By mimicking the behaviour of the optimal multiple copy detection, we assume that the decision signal $z(t)$ changes at any photon arrival at the counter. Thus, the optical signal at the photon counter has envelope either $\pm\psi - u_0(t)$ or $\pm\psi - u_1(t)$, depending on the value of $z(t)$. Moreover, $z(T)$ is assumed to be the final decision.

The mathematical problem is to choose the functions $u_0(t)$ and $u_1(t)$ that maximize the correct detection probability

$$P_c = P[z(t) = x]. \quad (2.65)$$

The problem can be solved by means of standard photon counting statistics.

Let us assume that $x = 0$, so that $\psi(t) = \psi e^{i2\pi f_0 T}$. Then, the process $z(t)$ can be interpret as a telegraph process [32] alternately driven by nonhomogeneous Poisson processes with rates

$$\lambda(t) = |\psi - u_0(t)|^2 \quad \text{ad} \quad \nu(t) = |\psi - u_1(t)|^2. \quad (2.66)$$

Following the multiple copies adaptive solution, we evaluate the time evolution of the conditional correct detection probability $p_{\hat{0}|0} = P[z = \hat{0}|x = 0]$. Define as $n(t, t + \Delta t)$ the number of photon arrivals at the counter in the interval $(t, t + \Delta t]$.

By the photon statistics,

$$p_{\hat{0}|0}(t + \Delta t) = P [z(t + \Delta t) = \hat{0} | x = 0] \quad (2.67)$$

$$\begin{aligned} &= P [z(t + \Delta t) = \hat{0}, z(t) = \hat{0} | x = 0] + P [z(t + \Delta t) = \hat{0}, z(t) = \hat{1} | x = 0] \\ &= P [z(t + \Delta t) = \hat{0} | z(t) = \hat{0}, x = 0] P [z(t) = \hat{0} | x = 0] \\ &\quad + P [z(t + \Delta t) = \hat{0} | z(t) = \hat{1}, x = 0] P [z(t) = \hat{1} | x = 0] \end{aligned} \quad (2.68)$$

$$\begin{aligned} &= P [n(t, t + \Delta t) = 0 | z(t) = \hat{0}, x = 0] p_{\hat{0}|0}(t) \\ &\quad + P [n(t, t + \Delta t) = 1 | z(t) = \hat{1}, x = 0] (1 - p_{\hat{0}|0}(t)) + o(\Delta t) \\ &= (1 - \lambda(t)\Delta t)p_{\hat{0}|0}(t) + \nu(t)\Delta t(1 - p_{\hat{0}|0}(t)) + o(\Delta t) \end{aligned} \quad (2.69)$$

$$= p_{\hat{0}|0}(t) + \nu(t)\Delta t - (\lambda(t) + \nu(t))\Delta t p_{\hat{0}|0}(t) + o(\Delta t) \quad (2.70)$$

In a similar way for $p_{\hat{1}|1} = P [z = \hat{1} | x = 1]$, we get

$$p_{\hat{1}|1}(t + \Delta t) = p_{\hat{1}|1}(t) + \tilde{\nu}(t)\Delta t - (\tilde{\lambda}(t) + \tilde{\nu}(t))\Delta t p_{\hat{1}|1}(t) \quad (2.71)$$

with

$$\tilde{\lambda}(t) = |-\psi - u_1(t)|^2 \quad \text{ad} \quad \tilde{\nu}(t) = |-\psi - u_0(t)|^2. \quad (2.72)$$

From (2.70) and (2.71), we get the evolution of $p_{\hat{0}|0}$ and $p_{\hat{1}|1}$ taking the limit as $\Delta t \rightarrow 0$,

$$\begin{aligned} \dot{p}_{\hat{0}|0} &= \lim_{\Delta t \rightarrow 0} \frac{p_{\hat{0}|0}(t + \Delta t) - p_{\hat{0}|0}(t)}{\Delta t} = \nu(t) - [\lambda(t) + \nu(t)]p_{\hat{0}|0}(t) \\ \dot{p}_{\hat{1}|1} &= \lim_{\Delta t \rightarrow 0} \frac{p_{\hat{1}|1}(t + \Delta t) - p_{\hat{1}|1}(t)}{\Delta t} = \tilde{\nu}(t) - [\tilde{\lambda}(t) + \tilde{\nu}(t)]p_{\hat{1}|1}(t) \end{aligned} \quad (2.73)$$

If our search is confined to symmetric solutions, i.e. $u_1(t) = -u_0(t)$, we get $\tilde{\lambda}(t) = \lambda(t)$ and $\tilde{\nu}(t) = \nu(t)$, and the correct detection probability satisfies the differential equation

$$\dot{P}_c = \dot{p}_{\hat{0}|0}p_0 + \dot{p}_{\hat{1}|1}p_1 = \nu(t) - [\lambda(t) + \nu(t)]P_c(t) \quad (2.74)$$

$$= |\alpha - u_1|^2 - 2(\alpha^2 + u_1^2)P_c(t) \quad (2.75)$$

On the basis of the results on multiple copy measurements we expect that, for some choice of the envelope of the feedback signal $u_0(t)$, the provisional correct

detection probability $P_c(t)$ is exactly equal to the Helstrom bound applied to the interval $(0, t)$, namely

$$P_c(t) = \frac{1}{2} \left[1 + \sqrt{1 - 4p_0p_1e^{-4\alpha^2t}} \right] \quad (2.76)$$

By substituting the above expression in (2.75), and defining $R(t) = \sqrt{1 - 4p_0p_1e^{-4\alpha^2t}}$, we get

$$\alpha^2 \frac{1 - R^2(t)}{R(t)} = \alpha^2 + u_1^2 + 2\alpha u_1(t) - [\alpha^2 + u_1^2(t)][1 + R(t)] \quad (2.77)$$

and after some algebra,

$$u_1(t) = \frac{\alpha}{R(t)} = \frac{\alpha}{\sqrt{1 - 4p_0p_1e^{-4\alpha^2t}}} \quad (2.78)$$

coinciding indeed with the Dolinar's solution.

2.6 A simple suboptimal receiver

The Dolinar receiver was initially proposed in 1973 [33], but due to the difficulties to implement a very precise control of the optical–electrical loop, only recently has been experimentally tested [34]. In fact, the coherent state to be added to the unknown incoming optical mode must be precisely shaped in the amplitude, following (2.78), and a high speed feedback is necessary to change the phase of this coherent state when a photon is observed.

On the other part, the Kennedy receiver and its improved version make use of a constant fixed displacement that leads to a much simpler implementation. Therefore, it is worthwhile to consider a simplified version of the Dolinar receiver where the feedback is constrained to have a constant fixed envelope $u_1(t) = \beta$. In such a setting only a phase modulation is required, that is the phase inversion to be applied to the displacing coherent state added when a photon is observed.

The expression (2.75) gives a first order differential equation in for $P_c(t)$ that can be solved for any possible function of the displacement $u_1(t)$. In particular, substituting the a constant value $u_1(t) = -\beta$ for the displacement and with

the initial condition $P_c(0) = p_0$, we get the following final probability of correct decision

$$P_c(T) = \frac{1}{2} + \frac{\psi\beta}{\psi^2 + \beta^2} + \left[p_0 - \frac{1}{2} - \frac{\psi\beta}{\psi^2 + \beta^2} \right] e^{-2(\psi^2 + \beta^2)T}. \quad (2.79)$$

The optimal value of β can be found by numerically solving the following transcendental equation, which is obtained by nulling the derivative of $P_c(T)$ with respect to β :

$$\beta T(\psi^2 + \beta^2) [(2p_0 - 1)(\psi^2 + \beta^2) - 2\psi\beta] e^{-(\psi^2 + \beta^2)T} = \psi(\psi^2 + \beta^2) \sinh(\psi^2 T + \beta^2 T). \quad (2.80)$$

In Figure 2.3 we note that the simplified Dolinar receiver slightly outperforms the improved Kennedy receiver. In Figure 2.4 the intensity of the displacement for different schemes is reported. It can be noted that, as confirmed by Figure 2.3, it is for very weak coherent signals that the simplified Dolinar receiver and the generalized Kennedy are particularly attractive since their performance approach the Helstrom bound, while with increasing signal strength, they both perform a displacement similar to the one applied by Kennedy's original proposal.

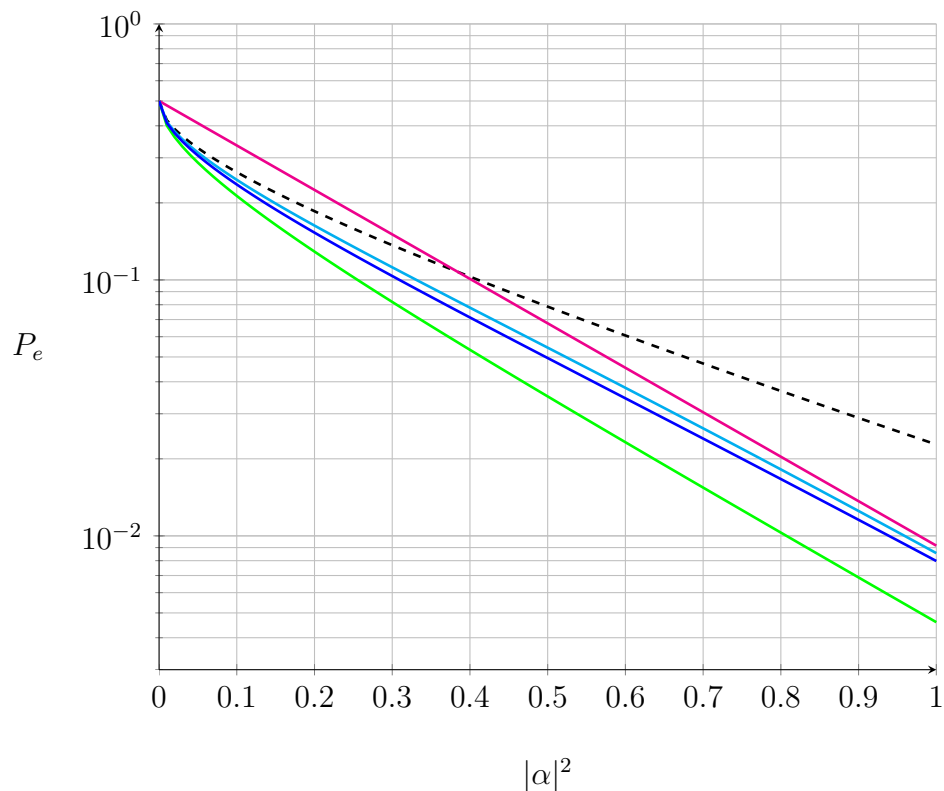


Figure 2.3: Performance comparison for different receiver schemes. In — the Helstrom limit, achieved with the Dolinar receiver. In --- the homodyne limit. In — the performance of the Kennedy receiver, which for $|\alpha|^2 > 0.5$ outperforms the classical limit given by the homodyne detection. In — the performance of the generalized Kennedy receiver, which slightly outperforms the Kennedy receiver. In — the performance of the simplified Dolinar receiver, that outperforms all the previous schemes except the actual Dolinar receiver.

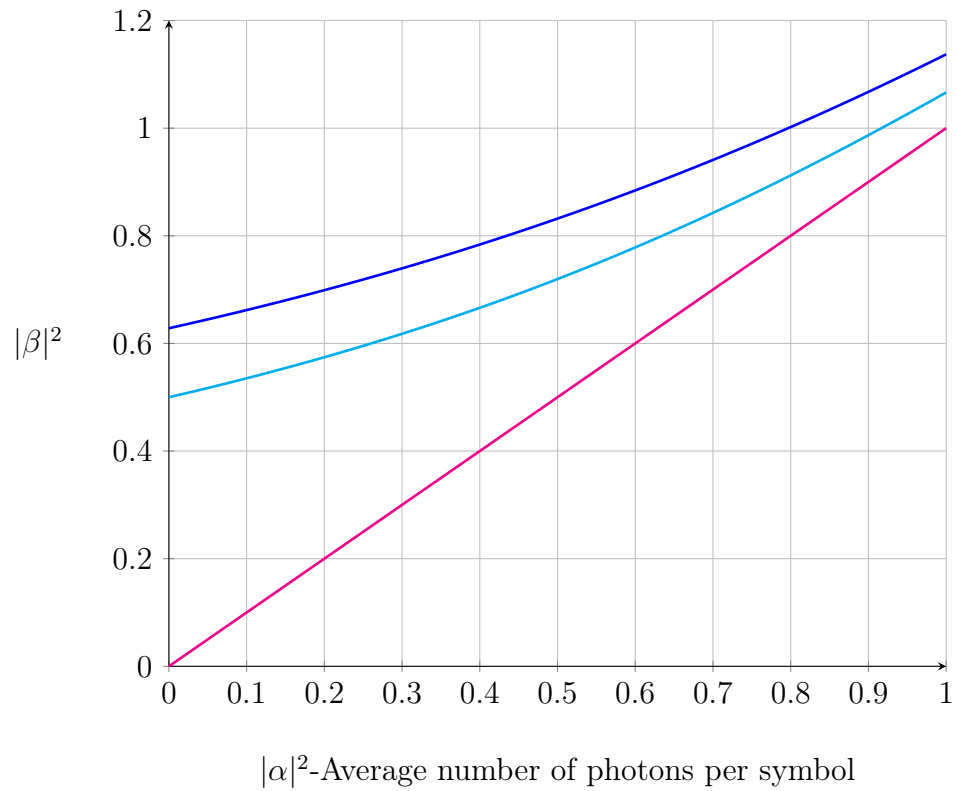


Figure 2.4: Intensity $|\beta|^2$ for the fixed displacement $D(-\beta)$ for different receiver schemes: — the Kennedy receiver, — the generalized Kennedy scheme, and — the simplified Dolinar. Equal a priori probabilities $p_0 = p_1 = 0.5$ and $T = 1$ are considered.

Chapter 3

Optimal Classical Encoding and Decoding over Noisy Qubit Channels

When the communication system includes a non-ideal quantum channel between transmitter and receiver, the transmitted quantum states are inevitably corrupted and distorted, and therefore the performance affected. The design of the communication system must involve considerations on the nature of the medium, and in order to achieve the best performance a joint optimization on the transmitted quantum states and on the measurement operators must be performed. In this chapter, we deal with this optimization in the general case of non-unital channels.

3.1 Introduction

In Chapter 2 we have studied the design of quantum receivers in the binary communication scenario, employing coherent states to represent the binary symbols. We have assumed that the quantum states transmitted by Alice would travel undisturbed to Bob, i.e. that there is an ideal channel between the two users, such that the discrimination is limited by the non-orthogonality of the coherent states.

In this chapter, we consider the problem of transmitting classical information over a *noisy* quantum channel, namely, a non-ideal channel described by a Completely Positive Trace Preserving (CPTP) map [35, 36]. In the effort of increasing the performance of the communication system, we aim at finding *optimal input states and output measurement* with respect to some performance index.

We limit our analysis to the binary case, namely where two symbol states can be transmitted,

$$x = 0, 1, \quad (3.1)$$

and two detection measurement are considered

$$\hat{x} = \hat{0}, \hat{1}. \quad (3.2)$$

In order to evaluate the quality of a digital communication system, we use the following functionals introduced in Section 1.4.1: symbol *error probability* and *channel capacity*.

3.1.1 Error Probability

In the literature, the main effort in the planning of a communication system is usually focused in the design of the receiver measurement *given* the constellation of possible transmitted quantum states. This is the case of the work by Helstrom [22], that found the optimal measurement operators given the transmitted quantum states and their a priori probability. In Section 2.1 we summarized the results of his analysis.

A different approach has been undertaken by Elron and Eldar in [37], who considered the problem of maximizing the probability of correct decision P_c when the cardinality M of the quantum states set is given, together with the symbols a priori probability and the Hilbert space with dimension N . Aiming at maximizing P_c , they do not consider any limitation on the set of allowed transmitted quantum states or measurement operator, that is equivalent to say that they assume an ideal channel between transmitter and receiver.

In the case of $N \geq M$, the problem becomes trivial, it suffices to use a set of orthogonal quantum states and employ the appropriate measurement projectors to attain perfect detection. On the contrary, when $N < M$ the quantum discrimination problem becomes non trivial.

The authors in [37] first show that, given the a priori probability $\{p_x\}$ and the set of measurement operators $\{P_{\hat{x}}\}$, the optimal quantum states ρ_x in the sense of maximizing P_c has its range included in the eigenspace of the maximal eigenvalue $\sigma_{P_{\hat{x}}}$ of the corresponding measurement operator $P_{\hat{x}}$ (necessary and sufficient condition). Then, they prove that it holds

$$P_c = \sum_{i=1}^M p_i \sigma_{P_i}. \quad (3.3)$$

Since at the transmitter side there is the possibility to assign the symbol $x = i$ to the quantum state ρ_i with just a relabeling of the i , it is convenient to associate the quantum states such that the order of $p_1 \geq p_2 \geq \dots \geq p_M$ reflects the relative order $\sigma_{P_1} \geq \sigma_{P_2} \geq \dots \geq \sigma_{P_M}$.

If $N < M$, the straightforward strategy for the encoding consists in discarding $M - N$ symbols and focus the discrimination only on the remaining N . In particular, if we consider a *tight frame setup*, that is a set of vectors $\{|\mu_i\rangle\}$ satisfying

$$\sum_{i=1}^M |\mu_i\rangle\langle\mu_i| = I_{\mathcal{H}}, \quad (3.4)$$

we can define a Tight Frame Encoding Setup (TFES) as an association between code states and measurement operators of the type

$$\begin{aligned} P_i &= |\mu_i\rangle\langle\mu_i| \\ \rho_i &= \frac{1}{\langle\mu_i|\mu_i\rangle} |\mu_i\rangle\langle\mu_i| \end{aligned} \quad (3.5)$$

with $\{|\mu_i\rangle\}$ verifying (3.4). In [37] the authors prove that all the optimal ensembles $\{\rho_i, P_i\}$ are TFES, and in particular that the maximal probability of correct decision

$$P_c = \sum_{i=1}^n p_i \quad (3.6)$$

is obtained with N quantum states defined from a TFES, associated with the first a priori probability in the ordering $p_1 \geq p_2 \geq \dots \geq p_N \geq \dots \geq p_M$.

The conclusion of the work by Elron and Eldar is that the maximum P_c can only be attained from a Tight Frame Encoding Setup, encoding the symbols with higher prior probability in N orthogonal quantum pure states that are recovered perfectly with the corresponding projector, and discarding the remaining $M - N$ symbols.

Although correct, the analysis is limited by the possibility to choose arbitrary quantum states and measurement operators in the TFES. This (implicit) assumption is equivalent to assume an ideal channel, but when a noisy non-unital channel is brought into the picture the quantum states at the receiver side are in general mixed states, with more complex constraints other than just (3.4).

In this chapter we start from the understanding of these constraints given by the channel structure, and develop a suitable reformulation of the discrimination problem that allows to jointly optimize with respect to the transmitted quantum states and the measurement operators.

3.1.2 Capacity

As we have seen in Chapter 1, the *capacity* of a channel indicates the maximal amount of information that can be reliably sent from the transmitter to the receiver. While in the classical case, in order to find the capacity a maximization over the input a priori symbol distribution is required, in the quantum case the situation is more complex depending on the possible strategies that transmitter and receiver can employ using product or entangled states and separable or entangled measurements.

An issue that comes with the maximization of Holevo capacity (1.30) and in general with the definition of the capacity is the cardinality of the symbol alphabet. The argument of the maximization considers any values of cardinality, both in the classical and in the quantum case.

However, in the quantum case, there is a maximal number n of measurement

operators that can be employed in a quantum system, fixed by its dimension d , and therefore there is a bound on the cardinality of the symbol set. This bound was first pointed out by Davies [38], who found that

$$d \leq n \leq d^2. \quad (3.7)$$

The left inequality must hold for the measurement operators in order to span all the Hilbert space \mathcal{H} of size d , while the right inequality comes from the size of the space of measurement operators over \mathcal{H} applying Caratheodory's theorem [38].

In the case of a qubit channel, the bounds (3.7) indicate that in order to achieve the Holevo capacity it may be required to employ from 2 to 4 quantum states. The question has been deeply investigated. In the case of *unital* qubit channels, that are channels where the completely mixed state is not modified by the channel (see Section 1.4.1 and eq. (1.13) for a formal definition), two orthogonal quantum states suffice. In the case of more general qubit channels, two [39], or three [40], or even four [41] non-orthogonal quantum states may be required. In reviewing all these results, Berry [42] gives simple criteria to determine, in the class of channels that require at most three inputs, whether only two states suffices, and if they need to be an orthogonal pair or not. Necessary and sufficient conditions on the channel parameters to achieve the classical capacity with only two states were also indicated in [43].

While the previous works rely on the saturation of the Holevo capacity, we pursue a different direction. If on one hand the Holevo–Schumacher–Westmoreland theorem completely describes the capacity $\mathcal{C}^{(1)}$ and in the corresponding maximization one seeks for the optimal quantum states to be transmitted, on the other hand it fails to indicate a practical way to define and implement the receiver measurement operators. In this chapter we consider the classical capacity of a binary channel induced by a qubit channel, and we look for the optimal quantum states and measurement operators to be employed. More precisely, we consider only two input quantum states and two measurement operators, and evaluate the capacity for each single use of the induced binary channel. This means that during

the transmission Alice employs product quantum states and Bob uses separable measurement operators.

While we are bound to obtain suboptimal performances, this assumption allows us to devise a constructive procedure for obtaining the optimal input states and receiver observables. Our approach to the optimization problem allows to develop some insight on the family of classical channels, represented by their transition probabilities, that can be obtained by properly engineering input states and output measurements for a given quantum channel.

3.2 Partial orderings for classical binary channels

A binary memoryless channel \mathcal{C} can be uniquely represented by its transition probability matrix

$$T_{\mathcal{C}} = \begin{bmatrix} p_{0|0} & p_{0|1} \\ p_{1|0} & p_{1|1} \end{bmatrix}$$

or, more compactly, by the pair of correct transition probabilities $(p_{1|1}, p_{0|0})$. In classical Information Theory literature, the problem of comparing discrete channels in terms of their transition probabilities has been studied extensively [44, 45, 46]. We are interested in the following (partial) orderings for binary memoryless channels

Product Ordering: in the standard product ordering, a channel \mathcal{C}' is dominated by another channel \mathcal{C} if $p'_{0|0} \leq p_{0|0}$ and $p'_{1|1} \leq p_{1|1}$.

Stochastic Degradedness [47]: a channel \mathcal{C}' is stochastically degraded with respect to another channel \mathcal{C} if \mathcal{C}' is equivalent to the cascade of \mathcal{C} , followed by a further channel \mathcal{C}'' , that is $T_{\mathcal{C}'} = T_{\mathcal{C}''}T_{\mathcal{C}}$.

Capability Ordering [47, 48]: a channel \mathcal{C}' is said to be less capable than another channel \mathcal{C} if, for any input x , by denoting with \hat{x}, \hat{x}' the corresponding outputs from $\mathcal{C}, \mathcal{C}'$, respectively, we have $H(x; \hat{x}') \leq H(x; \hat{x})$.

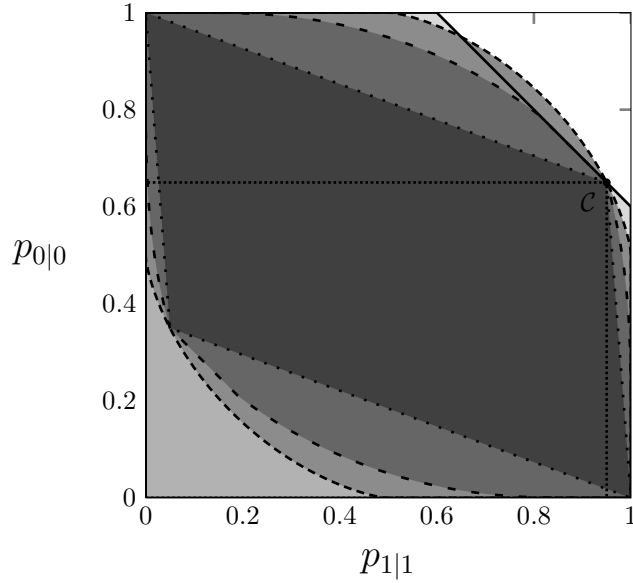


Figure 3.1: Illustration of the binary channel \mathcal{C} characterized by its correct transition probabilities $p_{1|1} = 0.95, p_{0|0} = 0.65$, and the sets of channels that: \square are dominated by \mathcal{C} ; \square have higher error rate than \mathcal{C} , with equally likely symbols; \square are stochastically degraded with respect to \mathcal{C} ; \square are less capable than \mathcal{C} ; \square have lower capacity than \mathcal{C} .

By using the representation of channels as points $(p_{1|1}, p_{0|0})$ in the unit square, Figure 3.1 shows the sets of channels that are dominated by, stochastically degraded with respect to, or less capable than some channel \mathcal{C} .

The following relations hold: i) if \mathcal{C}' is dominated by \mathcal{C} in the product ordering, then the probability of correct decision is not higher in \mathcal{C}' than in \mathcal{C} , with any input distribution; ii) if \mathcal{C}' is stochastically degraded with respect to \mathcal{C} then \mathcal{C}' is also less capable than \mathcal{C} by the data processing inequality [49]; iii) if \mathcal{C}' is less capable than \mathcal{C} , then \mathcal{C}' has a lower capacity than \mathcal{C} ; however, all the above inclusions are strict, as the converse statements are false, in general. Moreover, if we restrict our attention, without loss of generality, to the channels for which $p'_{0|0} + p'_{1|1} \geq 1$, we also have: iv) if \mathcal{C}' is dominated by \mathcal{C} in the product ordering, then \mathcal{C}' is also stochastically degraded with respect to \mathcal{C} ; v) if \mathcal{C}' is stochastically degraded with respect to \mathcal{C} , then the probability of correct decision with equally

likely inputs is not higher in \mathcal{C}' than in \mathcal{C} .

Due to the above implications, while we aim at optimizing channels in terms of correct decision probability or capacity, we will make use of both the product ordering and the stochastic degradedness notions, as they are simpler to assess in terms of the channel transition probabilities. In particular, when both $p_{0|0} + p_{1|1} \geq 1$ and $p'_{0|0} + p'_{1|1} \geq 1$, it can be seen that stochastic degradedness of a channel $(p'_{0|0}, p'_{1|1})$ with respect to another channel $(p_{0|0}, p_{1|1})$ is equivalent to the following system of inequalities

$$\begin{cases} p_{0|0}p'_{0|1} \geq p_{0|1}p'_{0|0} & (3.8) \\ p_{1|1}p'_{1|0} \geq p_{1|0}p'_{1|1} & (3.9) \end{cases}$$

that is to say that the point $(p'_{1|1}, p'_{0|0})$ lies in the triangle with vertices

$$\{(1, 0), (0, 1), (p_{1|1}, p_{0|0})\}.$$

The following statement shows that for binary channels orthogonal projectors are always optimal over all POVMs in terms of stochastic degradedness.

