

# Adaptive Discrimination Scheme for Quantum Pulse Position Modulation Signals

Nicola Dalla Pozza\* and Nicola Laurenti†

*Department of Information Engineering (DEI), University of Padova*

(Dated: January 9, 2014)

In the communication scenario, we consider the problem of the discrimination between the signals of the Quantum Pulse Position Modulation. We propose a receiver scheme that employs repeated local measurements in distinct temporal slots within the symbol time interval, adaptively chosen on the basis of the outcomes in the previous slots. By employing Dynamic Programming to optimize each measurement, we approach the theoretical performance limit much closer than existing receiver schemes.

PACS numbers: 03.67.Hk

## I. INTRODUCTION

The increasing demand for higher data rates in free space communications is driving system designers to consider optical frequencies, rather than radio carriers, for the satellite-to-satellite and satellite-to-earth links. In this switch of design paradigm, quantum optics plays an important role for the correct description of the physical phenomena involved in the transmission, propagation and detection stages of the communication.

Quantum optics supports the use of coherent states for the transmission of information over free space links with a long list of studies and results in both communication and information theory [1]. For example, in [2] it is proven that with a random coding over coherent states it is possible to achieve the capacity of a lossy bosonic link, that is the quantum model for the free space channel. Furthermore, binary coherent state discrimination has been theoretically solved [3] and experimentally tested [4], reaching the ultimate quantum limit (Helstrom bound [5]) for the performance in terms of error probability.

Different solutions have been proposed to encode information into coherent states, most notably amplitude modulation, phase modulation and pulse position modulation. The choice depends upon several factors to be taken into account in the design of the communication system, including the channel model, the target performance, and the complexity of the system at the transmitter and at the receiver side.

Pulse Position Modulation (PPM) encodes the information to be transmitted in the temporal position of a pulse within the symbol time length. For example, Figure 1 depicts the mapping from the symbols  $x = \{1, 2, 3, 4\}$  of a 4-PPM to the transmitted signal. The fact that the implementation of a PPM transmitter only requires an intensity modulator, and its efficiency in terms of average energy make this modulation a suitable candidate for communications from satellites and space-

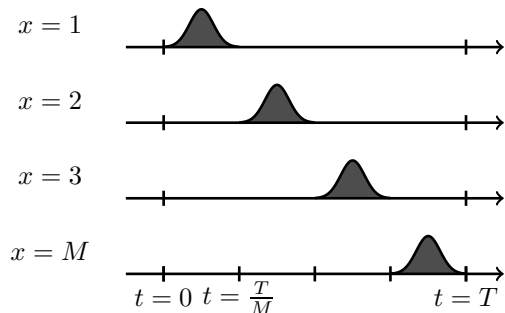


FIG. 1: Pulse Position Modulation, association between symbols  $x$  and the position of the pulse in the symbol time interval, with alphabet cardinality  $M = 4$ .

crafts.

The classical approach of the receiver for the discrimination of PPM signals is to detect the field intensity in the symbol time interval, and to estimate the transmitted symbol according to the slot where the maximum of energy has been measured. Dolinar [6] proposed an adaptive receiver scheme that exploits the possibility of nulling the PPM signal depending on the result of an intensity measurement. Recently, an improved version of this receiver scheme has been introduced [7], and has received a lot of attention due to the possibility to outperform the standard quantum limit given by the photon counting and approach the theoretical limit predicted by quantum discrimination theory (Helstrom bound). This solution has also been experimentally tested and tuned to face experimental non idealities [8].

In this paper we briefly review these receiver schemes and propose a similar feedback structure consisting of a sequence of  $M$  local measurements optimized by means of dynamic programming. This optimization allows to approach the Helstrom bound more closely than existing receiver schemes.

The paper is organized as follows. In Section II, we describe the communication setup and formally state the problem. In Section III we review the existing structures for PPM receivers, and propose our adaptive scheme that will be optimized later. Section IV briefly reviews some results in dynamic programming, which are then em-

\*Electronic address: nicola.dallapozza@dei.unipd.it

†Electronic address: nil@dei.unipd.it

ployed in Section V to derive the procedure for the optimization of the receiver scheme. In Section VI we show the performance of the receiver scheme and describe some numerical issues in the optimization procedure. Section VII summarizes the paper contribution and draws conclusions.

## II. COMMUNICATION SETUP

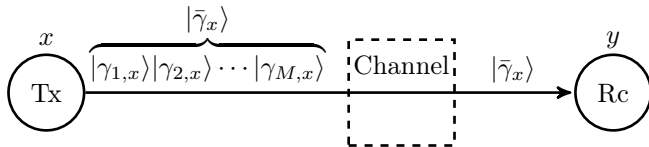


FIG. 2: Communication system setup, with ideal channel. With this assumption, the transmitted quantum states are received unaltered.