Proposition 3.1. *Let $\{\rho_0, \rho_1\}$ be any pair of input states, $\{P'_0, P'_1\}$ be any binary POVM, and denote by \mathcal{C}' the resulting binary channel. Then there exists a pair of orthogonal projections $\{P_0, P_1\}$ such that, denoting by \mathcal{C} the resulting binary channel with the same input states, \mathcal{C}' is stochastically degraded with respect to \mathcal{C} .*

Proof. Since $P'_0 + P'_1 = I$, it is easy to show that P'_0 and P'_1 must be simultaneously diagonalizable [50]. Choosing an appropriate basis we can thus write them as:

$$P'_0 = \text{diag}(q_a, q_b), \quad P'_1 = \text{diag}(1 - q_a, 1 - q_b)$$

with $0 \leq q_i \leq 1$, $i = a, b$. In the same basis, we can represent the channel output states $\tilde{\rho}_x = \mathcal{E}(\rho_x)$ as

$$\tilde{\rho}_0 = \begin{bmatrix} \lambda_0 & * \\ * & 1 - \lambda_0 \end{bmatrix}, \quad \tilde{\rho}_1 = \begin{bmatrix} \lambda_1 & * \\ * & 1 - \lambda_1 \end{bmatrix} \quad (3.10)$$

where entries denoted by $*$ are irrelevant for our analysis. Without loss of generality we assume $\lambda_0 \geq \lambda_1$. Moreover, we assume that $q_a \geq q_b$, so that

$$\begin{aligned} p'_{0|0} + p'_{1|1} &= \text{tr}(P'_0 \tilde{\rho}_0) + \text{tr}(P'_1 \tilde{\rho}_1) \\ &= 1 + (q_a - q_b)(\lambda_0 - \lambda_1) \geq 1. \end{aligned} \quad (3.11)$$

Now, let us consider the following projectors:

$$P_0 = \text{diag}(1, 0), \quad P_1 = \text{diag}(0, 1)$$

and observe that

$$\begin{aligned} p_{0|0} + p_{1|1} &= \text{tr}(P_0 \tilde{\rho}_0) + \text{tr}(P_1 \tilde{\rho}_1) \\ &= 1 + \lambda_0 - \lambda_1 \geq 1. \end{aligned} \quad (3.12)$$

In order to prove stochastic degradeness it is thus sufficient to prove that (3.8)-(3.9) hold. In fact,

$$\begin{aligned} p_{0|0} p'_{0|1} &= \text{tr}(P_0 \tilde{\rho}_0) \text{tr}(P'_0 \tilde{\rho}_1) = \lambda_0 (q_a \lambda_1 + q_b (1 - \lambda_1)) \\ &= q_a \lambda_0 \lambda_1 + q_b \lambda_0 - q_b \lambda_0 \lambda_1 \\ &\geq q_a \lambda_0 \lambda_1 + q_b \lambda_1 - q_b \lambda_0 \lambda_1 \\ &= \lambda_1 (q_a \lambda_0 + q_b (1 - \lambda_0)) \\ &= \text{tr}(P_0 \tilde{\rho}_1) \text{tr}(P'_0 \tilde{\rho}_0) = p_{0|1} p'_{0|0} \end{aligned}$$

where the inequality is due to $\lambda_0 \geq \lambda_1$, and similarly

$$\begin{aligned} p_{1|1} p'_{1|0} &= \text{tr}(P_1 \tilde{\rho}_1) \text{tr}(P'_1 \tilde{\rho}_0) \\ &= (1 - \lambda_1) [(1 - q_a) \lambda_0 + (1 - q_b) (1 - \lambda_0)] \\ &= (1 - \lambda_1) \lambda_0 (1 - q_a) + (1 - \lambda_0) (1 - \lambda_1) (1 - q_b) \\ &\geq (1 - \lambda_0) \lambda_1 (1 - q_a) + (1 - \lambda_0) (1 - \lambda_1) (1 - q_b) \\ &= (1 - \lambda_0) [(1 - q_a) \lambda_1 + (1 - q_b) (1 - \lambda_1)] \\ &= \text{tr}(P_1 \tilde{\rho}_0) \text{tr}(P'_1 \tilde{\rho}_1) = p_{1|0} p'_{1|1} \end{aligned}$$

which concludes the proof. \square

A similar conclusion, i.e. that orthogonal rank-1 measurement operators are optimal for the classical channel capacity functional, can be obtained as a corollary of the results presented in [38]. However, the approach we follow shows that for qubit channels this is a consequence of a stronger ordering property, namely stochastic degradeness.

A similar result holds for the probability of correct decision, as stated in the following proposition. This result can already be found in [22], but here we provide an alternative proof in the context of channel ordering on the base of proposition 3.1.

Proposition 3.2. *Let (ρ_0, ρ_1) be any pair of input states, (P'_0, P'_1) be any pair of POVM, and denote by \mathcal{C}' the resulting binary channel. For any input distribution (p_0, p_1) , there exists a pair of orthogonal projectors (P_0, P_1) such that, denoting by \mathcal{C} the resulting binary channel with the same input states, the probability of correct decision in \mathcal{C} is not lower than in \mathcal{C}' .*

Proof. Consider the pair (P_0, P_1) derived from (P'_0, P'_1) as in Proposition 3.1, yielding the transition probabilities $(p_{1|1}, p_{0|0})$. To this, add the trivial projector pairs $(I_{\mathcal{H}}, 0_{\mathcal{H}})$ and $(0_{\mathcal{H}}, I_{\mathcal{H}})$ which yield transition probabilities $(1, 0)$ and $(0, 1)$, respectively. By (3.8)-(3.9) the original $(p'_{1|1}, p'_{0|0})$ lie in the triangle of vertices $\{(1, 0), (p_{1|1}, p_{0|0}), (0, 1)\}$ and since the probability of correct decision (1.21) is a linear function of the transition probabilities, the proof follows from the fact that the extremal values of a linear function on a polytope are always found on vertices. \square

By combining Propositions 3.1 and 3.2 with implications ii) and iii) about channel orderings, it is easy to derive the following result:

Corollary 3.1. *The optimal measurements for either functional are always associated to a pair of orthogonal projectors.*

It was already recognized in [51] that the optimal measurement operators for the binary discrimination problem with respect to the error probability and the

mutual information are projectors. The authors also showed that if the output of the channel $\tilde{\rho}_x$ are pure states, the optimal projectors for the error probability coincide with the ones for the mutual information. In our work we prove the optimality of projectors in the context of stochastic degradedness, which leads to the same result, but is more general and establishes a direct link to classical channel hierarchy. Furthermore, in the next sections we shall show that the optimal projectors for the two functionals need not be the same, in general.

3.3 Coherence Vector Representation and Geometric Picture

In order to determine the optimal probabilities, it is crucial to understand how the channel transforms the transmitted quantum states, and then determine the achievable transition probabilities. Hence, we first focus on the characterization of the region of achievable transition probabilities within the unit square.

For our purpose, it is convenient to use a particular choice of basis for representing 2×2 complex matrices, associated to the unitary, self-adjoint operators $\{I, \sigma_x, \sigma_y, \sigma_z\}$, where σ_i are the Pauli operators, also called *coherence vector representation*. Input and output quantum states can then be represented as

$$\rho_x = \frac{1}{2}(I_{\mathcal{H}} + \vec{\rho}_x \cdot \vec{\sigma}), \quad \tilde{\rho}_x = \frac{1}{2}(I_{\mathcal{H}} + \vec{\tau}_x \cdot \vec{\sigma}) \quad (3.13)$$

where for any $\vec{v} \in \mathbb{R}^3$, $\vec{v} \cdot \vec{\sigma}$ is the shorthand notation for the linear combination of Pauli matrices

$$\vec{v} \cdot \vec{\sigma} = \vec{v}(1) \sigma_x + \vec{v}(2) \sigma_y + \vec{v}(3) \sigma_z. \quad (3.14)$$

All the valid (i.e., unit-trace, positive-semidefinite) states are associated to coherence vectors \vec{v} in the unit (Bloch's) sphere, with pure states lying on the surface [36].

As we showed in the previous section, the optimal choice of measurements for Bob is represented by a pair of projectors $\{P_0, P_1\}$. Leaving aside the trivial

projector pairs $\{I_{\mathcal{H}}, 0_{\mathcal{H}}\}$, we consider rank-1 projectors that admit coherence representation

$$P_{\hat{x}} = \frac{1}{2}(I_{\mathcal{H}} + \vec{\pi}_{\hat{x}} \cdot \vec{\sigma}) \quad (3.15)$$

with $\vec{\pi}_{\hat{x}}$ lying on the sphere surface, where the completeness relation $P_0 + P_1 = I_{\mathcal{H}}$ implies the constraint

$$\vec{\pi}_0 = -\vec{\pi}_1. \quad (3.16)$$

The qubit channel is described by a TPCP map \mathcal{E} acting on a two level system \mathcal{H} . In coherence vector representation, any TPCP map has an affine form [52]

$$\vec{\tau}_x = A \vec{\rho}_x + \vec{b}, \quad (3.17)$$

with A being a 3×3 real matrix associated to a not necessarily strict contraction, and \vec{b} a vector in the unit ball corresponding to the image of the completely mixed state through the channel. Geometrically, this means that the image E of the Bloch sphere S is an ellipsoid: if $A = U S V^T$ is the SVD decomposition of A , S is first rotated by V , squeezed along its axes by S , rotated again by U and then shifted by \vec{b} . Note that, as it has been pointed out in [52], not all maps of the form (3.17) mapping the Bloch sphere into itself yield a physical (i.e. CP) channel. Since in our work the channel is assumed to be physical and known, this is not a concern.

Any channel can be reduced, via change of basis for $\vec{\rho}$ and $\vec{\tau}$, to the case of a diagonal A ,

$$A = S = \text{diag}(a, b, c). \quad (3.18)$$

In fact, we can define

$$\vec{\psi} = V^T \vec{\rho} \quad (3.19)$$

$$\vec{\phi} = U^T \vec{\tau} \quad (3.20)$$

$$\vec{\xi} = U^T \vec{b} \quad (3.21)$$

so that (3.17) becomes

$$\vec{\phi} = S \vec{\psi} + \xi. \quad (3.22)$$

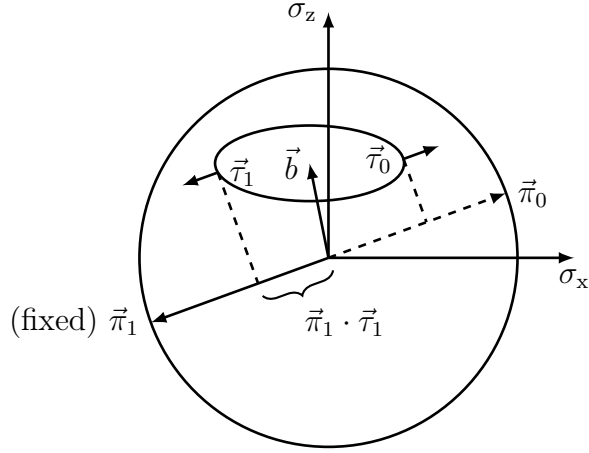


Figure 3.2: Inner product between $\vec{\pi}_1$ and $\vec{\tau}_1$, with the Bloch sphere projected onto the $\{\sigma_x, \sigma_z\}$ plane. For a *fixed* $\vec{\pi}_1$, the point $\vec{\tau}_1$ on the ellipsoid that maximizes $\vec{\pi}_1 \cdot \vec{\tau}_1$ has normal vector to the surface which is parallel to $\vec{\pi}_1$ (here depicted unnormalized).

From $\vec{\psi}$, we can get $\vec{\rho}$ of the original coordinate system by inversion of (3.19). In the coordinate system where A is diagonal, the ellipsoid E has axes parallel to those of the standard (x, y, z) coordinate system.

3.3.1 Optimization of input states for given projectors

Transition probabilities can be written in terms of the coherent representation of states and projectors by using the inner product in \mathbb{R}^3

$$\begin{aligned} p_{1|1} &= \frac{1 + \vec{\pi}_1 \cdot \vec{\tau}_1}{2}, \\ p_{0|0} &= \frac{1 + \vec{\pi}_0 \cdot \vec{\tau}_0}{2} = \frac{1 - \vec{\pi}_1 \cdot \vec{\tau}_0}{2}. \end{aligned} \quad (3.23)$$

If, as is often the case, $a, b, c < 1$ no point of the ellipsoid lies on the sphere surface. Consequently it is not possible to have $\vec{\pi}_{\hat{x}} \cdot \vec{\tau}_x = 1$, $\hat{x}, x = 0, 1$ and the region \mathcal{V} of admissible transition probability is strictly contained in the unit square.

Proposition 3.3. *Let $(\vec{\pi}_0, \vec{\pi}_1)$ be the coherence vector representation of a pair of orthogonal projectors. Denote with $\vec{\tau}_0, \vec{\tau}_1$ the points on the surface of E where*

the normal vector to the surface is parallel to $\vec{\pi}_0, \vec{\pi}_1$ respectively. Then the binary channel associated with $(\vec{\tau}_0, \vec{\tau}_1)$ dominates with respect to product ordering all the binary channel associated with other states pairs in the ellipsoid.

Proof. Consider $\vec{\pi}_1$ fixed as shown in Figure 3.2. By standard results in constrained optimization [53], the output vector $\vec{\tau}_1$ that corresponds to the maximum $p_{1|1}$ must identify a point on the surface of E , with normal vector parallel to $\vec{\pi}_1$. In fact, if we consider a plane normal to $\vec{\pi}_1$, all the points in the intersection with the ellipsoid correspond to vectors $\vec{\tau}$ with equal inner product with $\vec{\pi}_1$. Hence, they give the same transition probability $p_{1|1}$. Among the planes that are orthogonal to $\vec{\pi}_1$, the one that maximizes the inner product is thus the plane tangent to E and closer to $\vec{\pi}_1$. Analogously for $\vec{\tau}_0$. \square

Recalling that $\vec{\pi}_0 = -\vec{\pi}_1$, the vector $\vec{\tau}_0$ that maximizes $\vec{\pi}_0 \cdot \vec{\tau}_0$ is then the point on the surface of E with normal vector $-\vec{\pi}_1$, and is the “antipodal” point of $\vec{\tau}_1$ in the ellipsoid, that is

$$\vec{\tau}_0 + \vec{\tau}_1 = 2\vec{b}. \quad (3.24)$$

Note that the antipodal condition (3.24) on $\vec{\tau}_0, \vec{\tau}_1$ implies that the corresponding input vectors $\vec{\rho}_0, \vec{\rho}_1$ are also antipodal, on the Bloch sphere, meaning that the quantum states ρ_0, ρ_1 are orthogonal. Since (3.24) is a necessary condition for the optimization, we have the following result

Corollary 3.2. *If the optimal projectors $\{P_0, P_1\}$ have rank 1, the optimal quantum states to transmit for either functionals are orthogonal.*

The fact that an orthogonal alphabet of quantum states is a necessary condition for optimality for the classical channel capacity functional can already be found in [39], however here we examine in depth the optimization, deriving the relations between the optimal transmitted quantum state and the optimal receiver measurements.

We now derive the explicit expression of $\vec{\tau}'_0, \vec{\tau}'_1$ in an appropriate coordinate system and evaluate the corresponding transition probabilities starting from P_0, P_1 .

The origin of the coordinate system is the center of S and its axes are parallel to those of E . We can parametrize the point $\vec{\pi}_1$ on the surface of S by the angles $\alpha \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, $\beta \in [0, 2\pi)$ and the point $\vec{\tau}_1$ on the surface of E with $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, $\psi \in [0, 2\pi)$,

$$\vec{\pi}_1 = \begin{bmatrix} \cos \alpha \cos \beta \\ \cos \alpha \sin \beta \\ \sin \alpha \end{bmatrix}, \quad \vec{\tau}_1 = \begin{bmatrix} a \cos \theta \cos \psi + b_x \\ b \cos \theta \sin \psi + b_y \\ c \sin \theta + b_z \end{bmatrix}. \quad (3.25)$$

In order to find the desired $\vec{\tau}_1$, in Appendix 3.A we show that setting its gradient equal to $\vec{\pi}_1$, leads to the following conditions,

$$\begin{aligned} a \tan \psi &= b \tan \beta \\ \tan \theta \sqrt{a^2 \cos^2 \beta + b^2 \sin^2 \beta} &= c \tan \alpha \end{aligned} \quad (3.26)$$

and the resulting inner product is

$$\begin{aligned} \vec{\pi}_1 \cdot \vec{\tau}_1 &= (a \cos \theta \cos \psi + b_x) \cos \alpha \cos \beta + (b \cos \theta \sin \psi + b_y) \cos \alpha \sin \beta \\ &\quad + (c \sin \theta + b_z) \sin \alpha \\ &= \sqrt{a^2 \cos^2 \alpha \cos^2 \beta + b^2 \cos^2 \alpha \sin^2 \beta + c^2 \sin^2 \alpha} + \vec{\pi}_1 \cdot \vec{b}. \end{aligned} \quad (3.27)$$

Similarly, by (3.24),

$$\begin{aligned} \vec{\pi}_1 \cdot \vec{\tau}_0 &= -\vec{\pi}_1 \cdot (2\vec{b} - \vec{\tau}_1) = \vec{\pi}_1 \cdot \vec{\tau}_1 - 2\vec{\pi}_1 \cdot \vec{b} \\ &= \sqrt{a^2 \cos^2 \alpha \cos^2 \beta + b^2 \cos^2 \alpha \sin^2 \beta + c^2 \sin^2 \alpha} - \vec{\pi}_1 \cdot \vec{b}. \end{aligned} \quad (3.28)$$

3.3.2 Region of achievable transition probabilities

In this section we characterize the set of transition probabilities obtained as the vector $\vec{\pi}_1$ moves on the surface of the Bloch sphere, employing necessary conditions (3.26) for the optimal quantum states $(\vec{\tau}_0, \vec{\tau}_1)$ out of the channel.

Let us define the set \mathcal{V} in the unit square containing transition probabilities pair $(p'_{1|1}, p'_{0|0})$ corresponding to the binary channel \mathcal{C}' given by generic POVM and generic quantum states. Region \mathcal{V} shows evident properties of symmetry, with respect to the bisecting line $p_{1|1} = p_{0|0}$ and the anti-bisecting line $p_{1|1} + p_{0|0} = 1$.

In fact, given a pair $(p'_{1|1}, p'_{0|0}) \in \mathcal{V}$ obtained from the POVM pair (P'_0, P'_1) with the received quantum states $(\tilde{\rho}_0, \tilde{\rho}_1)$, swapping the measurement operators pair and the transmitted quantum states we obtain respectively transition probabilities $(1 - p'_{1|1}, 1 - p'_{0|0})$ and $(1 - p'_{0|0}, 1 - p'_{1|1})$, that are the points symmetrical to $(p'_{1|1}, p'_{0|0})$ with respect the central point $(0.5, 0.5)$ and with respect to line $p_{1|1} + p_{0|0} = 1$. Combining both swaps, we get the symmetry with respect to the bisecting line.

Proposition 3.4. *For all $k \in [-\|\vec{b}\|, \|\vec{b}\|]$, there exist a $\vec{\pi}_1$ with $\vec{\pi}_1 \cdot \vec{b} = k$ such that the channel \mathcal{C} with $\vec{\pi}_0 = -\vec{\pi}_1$ and $(\vec{\tau}_0, \vec{\tau}_1)$ as given by (3.26), dominates all the channels \mathcal{C}' similarly associated to any $\vec{\pi}'_1$ such that $\vec{\pi}'_1 \cdot \vec{b} = k$. Such $\vec{\pi}_1$ is given by*

$$\vec{\pi}_1 = \operatorname{argmax}_{\vec{\pi}'_1 \in S, \vec{\pi}'_1 \cdot \vec{b} = k} \max_{\vec{\tau}_1 \in E} \vec{\pi}'_1 \cdot \vec{\tau}_1, \quad (3.29)$$

where the solution of the inner maximization problem is given by (3.26).

Proof. From (3.24), we can rewrite transition probability $p_{0|0}$ (3.23) as

$$p_{0|0} = \frac{1 + \vec{\pi}_1 \cdot \vec{\tau}_1}{2} - \vec{\pi}_1 \cdot \vec{b} \quad (3.30)$$

so that with $\vec{\pi}_1 \cdot \vec{b}$ fixed, (3.29) maximizes both $p_{1|1}$ and $p_{0|0}$. \square

Figure 3.3 illustrates the relationship between the region of achievable transition probabilities when considering arbitrary POVM and transmitted states, orthogonal projectors and optimal states, respectively.

On the basis of Propositions 3.1–3.4, all the channels that are *maximal*¹ with respect to the stochastic degradedness ordering satisfy equation (3.29). We thus propose the following procedure for an efficient evaluation of \mathcal{V} :

1. For each value $k = \vec{\pi}_1 \cdot \vec{b}$, $k \in [0, \|\vec{b}\|_2]$, solve (3.29). This problem is equivalent to a quadratic problem with quadratic constraints (see Appendix

¹In a set with a partial ordering, an element is said to be *maximal* if it does not precede any other element.

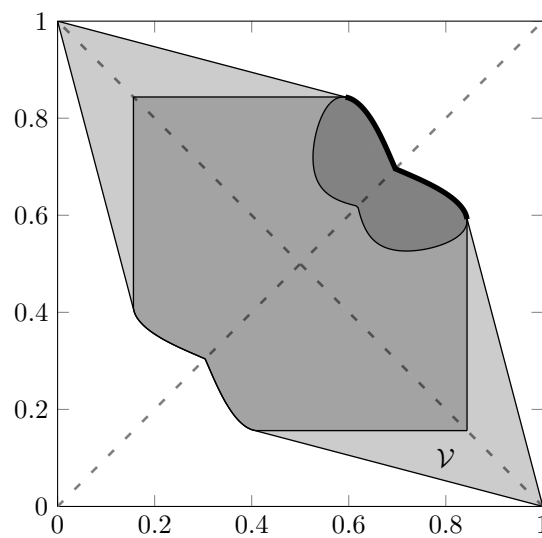



Figure 3.3: Region \mathcal{V} of admissible transition probabilities, for $A = \text{diag}([.35, .45, .20])$, $\vec{b} = [.15, .20, .10]$. Filled with \square , the set of binary channels $\{p_{1|1}, p_{0|0}\}$ with POVM and arbitrary quantum states. In \square , the set of binary channels with rank 1 projector measurement operator and arbitrary quantum states. Filled with \blacksquare , the set of binary channels with projector and optimal quantum states by (3.26). Line — is composed by the maximal binary channels.

3.B), that has no closed form solution but that can be easily solved via standard numerical methods [53].

2. Mirror the previous border with respect to the bisecting line, obtaining the set of channels

$$\mathcal{B}_{\mathcal{V}} = \{(p_{1|1}, p_{0|0}) \mid \vec{\pi}_1 \text{ solution of (3.29)}, k \in [-\|\vec{b}\|_2, \|\vec{b}\|_2]\}, \quad (3.31)$$

which contains all the maximal binary channels. Mirror $\mathcal{B}_{\mathcal{V}}$ with respect to the anti-bisecting line by symmetry relations.

3. Connect the edges to the points $(p_{1|1}, p_{0|0}) = (1, 0)$ and $(p_{1|1}, p_{0|0}) = (0, 1)$, and sweep all the region in between. The region \mathcal{V} is depicted in Figure 3.3 filled with .

3.4 Optimization and Numerical Methods

3.4.1 Probability of correct decision

In the case of probability of correct decision, we consider the a priori symbol probabilities p_0, p_1 as given. By exploiting the geometric representation of the previous section, we can rewrite the problem so that we can obtain a solution via standard numerical methods, as well as an insightful geometrical picture. Combining the definition of P_c with the relation of completeness, we get

$$P_c = (1 - p_1) + \text{tr}(P_1(p_1\tilde{\rho}_1 - p_0\tilde{\rho}_0)). \quad (3.32)$$

Following Helstrom [22], in order to find optimal solution for the problem of quantum binary discrimination it is convenient to introduce the difference operator $\Delta = p_1\tilde{\rho}_1 - p_0\tilde{\rho}_0$. We now use the coherence vector representation for P_1 , $\tilde{\rho}_1$, $\tilde{\rho}_0$ to get

$$\Delta = \frac{1}{2} \left((2p_1 - 1)I_{\mathcal{H}} + \vec{d} \cdot \vec{\sigma} \right),$$

$$\vec{d} = p_1\vec{\tau}_1 - p_0\vec{\tau}_0 = 2p_1\vec{b} - \vec{\tau}_0 := (d_x, d_y, d_z)^T$$

where the antipodal condition (3.24) has been used. From (3.32), we see that the optimal P_1 is the projection on the eigenspace of Δ associated to positive eigenvalues. The eigenvalues of $(2p_1 - 1)I + \vec{d} \cdot \sigma$ are

$$\begin{aligned}\lambda_0 &= \frac{2p_1 - 1 + \|\vec{d}\|_2}{2}, \\ \lambda_1 &= \frac{2p_1 - 1 - \|\vec{d}\|_2}{2}.\end{aligned}\tag{3.33}$$

Depending upon p_1 , $\vec{\pi}_1$, $\vec{\tau}_0$, the eigenvalues may be both positive, both negative or opposite in sign. Consequently, P_1 may be respectively the identity, the null observable on \mathcal{H} or a rank 1 projector as in Proposition 3.2. If P_1 is the identity or the null observable, it results

$$P_c = \max \{p_0, p_1\}\tag{3.34}$$

and performing a measurement does not increase the probability of correct discrimination with respect to our a priori information. In the case $2p - 1 < \sqrt{d_x^2 + d_y^2 + d_z^2}$, instead, we have $\lambda_0 > 0$ and $\lambda_1 < 0$. We rewrite

$$P_c = \frac{1}{2} + \|p_1 \tilde{\rho}_1 - p_0 \tilde{\rho}_0\|_1 = \frac{(1 + \vec{\pi}_1 \cdot \vec{d})}{2}.\tag{3.35}$$

This expression gives an immediate meaning to the optimal $\vec{\pi}_1$, which must be parallel to \vec{d} , and highlights that, in order to maximize P_c , \vec{d} must be taken of the maximum possible length. Hence, by (3.35) and (3.24), the optimization of P_c results to be the quadratic problem

$$\vec{d}_{opt} := \operatorname{argmax}_{\vec{\tau}_0 \in E} \|2p_1 \vec{b} - \vec{\tau}_0\|_2,\tag{3.36}$$

where $\|\cdot\|_2$ is the Euclidean norm. Numerical methods for convex optimization are well known [53] and can be employed to solve the quadratic problem (3.36) with quadratic constraints. Altogether, in the end we get

$$P_c = \max \left\{ \frac{(1 + \|\vec{d}_{opt}\|_2)}{2}, p_0, p_1 \right\}.\tag{3.37}$$

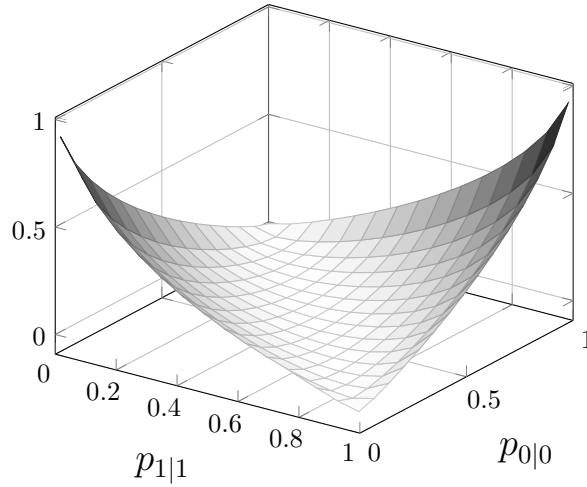


Figure 3.4: Classical capacity as a function of $p_{1|1}$ and $p_{0|0}$.

3.4.2 Capacity

Maximization of mutual information (1.24) requires optimization over many parameters: the a priori probability distribution p_x , the input states ρ_x and the receiver measurement $P_{\hat{x}}$. Also, due to nonlinear terms, explicit solutions are difficult to find. Instead, numerical maximization is viable thanks to convexity of the mutual information (see Figure 3.4).

From the geometric representation of states and measurement projectors, we can optimize (1.24) by a search over the region \mathcal{V} . In fact, we can split the maximization problem of (1.24) into two optimization problems: an inner maximization with respect to the *a priori* probability, and an outer maximization with respect to the transition probabilities:

$$\mathcal{C}_{max} = \max_{(p_{1|1}, p_{0|0}) \in \mathcal{V}} \max_{p_x} H(x; \hat{x}) \quad (3.38)$$

The inner maximization of (3.38) has an analytic closed form solution. Consider $(p_{1|1}, p_{0|0})$ as fixed, define the binary entropy

$$h_2(x) = -x \log_2 x - (1-x) \log_2(1-x). \quad (3.39)$$

with derivative

$$h'_2(x) = \log_2 \left(\frac{1}{x} - 1 \right), \quad (3.40)$$

and inverse function of the derivative

$$g(x) = (h'_2)^{-1} = \frac{1}{2^x + 1}. \quad (3.41)$$

We rewrite the mutual information as

$$\begin{aligned} H(x; \hat{x}) &= H(\hat{x}) - H(\hat{x}|x) \\ &= H(\hat{x}) - H(\hat{x}|x=1) p_1 - H(\hat{x}|x=0) p_0 \\ &= h_2(r) - h_2(p_{1|1}) p_1 - h_2(p_{0|0}) (1 - p_1), \end{aligned} \quad (3.42)$$

where $r = P[\hat{x}=0] = p_{0|0}(1 - p_1) + (1 - p_{1|1})p_1$. If we take the derivative of the above with respect to p_1 , we get

$$\begin{aligned} \frac{dH(x; \hat{x})}{dp_1} &= h'_2(r) \cdot \frac{dr}{dp_1} - h_2(p_{1|1}) + h_2(p_{0|0}) \\ &= h'_2(r)(1 - p_{1|1} - p_{0|0}) - h_2(p_{1|1}) + h_2(p_{0|0}). \end{aligned}$$

By imposing it be equal to zero, we have:

$$h'_2(r) = \frac{h_2(p_{1|1}) - h_2(p_{0|0})}{1 - p_{1|1} - p_{0|0}}. \quad (3.43)$$

Hence, by definition

$$\begin{aligned} r &= g\left(\frac{h_2(p_{1|1}) - h_2(p_{0|0})}{1 - p_{1|1} - p_{0|0}}\right), \\ p_1 &= \frac{r - p_{0|0}}{1 - p_{1|1} - p_{0|0}}, \end{aligned} \quad (3.44)$$

and the classical capacity is obtained if we substitute these values in (3.42).

Outer optimization with respect to $p_{1|1}, p_{0|0}$ can be performed on the edge of \mathcal{V} employing standard tools from constrained optimization. In the light of symmetry considerations, two or four optimal solutions can be found, or even a continuous arc of optimal solutions can be obtained if these points lay on a contour line.

Since a numerical optimization is required to solve (3.29) and hence to solve the outer optimization, we can do the maximization altogether: for a certain

$k \in [0, \|\vec{b}\|_2]$, first solve (3.29) to find the optimal $\vec{\pi}_1$, get $\vec{\tau}_1$ by (3.26) and $\vec{\tau}_0$ by (3.24). From $\vec{\pi}_1$, $\vec{\tau}_1$ and $\vec{\tau}_0$ obtain $(p_{1|1}, p_{0|0})$ by (3.23). With these transition probabilities, find p_1 from (3.44) and get $H(x; \hat{x})$. These steps need to be repeated for a finite set of values $\{k\}$ that discretizes $[0, \|\vec{b}\|_2]$, and by direct comparison we can get the maximum $H_k(x; \hat{x})$.

Of course, since a numerical procedure is required, the discretization of the range $[0, \|\vec{b}\|_2]$ is necessary in order to calculate and compare the values $H_k(x; \hat{x})$ for different k . However, due to smoothness nature of the functional, it is assured that it is possible to find a solution $H_{\bar{k}}$ arbitrarily close to the true one, i.e.

$$\forall \epsilon > 0 \quad \exists N, \bar{k} > 0 \quad \text{s.t.} \quad |H_{\bar{k}} - H_k| < \epsilon \quad (3.45)$$

where $H_{\bar{k}}$ is the maximal mutual information obtained on the N -step discretization of the range, while H_k is the true optimum on $[0, \|\vec{b}\|_2]$.

3.5 Examples

The procedure explained in Section 3.3 and described in more details in Appendix 3.B, allows us to find the region of transition probabilities given a description of the physical channel as in (3.17).

Once the numerical routine for the border calculation is set, we performed a Monte Carlo simulation on the parameters of A and \vec{b} , and discovered different region shapes induced by the corresponding channels. Of course, not all possible choices of entries of A and \vec{b} define a physical channels, so we have to check the necessary and sufficient conditions given in [52].

Figure 3.5 shows different kinds of shapes that can be obtained from the channel. The corresponding A and \vec{b} are reported in the caption. Both convex, e.g. region 1 and region 5, and concave, e.g. regions 2,3,4, are possible, with different types of concave border. Since there's no analytical solution for the borders, we cannot find a clear dependence between the ellipsoid parameters and the shape of \mathcal{V} . However, we can develop some intuitions and qualitative analysis on the shape of the region \mathcal{V} and on the position of the optimal points.

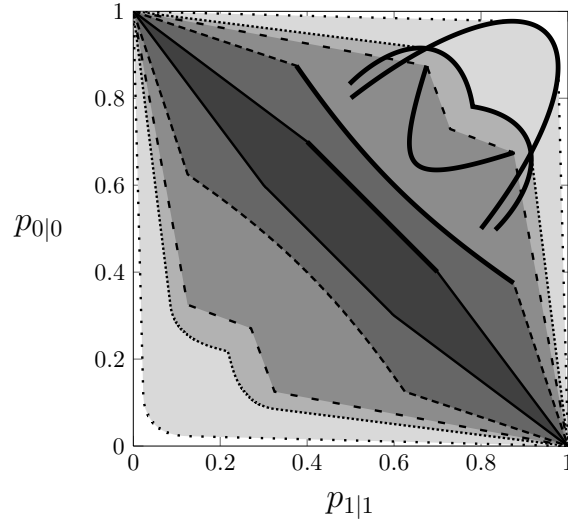







Figure 3.5: Different shapes of the border of \mathcal{V} . The corresponding $\mathcal{B}_{\mathcal{V}}$ are highlighted in thick black line. In the Table below it is reported the associated ellipsoid parameters. Notice that case 5 presents a border parallel to the anti-bisecting line, leading to a continuum of optimal points with respect to the error probability.

case	line	A	\vec{b}
1		$\text{diag}([0.3, 0.3, 0.9])$	$[0.3, 0, 0]$
2		$\text{diag}([0.2, 0.1, 0.62])$	$[0.3, 0, 0.15]$
3		$\text{diag}([0.55, 0.3, 0.3])$	$[0.2, 0, 0]$
4		$\text{diag}([0.25, 0.25, 0.2])$	$[0.5, 0, 0]$
5		$\text{diag}([0.1, 0.1, 0.1])$	$[0.3, 0, 0]$

For some particular cases we can interpret the shape of the region in the light of the ellipsoid parameters. For example, in the case of an ellipsoid with equal radii (sphere), we can see that due to the symmetry, the set of the binary channel with rank-1 projectors and arbitrary states is a stripe along the anti-bisecting line, whose thickness depends on the radius and whose length depends on the norm of \vec{b} . Another peculiar case is when the channel is unital and $\vec{b} = 0$. In this case the set of the binary channel with rank-1 projectors becomes a square centered in $(0.5, 0.5)$, with the side depending on the longest radius.

As seen in Section 3.4, the transition probabilities maximizing the probability of correct decision or the mutual information lie on the border of \mathcal{V} . While the contour curves of P_c are straight lines with slope depending on the a priori probabilities, the contour lines of \mathcal{C} are bent. In general, the optimal transition probabilities differ depending upon the functional considered.

In cases 1, 3, 4, 5 depicted in Figure 3.5, either the solutions coincide, or at least a pair of coinciding solutions exists. Notice that whenever the region \mathcal{V} presents a border parallel to the anti-bisecting line (as in case 5), the optimization of error probability with equally likely inputs leads to a continuum of solutions on the segment of the border.

Case 2 is clarified in Figure 3.6, where three different regions \mathcal{V} (in solid, dashed and dotted line) are depicted, and the optimal points for the different functionals exhibit significant difference. In the background (in thin grey solid lines) the contour lines of the mutual information are drawn to illustrate why the optimal points do not coincide: the local curvature of the border of \mathcal{V} is lower than the curvature of the mutual information contour line. In doing this comparison, we consider an equal a priori probability for the probability of correct decision.

This particular situation can arise for both concave and convex regions. However, a convex region with the point of maximal classical capacity along the bisecting line has the same point of maximal probability of correct decision with equal a priori probability.

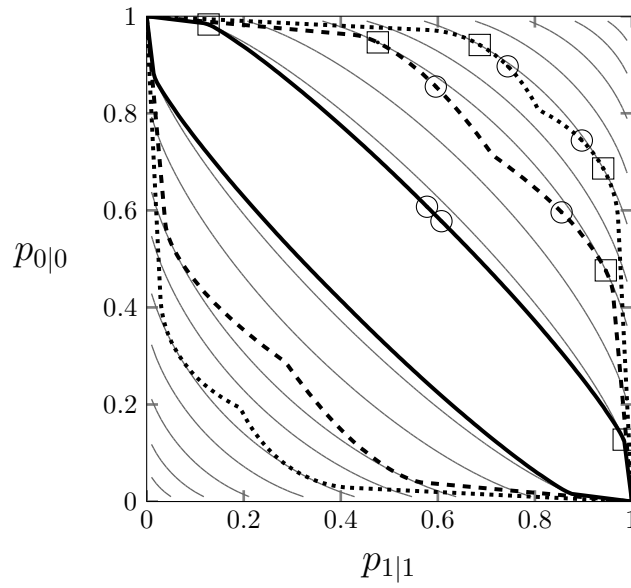


Figure 3.6: Examples of regions \mathcal{V} where the point of maximal mutual information \square does not coincides with the point of maximal probability of correct decision \circ . In the background, contour lines of mutual information maximized with respect to p_x .

line	A	\vec{b}	p_1
—	diag([0.14, 0.07, 0.19])	[0.46, 0.74, 0.03]	0.57
- - - - -	diag([0.34, 0.24, 0.45])	[-0.42, -0.27, -0.26]	0.55
.....	diag([0.11, 0.64, 0.07])	[-0.24, -0.15, 0.45]	0.54

Appendix

3.A Proof of the necessary condition for optimality

Consider the coherence vector representation for the quantum states in input and output of the channel $\vec{\rho}, \vec{\tau}$ and for the measurement operators $\vec{\pi}$. The affine relation between $\vec{\rho}$ and $\vec{\tau}$,

$$\vec{\tau} = A\vec{\rho} + \vec{b}, \quad (3.46)$$

maps the Bloch ball surface associated to

$$\vec{\rho}^T \vec{\rho} = 1 \quad (3.47)$$

into the ellipsoide E with equation

$$(\vec{\tau} - \vec{b})^T A^{-T} A^{-1} (\vec{\tau} - \vec{b}) = 1. \quad (3.48)$$

The (unnormalized) normal vector to the surface of E in the point located by $\vec{\tau}$ can be written as

$$\vec{\nabla}_{\vec{\tau}} = 2A^{-T} A^{-1} (\vec{\tau} - \vec{b}), \quad \vec{\tau} \in E. \quad (3.49)$$

In order to find the point in the ellipsoid with normal vector equal to $\vec{\pi}$, we set

$$\frac{A^{-T} A^{-1} (\vec{\tau} - \vec{b})}{\|A^{-T} A^{-1} (\vec{\tau} - \vec{b})\|} = \vec{\pi}. \quad (3.50)$$

After substituting (3.46), the expression becomes

$$\frac{A^{-T} \vec{\rho}}{\|A^{-T} \vec{\rho}\|} = \vec{\pi}. \quad (3.51)$$

In particular, by inversion of (3.51) we obtain the following relation

$$\vec{\rho}^T \vec{\rho} = 1 = \|\vec{\rho}\|^2 = \|A^{-T} \vec{\rho}\|^2 \|A^T \vec{\pi}\|^2, \quad (3.52)$$

and hence

$$\vec{\rho} = \|A^{-T} \vec{\rho}\| A^T \vec{\pi} = \frac{A^T \vec{\pi}}{\|A^T \vec{\pi}\|}, \quad (3.53)$$

which is equivalent to the necessary condition (3.26) between the projector $\vec{\pi}_1$ and the quantum state $\vec{\tau}_1$.

In addition, if we evaluate the inner product we get

$$\vec{\pi} \cdot \vec{\tau} = \vec{\pi} \cdot (A\vec{\rho} + \vec{b}) = \vec{\pi} \cdot A\vec{\rho} + \vec{\pi} \cdot \vec{b} \quad (3.54)$$

and using (3.51) and (3.52), we get

$$\vec{\pi} \cdot A\vec{\rho} = \frac{\vec{\rho}^T A^{-1} A \vec{\rho}}{\|A^{-T} \vec{\rho}\|} = \frac{\|\vec{\rho}\|^2}{\|A^{-T} \vec{\rho}\|} = \|A^T \vec{\pi}\|. \quad (3.55)$$

We finally obtain (3.27), that is

$$\vec{\pi} \cdot \vec{\tau} = \|A^T \vec{\pi}\| + \vec{\pi} \cdot \vec{b}. \quad (3.56)$$

3.B Quadratic Optimization Problems with Quadratic Constraints

Consider the problem (3.29), and define the cost function

$$f := \vec{\pi}_1 \cdot (\vec{\tau}_1 - \vec{b}) = \vec{\pi}_1 \cdot \Delta \vec{\tau}_1 \quad (3.57)$$

and the constraints

$$\vec{\tau}_1 \in E, \quad (3.58)$$

$$\vec{\pi}_1 \in S, \quad (3.59)$$

$$\vec{\pi}_1 \cdot \vec{b} = k, \quad (3.60)$$

for $k \in [0, \|\vec{b}\|_2]$. According to definitions (3.25), the maximization of (3.57) with constraints (3.58)-(3.60) requires an optimization with respect to four variables,

i.e. α , β , θ and ψ . As already pointed out previously in Section 3.3, the constraint (3.58) and the geometric interpretation of the optimization allow us to obtain the necessary conditions (3.26), so that we can substitute θ , ψ in terms of α , β , to get an expression of the cost function f similar to (3.27), i.e.

$$\tilde{f} := \sqrt{a^2 \cos^2 \alpha \cos^2 \beta + b^2 \cos^2 \alpha \sin^2 \beta + c^2 \sin^2 \alpha}. \quad (3.61)$$

Alternatively, if we substitute α , β in terms of θ , ψ we obtain

$$\tilde{f}' := \frac{abc}{\sqrt{a^2 b^2 \sin^2 \theta + b^2 c^2 \cos^2 \theta \cos^2 \psi + a^2 c^2 \cos^2 \theta \sin^2 \psi}}. \quad (3.62)$$

Depending on the choice of the variables, the optimization problem becomes the maximization or minimization of the square root term in \tilde{f} or \tilde{f}' . Also, we can simplify the formulation using \tilde{f}^2 or $(\tilde{f}')^2$ as functional, in order to get rid of the square root term.

The constraint (3.59), that can be rewritten as

$$\vec{\pi}_1^T \vec{\pi}_1 = 1, \quad (3.63)$$

has intersection with the plane (3.60) that defines a circle on S as region of optimization.

If we consider the points defined by the variables α, β in (3.25), the cost function \tilde{f} can be interpret as the norm of vector $\vec{\pi}_1$ in a non-normal coordinate system, and the problem

$$(\vec{\pi}_1)_{\max}(k) = \underset{\vec{\pi}_1^T \vec{\pi}_1 = 1, \vec{\pi}_1 \cdot \vec{b} = k}{\operatorname{argmax}} \tilde{f} \quad (3.64)$$

becomes a quadratic problem with quadratic constraints. The same type of problem can be obtained if we consider the square root term of \tilde{f}' to be minimized.

We develop a reformulation of (3.64) into the problem of finding the farthest point of an ellipse from a given point. We test this approach comparing it with other numerical optimization techniques, such as a “brute force” algorithm, that discretizes the ellipse and look for the best $\vec{\pi}_1$ in the discretization, and general numerical constrained optimization, that includes the constraints in the functional and finds the maximum with numerical iterative methods.

First, consider problem (3.64) not as function of variables α, β but as function of the coordinates system $\vec{\pi}_1 = [x, y, z]^T$ defined by (3.25). The cost function of the problem becomes

$$\tilde{f}^2 = a^2x^2 + b^2y^2 + c^2z^2, \quad (3.65)$$

with constraints for a given $k \in [0, \|\vec{b}\|_2]$

$$x^2 + y^2 + z^2 = 1, \quad (3.66)$$

$$x b_x + y b_y + z b_z = k. \quad (3.67)$$

The coordinates need first to be normalized ², with

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{1}{a} & 0 & 0 \\ 0 & \frac{1}{b} & 0 \\ 0 & 0 & \frac{1}{c} \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} := H_1 \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}, \quad (3.68)$$

and substitute z_1 by the constraint (3.67):

$$\begin{aligned} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{cb_x}{ab_z} & -\frac{cb_y}{bb_z} & 0 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \frac{ck}{b_z} \end{bmatrix} \\ &:= H_2 \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} + t_2. \end{aligned} \quad (3.69)$$

We then express the cost \tilde{f}^2 and (3.66) as a function of x_2, y_2 . Next, with the change of variables

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \frac{1}{\sqrt{b^2b_x^2 + a^2b_y^2}} \begin{bmatrix} bb_x & ab_y \\ ab_y & -bb_x \end{bmatrix} \begin{bmatrix} x_3 \\ y_3 \end{bmatrix} := H_3 \begin{bmatrix} x_3 \\ y_3 \end{bmatrix}, \quad (3.70)$$

and

$$\begin{aligned} \begin{bmatrix} x_3 \\ y_3 \end{bmatrix} &= \begin{bmatrix} \frac{ab}{R} & 0 \\ 0 & \frac{1}{b_z} \end{bmatrix} \begin{bmatrix} x_4 \\ y_4 \end{bmatrix} + \begin{bmatrix} \frac{kabc^2\sqrt{b^2b_x^2 + a^2b_y^2}}{R^2} \\ 0 \end{bmatrix} \\ &:= H_4 \begin{bmatrix} x_4 \\ y_4 \end{bmatrix} + t_4, \end{aligned} \quad (3.71)$$

²In the following, the subscript of variables refers to the step in the substitutions.

3.B. QUADRATIC OPTIMIZATION PROBLEMS WITH QUADRATIC CONSTRAINTS 85

where $R = \sqrt{a^2b^2b_z^2 + c^2b^2b_x^2 + c^2a^2b_y^2}$, we obtain a quadratic functional in the canonical form:

$$\tilde{f}^2 = x_3^2 + y_3^2. \quad (3.72)$$

The substitutions (3.68)-(3.71) applied to the constraint (3.66) give a shifted and rotated ellipse.

We can rotate the coordinate system so that the ellipse has axes parallel to the system's with the substitution

$$\begin{bmatrix} x_4 \\ y_4 \end{bmatrix} = H_5 \begin{bmatrix} x_5 \\ y_5 \end{bmatrix}, \quad (3.73)$$

where matrix H_5 is:

$$H_5 := \begin{bmatrix} \frac{v^2 - v(b_x^2 + b_y^2)c^2 + (b^2b_x^2 - a^2b_y^2)(b^2 - a^2)b_z^2 + \sqrt{S}}{n_1} & \frac{2b_xb_yb_z(b^2 - a^2)\sqrt{b^2b_x^2 + a^2b_y^2}}{n_2} \\ \frac{2b_xb_yb_z(b^2 - a^2)\sqrt{b^2b_x^2 + a^2b_y^2}}{n_1} & \frac{-v^2 + v(b_x^2 + b_y^2)c^2 - (b^2b_x^2 - a^2b_y^2)(b^2 - a^2)b_z^2 + \sqrt{S}}{n_2} \end{bmatrix} \quad (3.74)$$

where $\sqrt{S} = \sqrt{(b^2 - c^2)b_x^2 + (a^2 - c^2)b_y^2 + (a^2 - b^2)b_z^2 - 4b_x^2b_z^2(a^2 - b^2)(b^2 - c^2)}$, $v = b^2b_x^2 + a^2b_y^2$ and n_1, n_2 are coefficients introduced to normalize the first and second column of H_5 , respectively. Furthermore, define

$$A = \frac{(a^2 + b^2)b_z^2 + (b^2 + c^2)b_x^2 + (a^2 + c^2)b_y^2 + \sqrt{S}}{2R^2}, \quad (3.75a)$$

$$B = 0, \quad (3.75b)$$

$$C = \frac{(a^2 + b^2)b_z^2 + (b^2 + c^2)b_x^2 + (a^2 + c^2)b_y^2 - \sqrt{S}}{2R^2}, \quad (3.75c)$$

$$D = \frac{-2k(b^2 - a^2)b_xb_yb_zv}{n_1R^2} \left(b_z^2(a^2b^2 - c^2(a^2 + b^2)) + R^2 - c^4(b_x^2 + b_y^2) - c^2\sqrt{S} \right), \quad (3.75d)$$

$$E = \frac{-2k(b^2 - a^2)b_xb_yb_zv}{n_2R^2} \left(b_z^2(a^2b^2 - c^2(a^2 + b^2)) + R^2 - c^4(b_x^2 + b_y^2) + c^2\sqrt{S} \right), \quad (3.75e)$$

$$F = b_z^2 \left(\frac{k^2(a^4b^4b_z^2 + b^4c^4b_x^2 + a^4c^4b_y^2)}{R^4} - 1 \right). \quad (3.75f)$$

With this substitutions, the constraint (3.66) is rewritten in the quadratic form

$$A x_5^2 + B x_5y_5 + C y_5^2 + D x_5 + E y_5 + F = 0, \quad (3.76)$$

with coefficients in (3.75a)-(3.75f). The center and radii of the ellipse result to be

$$x_E = -\frac{D}{2A}, \quad (3.77)$$

$$y_E = -\frac{E}{2C}, \quad (3.78)$$

$$r_x = \sqrt{\frac{D^2C + E^2A - 4ACF}{4A^2C}}, \quad (3.79)$$

$$r_y = \sqrt{\frac{D^2C + E^2A - 4ACF}{4AC^2}}. \quad (3.80)$$

The problem (3.64) becomes

$$\operatorname{argmax}_{Ax_5^2 + Bx_5y_5 + Cy_5^2 + Dx_5 + Ey_5 + F = 0} x_5^2 + y_5^2 \quad (3.81)$$

that means to find the point on the ellipse described in (3.76) that is farthest from the origin (0,0). Problem (3.81) still require a numerical algorithm, but this formulation has been tested with respect to other methods mentioned above and results to be the most accurate.

The final vector $\vec{\pi}_1$ can be calculated reverting the changes of variables,

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = H_3 \left(H_4 H_5 \begin{bmatrix} x_5 \\ y_5 \end{bmatrix} + t_4 \right), \quad (3.82)$$

$$\vec{\pi}_1 = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = H_1 \left(H_2 \begin{bmatrix} x_2 \\ y_2 \\ 0 \end{bmatrix} + t_2 \right). \quad (3.83)$$

Chapter 4

Quantum Receivers for Pulse Position Modulation

Several modulation formats have been proposed to limit the distorting action of the channel on the transmitted signals. Modulation is the operation of encoding the symbol value to be transmitted into a property of the quantum states, such as phase, amplitude or polarization. In this Chapter we consider Pulse Position Modulation, where the transmitted symbol is encoded into the temporal position of a coherent state in the symbol time interval. Existing receiver schemes are presented as well as the theoretical limit of the error probability, and in the end an adaptive receiver scheme is proposed.

4.1 Introduction to Pulse Position Modulation

Optical Coherent Communications employ the properties of the coherent state of light as carriers for the messages sent from the transmitter to the receiver. Polarization, amplitude and phase are the main properties to be modulate in order to encode the symbol to be transmitted. Differently, Pulse Position Modulation (PPM) encodes the information to be transmitted in the temporal position of a pulse within the symbol time length.

In the classical optical implementation of the Pulse Position Modulation, the

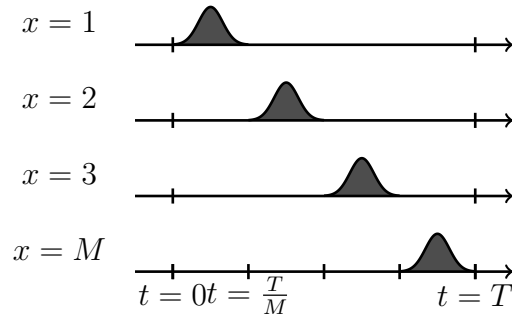


Figure 4.1: Classical representation of the PPM optical signals, with a laser pulse transmitted in the position corresponding to the symbol. T is the symbol time interval, and $M = 4$ the cardinality of the modulation.

symbol interval is virtually divided into M slots, with M the cardinality of the modulation. Each slot corresponds to a position, enumerated from 1 to M . A laser pulse is sent during the time interval in one of the possible positions identified by the symbol $x \in \{1, 2, \dots, M\}$ to be transmitted.

Pulse Position Modulation can also be interpreted as an intensity modulation, whose building block is the On Off Keying defined in 2.19. PPM is a coded version of OOK, where only the codewords with one “On” signal are considered, as shown in Figure 4.1. As a consequence, it is easily implemented with a pulse generator that is turned on or off according to whether the pulse is transmitted or not in the current slot.

This modulation is widely adopted for optical communication, and it is a candidate for satellite and deep space communications, because it requires low average power and attains reasonably high information efficiencies.

Quantum Pulse Position Modulation captures the idea of sequencing empty slots and pulses in order to define the quantum state associated with the symbol. We need to consider a Hilbert Space \mathcal{H}_0 for each slot, and in this Hilbert Space we define two possible quantum states $\rho_{0,i}$ and $\rho_{1,i}$ corresponding to the classical *empty slot* and *pulse*. Hence, the quantum state corresponding to the signal x

belongs in the composite Hilbert space \mathcal{H} , given by

$$\mathcal{H} = \mathcal{H}_0 \otimes \mathcal{H}_0 \otimes \cdots \otimes \mathcal{H}_0 = \mathcal{H}_0^{\otimes M}. \quad (4.1)$$

It \mathcal{H}_0 has dimension n , \mathcal{H} has size n^M . The association between symbol and transmitted quantum state is therefore

$$x = i, i \in \{1, \dots, M\} \iff \rho_i = \rho_{i,1} \otimes \rho_{i,2} \otimes \cdots \otimes \rho_{i,M} \quad (4.2)$$

with

$$\rho_{i,j} = \begin{cases} \rho_{0,i} & \text{if } i \neq j \\ \rho_{1,i} & \text{if } i = j \end{cases} \quad (4.3)$$

A straightforward implementation of Quantum Pulse Position Modulation from classic optics consider coherent states, and associates $\rho_{0,i} = |\gamma_{0,i}\rangle\langle\gamma_{0,i}|$ to the vacuum state $|0\rangle\langle 0|$ and $\rho_{1,i} = |\gamma_{1,i}\rangle\langle\gamma_{1,i}|$ with the coherent state $|\alpha\rangle\langle\alpha|$, $\alpha \neq 0$. For example, in a 4-PPM with this implementation the states associated with symbols are

$$\begin{aligned} x = 1 & \iff |\gamma_0\rangle = |\alpha\rangle \otimes |0\rangle \otimes |0\rangle \otimes |0\rangle \\ x = 2 & \iff |\gamma_1\rangle = |0\rangle \otimes |\alpha\rangle \otimes |0\rangle \otimes |0\rangle \\ x = 3 & \iff |\gamma_2\rangle = |0\rangle \otimes |0\rangle \otimes |\alpha\rangle \otimes |0\rangle \\ x = 4 & \iff |\gamma_3\rangle = |0\rangle \otimes |0\rangle \otimes |0\rangle \otimes |\alpha\rangle \end{aligned} \quad (4.4)$$

The inner product of two states, each generated by the tensor product of M component states, is given by the product of the M inner products of the component states, that is,

$$\langle\gamma_{i,j}|\gamma_{l,k}\rangle = \begin{cases} \langle 0|0\rangle = 1 & i = l = 0 \\ \langle 0|\alpha\rangle = e^{-\frac{|\alpha|^2}{2}} & i \neq l \end{cases} \quad (4.5)$$

and hence results

$$\chi = \langle\gamma_i|\gamma_j\rangle = \langle\alpha|0\rangle\langle 0|0\rangle\langle 0|\alpha\rangle\langle 0|0\rangle = e^{-\frac{|\alpha|^2}{2}} \cdot 1 \cdot e^{-\frac{|\alpha|^2}{2}} \cdot 1 = e^{-|\alpha|^2} \quad (4.6)$$

In the following Sections, we consider the problem of discrimination between the M -PPM signals defined with the coherent states, as in Eq. (4.4) for the case $M = 4$, assuming equal a priori probability $p_x = \frac{1}{M}$. The problem can be

framed in the Communication Scenario depicted in Figure 1.4, considering the association (4.4) between symbol and transmitted states and assuming an ideal channel.