We consider the communication system summarized in Figure 2. A transmitter encodes its message in a sequence of symbols  $x \in \{1, \dots, M\}$ , described as a random variable with distribution  $\{p_1, \dots, p_M\}$ . We assume an equal a priori distribution for the symbols  $x$ , i.e.  $p_x = \frac{1}{M} \forall x$ .

The transmitter maps each symbol of the sequence to a quantum state  $|\bar{\gamma}_x\rangle$  taken from the set  $\{|\bar{\gamma}_1\rangle, |\bar{\gamma}_2\rangle, \dots, |\bar{\gamma}_M\rangle\}$  and sends it through the channel. We assume the channel to be ideal, such that the quantum states at the channel output are exactly the ones that have been transmitted.

The receiver measures the output of the quantum channel and estimates which symbol  $x$  has been encoded. We denote with  $y$  the random variable associated with the symbol estimation. The figure of merit to evaluate the performance of the receiver scheme is the probability of correct decision, or equivalently, the error probability, defined from transmitted and estimated symbol as

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

$$P_e = 1 - P_c. \quad (2)$$

Pulse Position Modulation defines a particular structure for the quantum states  $|\bar{\gamma}_x\rangle$ . As we can see from Figure 1, the symbol time interval can be virtually divided in  $M$  temporal slots. A pulse in the  $i$ -th slot is associated to the symbol  $x = i$ . Quantum optics describes this signal as a sequence of coherent states in a tensor product, one in each slot, with all ground states  $|0\rangle$  except the coherent state  $|\alpha\rangle$  in the  $i$ -th position,  $\alpha \neq 0$ .

$$|\bar{\gamma}_x\rangle = |\gamma_{1,x}\rangle |\gamma_{2,x}\rangle \dots |\gamma_{M,x}\rangle, \quad \gamma_{i,x} = \begin{cases} \alpha & i = x \\ 0 & i \neq x \end{cases} \quad (3)$$

The resulting association is shown in Table I.

Symbol	Coherent States	Qubits
$x = 1$	$ \alpha\rangle 0\rangle 0\rangle \dots  0\rangle$	$ \xi_1\rangle \xi_0\rangle \xi_0\rangle \dots  \xi_0\rangle$
$x = 2$	$ 0\rangle \alpha\rangle 0\rangle \dots  0\rangle$	$ \xi_0\rangle \xi_1\rangle \xi_0\rangle \dots  \xi_0\rangle$
$\vdots$	$\vdots$	$\vdots$
$x = M$	$ 0\rangle 0\rangle 0\rangle \dots  \alpha\rangle$	$ \xi_0\rangle \xi_0\rangle \xi_0\rangle \dots  \xi_1\rangle$

TABLE I: Association between symbols, transmitted quantum states and qubits representation.

In place of defining the transmitted quantum states in the Fock space and work with coherent states, an equivalent way to define the PPM format is to describe each slot in a qubit framework with an Hilbert space  $\mathcal{H} \sim \mathbb{C}^2$ , where either of two quantum states are possible,  $|\xi_0\rangle$  and  $|\xi_1\rangle$ , corresponding to  $|0\rangle$  and  $|\alpha\rangle$  respectively. The representation provides the same inner product,

$$\chi := \langle \bar{\gamma}_i | \bar{\gamma}_j \rangle = |\langle 0 | \alpha \rangle|^2 = e^{-|\alpha|^2} = |\langle \xi_0 | \xi_1 \rangle|^2, \quad i \neq j, \quad (4)$$

because from the point of view of the slot measurement, the outcome probabilities are given by the operators in the subspace spanned by  $|0\rangle$  and  $|\alpha\rangle$ , that is isomorphic to  $\mathcal{H}$ .

The transmitted quantum states are then described in the tensor Hilbert space  $\mathcal{H}_0 = \mathcal{H}^{\otimes M}$ , by

$$|\xi_{1,x}\rangle |\xi_{2,x}\rangle \dots |\xi_{M,x}\rangle, \quad \xi_{i,x} = \begin{cases} \xi_1 & i = x \\ \xi_0 & i \neq x \end{cases} \quad (5)$$

As we shall see later, this abstract definition allows us to focus on the *consequences* of the measurements in a slot in terms of transition probabilities, rather than the actual physical implementation of the measurement. This is obtained thanks to the possibility of implementing arbitrary projective measurement in the subspace spanned by  $\{|0\rangle, |\alpha\rangle\}$ , as explained in Section II B.