The figure of merit for the performances of the discrimination are given by the error probability (or equivalently by the probability of correct decision) that can be written as

$$P_c = \sum_{i=1}^M P[\hat{x} = i, x = i] = \frac{1}{M} \sum_{i=1}^M P[\hat{x} = i | x = i], \quad (4.7)$$

$$P_e = 1 - P_c. \quad (4.8)$$

4.2 Optimal Performance

In this Section we review the optimal performances in term of error probability for PPM quantum states, as predicted by the quantum discrimination theory.

4.2.1 Geometrically Uniform Symmetry

A constellation of states with a symmetry property simplify the study of its performance and the search for optimal measurements. In particular, in many modulation format the *Geometrically Uniform Symmetry* (GUS) can be observed that we now define, that is also verified in PPM.

Definition 4.1. *A set of M pure states*

$$\{|\gamma_1\rangle, |\gamma_2\rangle, |\gamma_3\rangle, \dots, |\gamma_M\rangle\} \quad (4.9)$$

verifies the property of Geometrically Uniform Symmetry if there exists a unitary operator, called Symmetry Operator, such that all the states $|\gamma_i\rangle$ are obtained from a single reference state $|\gamma_M\rangle$ as

$$|\gamma_i\rangle = S^i |\gamma_M\rangle, \quad i = 1, 2, \dots, M \quad (4.10)$$

The Symmetry Operator is necessarily an M -th root of the identity operator, that is, it holds

$$S^M = I_{\mathcal{H}}. \quad (4.11)$$

By the definition of Quantum PPM, the presence of an underlying symmetry is evident. In fact, the state $|\gamma_{i+1}\rangle$ is obtained from $|\gamma_i\rangle$ through a shift of each kronecker factor by one position to the right (modulo M). However, to translate this concept into a symmetric operator is a non trivial problem since S is not separable into a tensor product of operators.

Cariolaro and Pierobon [54] have found the expression of such operator.

Proposition 4.1. *The matrix definition of the Symmetry Operator defining the PPM constellation is given by*

$$S = \sum_{k=1}^M \mathbf{e}_h(k) \otimes I_{\mathcal{H}}^{(n)} \otimes \mathbf{e}_h^T(k) = \sum_{k=1}^M \sum_{j=1}^n (\mathbf{e}_h(k) \otimes \mathbf{e}_n(j)) (\mathbf{e}_n(j) \otimes \mathbf{e}_h(k))^T, \quad (4.12)$$

where $\mathbf{e}_l(k)$ is the k -th element of the canonical basis¹ in a space of size l , $h = n^{M-1}$, and $I^{(n)}$ is the identity matrix of order n .

The proof is beyond the scope of this Chapter, the interested reader can find the details in [54, 25].

The symmetry operator generates a shift to the right

$$S(\rho_{i,1} \otimes \cdot \rho_{i,M-1}) \otimes \rho_{i,M} S^{-1} = \rho_{i,M} \otimes (\rho_{i,1} \otimes \cdot \rho_{i,M-1}), \quad (4.13)$$

that in the case of pure states becomes

$$S(|\gamma_{i,1}\rangle \otimes \cdot |\gamma_{i,M-1}\rangle \otimes |\gamma_{i,M}\rangle) = |\gamma_{i,M}\rangle \otimes |\gamma_{i,1}\rangle \otimes \cdot |\gamma_{i,M-1}\rangle, \quad (4.14)$$

Applying (4.14) iteratively, we get

$$|\gamma_i\rangle = S|\gamma_{i-1}\rangle = S(S|\gamma_{i-2}\rangle) = S^i|\gamma_M\rangle. \quad (4.15)$$

¹The notation $\mathbf{e}_l(k)$ represents a column vector of length l , whose elements are all 0's with the exception of a 1 in the k -th position.

As an example, the symmetry operator for 4-PPM takes the form

$$S = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.16)$$

4.2.2 Least Square Measurements

In order to introduce the least square measurements, we pose a slightly different problem rather than the minimization of the error probability.

Consider the case where the transmitted quantum states are pure $|\gamma_i\rangle \in \mathcal{H}'$,

$$|\gamma_i\rangle \in \{|\gamma_1\rangle, |\gamma_2\rangle, \dots, |\gamma_M\rangle\}, \quad (4.17)$$

and the random variable x has a uniform distribution

$$p_x = \frac{1}{M}, \quad x \in \{1, \dots, M\}. \quad (4.18)$$

We consider the following problem statement

Problem 4.1. Find the set of operator measurements $\{|\pi_i\rangle\langle\pi_i|, i = 1, \dots, M\}$ that satisfies the completeness relation (1.15) and minimizes

$$\mathcal{E} = \sum_{i=1}^M \langle e_i | e_i \rangle, \quad |e_i\rangle = \gamma[i] - |\pi_i\rangle \quad (4.19)$$

This problem is known as *Least Squared Error* (LSE), and differs from the problem statement in Section 1 for the figure of merit to minimize.

Intuitively, the LSE criterium look for measurement operators that are “close” to the state transmitted. As we will shortly see, the optimal solution for this problem coincides with the optimal solution when the figure of merit to optimize is the error probability for a communication system in which pure states with geometric uniform symmetry are employed with equal a priori probability.

The statement of the problem 4.1 considers the use of rank-1 measurement projectors $|\pi_i\rangle\langle\pi_i|$. This assumption is motivated by the following theorem.

Proposition 4.2 (Kennedy’s Theorem [23, 25]). *In the discrimination problem between M possible pure states, the optimal measurement operators that minimize the error probability are rank-1 projectors,*

$$P_k = |\pi_k\rangle\langle\pi_k|, \quad k = 1, 2, \dots, M \quad (4.20)$$

and the measurement vectors $|\pi_k\rangle$ are orthonormal.

The assumption can be further refined considering the subspace spanned by the quantum states $\{|\gamma_i\rangle\}$

$$U = \text{span}(|\gamma_1\rangle, |\gamma_2\rangle, \dots, |\gamma_M\rangle) \subseteq \mathcal{H}'. \quad (4.21)$$

In the case of a set of M linearly independent quantum states (not necessarily orthonormal), the dimension r of U is

$$r = \dim U \leq M \leq n = \dim \mathcal{H}' \quad (4.22)$$

and it is not restrictive to suppose that the measurement vectors $|\pi_k\rangle$ lies in the space U , since any other component of $|\pi_k\rangle$ that belongs to the complementary space U^\perp gives zero contribution to the outcomes probability.

Therefore, we see that both the quantum states $|\gamma_i\rangle$ and the measurement vectors $|\pi_k\rangle$ lies in the same space U . As a consequence, we can replace the constraint of the completeness relation (1.15) with the constraint

$$\sum_{k=1}^M |\pi_k\rangle\langle\pi_k| = \mathbf{P}_U \quad (4.23)$$

where \mathbf{P}_U is the projector of the original Hilbert Space \mathcal{H}' onto U .

In addition, since both the sets $\{|\gamma_i\rangle\}_i$ and $\{|\pi_l\rangle\}_k$ generate U , we can write the elements of one set as linear combination of elements of the other

$$|\pi_k\rangle = \sum_{i=1}^M a_{k,i} |\gamma_i\rangle \quad (4.24)$$

$$|\gamma_i\rangle = \sum_{k=1}^M b_{i,k} |\pi_k\rangle \quad (4.25)$$

with complex coefficients $a_{k,i}$ and $b_{i,k}$ in general.

Using a vector representation for the pure states, we can define the *state matrix* as the matrix whose columns are the transmitted quantum states

$$\Gamma = [|\gamma_1\rangle \mid |\gamma_2\rangle \mid \dots \mid |\gamma_M\rangle]. \quad (4.26)$$

This matrix is related to *Gram's matrix* \mathbf{G} and the *Gram's operator* \mathbf{T} by the definition

$$\mathbf{G} = \Gamma^H \Gamma = \begin{bmatrix} \langle\gamma_1|\gamma_1\rangle & \cdots & \langle\gamma_1|\gamma_M\rangle \\ \vdots & \ddots & \vdots \\ \langle\gamma_M|\gamma_1\rangle & \cdots & \langle\gamma_M|\gamma_M\rangle \end{bmatrix} \quad (4.27)$$

$$\mathbf{T} = \sum_{i=1}^M |\gamma_i\rangle\langle\gamma_i| \quad (4.28)$$

Note that the size of \mathbf{G} is $M \times M$, while the size of \mathbf{T} is $n \times n$, with n the size of the Hilbert Space \mathcal{H} of $|\gamma_i\rangle$.

If we consider the singular value decomposition (SVD) of the state matrix Γ ,

$$\Gamma = \mathbf{U} \mathbf{S} \mathbf{V}^H \quad (4.29)$$

it is easy to find the relation between \mathbf{G} and \mathbf{T} . The matrix \mathbf{U} is an $n \times n$ orthonormal matrix that constitutes a basis Hilbert Space \mathcal{H}' . The matrix \mathbf{S} is a $n \times M$ rectangular matrix with rank $r \leq M \leq n$ and zero off-diagonal elements. The non-zero diagonal elements σ_i are called singular values of Γ . Finally, the matrix \mathbf{V} is a $M \times M$ orthonormal matrix composed by the *kernel* of Γ and its orthogonal space.

If we rearrange the columns of the matrices such that the first diagonal elements $[\mathbf{S}]_{i,i}$ are the singular values σ_i , we can recognise

- the range space of Γ , that is the subspace U , is generated by the first r columns of \mathbf{U} , indicated by \mathbf{U}_r

$$U = \text{span} \{ \Gamma \} = \text{span} \{ u_1, u_2, \dots, u_r \} = \text{span} \{ \mathbf{U}_r \} \quad (4.30)$$

- the null space of Γ is generated by the last $M - r$ columns of \mathbf{V} , and hence

$$\ker(\Gamma) = \text{span} \{ v_{r+1}, \dots, v_M \}, \quad \ker^\perp(\Gamma) = \text{span} \{ v_1, \dots, v_r \} = \text{span} \mathbf{V}_r \quad (4.31)$$

- if we eliminate the unnecessary columns, we can decompose Γ into its *reduced* SVD

$$\Gamma = \mathbf{U}_r \mathbf{S}_r \mathbf{V}_r^H \quad (4.32)$$

where \mathbf{U}_r , \mathbf{S}_r , \mathbf{V}_r has dimensions $n \times r$, $r \times r$, $r \times M$.

As a consequence of this decomposition, we can hence rewrite \mathbf{G} and \mathbf{T} as

$$\mathbf{G} = \Gamma^H \Gamma = \mathbf{V} \mathbf{S}^H \mathbf{U}^H \mathbf{U} \mathbf{S} \mathbf{V}^H = \mathbf{V} \mathbf{S}^2 \mathbf{V}^H \quad (4.33)$$

$$\mathbf{T} = \Gamma \Gamma^H = \mathbf{U} \mathbf{S} \mathbf{V}^H \mathbf{V} \mathbf{S}^H \mathbf{U}^H = \mathbf{U} \mathbf{S}^2 \mathbf{U}^H \quad (4.34)$$

or, using the (4.32)

$$\mathbf{G} = \Gamma^H \Gamma = \mathbf{V}_r \mathbf{S}_r^H \mathbf{U}_r^H \mathbf{U}_r \mathbf{S}_r \mathbf{V}_r^H = \mathbf{V}_r \mathbf{S}_r^2 \mathbf{V}_r^H \quad (4.35)$$

$$\mathbf{T} = \Gamma \Gamma^H = \mathbf{U}_r \mathbf{S}_r \mathbf{V}_r^H \mathbf{V}_r \mathbf{S}_r^H \mathbf{U}_r^H = \mathbf{U}_r \mathbf{S}_r^2 \mathbf{U}_r^H \quad (4.36)$$

The matrices \mathbf{G} and \mathbf{T} have the same set of nonzero eigenvalues, v_i^2 , same rank, but different size, range and null space.

In the same way of (4.26), we can define the *measurement matrix*, i.e. the matrix with the measurement vector as column

$$\mathbf{M} = [|\pi_1\rangle \mid |\pi_2\rangle \mid \dots \mid |\pi_M\rangle] \quad (4.37)$$

We can hence rewrite the LSE criterium as

$$\mathcal{E} = \text{tr}((\Gamma - \mathbf{M})(\Gamma - \mathbf{M})^H) \quad (4.38)$$

and the constraints of the completeness relation as

$$\mathbf{M}\mathbf{M}^H = I_{\mathcal{H}} \quad (4.39)$$

The solutions to the problem (4.1) are called *Least Squared Measurement* or *Squared Root Measurement*, and are defined by the following proposition

Proposition 4.3 (Square Root Measurement, [55, 56]). *The measurement matrix \mathbf{M} that minimizes the quadratic error (4.38) with the constraint of the completeness relation (4.39) is given by*

$$\mathbf{M}_{opt} = \sum_{k=1}^M |u_k\rangle \langle v_k| = \mathbf{U}_r \mathbf{V}_r^H, \quad (4.40)$$

where \mathbf{U}_M , \mathbf{V}_M are the matrices of the Singular Value Decomposition of Γ . In addition, the minimum quadratic error is evaluated as

$$\mathcal{E} = \sum_{i=1}^M (1 - \sigma_i)^2. \quad (4.41)$$

For a proof of the proposition, see [55, 56, 25].

As corollary of this proposition, we have that the optimal measurement matrix $\mathbf{M}_{opt} = \mathbf{U}_r \mathbf{V}_r^H$ can be calculated with the following expressions

$$\mathbf{M}_{opt} = \Gamma(\Gamma^H \Gamma)^{-1/2} = \Gamma \mathbf{G}^{-1/2}, \quad (4.42)$$

$$\mathbf{M}_{opt} = (\Gamma \Gamma^H)^{-1/2} \Gamma = \mathbf{T}^{-1/2} \Gamma. \quad (4.43)$$

that uses the inverse of the square root of \mathbf{G} and \mathbf{T} , calculated as

$$\mathbf{G}^{-1/2} = \mathbf{V}_r \mathbf{S}_r^{-1} \mathbf{V}_r^H, \quad \mathbf{T}^{-1/2} = \mathbf{U}_r \mathbf{S}_r^{-1} \mathbf{U}_r^H. \quad (4.44)$$

By (4.44), different methods have been proposed to evaluate the optimal measurement matrix \mathbf{M}_{opt} , using the SVD of the state matrix or calculating the inverse of the square root of \mathbf{T} or \mathbf{G} . For a detailed description of the methods we refer to [25].

We can compute the transition probability directly from \mathbf{G} . In fact, by definition the transition probabilities are given by

$$p_{i,j} = \left| [\mathbf{M}^H \Gamma]_{i,j} \right|^2, \quad (4.45)$$

where the matrix $\mathbf{M}^H \Gamma$ can be evaluated from

$$\mathbf{M}^H \Gamma = (\Gamma \mathbf{G}^{-1/2})^H \Gamma = (\Gamma (\Gamma^H \Gamma)^{-1/2})^H \Gamma = (\Gamma^H \Gamma)^{-1/2} \Gamma^H \Gamma = (\Gamma^H \Gamma)^{1/2} = \mathbf{G}^{1/2} \quad (4.46)$$

resulting in

$$p_{i,j} = \left| \mathbf{G}_{i,j}^{1/2} \right|^2. \quad (4.47)$$

This method is also referred to *Squared Root Measurement*, because it involves the use of the square root of matrix \mathbf{G} .

4.2.3 Least Square Measurements and Geometric Uniform Symmetry

In the presence of a set of transmitted quantum states with Geometrically Uniform symmetry, the evaluation of the optimal measurement operators can be simplified. In fact, the GUS symmetry is reflected in the measurement operators, as indicated by the next proposition.

Proposition 4.4 (Eldar and Forney, [56, 25]). *If the set of transmitted quantum states has the property of Geometrically Uniform Symmetry, it is not restrictive to suppose that the optimal measurement operators have the same symmetry, that is*

$$P_k = S^k P_M S^{-k}, \quad k = 1, \dots, M. \quad (4.48)$$

In the case that the measurement operators are just rank-1 projectors, the Geometrically Uniform Symmetry becomes

$$|\pi_k\rangle = S^k |\pi_M\rangle, \quad k = 1, 2, \dots, M. \quad (4.49)$$

In Section 4.2.2 we have seen that we can obtain the performance and the optimal measurement operators for the LSE problem given the state matrix and the Gram matrix. When the set of quantum states possesses the property of GUS, the Gram matrix \mathbf{G} has a particular structure

$$\mathbf{G}_{i,j} = \langle \gamma_i | \gamma_j \rangle = \langle \gamma_M | S^{-i} S^j | \gamma_M \rangle = \langle \gamma_M | S^{j-i} | \gamma_M \rangle = g_{j-i \pmod{M}} \quad (4.50)$$

and hence the entry $\mathbf{G}_{i,j}$ depends only on the difference $j - i \pmod{M}$. A matrix with such a structure is called *circulant*, and is completely specified by its first row $[g_0, g_1, \dots, g_{M-1}]$. For example, if $M = 4$ we have the following matrix

$$\mathbf{G} = \begin{bmatrix} g_0 & g_1 & g_2 & g_3 \\ g_3 & g_0 & g_1 & g_2 \\ g_2 & g_3 & g_0 & g_1 \\ g_1 & g_2 & g_3 & g_0 \end{bmatrix} \quad (4.51)$$

For circulant matrix, the eigenvectors $|w_p\rangle$ are the column of the matrix of the Discrete Fourier Transform with the same size

$$|w_p\rangle = [1 \ w^{-p} \ w^{-2p} \ \dots \ w^{-p(M-1)}]^T, \quad p = 0, 1, \dots, M-1 \quad (4.52)$$

with $w = e^{i\frac{2\pi}{M}}$, such that

$$\mathbf{G} = W S W^H = \sum_{p=0}^{M-1} \lambda_p |w_p\rangle \langle w_p| \quad (4.53)$$

with

$$W = \frac{1}{\sqrt{M}} \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & w^{-1} & w^{-2} & w^{-2(M-1)} & \\ \vdots & & \vdots & \ddots & \vdots \\ 1 & w^{-(M-1)} & w^{-2(M-1)} & \dots & w^{-(M-1)(M-1)} \end{bmatrix} \quad (4.54)$$

The eigenvalues corresponding to $|w_p\rangle$ are

$$\lambda_p = g_0 + g_1 w^p + g_2 w^{2p} + g_3 w^{3p} + \dots + g_{M-1} w^{M-1} = \sum_{q=0}^{M-1} g_q w^{-pq} \quad (4.55)$$

From (4.53), it is easy to find the square root of \mathbf{G}

$$\mathbf{G}^{\pm 1/2} = \sum_{p=0}^{M-1} \lambda_p^{\pm \frac{1}{2}} |w_p\rangle \langle w_p| \quad (4.56)$$

and hence

$$(\mathbf{G}^{\pm 1/2})_{i,j} = \frac{1}{M} \sum_{p=0}^{M-1} \lambda_p^{\pm \frac{1}{2}} w^{-p(i-j)} \quad (4.57)$$

Following (4.47), we can calculate the transition probabilities

$$p_{j|i} = \left| \frac{1}{M} \sum_{p=0}^{M-1} \lambda_p^{\pm \frac{1}{2}} w^{-p(i-j)} \right|^2, \quad i, j = 1, 2, \dots, M \quad (4.58)$$

and in particular

$$p_{i|i} = \left| \frac{1}{M} \sum_{p=0}^{M-1} \lambda_p^{\frac{1}{2}} \right|^2 \quad (4.59)$$

that gives the probability of correct decision and the error probability

$$P_c = \frac{1}{M} \sum_{i=1}^M p_{i|i} = \frac{1}{M^2} \left| \sum_{p=0}^{M-1} \lambda_p^{\frac{1}{2}} \right|^2, \quad (4.60)$$

$$P_e = 1 - P_c = 1 - \frac{1}{M^2} \left| \sum_{p=0}^{M-1} \lambda_p^{\frac{1}{2}} \right|^2. \quad (4.61)$$

4.2.4 Theoretical Limit Performances of PPM

When this method is applied to the Pulse Position Modulation, we easily manage to evaluate its performances and find the optimal measurement operator. In the case of PPM, define as in (4.110) the inner product

$$\chi = \langle \gamma_i | \gamma_j \rangle, \quad i \neq j \quad (4.62)$$

such that the Gram's matrix becomes

$$\mathbf{G} = \begin{bmatrix} 1 & \chi & \chi & \dots & \chi \\ \chi & 1 & \chi & \dots & \chi \\ \vdots & & & \ddots & \\ \chi & \chi & \chi & \dots & 1 \end{bmatrix} \quad (4.63)$$

The matrix (4.63) has eigenvalues

$$\lambda_p = \sum_{q=0}^{M-1} g_q w^{-pq} = 1 + \chi \sum_{q=1}^{M-1} w^{-pq} \quad (4.64)$$

and using the orthogonality condition

$$\sum_{q=0}^{M-1} w^{-pq} = \begin{cases} M, & \text{if } p = 0 \\ 0, & \text{if } p \neq 0 \end{cases} \quad (4.65)$$

we get

$$\lambda_p = \begin{cases} 1 + (M-1)\chi, & p = 0 \\ 1 - \chi, & p \neq 0 \end{cases} \quad (4.66)$$

The transition probabilities can be evaluated as

$$p_{j|i} = \frac{1}{M^2} \left[\sum_{p=0}^{M-1} \lambda_p^{\frac{1}{2}} w^{-p(i-j)} \right]^2 = \frac{1}{M^2} \left[\lambda_0^{\frac{1}{2}} + \lambda_1^{\frac{1}{2}} \sum_{p=1}^{M-1} w^{-p(i-j)} \right]^2 \quad (4.67)$$

that depending on the values of i, j are calculated as

$$p_{j|i} = \begin{cases} \frac{1}{M^2} \left(\lambda_0^{\frac{1}{2}} - \lambda_1^{\frac{1}{2}} \right)^2 & i \neq j \\ \frac{1}{M^2} \left(\lambda_0^{\frac{1}{2}} + (M-1)\lambda_1^{\frac{1}{2}} \right)^2 & i = j \end{cases} \quad (4.68)$$

Finally, we can evaluate the probability of correct decision by (4.60), obtaining the theoretical quantum limit

$$P_c^{(theoretical)} = \frac{1}{M^2} \left(\sqrt{1 + (M-1)\chi} + (M-1)\sqrt{1-\chi} \right)^2, \quad (4.69)$$

$$P_e^{(theoretical)} = \frac{M-1}{M^2} \left(\sqrt{1 + (M-1)\chi} - \sqrt{1-\chi} \right)^2. \quad (4.70)$$

4.3 Existing Receiver Schemes for PPM

In this Section, we review the existing receiver schemes for PPM, in particular a class of receivers called “*conditionally nulling receivers*”, proposed by Dolinar in 1982 in [57], that combine an adaptive nulling scheme with a energy photodetection, and approaching very closely the error probability limit.

4.3.1 Classical Receivers

In classical optical communications, the Pulse Position Modulation is described as a coded version of the binary OOK modulation. Considering all the possible binary sequences with digits $|\alpha\rangle$ and $|0\rangle$ of length M , PPM is the modulation associated to a codebook composed by word with only one state $|\alpha\rangle$.

The classical receiver revises these considerations, and it implements an energy photodetection in each slot k , looking for any photons, that are the signature of the quantum states $|\alpha\rangle$. Since this receiver requires only photon counting, it is also referred to as *Direct Detection* (DD).

In the assumptions of ideal channel and ideal photon counting, that is with no dark counts, no dead time and unit efficiency, if at least one photon is observed in a slot the receiver estimates the corresponding symbol. On the contrary, if no photons are observed in the whole symbol time interval, the receiver outputs a symbol value at random. In the following we see how this estimation scheme is the result of an optimization on the outcome sequences with the *Maximum a Posteriori* (MAP) criterium.

More precisely, in each slot k a measurement is performed associated with the outcome z_k , with $z_k = \hat{0}$ indicating the empty slot and outcome $z_k = \hat{1}$ associated with the presence of at least one photon. The POVM of this measurement are

$$\Pi_{\hat{0}} = |0\rangle\langle 0|, \quad \Pi_{\hat{1}} = I - \Pi_{\hat{0}}. \quad (4.71)$$

Usually, the outcome $z_k = \hat{1}$ is associated to a “click” in the detector for each observed photons. The transition probabilities read

$$P [z_k = \hat{0} | |\gamma_{k,i}\rangle = |0\rangle] = 1, \quad (4.72)$$

$$P [z_k = \hat{0} | |\gamma_{k,i}\rangle = |\alpha\rangle] = e^{-|\alpha|^2}, \quad (4.73)$$

$$P [z_k = \hat{1} | |\gamma_{k,i}\rangle = |\alpha\rangle] = 1 - e^{-|\alpha|^2}. \quad (4.74)$$

Note that the transition probability

$$P [z_k = \hat{1} | |\gamma_{k,i}\rangle = |0\rangle] \quad (4.75)$$

equals zero as the corresponding transition is not allowed. If we take into account some imperfections of the photon detection, such as the *Dark Counts*, this transition probability may be allowed.

The projectors associated with the measurement on the whole time symbols are tensor product of (4.71). Consider the string of length M , \bar{z}_M , with digits $z_k \in \{\hat{0}, \hat{1}\}, k \in \{1, \dots, M\}$. The projector associated to \bar{z}_M has the tensor product of (4.71) in the order that the letters appear in the string,

$$\Pi_{\bar{z}_M} = \Pi_{[\hat{0}\hat{1}\hat{1}\dots\hat{0}\hat{1}\hat{0}]} = \Pi_{\hat{0}} \otimes \Pi_{\hat{1}} \otimes \Pi_{\hat{1}} \otimes \dots \otimes \Pi_{\hat{0}} \otimes \Pi_{\hat{1}} \otimes \Pi_{\hat{0}} \quad (4.76)$$

For example, for $M = 4$, some of the measurement projectors are

$$\begin{aligned} \Pi_{[\hat{0}\hat{0}\hat{0}\hat{0}]} &= \Pi_{\hat{0}} \otimes \Pi_{\hat{0}} \otimes \Pi_{\hat{0}} \otimes \Pi_{\hat{0}} \\ \Pi_{[\hat{0}\hat{0}\hat{0}\hat{1}]} &= \Pi_{\hat{0}} \otimes \Pi_{\hat{0}} \otimes \Pi_{\hat{0}} \otimes \Pi_{\hat{1}} \\ \Pi_{[\hat{0}\hat{0}\hat{1}\hat{0}]} &= \Pi_{\hat{0}} \otimes \Pi_{\hat{0}} \otimes \Pi_{\hat{1}} \otimes \Pi_{\hat{0}} \\ \Pi_{[\hat{0}\hat{0}\hat{1}\hat{1}]} &= \Pi_{\hat{0}} \otimes \Pi_{\hat{0}} \otimes \Pi_{\hat{1}} \otimes \Pi_{\hat{1}} \\ &\vdots \end{aligned} \quad (4.77)$$

We can calculate the transition probabilities between the transmitted and the estimated symbols from the transition probabilities slot by slot (4.73).

Firstly, all symbols can give the outcome $z = [\hat{0}\hat{0}\hat{0}\dots\hat{0}]$ if the state $|\gamma_{k,i}\rangle = |\alpha\rangle$ gives outcome $\hat{0}$ (as well as all the other slot), with the probability

$$P[z = [\hat{0}\hat{0}\hat{0}\dots\hat{0}] | x = i] = \quad (4.78)$$

$$\begin{aligned} &= (\langle 0 | \langle 0 | \dots \langle \alpha | \langle 0 |) (\Pi_{\hat{0}} \dots \Pi_{\hat{0}}) (|0\rangle |0\rangle \dots |\alpha\rangle |0\rangle) \\ &= \langle 0 | \Pi_{\hat{0}} | 0 \rangle \langle 0 | \Pi_{\hat{0}} | 0 \rangle \dots \langle \alpha | \Pi_{\hat{0}} | \alpha \rangle \dots \langle 0 | \Pi_{\hat{0}} | 0 \rangle \langle 0 | \Pi_{\hat{0}} | 0 \rangle \end{aligned} \quad (4.79)$$

$$= 1 \cdot 1 \cdot \dots \cdot e^{-|\alpha|^2} \cdot \dots \cdot 1 \cdot 1 = 1 - e^{-|\alpha|^2}, \quad \forall i = 1, \dots, M. \quad (4.80)$$

Secondly, all symbols can be detected correctly, that means with outcome $\hat{0}$ when the slot is $|\gamma_{k,i}\rangle = |0\rangle$ and with outcome $\hat{1}$ where there is $|\gamma_{k,i}\rangle = |\alpha\rangle$. The probability of the outcome is

$$P[z = [\hat{0}\hat{1}\hat{0}\dots\hat{0}] | x = i] = \begin{cases} 1 - e^{-|\alpha|^2} & \text{if } z_k = \hat{1} \\ 0 & \text{otherwise} \end{cases} \quad (4.81)$$

Thirdly, all other outcomes sequence \bar{z} has at least two outcomes $z = \hat{1}$, and therefore they cannot be observed because of the probability (4.75),

$$P [\bar{z} = [\hat{0}\hat{1}\hat{1}\dots\hat{0}]|x = i] = 0, \quad \forall i = 1, \dots, M. \quad (4.82)$$

The classical receiver chooses the estimate with a MAP criterium. The a posteriori probabilities for the outcome sequences read

$$P [x = i|z_j = \hat{1}, z_k = \hat{0} \text{ for } k \neq j] = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \quad (4.83)$$

and hence

$$\operatorname{argmax}_{i=1,\dots,M} P [x = i|z_j = \hat{1}, z_k = \hat{0} \text{ for } k \neq j] = j \quad (4.84)$$

while for the outcome $\bar{z} = [\hat{0}\dots\hat{0}]$,

$$P [x = i|z_j = \hat{0} \quad \forall j] = \frac{1}{M} \quad \forall i = 1, \dots, M \quad (4.85)$$

and hence the estimation is chosen at random,

$$\operatorname{argmax}_{i=1,\dots,M} P [x = i|z_j = \hat{0} \quad \forall j] = \{1, \dots, M\}. \quad (4.86)$$

The possible errors in the detection happen if no outcome $\hat{1}$ is measured from the slots and when choosing at random the receiver fails in the guessing. Hence, the transition probabilities between transmitted and estimated symbols are

$$P [\hat{x} = i|x = j] = \begin{cases} 1 - e^{-|\alpha|^2} + \frac{e^{-|\alpha|^2}}{M} & \text{if } i = j \\ \frac{e^{-|\alpha|^2}}{M} & \text{if } i \neq j \end{cases} \quad (4.87)$$

Finally, the error probability results to be

$$P_e^{(classical)} = \frac{M-1}{M}e^{-|\alpha|^2} = \frac{M-1}{M}\chi. \quad (4.88)$$

4.3.2 Unconditional Nulling Receiver

The direct detection does not distinguish symmetrically between the quantum states $|\alpha\rangle$ and $|0\rangle$. While the outcome $\hat{1}$ from one slot assures that in the slot

there was the quantum state $|\alpha\rangle$, in presence of the outcome $\hat{0}$ we only have a *partial* confirmation that in the slot there was $|0\rangle$.

As a consequence of the direct detection, only a few outcome sequences \bar{z} are possible, the M strings with one $z_k = \hat{1}$ and $\bar{z} = [\hat{0} \dots \hat{0}]$. It seems reasonable to ask if a receiver that consider more outcomes can obtain an improved performance.

This can be achieved considering the opposite situation in terms of words in the codebook, that is when the symbol $x = i$ is associated a quantum state that is the tensor product of states $|\alpha\rangle$ in all slots except for the one in position i that is $|0\rangle$.

This new situation can be achieved sending PPM signals from the transmitter and performing a perfect *nulling* at the receiver side. The receiver performs a displacement² of $|\alpha\rangle$ to the incoming state $|\gamma_{k,i}\rangle$ in the slot. If the state was $|\gamma_{k,i}\rangle = |0\rangle$, the receiver displace it to the state $|\alpha\rangle$, otherwise if the quantum state was $|\gamma_{k,i}\rangle = |\alpha\rangle$, it is displaced to the ground state $|0\rangle$.

In this variation of the PPM modulation, we can still perform a classical discrimination with measurement projectors (4.71) in each slot, obtaining a global POVM in the whole symbol time interval that is constructed through tensor products as in (4.76).

The transition probabilities for the direct detection are the same as (4.73), since the energy detection does not distinguish between $|\alpha\rangle$ and $|\alpha\rangle$, therefore including the nulling operation we achieve the following transition probabilities

$$\begin{aligned} P [z_k = \hat{0} | |\gamma_{k,i}\rangle = |0\rangle] &= e^{-|\alpha|^2}, \\ P [z_k = \hat{0} | |\gamma_{k,i}\rangle = |\alpha\rangle] &= 1, \\ P [z_k = \hat{1} | |\gamma_{k,i}\rangle = |0\rangle] &= 1 - e^{-|\alpha|^2}, \\ P [z_k = \hat{1} | |\gamma_{k,i}\rangle = |\alpha\rangle] &= 0. \end{aligned} \tag{4.89}$$

²The displacement operation $D(\alpha)$ allows to shift a coherent state $|\gamma\rangle$ by an amount of α such that

$$D(\alpha) |\gamma\rangle = |\gamma + \alpha\rangle.$$

For more details, we suggest the reading of [58].

However with this scheme, now more outcomes \bar{z} are possible. In particular, we can classify the outcomes in two categories:

- \bar{z} has only one digit $\hat{0}$. In displacing the symbol $x = i$, the state $|\alpha\rangle$ becomes $|0\rangle$ that is detected as $\hat{0}$ from the photon counting. All the other state $|0\rangle$, displaced to $|\alpha\rangle$, are detected as $\hat{1}$. The probability of this outcome sequence is

$$P[\bar{z} = [\hat{1}\hat{0}\hat{1} \dots \hat{1}] | x = i] = 1 \cdot (1 - e^{-|\alpha|^2})^{M-1} \quad (4.90)$$

- \bar{z} has k digits equal to $\hat{0}$ in the set $\mathcal{Z}, \mathcal{Z} \subset \{1, \dots, M\}$. One of them comes from the detection of $|\alpha\rangle$ displaced to $|0\rangle$, the other $k - 1$ from $|0\rangle$ displaced to $|\alpha\rangle$ misdetected by the photon counting. The probability of this outcome sequence is

$$P[z_k = \hat{0} \text{ for } k \in \mathcal{Z}, z_j = \hat{1} \text{ for } k \notin \mathcal{Z}, |x = i] = 1 \cdot (1 - e^{-|\alpha|^2})^{M-k} \cdot (e^{-|\alpha|^2})^{k-1} \quad (4.91)$$

The estimation \hat{x} is again chosen with the MAP criterion. The a posteriori probabilities are

$$P[x = i | z_j = \hat{0}, z_k = \hat{1} \text{ for } k \neq j] = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \quad (4.92)$$

if there is only one outcome $\hat{0}$, otherwise

$$P[x = i | z_k = \hat{0} \text{ for } k \in \mathcal{Z}, z_j = \hat{1} \text{ for } k \notin \mathcal{Z}] = \begin{cases} \frac{1}{k} & \text{if } i \in \mathcal{Z} \\ 0 & \text{otherwise} \end{cases} \quad (4.93)$$

In the former case, (4.92), the MAP criterion gives

$$\arg \max_{i=1, \dots, M} P[x = i | z_j = \hat{0}, z_k = \hat{1} \text{ for } k \neq j] = j, \quad (4.94)$$

while in the latter case it is

$$\arg \max_{i=1, \dots, M} P[x = i | z_k = \hat{0} \text{ for } k \in \mathcal{Z}, z_j = \hat{1} \text{ for } k \notin \mathcal{Z}] = \mathcal{Z}, \quad (4.95)$$

which means that when the receiver sees outcomes $\hat{0}$ in the set \mathcal{Z} , it chooses at random one symbol that set.

Finally, the probability of correct decision is

$$P_c^{(nulling)} = \frac{1}{M} \sum_i P[\hat{x} = i | x = i] = P[\hat{x} = 1 | x = 1] = \quad (4.96)$$

$$= \sum_{j=0}^{M-1} \frac{1}{j+1} \binom{M-1}{i} (1-p)^{M-1-j} p^j \quad (4.97)$$

$$= \sum_{j=0}^{M-1} \frac{1}{M} \binom{M}{i+1} (1-p)^{M-1-j} p^j \quad (4.98)$$

$$= \frac{1}{Mp} \sum_l = 1^M \binom{M}{l} (1-p)^{M-l} p^l \quad (4.99)$$

$$= \frac{1}{Mp} [1 - (1-p)^M] \quad (4.100)$$

$$(4.101)$$

with $p = e^{-|\alpha|^2} = \chi$. The error probability can be written as

$$P_e^{(nulling)} = \frac{1}{M\chi} [(1-\chi)^M - 1 + M\chi]. \quad (4.102)$$

It can be shown that

$$P_e^{(classical)} \leq P_e^{(nulling)} \quad (4.103)$$

for all $p \in [0, 1]$, and hence this unconditional nulling strategy reduces the performance. However, a modification of the nulling strategy brings to a great improvement of the error probability, as we see in Section 4.3.3.

4.3.3 Conditional Nulling Receiver

Both direct detection and the unconditional nulling receiver do not symmetrically discriminate between the quantum states $|0\rangle$ and $|\alpha\rangle$. With a direct detection we can identify with probability 1 the presence of the quantum state $|\alpha\rangle$. On the contrary with the nulling operation, we can identify for sure the slot with $|0\rangle$, when we get the outcome $\hat{1}$.

The Conditional Nulling Receiver takes advantage of both situations [57]. Since the nulling operation is performed at the receiver side, the receiver can choose at each slot whether to displace or not the incoming unknown quantum state.

The nulling strategy is described as follows. In the first slot, the unknown quantum state is displaced by an amount $-\alpha$, and a direct detection is performed. If the state is $|\alpha\rangle$, it is displaced to $|0\rangle$ and the outcome is definitely $\hat{0}$. On the contrary, if the quantum state was $|0\rangle$, both outcomes $\hat{1}$ and $\hat{0}$ are possible. The receiver assigns to the outcome $z_1 = \hat{0}$ a temporary estimation $\hat{x} = 0$, while the outcome $\hat{1}$ is a definite claim that the state was $|0\rangle$.

In the former case, the receiver proceeds performing a direct detection on the remaining slots, continuing to believe in the temporary estimation unless it obtains conclusive evidence of an other hypothesis observing any photons in the remaining slots. In this case, the receiver neglects the temporary hypothesis and takes the one corresponding to the slot where at least one photon has been observed.

In the latter case, that is when the outcome $z_1 = \hat{1}$ is observed, the receiver is certain that the transmitted symbol was not $x = 1$, and the discrimination problem reduces to its $M - 1$ -ary version, and the receiver applies its strategy recursively.

To resume, the strategy of the receiver simply consists in displacing the quantum state until it obtains partial confirmation of the hypothesis k with the outcome $z_k = \hat{0}$. Afterward, it performs a photon counting in the remaining time slots. If during this direct detection an outcome $\hat{1}$ is measured, the previous hypothesis is neglected and the receiver takes the hypothesis corresponding to the slot where the photons has been observed.

The receiver strategy can be well represented with a binary tree graph, where each node represents a measurement on a slot, and the edges the possible outcomes $\hat{1}$ or $\hat{0}$. After the last measurements that correspond to the lower level of the tree, the final estimation \hat{x} is made. In Figure 4.1 it is represented the tree

associated with the receiver strategy with $M = 4$. The tree grows from left to right, and each level represents a measurement. Note that some branches are cut since some transition probabilities are null.

Having defined the receiver strategy, we can evaluate its performance in terms of error probability.

Consider first the symbol $x = 1$. Since in the first slot the quantum state is nulled, and afterwards a direct detection is performed on the remaining $|0\rangle$ states, in each slot the outcome is certainly $z_k = \hat{0}$, and the symbol is always correctly detected,

$$P[\hat{x} = 1|x = 1] = 1. \quad (4.104)$$

On the contrary, a symbol $x = i$, $i > 1$ can be wrongly estimated into $\hat{x} = k$, $k < i$. This case may happen if the receiver measures two wrong outcomes, z_k and z_i . The receiver starts to displace the quantum state $|0\rangle$ into $|\alpha\rangle$, measuring $z_l = \hat{1}$, $l < k$ with probability $1 - e^{-|\alpha|^2} = 1 - \chi$. If in the k -th slot, the displaced state $|\alpha\rangle$ gives $z_k = \hat{0}$, and the receiver starts to photon counting the remaining slot without applying the displacement operation. If in the i -th slot the outcome is $z_i = \hat{0}$, the receiver misses the estimation.

The error probability can therefore be evaluated as

$$\begin{aligned} P_e^{(cond.nulling)} &= \frac{1}{M} \sum_{i=1}^M P[\hat{x} \neq i|x = i] = \frac{1}{M} \sum_{i=2}^M \sum_{l=1}^{i-1} P[\hat{x} = l|x = i] = \\ &= \frac{1}{M} \sum_{i=2}^M \sum_{l=1}^{i-1} (1 - \chi)^{l-1} \chi^2 = \frac{\chi^2}{M} \sum_{i=2}^M \frac{(1 - \chi)^{i-1} - 1}{((1 - \chi) - 1)} \\ &= \frac{\chi}{M} \left[M - 1 - \left[\frac{(1 - \chi)^M - 1}{((1 - \chi) - 1)} - 1 \right] \right] \\ &= \frac{(1 - \chi)^M - 1 + M\chi}{M} \end{aligned} \quad (4.105)$$

The Conditional Nulling Receiver drastically improves the performance for the Pulse Position Modulation, approaching the theoretical limit for all the cardinality M and the values of $\chi = e^{-|\alpha|^2}$. In particular, it can be numerically shown that the deviation does not exceed the optimal limit by a multiplicative

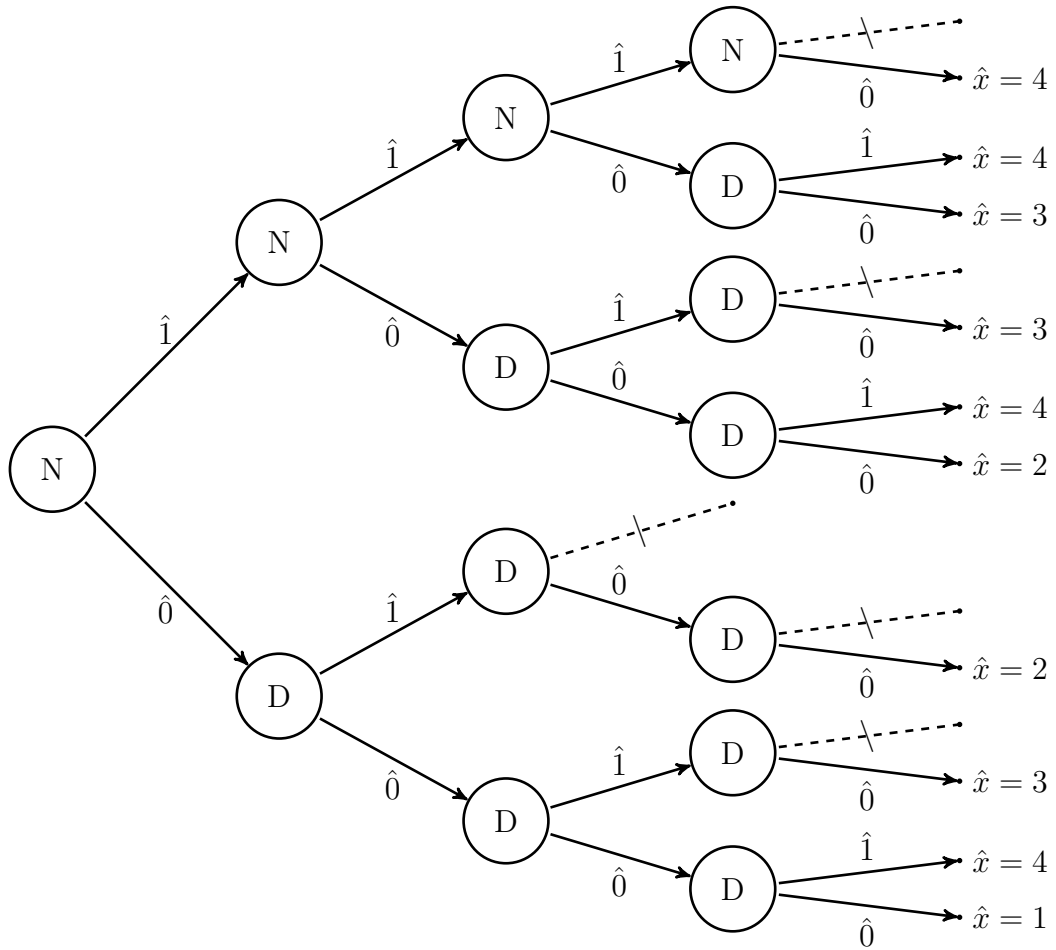


Figure 4.1: Conditional Receiver strategy, depicted as a binary tree, growing from left to right, for $M = 4$. Each node indicates if the nulling operation is applied (N) or if just direct detection (D) is performed. On the edge the outcomes $\hat{0}$ or $\hat{1}$ of the photon counting. After the last outcomes, the final estimation \hat{x} of the symbol $x \in \{1, 2, 3, 4\}$ is performed. Dashed edges indicate impossible outcomes.

factor of 2.15 [57],

$$1 \leq \frac{P_e^{(theoretical)}}{P_e^{(cond.nulling)}} \leq 2.15 \quad \forall \chi, M. \quad (4.106)$$

4.3.4 Improved Conditional Nulling Receiver

Two architectures have been proposed by Guha *et al.* [59] to improve the *Conditional Nulling* scheme. The key idea is that a non-exact nulling of the signal can lead to better performances, just as the *Improved Kennedy Receiver* described in Section 2.3 uses the same concept to enhance the Kennedy receiver performance [24].

Both architectures follow the adaptive strategy of the Conditional Nulling Receiver, depicted in Figure 4.1, but with a slightly modification in the local measurements performed in each slots in place of the nulling operation.

The first architecture, denoted in [59] as *Type I*, applies a constant displacement $D(\beta)$, with $\beta \neq -\alpha$ in place of the exact nulling. The second architecture, called *Type II*, in addition to the non-exact nulling includes an optical phase amplifier with gain G to squeeze the partially-nulled coherent state, further improving the performance. A photon counting follows these operations, and gives the outcomes $z_k = \hat{0}$ and $z_k = \hat{1}$ that drive the receiver strategy.

The error probability for the *Improved Conditional Nulling* Type II scheme, in our ideal assumptions of unit efficiency, no dark current and no death time, becomes

$$P_e^{(impr.cond.nulling)} = \frac{q_1(1-(1-q_0)^{M-1}) + e^{-|\alpha|^2}(Mq_0 - 1 + (1-q_0)^M)}{Mq_0} \quad (4.107)$$

with

$$\begin{aligned} q_0 &= p_{\hat{0}|0} = \frac{\exp\left[-\frac{(\sqrt{G} + \sqrt{G-1})^2 |\beta|^2}{1 + \sqrt{G-1}(\sqrt{G} + \sqrt{G-1})}\right]}{\sqrt{G}}, \\ q_1 &= p_{\hat{1}|1} = 1 - \frac{\exp\left[-\frac{(\sqrt{G} + \sqrt{G-1})^2 |\alpha - \beta|^2}{1 + \sqrt{G-1}(\sqrt{G} + \sqrt{G-1})}\right]}{\sqrt{G}}. \end{aligned} \quad (4.108)$$

Substituting $G = 1$ gets the expression for the Type I architecture. The displacement β and the phase-sensitive amplifier G are numerically optimized to reach the maximum performance (4.107) of this receiver scheme.

The Improved Conditional Nulling receiver outperforms the Conditional Nulling and the Classical receiver, as represented in Figures 4.1, 4.3 and 4.4. The gain is more evident for weak values of mean photon number, $|\alpha|^2 < 2$, while for $|\alpha|^2 \approx 2$ the receiver approach the performance of the Conditional Nulling.

4.4 Qubit Framework Representation

We shall introduce a useful representation for the PPM quantum states that simplifies the description of the receiver schemes. As we have already seen in (4.2), the global PPM coherent states in \mathcal{H} are composed by a sequence of coherent states in \mathcal{H}_0 in a tensor product.

Since in each \mathcal{H}_0 only two quantum states are possible, we can consider the subspace of the coherent states spanned by $\{|0\rangle, |\alpha\rangle\}$, that is isomorphic to a qubit Hilbert space \mathcal{H}'_0 . We can associate to each quantum state $|0\rangle$ and $|\alpha\rangle$ a qubit in \mathcal{H}'_0 , respectively

$$|\tilde{\gamma}_0\rangle = \cos\theta |x\rangle + \sin\theta |\hat{x}\rangle, \quad |\tilde{\gamma}_1\rangle = \cos\theta |x\rangle - \sin\theta |\hat{x}\rangle, \quad (4.109)$$

where $\{|x\rangle, |\hat{x}\rangle\}$ is a basis of \mathcal{H}'_0 , and without loss of generality $\theta \in [0, \pi/4]$. For a correct representation in terms of qubit, the inner product between $|0\rangle$ and $|\alpha\rangle$ must equal the inner product between $|\tilde{\gamma}_0\rangle$ and $|\tilde{\gamma}_1\rangle$,

$$\chi = |\langle 0|\alpha\rangle|^2 = e^{-|\alpha|^2} = |\langle \tilde{\gamma}_0|\tilde{\gamma}_1\rangle|^2 = \cos^2 2\theta. \quad (4.110)$$

In this representation, the quantum state corresponding to the coherent state $|\gamma_i\rangle$ is then defined in $\mathcal{H}'_0^{\otimes M}$ with the tensor product of qubit (4.109) following the definitions (4.2)-(4.3).

A projective measurement on \mathcal{H}'_0 with outcomes $z_k = \hat{0}, \hat{1}$ can be written without loss of generality with operators,

$$|\mu_{k,\hat{0}}\rangle = \cos\phi_k |x\rangle + \sin\phi_k |\hat{x}\rangle, \quad |\mu_{k,\hat{1}}\rangle = \sin\phi_k |x\rangle - \cos\phi_k |\hat{x}\rangle \quad (4.111)$$

described by the angle $\phi_k \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, where the subscript indicate a generic dependances of the angle ϕ by some variables such as the slot number or previous outcomes.

Transition probabilities can be evaluated using the Born Rule (1.19), in this case resulting in

$$P [z_k = \hat{0} | |\gamma_{k,i}\rangle = |\tilde{\gamma}_0\rangle] = p_{\hat{0}|0} = |\langle m_{\hat{0}} | \tilde{\gamma}_0 \rangle|^2 = \cos^2(\theta - \phi_k), \quad (4.112)$$

$$P [z_k = \hat{1} | |\gamma_{k,i}\rangle = |\tilde{\gamma}_0\rangle] = p_{\hat{1}|0} = |\langle m_{\hat{1}} | \tilde{\gamma}_0 \rangle|^2 = \sin^2(\theta - \phi_k), \quad (4.113)$$

$$P [z_k = \hat{0} | |\gamma_{k,i}\rangle = |\tilde{\gamma}_1\rangle] = p_{\hat{0}|1} = |\langle m_{\hat{0}} | \tilde{\gamma}_1 \rangle|^2 = \cos^2(\theta + \phi_k), \quad (4.114)$$

$$P [z_k = \hat{1} | |\gamma_{k,i}\rangle = |\tilde{\gamma}_1\rangle] = p_{\hat{1}|1} = |\langle m_{\hat{1}} | \tilde{\gamma}_1 \rangle|^2 = \sin^2(\theta + \phi_k). \quad (4.115)$$

In this Section we analyze the operators (4.111) associated with the local measurement in each slot of the receiver scheme presented previously, to highlight their limitations.

The measurement operator corresponding to the direct detection (4.71) are

$$\begin{aligned} |\mu_{k,\hat{0}}\rangle &= |\tilde{\gamma}_0\rangle = \cos \theta |x\rangle + \sin \theta |\hat{x}\rangle \\ |\mu_{k,\hat{1}}\rangle &= \sin \theta |x\rangle - \cos \theta |\hat{x}\rangle \end{aligned} \quad (4.116)$$

such that we obtain the transition probabilities (4.72)-(4.74)

$$\begin{aligned} p_{\hat{0}|0} &= 1 \\ p_{\hat{0}|1} &= \chi \\ p_{\hat{1}|0} &= 0 \\ p_{\hat{1}|1} &= 1 - \chi \end{aligned} \quad (4.117)$$

In the qubit framework, the nulling operation inverts the role of the quantum states $|0\rangle$ and $|\alpha\rangle$, therefore changing the transition probabilities of the outcomes z_k . The corresponding projectors in the qubit representation are

$$\begin{aligned} |\mu_{k,\hat{1}}\rangle &= |\tilde{\gamma}_1\rangle = \cos \theta |x\rangle - \sin \theta |\hat{x}\rangle \\ |\mu_{k,\hat{0}}\rangle &= \sin \theta |x\rangle + \cos \theta |\hat{x}\rangle \end{aligned} \quad (4.118)$$

that give transition probabilities

$$\begin{aligned}
 p_{\hat{0}|0} &= \chi \\
 p_{\hat{0}|1} &= 0 \\
 p_{\hat{1}|0} &= 1 - \chi \\
 p_{\hat{1}|1} &= 1
 \end{aligned} \tag{4.119}$$

exactly matching those in equations (4.89).

In the case of Direct Detection, the measurement projectors (4.116) are measured in each slot. Instead, the Unconditional Nulling receiver implements in each slot the measurements associated with (4.118). The Conditional Nulling receiver uses an adaptive strategy, measuring with (4.116) or with (4.118) depending on the previous outcome.

The Improved Conditional Nulling receiver uses the same adaptive strategy as the Conditional Nulling. When the displacement β and the phase amplification G are employed, the angle ϕ_k that defines the measurement operators (4.111) in the slot can be obtained by inversion of the transition probabilities $p_{\hat{0}|0}, p_{\hat{1}|1}$ given in (4.108), resulting in

$$\phi = \frac{1}{2} \angle \frac{\sin 2\theta(p_{\hat{0}|0} - p_{\hat{1}|1}) + j \cos 2\theta(p_{\hat{0}|0} + p_{\hat{1}|1} - 1)}{\sin 2\theta(p_{\hat{0}|0} - p_{\hat{1}|1}) + j \cos 2\theta(p_{\hat{0}|0} + p_{\hat{1}|1} - 1)}. \tag{4.120}$$

where \angle is the four-quadrant inverse tangent³.