The quantum limit performance for the Pulse Position Modulation is a well known result by quantum discrimination theory [9], and the optimal measurement operator has been characterized by means of square root measurement, exploiting the geometric uniform symmetry of the constellation [10]. The resulting error probability is

$$P_e^{theo} = \frac{M-1}{M^2} \left( \sqrt{1 + (M-1)\chi} - \sqrt{1-\chi} \right)^2. \quad (6)$$

### A. Quantum States and Operators in $\mathcal{H}$

Let us define a basis  $\{|x\rangle, |y\rangle\}$  in  $\mathcal{H}$  such that without loss of generality we can write

$$|\xi_0\rangle = \cos\theta|x\rangle + \sin\theta|y\rangle, \quad |\xi_1\rangle = \cos\theta|x\rangle - \sin\theta|y\rangle, \quad (7)$$

with  $\theta \in [0, \pi/4]$ . The inner product (4) becomes

$$\chi = |\langle \xi_0 | \xi_1 \rangle|^2 = \cos^2 2\theta. \quad (8)$$

A measurement is described by a pair of POVM ( $P_{k,0}, P_{k,1}$ ), each one associated with the outcome  $z_k \in \{0, 1\}$ , which must verify the completeness relation

$$P_{k,0} + P_{k,1} = I. \quad (9)$$

In general, the POVM can depend upon some variables, e.g. the time slot, that are indicated in the dependency upon  $k$ . In addition, with the notation  $\bar{z}_k$  we indicate the sequence of outcomes  $[z_1 \dots z_k]$ .

In the binary scenario, the POVMs can be represented in the space of the Operators on  $\mathcal{H}$  with the matrices

$$P_{k,0} = \frac{1}{2} \begin{bmatrix} 1 + a_k & b_k \\ b_k & 1 - a_k \end{bmatrix}, \quad (10)$$

$$P_{k,1} = \frac{1}{2} \begin{bmatrix} 1 - a_k & -b_k \\ -b_k & 1 + a_k \end{bmatrix},$$

with  $a_k^2 + b_k^2 \leq 1$ . In the particular case of  $a_k^2 + b_k^2 = 1$ , the POVMs are rank-1 orthogonal projectors ( $\Pi_0, \Pi_1$ ), and can be written without loss of generality with operators

$$P_{k,0} = \Pi_{k,0} := |\mu_{k,0}\rangle\langle\mu_{k,0}|, \quad (11)$$

$$P_{k,1} = \Pi_{k,1} := |\mu_{k,1}\rangle\langle\mu_{k,1}|,$$

with

$$|\mu_{k,0}\rangle = \cos \phi_k |x\rangle + \sin \phi_k |y\rangle, \quad (12)$$

$$|\mu_{k,1}\rangle = \sin \phi_k |x\rangle - \cos \phi_k |y\rangle,$$

where  $\phi_k = \frac{1}{2} \angle(a_k + jb_k)$ ,  $\phi_k \in [-\frac{\pi}{2}, \frac{\pi}{2}]$  and  $\angle$  the four-quadrant inverse tangent [18].

The conditional probabilities of  $z_k$  given the quantum state  $|\xi_j\rangle$  in the slot can be calculated from (7) with (10) as

$$P[z_k = i | \xi_{k,x} = |\xi_j\rangle] = p_{i|j} = \text{Tr}[P_{k,i} |\xi_j\rangle\langle\xi_j|] =$$

$$= \frac{1 + (-1)^{i+j} a_k \sin 2\theta + (-1)^i b_k \cos 2\theta}{2} \quad (13)$$

In the specific case of rank-1 projectors, we denote the conditional probability evaluated with (12) as

$$\hat{p}_{i|j} = |\langle\mu_{k,i} | \xi_j\rangle|^2 = \cos^2(\theta - (-1)^j \phi_k + i \frac{\pi}{2}) \quad i, j = 0, 1. \quad (14)$$

We will use the shorthand notation  $p_{i|j}$ ,  $[p_{i,j}]$  for the conditional [joint] probability of  $z_k = i$  with the qubit  $|\xi_{k,x}\rangle = |\xi_j\rangle$  when it is clear from the context which random variables the realizations  $i, j$  refer to. The same notation will be used with  $\hat{p}_{i|j}$   $[\hat{p}_{i,j}]$  in the case of rank-1 projectors employed for the measurement.

## B. Implementation of a binary POVM pair

In this section we state that any arbitrary POVM pair ( $P_{k,0}, P_{k,1}$ ) can be implemented on the subspace spanned by  $\{|0\rangle, |\alpha\rangle\}$ . This fact simplifies the description and the optimization of the receiver scheme for PPM signals.