4.5 An Adaptive Receiver for PPM

The Conditional Nulling scheme and its improved versions have some limitations in both the strategy and in the local measurements. First, in both architectures the choice of whether to apply the displacement or not in the next measurement depends only upon the last outcome, while in a more general scheme it would depend upon all the previous outcomes. Second, the local transition probabilities obtained with a photon counting, possibly preceded by a displacement

³The function $\phi = \angle a + jb$ returns the argument of the complex number $a + jb = \sqrt{a^2 + b^2} e^{j\phi}$, with $j = \sqrt{-1}$ the imaginary unit.

and a squeezing operation, are more restricted than the one described by the counterpart of (4.111) in the coherent space. It was pointed out in [59] that further performance improvements can be obtained by considering different displacements β_k for each slot $k = 1 \dots M$. Moreover, further generalization leads to time varying displacements $\beta_k(t)$. Third, the decision tree of these architecture is not symmetrical, and to direct detection of all the slots after outcome $z_k = \hat{1}$ has been observed in a nulled slot may not be the best strategy.

In this Section we present a receiver scheme that overcomes these limitations. We propose a general structure for an adaptive receiver, where the next measurements are decided upon *all* the previous outcomes \bar{z}_k . The receiver algorithm defines a perfect binary tree with M levels, where each node corresponds to a slot measurement and each edge to an outcome (see Figure 4.1). In order to focus on the transition probabilities, we use the qubit representation to describe the measurement in the $k + 1$ -th slot with the angle ϕ_k , i.e. after k outcomes has been observed. In fact, due to the Dolinar receiver on the binary discrimination of coherent states and the results by Takeoka *et al.* [60], with linear optics and photon counting any projective measurements can be implemented on \mathcal{H}_0 . We specify the function $\phi_k = \pi_k(\bar{z}_k)$ to define the adaptive strategy.

The receiver starts with the first measurement specified by ϕ_1 . Then, depending on the outcome $z_1 = \hat{0}$ or $z_1 = \hat{1}$, it proceeds with a measurement in the second slot defined by $\phi_2 = \pi_2(\hat{0})$ or $\phi_2 = \pi_2(\hat{1})$ respectively. In general it results $\pi_2(\hat{0}) \neq \pi_2(\hat{1})$. The receiver continues to perform measurements following the path indicated by the outcomes sequence. After the last measurement, the final estimation is taken based on the whole outcome sequence \bar{z}_M .

This receiver structure is a generalization of the previously seen adaptive receiver. In order to achieve optimal performance, an optimization of all the angles $\pi_k(\bar{z}_k)$, for all k and \bar{z}_k is necessary. Since the number of angles grows exponentially in the number of levels, that is the PPM cardinality, the optimization of the final probability of correct decision is highly demanding. However, we can sim-

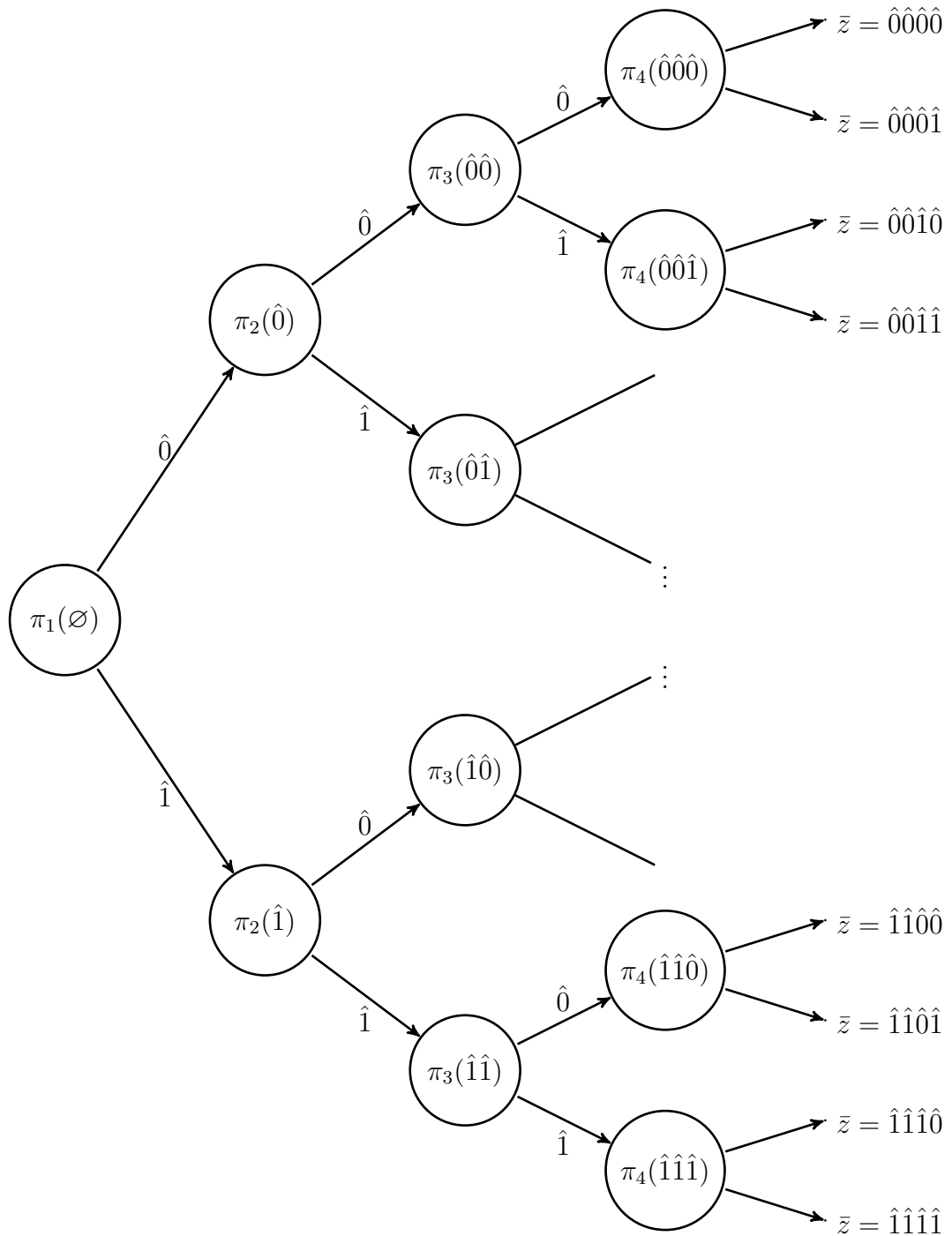


Figure 4.1: Strategy tree for an adaptive receiver algorithm, in the case of a 4-PPM. Each node represents a measurements, identified by the angle $\phi_k = \pi_k(\bar{z}_k)$ employed at the measurement $k + 1$ after observing the outcomes sequence \bar{z}_k , and each branch a possible outcome. The algorithm proceeds from the root on the left to the leaves on the right following the path indicated by the outcomes.

plify the optimization problem by applying the dynamic programming algorithm, that is the topic of the next Section.

In order to achieve an optimal performance, an optimization of all the angles ϕ_k is necessary. Since the number of angles grows exponentially in the number of levels, that is the PPM cardinality, an optimization of the final probability of correct decision is demanding, but applying dynamic programming we can define an optimization algorithm. For a review of the dynamic programming optimization, see Appendix 4.A.

4.5.1 State of the Algorithm

The dynamic programming algorithm applies to dynamic systems, whose time evolution is described by a system of equations involving its system state. In our case, the time evolution occurs in discrete time steps.

We refer to the iteration k of the algorithm as the moment just *after* the sequence \bar{z}_k has been observed, that is *after* the k -th measurement and *before* the $k + 1$ -th one. When $k = M$, no more measurements are performed but a final estimation is made in order to choose \hat{x} . By extension, we can define $k = 0$ the time step before the measurement in the first slot.

The final performance index (1.21) is a function of the joint probabilities between the possible outcomes and the transmitted symbols. Therefore, it seems natural to consider how these probabilities evolve during the execution of the receiver scheme. The information acquired at slot k can be described by the evolution of the vector of the joint probabilities

$$\bar{p}_{\bar{z}_k} = \begin{bmatrix} p_{1,\bar{z}_k} \\ p_{2,\bar{z}_k} \\ \vdots \\ p_{M,\bar{z}_k} \end{bmatrix}, \quad p_{i,\bar{z}_k} = P[x = i, \bar{z}_k]. \quad (4.121)$$

However, not all the entries in (4.121) are necessary in order to define the *system state* of the receiver algorithm. In fact, given the outcome sequence \bar{z}_k , it turns

out that the system state $s_k \in S_k$ of the algorithm at step k corresponds to the triple

$$s_k(\bar{z}_k) = (p_{m, \bar{z}_k}, p_{M, \bar{z}_k}, m(\bar{z}_k)), \quad (4.122)$$

$$S_k = \{(u, v, i) : 0 \leq u \leq 1, 0 \leq v \leq 1, \\ u + v \leq 1, i \in \{1, \dots, k\}\}, \quad (4.123)$$

where $m(\bar{z}_k)$ is the maximum a posteriori (MAP) estimate of x given the observed sequence \bar{z}_k among the symbols $1, \dots, k$, i.e.

$$m(\bar{z}_k) = \operatorname{argmax}_{i \leq k} p_{i, \bar{z}_k}, \quad (4.124)$$

with $p_{m(\bar{z}_k), \bar{z}_k}$ indicating the joint probability

$$p_{m(\bar{z}_k), \bar{z}_k} = \max_{i \leq k} p_{i, \bar{z}_k}. \quad (4.125)$$

We highlight that the system state s_k is a random variable, that depends upon the realization of the outcomes sequence \bar{z}_k . However, to shorten the notation, when assuming a given \bar{z}_k , we drop its dependency.

We show that by definition (4.122) we can describe the evolution of the system state with an update equation. Later, in the next Section, we show that we can write the probability of correct decision as a function of $s_M(\bar{z}_M)$.

Consider the system state $s_k(\bar{z}_k)$ and the outcome sequences that can be generated from \bar{z}_k with $z_{k+1} = \hat{0}$ or $z_{k+1} = \hat{1}$, employing the angle $\phi := \pi_k(\bar{z}_k)$. The joint probabilities $p_{M, [\bar{z}_k \hat{0}]}, p_{M, [\bar{z}_k \hat{1}]}$ can be easily obtained from the transition probabilities in equations (4.115), in the case $k < M - 1$ by

$$\begin{aligned} p_{M, [\bar{z}_k \hat{0}]} &= \cos^2(\theta - \phi) p_{M, \bar{z}_k}, \\ p_{M, [\bar{z}_k \hat{1}]} &= \sin^2(\theta - \phi) p_{M, \bar{z}_k}, \end{aligned} \quad (4.126)$$

while if $k = M - 1$ by

$$\begin{aligned} p_{M, [\bar{z}_k \hat{0}]} &= \cos^2(\theta + \phi) p_{M, \bar{z}_k}, \\ p_{M, [\bar{z}_k \hat{1}]} &= \sin^2(\theta + \phi) p_{M, \bar{z}_k}. \end{aligned} \quad (4.127)$$

Note that given the system state $s_k(\bar{z}_k)$, the probability p_{M, \bar{z}_k} equals the joint probability

$$p_{i, \bar{z}_k} = P[x = i, \bar{z}_k], \quad i = k + 1, k + 2, \dots, M$$

due to the equal a priori probability and the same quantum states $|\tilde{\gamma}_0\rangle$ in the first k slots for all symbols $x = k + 1, \dots, M$ (see Lemma 4.3 in Appendix 4.5.2).

In the case of $p_{m(\bar{z}_k \hat{0}), [\bar{z}_k \hat{0}]}$ and $p_{m(\bar{z}_k \hat{1}), [\bar{z}_k \hat{1}]}$, in order to obtain the update equation we apply the definition (4.125),

$$\begin{aligned} p_{m(\bar{z}_k \hat{0}), [\bar{z}_k \hat{0}]} &= \max_{i \leq k+1} \{p_{i, [\bar{z}_k \hat{0}]}\} \\ &= \max \{ \cos^2(\theta + \phi) p_{k+1, \bar{z}_k}, \max_{i \leq k} \{ \cos^2(\theta - \phi) p_{i, \bar{z}_k} \} \} \\ &= \max \{ \cos^2(\theta + \phi) p_{M, \bar{z}_k}, \cos^2(\theta - \phi) p_{m(\bar{z}_k), \bar{z}_k} \}, \end{aligned} \quad (4.128)$$

$$\begin{aligned} p_{m(\bar{z}_k \hat{1}), [\bar{z}_k \hat{1}]} &= \max_{i \leq k+1} \{p_{i, [\bar{z}_k \hat{1}]}\} \\ &= \max \{ \sin^2(\theta + \phi) p_{k+1, \bar{z}_k}, \max_{i \leq k} \{ \sin^2(\theta - \phi) p_{i, \bar{z}_k} \} \} \\ &= \max \{ \sin^2(\theta + \phi) p_{M, \bar{z}_k}, \sin^2(\theta - \phi) p_{m(\bar{z}_k), \bar{z}_k} \}. \end{aligned} \quad (4.129)$$

Thereby, the symbol $m(\bar{z}_k)$ is updated into

$$m([\bar{z}_k \ z_{k+1}]) \in \{m(\bar{z}_k), k + 1\} \quad (4.130)$$

accordingly with the term maximizing (4.128) and (4.129). This means that at each update of the system state, the symbol $m(\bar{z}_k)$ can be replaced only by the symbol corresponding to the current slot.

Given the update equations, we have to specify the initial system state s_0 of the algorithm before the first measurement, that does not depend upon any outcomes, $\bar{z}_0 = \emptyset$. This system state collects the a priori informations, and due to the equal a priori distribution for the symbols, we can define

$$s_0 = \left(\frac{1}{M}, \frac{1}{M}, \dots \right), \quad (4.131)$$

where it is unnecessary to specify $m(\emptyset)$, since the first update of the system state leads to the trivial maximization

$$m(z_1) = \operatorname{argmax}_{i \leq 1} p_{i, z_1} = 1. \quad (4.132)$$

The update equations (4.126)-(4.130) suggest the role of the angles $\phi_k = \pi_k(\bar{z}_k)$ as the control variables we can use to tune the transition probabilities. In the next Section we formulate the corresponding expected reward function, and following the dynamic programming we define ϕ_k as function of the state s_k ,

$$\phi_k = \pi_k(s_k(\bar{z}_k)) \quad (4.133)$$

optimizing for each \bar{z}_k .

4.5.2 Useful Lemmas

In this Section we give some useful Lemmas that helps to better understand the optimization algorithm.

Lemma 4.1. *Let $p_0, p_1, p_0 + p_1 \leq 1$ be the a priori probability of two symbols $i = 0, 1$ associated to the quantum states $|\tilde{\gamma}_0\rangle, |\tilde{\gamma}_1\rangle$ respectively. The maximum probability of correct discrimination between $i = 0$ and $i = 1$ achievable with measurement operators (4.111) defined by ϕ is*

$$\frac{1}{2} \left[p_0 + p_1 + \sqrt{(p_0 + p_1)^2 - 4p_0p_1\chi} \right], \quad (4.134)$$

and the optimal angle defined by

$$\phi^* = \frac{1}{2} \arccos \left[\cos 2\theta(p_0 - p_1) + \chi \sin 2\theta(p_0 + p_1) \right]. \quad (4.135)$$

Proof. Consider the measurement operators (4.111), and without loss of generality associate the outcomes $z = \hat{0}$ with the estimation of $i = 0$ and $z = \hat{1}$ with $i = 1$. The probability of correct discrimination can be written as

$$\begin{aligned} P_c &= P[z = \hat{0}|i = 0] p_0 + P[z = \hat{1}|i = 1] p_1 \\ &= \cos^2(\theta - \phi)p_0 + \sin^2(\theta + \phi)p_1 \\ &= \frac{1 + \cos(2\theta - 2\phi)}{2} p_0 + \frac{1 - \cos(2\theta + 2\phi)}{2} p_1 \end{aligned} \quad (4.136)$$

with transition probabilities defined by (4.115). The maximization of (4.136) with respect to the angle ϕ leads to the relation

$$\tan 2\phi = \tan 2\theta \frac{p_0 + p_1}{p_0 - p_1}, \quad (4.137)$$

solved by the angles verifying

$$\begin{aligned}\sin 2\phi &= \pm \frac{p_0 + p_1}{\sqrt{R}} \sin 2\theta, \\ \cos 2\phi &= \pm \frac{p_0 - p_1}{\sqrt{R}} \cos 2\theta,\end{aligned}\tag{4.138}$$

with R a normalization term,

$$R = \cos^2 2\theta (p_0 - p_1)^2 + \sin^2 2\theta (p_0 + p_1)^2.\tag{4.139}$$

Expression (4.138) with the plus sign corresponds to the point of maximum, and the thesis (4.135) follows. Substituting (4.135) in (4.136) gives

$$\begin{aligned}P_c &= \frac{1}{2} \left[p_0 + p_1 + \sqrt{R} \right] \\ &= \frac{1}{2} \left[p_0 + p_1 + \sqrt{(p_0 + p_1)^2 - 4p_0p_1 \cos^2 2\theta} \right].\end{aligned}\tag{4.140}$$

□

Corollary 4.1. *If the quantum states associated to the two symbols $i = 0, 1$ is the same, i.e. $|\tilde{\gamma}_0\rangle = |\tilde{\gamma}_1\rangle$, the maximal probability of correct discrimination between $i = 0$ and $i = 1$ is the maximum of their a priori probability,*

$$P_c = \max\{p_0, p_1\}.\tag{4.141}$$

The Corollary follows from $\chi = |\langle \tilde{\gamma}_0 | \tilde{\gamma}_1 \rangle|^2 = 1$. Intuitively, this means that in the case of the same quantum states, we cannot discriminate the symbols $i = 0, 1$ better than their a priori distribution.

Moreover, since (4.134) is always non lower than (4.141), at the last measurement it is always better to discriminate between the last symbol $x = M$ and a previous one $x = i < M$. This is reasonable since the last slot would eventually deliver information about the last symbol, and it is useless to discriminate previous ones.

In addition, expression (4.134) is monotonically increasing with the probabilities p_0, p_1 , such that it is always better to compare the symbol $x = i < M$ with highest a priori probability.

Lemma 4.2. *The relative ordering of the a priori probabilities of symbols $i, j \neq k$ before the k -th measurement is maintained in the a posteriori distribution, independently of the outcome z_k .*

Proof. Consider two symbols $x = i$ and $x = j$, $i, j \neq k$ with a priori probabilities $p_{i,z_{k-1}}, p_{j,z_{k-1}}$ before the k -th measurement, with $p_{i,z_{k-1}} > p_{j,z_{k-1}}$. Both symbols i, j has a quantum state $|\tilde{\gamma}_0\rangle$ in the k -th position of $|\gamma_i\rangle, |\gamma_j\rangle$. Hence, the transition probability is the same,

$$P[z_k|i] = |\langle m_{z_k} | \tilde{\gamma}_0 \rangle|^2 = P[z_k|j] \quad (4.142)$$

and the joint probabilities are multiplied by the same factor

$$p_{i,z_k} = |\langle m_{z_k} | \tilde{\gamma}_0 \rangle|^2 p_{i,z_{k-1}}, \quad p_{j,z_k} = |\langle m_{z_k} | \tilde{\gamma}_0 \rangle|^2 p_{j,z_{k-1}}, \quad (4.143)$$

and hence the a posteriori distribution of $x = i, j \neq k$ reflects the same relative ordering of the priori. \square

This Lemma is a consequence of the fact that $i, j \neq k$ has the same quantum states $|\tilde{\gamma}_0\rangle$ in position k . The k -th measurement does not give any information about the discrimination between symbols $i, j \neq k$ because they all behave in the same way in this slot. The k -th slot can give informations only for the the discrimination of the k -th symbol, whether it is more likely or not respect the other.

Corollary 4.2. *Given an outcome sequence \bar{z}_k , the relative ordering of the joint probabilities of symbols $x = 1, \dots, k$ is maintained in the joint probabilities after measurement $k + 1, \dots, M$, independently of the outcomes z_{k+1}, \dots, z_M .*

Proof. The quantum states of symbols $x = 1, \dots, k$ in position $k + 1, \dots, M$ of the tensor product are $|\tilde{\gamma}_0\rangle$, so the joint probabilities after measurement $l > k$ is

$$P_{i, [\bar{z}_k z_{k+1} \dots z_l]} = P[z_l | z_{l-1}, \dots, \bar{z}_k, |\tilde{\gamma}_0\rangle] \cdot \dots \cdot P[z_{k+1} | \bar{z}_k, |\tilde{\gamma}_0\rangle] p_{i, \bar{z}_k} \quad (4.144)$$

that is, all the joint probabilities are multiplied by the same factors. \square

Lemma 4.3. *The joint probabilities of symbols $x = k + 1, k + 2, \dots, M$ with \bar{z}_l after any measurement in the l -th slot, $l = 1, \dots, k$, verify $p_{k+1, \bar{z}_l} = p_{k+2, \bar{z}_l} = \dots = p_{M, \bar{z}_l}$.*

Proof. Consider a measurement $l \in 1, \dots, k$. The joint probabilities of symbols $x = k + 1, k + 2, \dots, M$ with the outcomes vector \bar{z}_l can be calculated as

$$p_{x, \bar{z}_l} = P[z_l | \bar{z}_{l-1}, x] P[z_{l-1} | \bar{z}_{l-2}, x] \cdots P[z_1 | x] \frac{1}{M}, \quad (4.145)$$

but since each symbol $x = k + 1, \dots, M$ has a quantum state $|\tilde{\gamma}_0\rangle$ in position $1, \dots, l$, $l \leq k$, the conditional probabilities give

$$p_{x, \bar{z}_l} = |\langle m_{\bar{z}_l} | \tilde{\gamma}_0 \rangle|^2 |\langle m_{\bar{z}_{l-1}} | \tilde{\gamma}_0 \rangle|^2 \cdots |\langle m_{\bar{z}_1} | \tilde{\gamma}_0 \rangle|^2 \frac{1}{M}, \quad (4.146)$$

and hence $p_{k+1, \bar{z}_l} = p_{k+2, \bar{z}_l} = \dots = p_{M, \bar{z}_l}$. \square

Intuitively, since all symbols $x = k + 1, k + 2, \dots, M$ in position $1, \dots, k$ has quantum states equal to $|\tilde{\gamma}_0\rangle$, we cannot gain any information about their discrimination on the base of the first k outcomes, and therefore their joint probabilities are the same.

Lemma 4.4. *Joint probabilities are non increasing in subsequent measurements, and always lower than the a priori probability $\frac{1}{M}$.*

Proof. Writing the joint probability with the conditional chain rule

$$p_{i, \bar{z}_k} = P[z_k | \bar{z}_{k-1}, i] P[z_{k-1} | \bar{z}_{k-2}, i] \cdots P[z_1 | i] \frac{1}{M} \quad (4.147)$$

we see that after each measurement, the joint probabilities are updated with the transition probabilities depending upon the outcome. Since the transition probabilities are not greater than 1, they are non increasing, and it is clear that $p_{i, \bar{z}_k} \leq \frac{1}{M}$. \square

4.5.3 Expected reward function

In this section we rewrite the probability of correct decision as a function of the system state and find a suitable definition in terms of the reward-to-go functions.

In general, we can identify two stages in a receiver strategy: a measurement stage, that is the task of acquiring information about the system under investigation, and an estimation stage, that is the elaboration of the information gained in order to formulate the answer to our detection problem. An optimization of the receiver scheme requires of course the optimization of both stages.

The estimation stage can be defined by a map h that assigns to each measurement outcome sequences \bar{z}_M an estimate \hat{x} of the transmitted symbol x . In particular, the map defines a partition of the space \mathcal{Z}_M of all possible outcomes sequences in disjoint subsets $\mathcal{R}_{\hat{x}}$, $\hat{x} \in 1, \dots, M$, each one associated to an estimation,

$$\begin{aligned} h : \mathcal{Z}_M &\longrightarrow \{1, \dots, M\} \\ z &\longmapsto \hat{x} : z \in \mathcal{R}_{\hat{x}} \end{aligned} \quad (4.148)$$

The final probability of correct decision is therefore rewritten as

$$\begin{aligned} P_c &= P[\hat{x} = x] = \sum_{i=1}^M P[\hat{x} = i, x = i] \\ &= \sum_{i=1}^M \sum_{z \in \mathcal{Z}_M} P[\hat{x} = i, x = i, \bar{z}_M = z] \\ &= \sum_{i=1}^M \sum_{z \in \mathcal{Z}_M} P[\hat{x} = i | x = i, \bar{z}_M = z] p_{i,z} \end{aligned} \quad (4.149)$$

and by definition (4.148), given the observation \bar{z}_M the estimate \hat{x} is deterministic,

$$P[\hat{x} = i | x = i, \bar{z}_M = z] = \begin{cases} 1 & \text{if } z \in \mathcal{R}_i \\ 0 & \text{otherwise} \end{cases} \quad (4.150)$$

and the probability of correct decision is written as function of joint probabilities between transmitted symbols and outcome sequences,

$$P_c = \sum_{z \in \mathcal{Z}_M} P[x = h(z), \bar{z}_M = z]. \quad (4.151)$$

After the measurement stage, for any measurement strategy employed from slot 1 to slot M , we can evaluate the joint probabilities p_{x, \bar{z}_M} for each x and \bar{z}_M . Given

these, in order to maximize (4.151), we define h as the indicator function (4.150) that selects the maximum joint probability (MAP criterium),

$$h(z) = \operatorname{argmax}_{i \in \{1, \dots, M\}} p_{i,z}, \quad (4.152)$$

and the probability of correct decision reads

$$P_c = \sum_{z \in \mathcal{Z}_M} p_{h(z),z} = \sum_{z \in \mathcal{Z}_M} \max_{i \leq M} p_{i,z} = \sum_{z \in \mathcal{Z}_M} p_{m(z),z} \quad (4.153)$$

as a function of the system state s_M .

Expression (4.153) corresponds to the expression (4.183), and suggests to define

$$g(s_M(z)) = p_{m(z),z}, \quad (4.154)$$

$$\mathbb{E}_{s_M} [g(s_M)] = \sum_{z \in \mathcal{Z}_M} p_{m(z),z}. \quad (4.155)$$

Consider the system state $s_k(z)$ before the $k + 1$ -th measurement, and define the expected reward function

$$\begin{aligned} J_k(s_k(z), \phi_k, \pi_{k+1}, \dots, \pi_{M-1}) &= \\ &= \sum_{z' \in \mathcal{Z}_{M-k}} p_{m([z \ z'], [z \ z'])(s_k(z), \phi_k, \pi_{k+1}, \dots, \pi_{M-1})} \end{aligned} \quad (4.156)$$

where the set \mathcal{Z}_{M-k} contains all the possible sequences $[z_{k+1} \dots z_M]$ composed by $M - k$ outcomes. The dependency upon the variables $s_k(z), \phi_k, \pi_{k+1}, \dots, \pi_{M-1}$ in the RHS of (4.156) is included the transition probabilities of the terms in the sum,

$$p_{i, \bar{z}_M} = p_{z_M | \bar{z}_{M-1}, i}(\pi_{M-1}) \cdots p_{z_{k+1} | \bar{z}_M, i}(\phi_k) p_{i, \bar{z}_M} \quad (4.157)$$

where $i \in \{m(z), k + 1, \dots, M\}$ by the update equation (4.130).

It is trivial to see that

$$J_0(s_0, \pi_0(s_0), \dots, \pi_{M-1}) = P_c \quad (4.158)$$

and

$$\begin{aligned}
J_{M-1}(s_{M-1}(z), \phi_{M-1}) &= \sum_{z_M=\hat{0}, \hat{1}} p_{m([z \ z_M]), [z \ z_M]}(\phi_{M-1}) \\
&= p_{m([z \ \hat{0}], [z \ \hat{0}])}(\phi_{M-1}) + p_{m([z \ \hat{1}], [z \ \hat{1}])}(\phi_{M-1}) \\
&= P [z_M = \hat{0} | x = i] p_{i,z} + P [z_M = \hat{1} | x = j] p_{j,z}.
\end{aligned} \tag{4.159}$$

with $i = h([z \ \hat{0}])$ and $j = h([z \ \hat{1}])$. As explained in the Lemma 4.1 in Appendix 4.5.2, expression (4.159) is the probability of correct decision of the binary discrimination problem between symbols $i = m(z)$ and $j = M$, with the joint probabilities $p_{i,z}, p_{j,z}$ provided by the system state $s_{M-1}(z)$. The expression admits an analytical maximisation

$$J_{M-1}^*(s_{M-1}) = \frac{1}{2} \left[p_m + p_M + \sqrt{(p_m + p_M)^2 - 4p_m p_M \chi} \right] \tag{4.160}$$

obtained employing the angle in the M -th measurement

$$\pi_{M-1}^*(s_{M-1}) = \frac{1}{2} \angle \cos 2\theta(p_m - p_M) + j \sin 2\theta(p_m + p_M) \tag{4.161}$$

where in both (4.160) and (4.161) we drop the dependency from \bar{z}_{M-1} .

Moreover, we can easily write down the update equation for the expected reward

$$\begin{aligned}
J_k(s_k(z), \phi_k, \pi_{k+1}, \dots, \pi_{M-1}) &= \sum_{\substack{z' \in \\ \mathcal{Z}_{M-k}}} p_{m([z \ z'], [z \ z'])} \\
&= \sum_{z_{k+1}} \sum_{\substack{z'' \in \\ \mathcal{Z}_{M-k-1}}} p_{m([z \ z_{k+1} \ z''], [z \ z_{k+1} \ z''])} \\
&= J_{k+1}(s_{k+1}([z \ \hat{0}]), \pi_{k+1}, \pi_{k+2}, \dots, \pi_{M-1}) \\
&\quad + J_{k+1}(s_{k+1}([z \ \hat{1}]), \pi_{k+1}, \pi_{k+2}, \dots, \pi_{M-1})
\end{aligned} \tag{4.162}$$

In equation (4.162), the role of ϕ_k comes into the update of the system state. In fact, for each outcome z_{k+1} , two evolutions of $s_k([z \ z_{k+1}])$ are possible, that depend on $\phi = \phi_k$ as indicated in (4.128) and (4.129). Therefore, four possible

evolutions must be considered in evaluating (4.162), indicated with A, B, C and D in equations (4.163)-(4.166). Since we want to maximize the probability of correct decision, $J_k^*(s_k)$ is the maximum between these possibilities, as in (4.167).

$$J_{k,A}(s_k, \phi) = J_{k+1}^* \left(\overbrace{\cos^2(\theta - \phi)p_m, \cos^2(\theta - \phi)p_M, m}^{s_{k+1}([\bar{z}_k \hat{0}])} \right) + J_{k+1}^* \left(\overbrace{\sin^2(\theta - \phi)p_m, \sin^2(\theta - \phi)p_M, m}^{s_{k+1}([\bar{z}_k \hat{1}])} \right), \quad (4.163)$$

$$J_{k,B}(s_k, \phi) = J_{k+1}^* \left(\cos^2(\theta - \phi)p_m, \cos^2(\theta - \phi)p_M, m \right) + J_{k+1}^* \left(\sin^2(\theta + \phi)p_M, \sin^2(\theta - \phi)p_M, k + 1 \right), \quad (4.164)$$

$$J_{k,C}(s_k, \phi) = J_{k+1}^* \left(\cos^2(\theta + \phi)p_M, \cos^2(\theta - \phi)p_M, k + 1 \right) + J_{k+1}^* \left(\sin^2(\theta - \phi)p_m, \sin^2(\theta - \phi)p_M, m \right), \quad (4.165)$$

$$J_{k,D}(s_k, \phi) = J_{k+1}^* \left(\cos^2(\theta + \phi)p_M, \cos^2(\theta - \phi)p_M, k + 1 \right) + J_{k+1}^* \left(\sin^2(\theta + \phi)p_M, \sin^2(\theta - \phi)p_M, k + 1 \right) \quad (4.166)$$

$$J_k^*(s_k) = \max_{\phi} \{J_{k,A}(s_k, \phi), J_{k,B}(s_k, \phi), J_{k,C}(s_k, \phi), J_{k,D}(s_k, \phi)\} \quad (4.167)$$

$$\pi_k^*(s_k) = \operatorname{argmax}_{\phi} \{J_{k,AC}(\phi), J_{k,AD}(\phi), J_{k,BC}(\phi), J_{k,BD}(\phi)\} \quad (4.168)$$

Along with the reward-to-go function, we define the function π_k^* that represents the optimal value of the control variable corresponding to the current system state, $\phi_k = \pi_k^*(s_k)$ to employ in the measurement $k + 1$ given that the outcome sequence \bar{z}_k has been observed.

4.5.4 Dynamic Programming Algorithm

In the dynamic programming algorithm, we have to evaluate the reward-to-go function at iteration k for each possible values of the system state s_k . The expression of J_k depends on the variables p_m, p_M and not upon the particular sequence \bar{z}_k . Therefore, in the following we drop the dependence of the system state from \bar{z}_k .

The optimization algorithm to evaluate the performance of the adaptive receiver can be summarized by the following step by step procedure:

1. Evaluate the reward-to-go function J_{M-1}^* and the angle function π_{M-1}^* for each (p_m, p_M) , $p_m + p_M \leq 1$ as in (4.160) and (4.161).
2. From J_{k+1}^* , for each (p_m, p_M) , $p_m + p_M \leq 1$ evaluate J_k^* and π_k^* as in (4.167) and (4.168) respectively.
3. For each (p_m, p_M) , $p_m + p_M \leq 1$, depending on the association A, B, C or D of (4.163)-(4.166) used in the previous step, define the children nodes of $s_k(z_k)$ generated with outcome $z_{k+1} = \hat{0}$ and $z_{k+1} = \hat{1}$

$$\text{children}(s_k) = \{s_{k+1}([\bar{z}_k \hat{0}]), s_{k+1}([\bar{z}_k \hat{1}])\} \quad (4.169)$$

Note that in $s_{k+1}([\bar{z}_k \hat{0}])$ and in $s_{k+1}([\bar{z}_k \hat{1}])$ we can define $m([\bar{z}_k \hat{0}])$ and $m([\bar{z}_k \hat{1}])$ only in the case it is equal to $k+1$, while in the case $m([\bar{z}_k z_{k+1}]) = m(\bar{z}_k)$ we cannot assign an exact value, because $m(\bar{z}_k) \in \{1, \dots, k\}$. Instead, we can assign the label “previous” that indicates that it is a value $m \leq k$ that will be defined in later iterations of the optimization.

4. Repeat step 2. and 3. for $k = M - 1, \dots, 1$.
5. Evaluate the angle in the first measurement and the performances of the adaptive algorithm from s_0 as

$$\begin{aligned} \pi_0^* = \operatorname{argmax}_{\phi} J_1^* & \left(\frac{\cos^2(\theta + \phi)}{M}, \frac{\cos^2(\theta - \phi)}{M}, 1 \right) \\ & + J_1^* \left(\frac{\sin^2(\theta + \phi)}{M}, \frac{\sin^2(\theta - \phi)}{M}, 1 \right) \end{aligned} \quad (4.170)$$

$$\begin{aligned} P_c = J_1^* & \left(\frac{\cos^2(\theta + \pi_0^*)}{M}, \frac{\cos^2(\theta - \pi_0^*)}{M}, 1 \right) \\ & + J_1^* \left(\frac{\sin^2(\theta + \pi_0^*)}{M}, \frac{\sin^2(\theta - \pi_0^*)}{M}, 1 \right) \end{aligned} \quad (4.171)$$

In order to reconstruct the binary tree angles and find the estimation region, we need to retrace the optimization steps forward. In the following procedure, two binary trees are built, one with nodes the system states s_k and the other with nodes corresponding to the angle ϕ_k . The levels $k = 0, 1, \dots, M - 1$ of the trees

represent the system state before $k + 1$ -th measurement, the edges between the nodes correspond to a measurement outcome $z_k = \hat{0}$ or $z_k = \hat{1}$. The path from the root to the node gives the outcomes sequence.

In particular, retracing the path of the binary tree we can fill up the system state substituting the labels “previous” with the correct symbol $m(\bar{z}_k)$.

The construction of the binary trees is given by the following steps:

6. Define the initial system state s_\emptyset as the root of the binary tree of the system states.
7. Define $\phi_0 = \pi_0^*(s_0)$ as the root of the tree of the angles.
8. Define the children nodes of the system state s_0 , the one corresponding to the outcome $z_1 = \hat{0}$,

$$s_1(\hat{0}) = \left(\frac{\cos^2(\theta + \pi_0^*)}{M}, \frac{\cos^2(\theta - \pi_0^*)}{M}, 1 \right), \quad (4.172)$$

and the other corresponding to $z_1 = \hat{1}$,

$$s_1(\hat{1}) = \left(\frac{\sin^2(\theta + \pi_0^*)}{M}, \frac{\sin^2(\theta - \pi_0^*)}{M}, 1 \right). \quad (4.173)$$

9. For each node $s_k(\bar{z}_k)$ of the level k in the binary angles tree, the angle corresponding to the *next* measurement is

$$\pi_k^*(s_k(\bar{z}_k)) \quad (4.174)$$

and in the next level of the system state tree add $s_{k+1}([\bar{z}_k\hat{0}])$ and $s_{k+1}([\bar{z}_k\hat{1}])$, replacing, if present, the label “previous” with the symbol $m(s_k(\bar{z}_k))$.

10. Repeat step 9. for $k = 2, \dots, M - 1$.

Once completed these trees, following the outcome sequence through the angle tree we get the angle $\phi_k = \pi_k(s_k(\bar{z}_k))$ to employ in the $k + 1$ -th measurement. The region of estimation are defined in this way: if the sequence ends in $z_M = \hat{1}$, it is attributed to $\hat{x} = M$, otherwise for $z_M = \hat{0}$ it is assigned to $\hat{x} = m(s_{M-1}(\bar{z}_{M-1}))$.

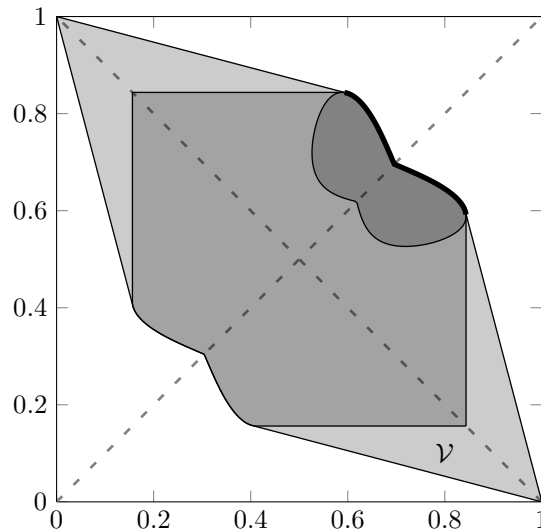


Figure 4.1: (Color online) Performances comparison of different receiver schemes, for 2-PPM: — classical direct detection, — quantum theoretical limit and adaptive scheme (overlapped), — conditional nulling receiver, - - - type I improved conditional nulling scheme, — type II improved conditional nulling scheme.

4.6 Results and Numerical Issues

In the previous Section we described the algorithm to optimize the sequences of angles ϕ_k used by the adaptive receiver. We run the optimization algorithm for different cardinalities M of the PPM and for different values of the inner product χ . For a fair comparison with respect to the other existing schemes, we compare the performances of the receiver architectures on the base of mean photon number in the coherent state $|\alpha\rangle$, i.e. $|\alpha|^2$, obtained by inversion of (4.110).

A first result is that in the case of $M = 2$ this adaptive receiver reaches the theoretical quantum limit. This is not surprising, because as already pointed out in [28] in the case of binary discrimination of pure states an optimized sequence of local measurements suffices to implement the POVM for the optimal discrimination. In Figure 4.1 the performances of the classical receiver, conditional nulling, type I and type II schemes are compared with respect to the theoretical limit. The

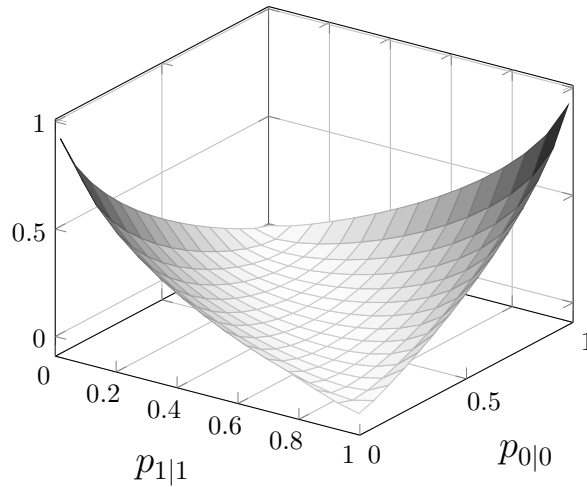


Figure 4.2: (Color online) Performances comparison of different receiver schemes, for 3-PPM: — classical direct detection, — quantum theoretical limit, — adaptive scheme, - - - retraced forward path, — conditional nulling receiver, - - - type I improved conditional nulling scheme, — type II improved conditional nulling scheme.

performance of the adaptive receiver overlaps with the theoretical one.

As the cardinality M increases, the performance of the adaptive receiver slightly moves away from the theoretical optimum. In Figure 4.2, 4.3 and 4.4 the performance of the existing receivers and the optimized adaptive one are compared for $M = 3$, $M = 4$ and $M = 8$ respectively. The trend is the same in all the figures: the adaptive scheme outperforms the existing conditional nulling, type I and type II receiver, placing the error probability curves between these and the theoretical limit. The adaptive scheme maintains the gap with respect to type I and type II even around $|\alpha|^2 = 2$, where these schemes get close to the conditional nulling performances. In addition, the performance of our scheme get really close to the theoretical limit for low mean photon number.

The evaluation of the dynamic programming algorithm can be really demanding, in particular the evaluation of J_k^* for all possible system states s_k may require a non trivial amount of computational time and memory. In addition, this eval-

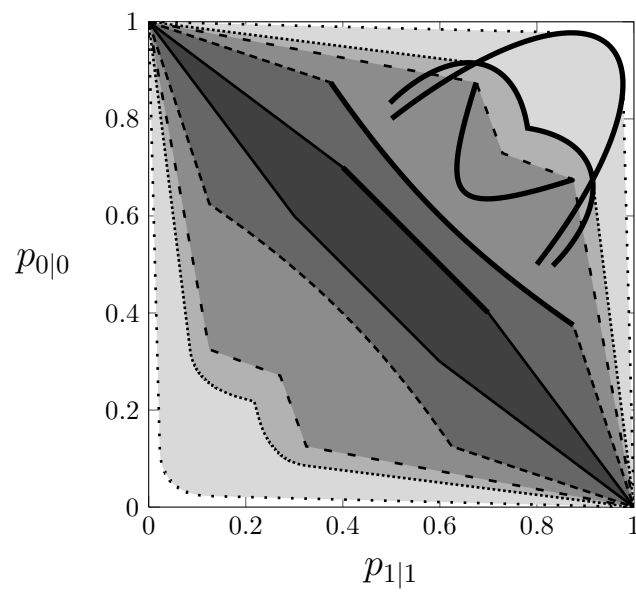


Figure 4.3: Performances comparison of different receiver schemes, for 4-PPM: — classical direct detection, — quantum theoretical limit, — adaptive scheme, retraced forward path, — conditional nulling receiver, type I improved conditional nulling scheme, — type II improved conditional nulling scheme.

uation must be repeated from $k = M - 1$ down to $k = 1$.

Since a numerical procedure is required to evaluate J_k^* at each step, the set $\{(u, v), 0 \leq u \leq 1, 0 \leq v \leq 1, u + v \leq 1\}$ of the system state space S_k needs to be discretized in a two dimensional grid. As a consequence, the search of the optimal angles π_k^* in (4.168) and the evaluation of J_k^* in (4.167) requires to approximate the system state s_{k+1} in the grid when considering $J_{k+1}^*(s_{k+1})$. The issue of this approximation spread out in successive evaluations of J_k^* , especially in the case of poorly discretized grid, where we encounter bad (even unfeasible) results for high values of M and $|\alpha^2|$. In our optimization, we use a discretization with at least a grid step of 10^{-3} for each side of the unit square that includes the set of (u, v) in (4.122).

However, some considerations can be done in order to lighten the computation. The first consideration is that for different cardinality M , the sequence of tables J_k^* to be calculated are the same. This means that if we want to evaluate the performances of the adaptive receiver for a PPM with cardinality $M_1 < M_2 < \dots < \tilde{M}$, we can calculate the table sequence $J_{\tilde{M}}^*, J_{\tilde{M}-1}^*, \dots, J_1^*$ for the maximum cardinality \tilde{M} . In evaluating the performances of the other modulation cardinality M_i , we only need to start from a different initial system state, that is

$$s_0 = \left(\frac{1}{M_i}, \frac{1}{M_i}, \cdot \right) \quad (4.175)$$

end evaluate the probability of correct decision as

$$P_c = \max_{\phi} J_{\tilde{M}-M_i+1}^* \left(\frac{\cos^2(\theta + \phi)}{M_i}, \frac{\cos^2(\theta - \phi)}{M_i}, 1 \right) + J_{\tilde{M}-M_i+1}^* \left(\frac{\sin^2(\theta + \phi)}{M_i}, \frac{\sin^2(\theta - \phi)}{M_i}, 1 \right) \quad (4.176)$$

In addition, as already pointed out and proved in Lemma 4.3, before the $k+1$ -th measurement the joint probability of the $M-k$ symbols $x = k+1, k+2, \dots, M$ are the same. This means that considering the variables p_m, p_M of the system state s_k that define the entries of J_k^* , it results $p_M \leq \frac{1}{M-k+1}$, therefore reducing the elements of the set $\{(u, v)\}$ of S_k to evaluate in

$$\left\{ (u, v), u + v \leq 1, 0 \leq u \leq 1, 0 \leq v \leq \frac{1}{M-k+1} \right\}. \quad (4.177)$$

Furthermore, if we are interested in the performance for a single value of the cardinality M , we can take advantage of Lemma 4.4 in Appendix 4.5.2. Since the elements p_m and p_M are *joint* probabilities of symbols with the outcome sequence, and since they start from the value $1/M$, we can restrict the grid to evaluate for table J_k^* to consider only the set

$$\left\{ (u, v), 0 \leq u \leq \frac{1}{M}, 0 \leq v \leq \frac{1}{M} \right\}. \quad (4.178)$$

In order to understand the consequence of the approximation of the system state space S_k , we check the performances of the dynamic programming retracing all the angles path for each measurement, evaluating the final joint probabilities and calculating the probability of correct decision as the sum in (4.151). Due to the discretization of S_k as a grid, the performance obtained retracing the angles tree can be slightly different with respect to the performances of dynamic programming. The performances of this *forward* path retracing are depicted in Figures 4.1, 4.2, 4.3 and 4.4 in a blue dashed line. As you can see, for lower cardinality it coincides with the prediction, but the gap spreads out as M increases, especially for high $|\alpha|^2$ (see for example Figure 4.3).

In Figure 4.4, we managed to keep the performances of the forward retracing close to the predicted one for $M = 8$ by discretizing the grid $[0, 1/8] \times [0, 1/8]$ with a grid step of 10^{-3} .

4.7 Conclusion

In the present work we have studied the design of quantum receivers for Pulse Position Modulation.

By the PPM signal structure, we could describe the overall transmitted quantum states in the symbol time interval as sequences of quantum states in shorter temporal slots in a tensorial product. The signal measurement is then reformulated as a sequence of shorter measurements, one in each slot, that allows to design adaptive receiver scheme.

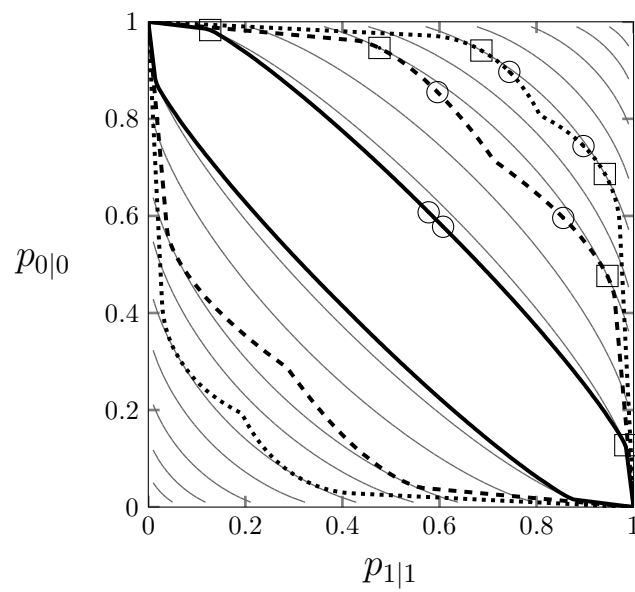


Figure 4.4: (Color online) Performances comparison of different receiver schemes, for 8-PPM: — classical direct detection, — quantum theoretical limit, — adaptive scheme, - - - retraced forward path, — conditional nulling receiver, - - - type I improved conditional nulling scheme, — type II improved conditional nulling scheme.

We move to an isomorphic representation of the quantum states in terms of qubits. The description of the existing receiver architecture in this framework highlights the limitations in terms of outcomes probabilities. We propose a more general adaptive receiver structure, where the measurement in each slot is a function of all the previous outcomes and the time evolving joint probabilities of the symbols with the outcomes sequence.

We propose an optimization of such adaptive scheme by means of dynamic programming, providing a description of the algorithm to evaluate the performance of the adaptive receiver and to calculate the measurement in each slot. The probability of error, although it does not reach the theoretical quantum limit except for $M > 2$, significantly outperforms the existing receiver schemes.

As a concluding remark, adaptive receiver seem to be the way to follow to achieve better performances for communication purpose, thanks to the possibility to embed the information of previous outcomes and improve subsequent measurements.

Appendix

4.A Review of Dynamic Programming

In this Section we introduce the (discrete time) dynamic programming framework and its basic algorithm. For a more detailed review, see [30].

Consider a discrete time dynamic system described by the update equation

$$s_{k+1} = f_k(s_k, u_k, w_k), \quad k = 0, \dots, N - 1, \quad (4.179)$$

with given initial system state s_0 , where

- k is the step index corresponding to the time.
- $s_k \in S_k$ is the *system state*, that is the collection of past informations up to time k useful to describe the evolution of the system and relevant for the optimization problem. To avoid misunderstanding, in the following we will use the term *system state* and *quantum state*, to discriminate the description of a system as in (4.179) and the physical description given by the density operator.
- $u_k \in U_k$ is the control, that is the physical variable or quantity we can use to drive the system evolution. Since we can impose the value of u_k in order to control the system, it is not described by a random variable.
- $w_k \in W_k$ is a random parameter out of our control, sometimes referred to as disturbance or noise. It can be related to s_k and u_k , i.e. its probability description can depend upon s_k and u_k as in $P[\cdot|s_k, u_k]$.

A reward function ⁴ is associated with the system evolution, that in our case we can write as

$$g(s_N), \quad (4.180)$$

and depends upon the final system state s_N . Since the evolution (4.179) is influenced by the random variables w_0, \dots, w_{N-1} , the final system state s_N is a random variable and the reward function we want to maximize is

$$\mathbb{E}_{s_N} [g(s_N)] = \int_{S_N} d\sigma g(\sigma) \wp_{s_N}(\sigma), \quad (4.181)$$

where with the notation $\wp_r(\cdot)$ we indicate the probability density function of the random variable r , in this case the system state s_N .

Considering the update equation (4.179) for $k = N - 1$, the reward function can be rewritten as

$$\begin{aligned} \mathbb{E}_{s_{N-1}, w_{N-1}} [g(f_{N-1}(s_{N-1}, u_{N-1}, w_{N-1}))] &= \\ &= \int_{S_{N-1}} d\rho \int_{W_{N-1}} d\omega g(\sigma) \wp_{s_{N-1}, w_{N-1}}(\rho, \omega). \end{aligned} \quad (4.182)$$

with the system state $s_N = \sigma$ evaluated from the evolution $\sigma = f_{N-1}(\rho, u_k, \omega)$.

By (4.179) we can continue the substitutions backward in time to obtain a reformulation of the reward function (4.181) in terms of the controls and the initial system state s_0 ,

$$\int_{W_0} dw_0 \cdots \int_{W_{N-1}} dw_{N-1} g(\sigma) \wp_{w_0, \dots, w_{N-1}}(\omega_0, \dots, \omega_{N-1}) \quad (4.183)$$

with $s_N = \sigma$ obtained by the composition of the update equations $f_{N-1}, f_{N-2}, \dots, f_0$ from s_0 with the variables u_0, \dots, u_{N-1} , $w_0 = \omega_0, \dots, w_{N-1} = \omega_{N-1}$.

Seeking the maximization of (4.183), we can employ different strategies. For example, the values of the control u_0, \dots, u_{N-1} can be determined before the

⁴Dynamic programming is usually formulated for a minimum optimization problem, but in our case a maximization problem is more suitable since we aim at maximize the probability of correct decision. Therefore, we introduce the concepts of reward and reward-to-go function in place of cost and cost-to-go function.

system starts, and then applied during the evolution, or we can postpone the choice of u_k at time k since there's no penalties in delaying the decision. In particular, this latter strategy allows to define u_k as a *function* of the system state s_k ,

$$u_k = \pi_k(s_k), \quad s_k \in S_k, \quad u_k \in U_k, \quad (4.184)$$

leading to an adaptive control algorithm. Its performance are not worse than the fixed control, and we can take advantage of the information gained from time 0 to k .

The set of functions $\bar{\pi} = (\pi_0, \pi_1, \dots, \pi_{N-1})$ is called a *policy*. We can define the expected reward considering the current system state, the current input and the policy for the future inputs at time step $k = 0, \dots, N - 1$ as the function

$$J_k : S_k \times U_k \times \Pi_{k+1} \times \dots \times \Pi_{N-1} \mapsto \mathbb{R} \quad (4.185)$$

specified as

$$J_k(\sigma, \nu, \pi_{k+1}, \dots, \pi_{N-1}) = \int_{W_k} dw_k \cdots \int_{W_{N-1}} dw_{N-1} g(\sigma) \wp_{s_k, w_k, \dots, w_{N-1}}(\rho, \omega_k, \dots, \omega_{N-1}) \quad (4.186)$$

with $\Pi_k = U_k^{S_k}$ the set of all possible functions $\pi_k : S_k \mapsto U_k$, and $s_N = \sigma$ the system state calculated with the composition of f_{N-1}, \dots, f_k employing $s_k = \rho$, $u_k = \nu$, $u_{k+1} = \pi_{k+1}(s_{k+1}), \dots, u_{N-1} = \pi_{N-1}(s_{N-1})$.

Note that integrating (4.186) with respect to s_k ,

$$\int_{S_k} d\sigma J_k(\sigma, \nu, \pi_{k+1}, \dots, \pi_{N-1}) \quad (4.187)$$

we can define a *reward-to-go* function at time k from s_k , which uses $u_k = \nu$, $u_{k+1} = \pi_{k+1}(s_{k+1}), \dots, u_{N-1} = \pi_{N-1}(s_{N-1})$.

Define $\bar{\pi}^* = (\pi_0^*, \pi_1^*, \dots, \pi_{N-1}^*)$ the optimal policy, that is the one that maximize J_0

$$\bar{\pi}^* := \operatorname{argmax}_{\bar{\pi}} J_0(s_0, \bar{\pi}(s_0)), \quad (4.188)$$

and define

$$J_0^*(s_0) := J_0(s_0, \bar{\pi}^*(s_0)) \quad (4.189)$$

the optimal reward from s_0 .

The dynamic programming algorithm relies on the following rather obvious idea.

Principle of Optimality [30]

Let $\bar{\pi}^* = (\pi_0^*, \pi_1^*, \dots, \pi_{N-1}^*)$ be the optimal policy that maximizes the reward J_0 and let s_1, s_2, \dots, s_N be the corresponding system state evolution. Consider the subproblem of the maximization of the reward-to-go function from time k ,

$$\max_{\pi_k, \dots, \pi_{N-1}} \int_{S_k} d\sigma J_k(\sigma, \pi_k(\sigma), \dots, \pi_{N-1}) \quad (4.190)$$

The optimal policy for this subproblem is the truncated sequence $(\pi_k^*, \pi_{k+1}^*, \dots, \pi_{N-1}^*)$.

The maximization of J_0 with respect to the policy $(\pi_0, \dots, \pi_{N-1})$ with multivariate calculus requires the solution of a complicated system with equations in all the variables π_k . Instead, the dynamic programming algorithm decomposes the main problem into a sequence of subproblems.

Dynamic Programming Algorithm [30]

The optimal reward J_0^* is the last step of the following algorithm, which proceeds backwards from N to 0,

1. define the initial condition

$$J_N^*(\sigma) = g(\sigma), \quad \sigma \in S_N \quad (4.191)$$

2. for $k = N - 1, \dots, 0$, for all $\sigma \in S_k$ evaluate the optimal control and the optimal expected reward function at time k , namely

$$\pi_k^*(\sigma) = \operatorname{argmax}_{\nu \in U_k} J_k(\sigma, \nu, \pi_{k+1}^*, \dots, \pi_{N-1}^*) \quad (4.192)$$

$$= \operatorname{argmax}_{\nu \in U_k} \int_{W_k} d\omega J_{k+1}^*(f_k(\sigma, \nu, \omega)), \quad (4.193)$$

$$J_k^*(\sigma) = J_k(\sigma, \pi_k^*(\sigma), \pi_{k+1}^*, \dots, \pi_{N-1}^*). \quad (4.194)$$

3. the optimal reward and the optimal policy are

$$J_0^*(s_0) = J_0(s_0, \pi_0^*(s_0), \dots, \pi_{N-1}^*), \quad (4.195)$$

$$\bar{\pi}^* = (\pi_0^*, \pi_1^*, \dots, \pi_{N-1}^*). \quad (4.196)$$

At each step k , assuming to know by induction the optimal policy $(\pi_{k+1}^*, \dots, \pi_{N-1}^*)$ and the optimal reward $J_{k+1}^*(\cdot)$, the algorithm considers $s_k = \sigma$ as the initial system state for the time evolution from k to N , and maximize with respect to the control $u_k = \nu$. This maximization is solved for every possible $s_k \in S_k$, in order to define the function $\pi_k^*(s_k)$ and the optimal reward $J_k^*(s_k)$. This step is repeated until $k = 0$. Note that at each step, only one control variable u_k is involved in the maximization, simplifying its optimization.

Chapter 5

Conclusions

We have considered the scenario of communication of classical information through quantum medium and described the transmission and detection task of a communication system as a discrimination problem in the set of the possible quantum states sent.

The transmitter encodes the message in a sequence of quantum states, each one described by a density matrix in the set

$$\rho_x \in \{\rho_1, \rho_2, \rho_3, \dots, \rho_M\}. \quad (5.1)$$

The channel interposed between transmitter and receiver modifies the transmitted density operator ρ_x , and the output density operator $\tilde{\rho}_x$ can be described by a Completely Positive Trace Preserving linear map $\mathcal{E} : \rho_x \mapsto \tilde{\rho}_x$ that acts as in

$$\tilde{\rho}_x = \sum_k E_k^\dagger \rho_x E_k. \quad (5.2)$$

The receiver performs a measurement on the density operator $\tilde{\rho}_x$ in order to discriminate the original quantum states, indicated by x , as best as possible. This is evaluated through some figures of merit for the communication, that may be the probability of correct decision or the capacity of the system. The Positive Operator Values Measurement defining the measurement must therefore be optimized in order to maximize the performances.

In Chapter 2 we first study the problem of implementing binary communication with coherent states and Binary Phase Shift Keying modulation. In this case, the two possible quantum states transmitted are

$$|\gamma_x\rangle = |e^{ix\pi}\alpha\rangle, \quad x = 0, 1. \quad (5.3)$$

Optimal discrimination performance in terms of probability of correct detection, as predicted by the quantum discrimination theory, gives the Helstrom bound

$$P_c^{(Helstrom)} = \frac{1}{2} \left[1 + \sqrt{1 - 4p_0p_1\chi^2} \right], \quad (5.4)$$

with $\chi = \langle -\alpha|\alpha\rangle = e^{-2|\alpha|^2}$ the inner product of the quantum states (5.3).

Several receiver schemes have been proposed for this modulation: the Homodyne detection, the Kennedy receiver, the Generalized Kennedy receiver and the Dolinar scheme, but only the last one achieves the Helstrom limit precisely.

We reformulate this scheme in a multiple-copies framework. Rather than considering the coherent state $|\gamma_x\rangle$ of duration T , we virtually divide the time interval in N sub-intervals such that the whole quantum states is the composition of N weaker states in a tensor products,

$$|\gamma_x\rangle = \left| e^{ix\pi\frac{\alpha}{\sqrt{N}}} \right\rangle \left| e^{ix\pi\frac{\alpha}{\sqrt{N}}} \right\rangle \cdots \left| e^{ix\pi\frac{\alpha}{\sqrt{N}}} \right\rangle, \quad x = 0, 1. \quad (5.5)$$

The original discrimination problem becomes the discrimination between two coherent states from the observation of multiple copies of them. Since in the case of two pure states, an adaptive strategy of local measurement on the single copies with a bayesian update on successive measurements suffices to reach the minimum error probability, we apply the same result for the multiple copy coherent discrimination. As N goes to infinity, the adaptive strategy coincides with the Dolinar receiver.

In addition, with the same theory we investigate the performance of a simplified version of the Dolinar receiver, using a displacement that is constant in time rather than the time varying receiver required by the Dolinar's. Although resulting in suboptimal performances, this architecture is much simpler to implement and still outperforms other near optimal receiver schemes.

In Chapter 3, we study the degrading action of the channel on the transmitted quantum states. We consider a *given* description of a general qubit channel, and seek for the optimal binary input states and output measurements that define a classical binary channel.

We consider both the error probability and the capacity of the binary channel as figures of merit for the optimization. In particular, since both functionals can be defined on the induced transition probability, we focus on the characterization of region of the admissible transition probabilities.

We first introduce partial orderings for the binary channels, that allow us to prove that the optimal POVM are projectors, for both functionals. Then, leveraging the coherence vector representation of the qubit states and channels, we are able to find necessary conditions on the optimal transmitted input states, which must be orthogonal.

Including this conditions in the optimization problem, we are able to find the region of admissible transition probabilities. This is shown to be a quadratic problem with quadratic constraints, and therefore solvable with standard numerical methods. A particularly suitable numerical procedure for the quadratic problem is given.

For the maximization of the channel capacity, a two step procedure is given. The “inner” maximization with respect to the *a priori* probability admits a closed-form solution that can be used to simplify optimization with respect to the transition probability. The “outer” maximization with respect to the *transition* probability relies on the numerical optimization used to find the boundary of the admissible transition probability region.

For the minimization of the error probability, a closed form solution for the optimization is obtained, recovering previous results by discrimination theory.

Numerical results as well as qualitative analysis of the contour plots of the functionals suggest that, even if typically the solutions for the two functionals exhibit are very close, considerable differences can emerge in particular cases, depending on the curvature of the boundaries of the admissible transition prob-

abilities.

In Chapter 4, we consider a more complex modulation format. Pulse Position Modulation is a suitable candidate for the satellite and deep space communications, and several studies have investigated its performance for a communication link both theoretically and experimentally.

Pulse Position Modulation defines the quantum state to transmit (5.3) as the tensor product of M coherent states, all in the vacuum state except the one in the x -th position.

The theoretical limit for the performance of PPM in terms of error probability is known by means of Quantum Discrimination Theory, exploiting the Geometric Uniform Symmetry through the least square measurement.

Many receiver schemes have been proposed to approach such limit: the classical receiver, the Dolinar PPM receiver, and its improved versions by Guha *et al.*

We reformulate the PPM discrimination problem in a qubit framework, and we review these schemes in this representation. These receivers employ local measurements in each sub-interval, and use an adaptive strategy to choose which measurements perform in successive time slots.

However, all these existing architectures present some limitation in both the local measurements and in the adaptive strategy. The local measurements are performed with a combination of (constant) displacement, phase amplification and photon counting, and in the adaptive strategy the next measurement is decided from the previous results.

On the contrary, we consider a more general adaptive receiver, allowing for the a general projective measurement in the sub-intervals and an adaptive strategy where the measurements depends upon all the previous outcomes. The resulting receiver depends on a set of parameter that increase exponentially in the cardinality M . In order to perform an optimization on these parameters that has limited complexity, we use dynamic programming, and define both the optimization and the receiver algorithm.

We evaluate the performance of this receiver for different cardinality M and different values of the average number of photons in the coherent state of the PPM. In particular, we focus on the region of weak values of coherent states, that simulates a high loss in the communication link or a transmission with weak pulses.

Due to the greater generality of the scheme, the optimized adaptive architecture outperforms all the other existing receivers. In the case of $M = 2$ the receiver precisely achieves the Helstrom bound, confirming the known result that in the binary case an adaptive receiver with local measurements suffices to achieve the quantum limit. For other values of M , the error probability slightly moves away from the quantum limit, but still outperforms the other receivers.

In conclusion, the challenges proposed by the discrimination problem in the communication scenario have been investigated with particular attention to the optimization process for the definition of the measurement operators. Adaptive receivers seem to be the way to follow to achieve better performances for communication purposes, due to the possibility to embed the information of previous outcomes in the choice of successive measurements.

Bibliography

- [1] Michael Westmoreland and Benjamin Schumacher. Capacities of Quantum Channels and Quantum Coherent Information. *Quantum Computing and Quantum Communications, Lecture Notes in Computer Science*, 1509:285–295, May 1999.
- [2] Hans-Albert Bachor and Timothy C. Ralph. *A Guide to Experiments in Quantum Optics*. Wiley-VCH Verlag GmbH, Weinheim, Germany, second edition, 2004.
- [3] Roy Glauber. Coherent and Incoherent States of the Radiation Field. *Physical Review*, 131(6):2766–2788, September 1963.
- [4] Werner Karl Heisenberg. Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift für Physik*, 43(3-4):172–198, March 1927.
- [5] John Archibald Wheeler and Wojciech Hubert Zurek. *Quantum Theory and Measurement*. Princeton Series in Physics. Princeton University Press, Princeton, NJ, 1983.
- [6] Dennis Dieks. Communication by EPR devices. *Physics Letters A*, 92(6):271–272, November 1982.
- [7] William Kent Wootters and Wojciech Hubert Zurek. A single quantum cannot be cloned. *Nature*, 299(5886):802–803, October 1982.

- [8] Nicolas Gisin, Grégoire Ribordy, Wolfgang Tittel, and Hugo Zbinden. Quantum cryptography. *Reviews of Modern Physics*, 74(1):145–195, March 2002.
- [9] Hamid Hemmati. *Deep Space Optical Communications*. Wiley, Hoboken, New Jersey, 2006.
- [10] Claude Elwood Shannon. A Mathematical Theory of Communication. *The Bell System Technical Journal*, 27:379–423, 1948.
- [11] Vittorio Giovannetti, Saikat Guha, Seth Lloyd, Lorenzo Maccone, Jeffrey H. Shapiro, and Horace P. Yuen. Classical Capacity of the Lossy Bosonic Channel: The Exact Solution. *Physical Review Letters*, 92(2):027902, January 2004.
- [12] Carlton Caves and Peter D. Drummond. Quantum limits on bosonic communication rates. *Reviews of Modern Physics*, 66(2):481–537, April 1994.
- [13] James Power Gordon. Quantum Effects in Communications Systems. *Proceedings of the IRE*, 50(9):1898–1908, 1962.
- [14] Graeme Smith. Quantum Channel Capacities. In *Information Theory Workshop (ITW), 2010 IEEE*, pages 1–5, 2010.
- [15] Christopher King and Mary Beth Ruskai. Minimal entropy of states emerging from noisy quantum channels. *Information Theory, IEEE Transactions on*, 47(1):192–209, 2001.
- [16] Charles Henry Bennett and Peter Williston Shor. Quantum information theory. *Information Theory, IEEE Transactions on*, 44(6):2724–2742, 1998.
- [17] Paul Hausladen, Richard Jozsa, Benjamin Schumacher, Michael Westmoreland, and William Wootters. Classical information capacity of a quantum channel. *Physical Review A*, 54(3):1869–1876, September 1996.

- [18] Alexander Semenovich Holevo. The capacity of the quantum channel with general signal states. *Information Theory, IEEE Transactions on*, 44(1):269–273, 1998.
- [19] Benjamin Schumacher and Michael Westmoreland. Sending classical information via noisy quantum channels. *Physical Review A*, 56(1):131–138, July 1997.
- [20] Alexander Semenovich Holevo. On Capacity of a Quantum Communications Channel. *Probl. Peredachi Inf.*, 15(4):3–11, 1979.
- [21] Christopher Fuchs and Carlton Caves. Ensemble-Dependent Bounds for Accessible Information in Quantum Mechanics. *Physical Review Letters*, 73(23):3047–3050, December 1994.
- [22] Carl W. Helstrom. *Quantum Detection and Estimation Theory*. Academic Press, New York, 1976.
- [23] Robert S. Kennedy. A Near-Optimum Receiver for the Binary Coherent State Quantum Channel. *Research Laboratory of Electronics, MIT Quarterly Progress Report No. 108*, pages 219–225, 1973.
- [24] Masahiro Takeoka and Masahide Sasaki. Discrimination of the binary coherent signal: Gaussian-operation limit and simple non-Gaussian near-optimal receivers. *Physical Review A*, 78(2):022320, August 2008.
- [25] Gianfranco Cariolaro. *Comunicazioni Quantistiche*. Edizione Copisteria Portello, Padova, 2009.
- [26] Victor A. Vilnrotter and Eugene R. Rodemich. A generalization of the near-optimum binary coherent state receiver concept (Corresp.). *Information Theory, IEEE Transactions on*, 30(2):446–450, 1984.
- [27] Antonio Assalini, Nicola Dalla Pozza, and Gianfranco Pierobon. Revisiting the Dolinar receiver through multiple-copy state discrimination theory. *Physical Review A*, 84(2):022342, August 2011.

- [28] Antonio Acín, Emili Bagan, Marià Baig, Lluís Masanes, and Ramon Muñoz Tapia. Multiple-copy two-state discrimination with individual measurements. *Physical Review A*, 71(3):032338, March 2005.
- [29] Dorje Brody and Bernhard Meister. Minimum Decision Cost for Quantum Ensembles. *Physical Review Letters*, 76(1):1–5, January 1996.
- [30] Dimitri Panteli Bertsekas. *Dynamic Programming and Optimal Control, Vol. I-II*. Athena Scientific, 3rd edition, 2007.
- [31] Howard M. Wiseman, Dominic W. Berry, Stephen D. Bartlett, Brendon L. Higgins, and Geoff J. Pryde. Adaptive Measurements in the Optical Quantum Information Laboratory. *IEEE Journal of Selected Topics in Quantum Electronics*, 15(6):1661–1672, 2009.
- [32] Emanuel Parzen. *Stochastic Process*. Holden-Day, Inc., San Francisco, 1962.
- [33] Samuel Joseph Dolinar. Quarterly Progress Report No.111, Research Laboratory of Electronics, MIT. *Research Laboratory of Electronics, MIT Quarterly Progress Report No. 111*, 115, 1973.
- [34] Robert L Cook, Paul J Martin, and John M Geremia. Optical coherent state discrimination using a closed-loop quantum measurement. *Nature*, 446(7137):774–7, April 2007.
- [35] Karl Kraus, Arno Böhm, John. D. Dollard, and William K. Wootters. *States, Effects, and Operations Fundamental Notions of Quantum Theory*, volume 190 of *Lecture Notes in Physics*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1983.
- [36] Michael Nielsen and Isaac Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, New York, 2004.
- [37] Noam Elron and Yonina C. Eldar. Optimal Encoding of Classical Information in a Quantum Medium. *Information Theory, IEEE Transactions on*, 53(5):1900–1907, 2007.

- [38] Edward Davies. Information and quantum measurement. *Information Theory, IEEE Transactions on*, 24(5):596–599, 1978.
- [39] Christopher Alan Fuchs. Nonorthogonal Quantum States Maximize Classical Information Capacity. *Physical Review Letters*, 79(6):1162–1165, August 1997.
- [40] Christopher King, Michael Nathanson, and Mary Beth Ruskai. Qubit Channels Can Require More Than Two Inputs to Achieve Capacity. *Physical Review Letters*, 88(5):057901, January 2002.
- [41] Masahito Hayashi, Hiroshi Imai, Keiji Matsumoto, Mary Beth Ruskai, and Toshiyuki Shiono. Qubit channels which require four inputs to achieve capacity: implications for additivity conjectures. *Quantum Information & Computation*, 5(1):13–31, January 2005.
- [42] Dominic William Berry. Qubit channels that achieve capacity with two states. *Physical Review A*, 71(3):032334, March 2005.
- [43] Akio Fujiwara and Hantarou Nagaoka. Operational capacity and pseudo-classicality of a quantum channel. *Information Theory, IEEE Transactions on*, 44(3):1071–1086, 1998.
- [44] Claude E. Shannon. A note on a partial ordering for communication channels. *Information and Control*, 1(4):390–397, December 1958.
- [45] Norman Abramson. A partial ordering for binary channels. *Information Theory, IEEE Transactions on*, 6(5):529–539, 1960.
- [46] Hermann Helgert. A partial ordering of discrete, memoryless channels. *Information Theory, IEEE Transactions on*, 13(3):360–365, 1967.
- [47] János Körner and Katalin Marton. The comparison of two noisy channels. In I Csiszár and P Elias, editors, *Topics in Information Theory*. Coll. Math. Soc. J. Bolyai, No. 16, North Holland, Amsterdam, 1977.

- [48] Marten van Dijk. On a special class of broadcast channels with confidential messages. *Information Theory, IEEE Transactions on*, 43(2):712–714, 1997.
- [49] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley-Interscience, New York, 2006.
- [50] Francesca Albertini and Francesco Ticozzi. Discrete-time controllability for feedback quantum dynamics. *Automatica*, 47(11):2451–2456, November 2011.
- [51] Norberto Tomassoni and Matteo G.A. Paris. Quantum binary channels with mixed states. *Physics Letters A*, 373(1):61–64, December 2008.
- [52] Mary Beth Ruskai, Stanislaw Szarek, and Elisabeth Werner. An analysis of completely-positive trace-preserving maps on. *Linear Algebra and its Applications*, 347(1-3):159–187, May 2002.
- [53] Stephen Boyd and Lieven Vandenbergh. *Convex Optimization*. Cambridge University Press, New York, 2004.
- [54] Gianfranco Cariolaro and Gianfranco Pierobon. Theory of quantum pulse position modulation and related numerical problems. *IEEE Transactions on Communications*, 58(4):1213–1222, April 2010.
- [55] Paul Hausladen and William K. Wootters. A ‘Pretty Good’ Measurement for Distinguishing Quantum States. *Journal of Modern Optics*, 41(12):2385–2390, December 1994.
- [56] Yonina C. Eldar and G. David Jr. Forney. On quantum detection and the square-root measurement. *Information Theory, IEEE Transactions on*, 47(3):858–872, 2001.
- [57] Samuel Joseph Dolinar. A near-optimum receiver structure for the detection of M-ary optical PPM signals. *The Telecommunications and Data Acquisition Report*, pages 30–42, February 1983.

- [58] Christopher Gerry and Peter Knight. *Introductory Quantum Optics*. Cambridge University Press, Cambridge, 2004.
- [59] Saikat Guha, Jonathan L. Habif, and M. Takeoka. Approaching Helstrom limits to optical pulse-position demodulation using single photon detection and optical feedback. *Journal of Modern Optics*, 58(3-4):257–265, February 2011.
- [60] Masahiro Takeoka, Masahide Sasaki, Peter van Loock, and Norbert Lütkenhaus. Implementation of projective measurements with linear optics and continuous photon counting. *Physical Review A*, 71(2):022318, February 2005.

Corrections for
RECEIVER DESIGN FOR QUANTUM
COMMUNICATION
Nicola Dalla Pozza

September 16, 2013

The following are corrections to known errata within *Receiver Design for Quantum Communication*.

Page 25, eq. 1.19. Missing channel action on the quantum state.

ERRATUM:

$$\begin{aligned} &= \text{tr}(P_j \rho_k) & (1.19) \\ \text{(if } \rho_k \text{ is a pure state)} &= \langle \gamma_k | P_j | \gamma_k \rangle. \end{aligned}$$

CORRIGE:

$$\begin{aligned} &= \text{tr}(P_j \tilde{\rho}_k) & (1.19) \\ \text{(if } \tilde{\rho}_k \text{ is a pure state)} &= \langle \tilde{\gamma}_k | P_j | \tilde{\gamma}_k \rangle. \end{aligned}$$

Page 28, line 7 from bottom. Typo.

ERRATUM: “A reader interested in the transmission of quantum information can found a brief introduction ...”

CORRIGE: “A reader interested in the transmission of quantum information can find a brief introduction ...”

Page 28, line 7 from bottom. Missing definition of σ_i . Add the line

where $\{\sigma_i\}$ are the singular values of the decomposition of Γ .

after eq. (4.41) and the sentence “For a proof of the proposition, see [55, 56, 25].”

Page 116, repeated paragraph. The last paragraph of page 114, which ends at page 116, is repeated as the first paragraph at page 116. The correct section reference is not “...next Section.” but “... Appendix 4.A”.

Page 129, figure 4.1. Wrong graphic.
CORRIGE:

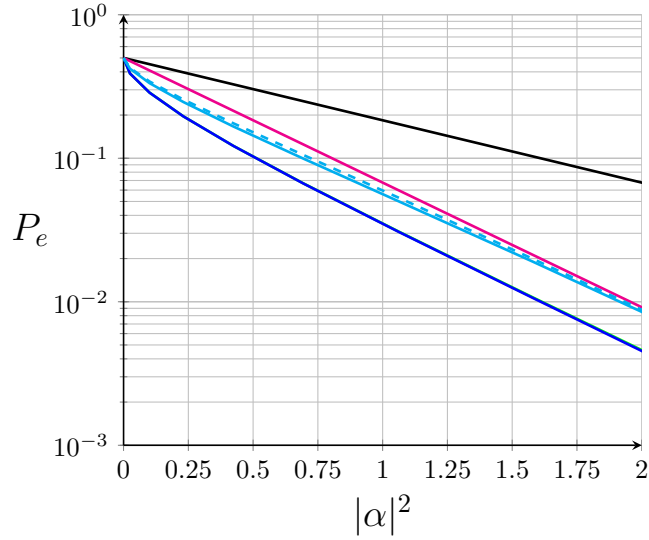


Figure 4.1: Performances comparison of different receiver schemes, for 2-PPM: — classical direct detection, — quantum theoretical limit and adaptive scheme (overlapped), — conditional nulling receiver, - - - type I improved conditional nulling scheme, — type II improved conditional nulling scheme.

Page 130, figure 4.2. Wrong graphic.
CORRIGE:

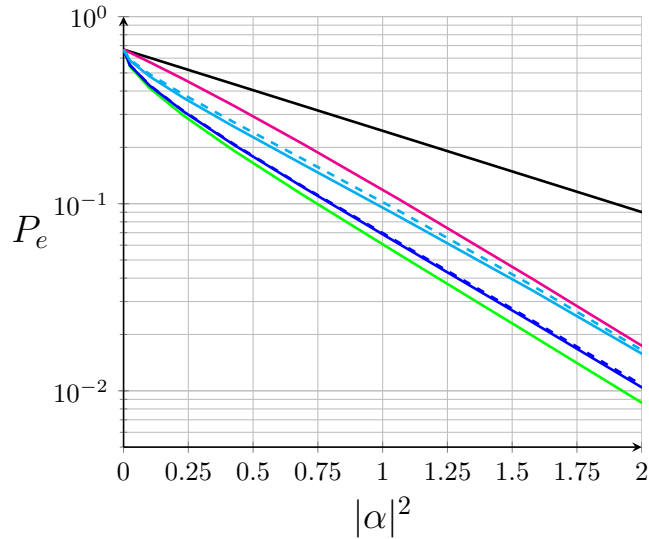


Figure 4.2: Performances comparison of different receiver schemes, for 3-PPM: — classical direct detection, — quantum theoretical limit, — adaptive scheme, - - - retraced forward path, — conditional nulling receiver, - - - type I improved conditional nulling scheme, — type II improved conditional nulling scheme.

Page 131, figure 4.3. Wrong graphic.
CORRIGE:

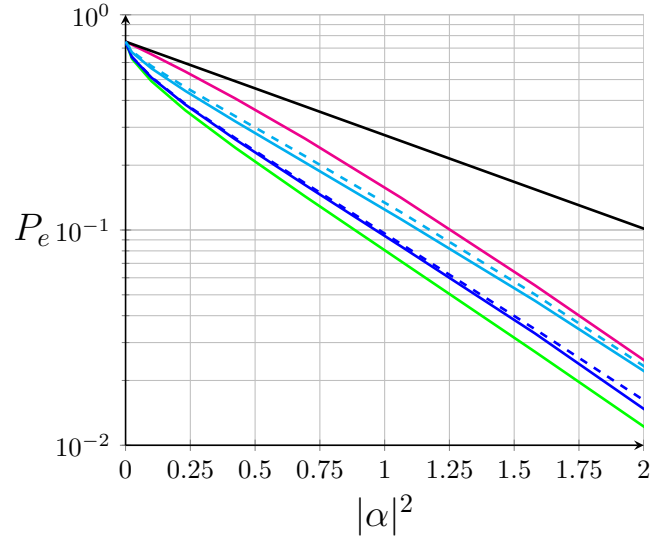


Figure 4.3: Performances comparison of different receiver schemes, for 4-PPM: — classical direct detection, — quantum theoretical limit, — adaptive scheme, - - - retraced forward path, — conditional nulling receiver, - - - type I improved conditional nulling scheme, — type II improved conditional nulling scheme.

Page 134, figure 4.4. Wrong graphic.
CORRIGE:

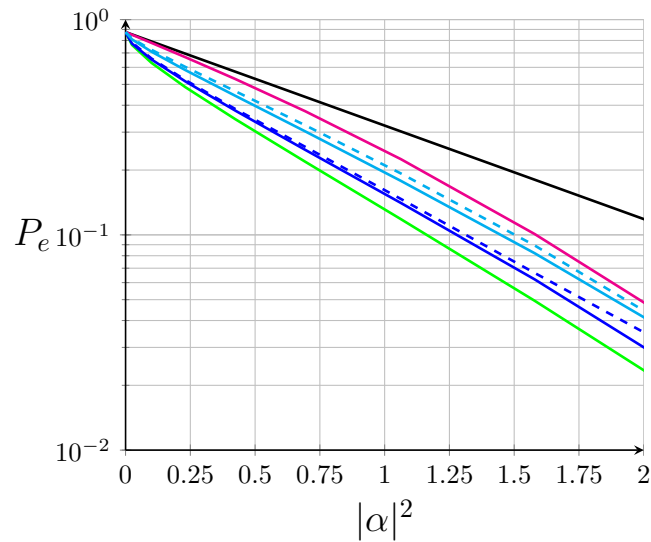


Figure 4.4: Performances comparison of different receiver schemes, for 8-PPM: — classical direct detection, — quantum theoretical limit, — adaptive scheme, - - - retraced forward path, — conditional nulling receiver, - - - type I improved conditional nulling scheme, — type II improved conditional nulling scheme.