It is well known [5, 11, 12] that in the case of binary discrimination with minimum error probability, the optimal measurement operators are orthogonal projectors ( $\Pi'_0, \Pi'_1$ ) on the subspace spanned by the quantum states. It is then reasonable to assume that the optimization process on the whole PPM symbol detection requires projectors (and not the more general POVMs) for the local measurements in each slot.

An optimal detection scheme for the binary coherent discrimination problem, e.g. the Dolinar receiver [3, 4], necessarily implements such measurement projectors on the subspace spanned by  $\{|0\rangle, |\alpha\rangle\}$ . In addition, the Dolinar receiver can implement the projectors *for every possible value* of the a priori probability  $\{p'_0, p'_1\}$  and coherent state parameter  $\alpha$ .

Since for a given  $\alpha$ , each value of  $p'_0$  in the continuous range  $[0, 1]$  corresponds to a pair of optimal projectors ( $\Pi'_0, \Pi'_1$ ), there exists a version of the Dolinar receiver that from parameters  $\alpha$  and  $p'_0$  implements the corresponding optimal projectors. Note that in general,  $p'_0$  does not correspond to the a priori probability  $p_0$ , but it is a *fictitious* a priori probability assumed for the design of this version of the Dolinar's.

From the pairs of rank-1 projectors we can design any arbitrary projectors pairs by linear combination. In fact, consider the spectral decomposition of the POVM pair (10),

$$P_{k,0} = \lambda_0 v_0 v_0^T + \lambda_1 v_1 v_1^T, \quad (15)$$

$$P_{k,1} = \lambda_1 v_0 v_0^T + \lambda_0 v_1 v_1^T,$$

where  $\lambda_0, \lambda_1$  and  $v_0, v_1$  represent the appropriate eigenvalues and eigenvectors. Any POVM pair of the type (10) is implemented by performing a measurement with the projectors pair

$$\Pi_0 = v_0 v_0^T, \quad \Pi_1 = v_1 v_1^T, \quad (16)$$

and then relabeling the outcome  $\hat{0}$  into 0 [1] with probability  $\lambda_0$   $[\lambda_1]$  and the outcome  $\hat{1}$  with probability  $\lambda_1$   $[\lambda_0]$  respectively.

The Dolinar receiver scheme can be described both in the temporal and in the spatial domain [13, 14]. In the former case, the scheme requires a time variant displacement operation driven by a photon counter in a feedforward fashion. In the latter case, an infinite sequence of beam splitters are required [14], each of them provided with a block containing a displacement operation, a photon counting and a feedforward from the previous block.

If setup limitations or constraints, e.g. the availability of only fixed (time or spatial invariant) displacements, prevents the exact realization of a Dolinar scheme, the PPM receiver design algorithm suffers from these limitations in implementing the local measurement projectors, leading to suboptimal performances.

In the present work we consider both the above scenarios, i.e. the availability of arbitrary local projective measurements in each slot, given for example by a Dolinar

receiver, and the limitation to perform only fixed displacements. If on one hand the former solution allows greater possibilities in the design, the latter one is more practical and suitable for an experimental setup.

### III. STRUCTURE OF AN ADAPTIVE RECEIVER

The qubit description of the transmitted quantum state given by the tensor product of  $|\xi_i\rangle$ ,  $i = 0, 1$  reflects the structure of the PPM signals as a temporal sequence of  $M$  coherent states. Current classical receiver schemes measure the entire signal as a whole, leaving the estimation of the symbol after the measurement is concluded. However, due to the temporal sequence of the slot pulses, this measurement can be interpreted as a sequence of  $M$  measurements.

In each slot, the classical receiver performs a photon counting, also referred to as *direct detection* (DD), that is an energy measurement with outcomes corresponding to the presence of any or no photons. An ideal photon counter, with unit efficiency and no dark counts, is associated to operators

$$\Pi_0 = |0\rangle\langle 0|, \quad \Pi_1 = I - |0\rangle\langle 0|, \quad (17)$$

with conditional probabilities on the Fock space

$$q_{i|j} = P[z_k = i | \gamma_{k,x} = |j\rangle], \quad i = 0, 1, j = 0, \alpha \quad (18)$$

calculated by the Born's rule

$$\begin{aligned} q_{0|0} &= \text{Tr}[\Pi_0|0\rangle\langle 0|] = |\langle 0|0\rangle|^2 = 1, \\ q_{1|\alpha} &= \text{Tr}[(I - |0\rangle\langle 0|)|\alpha\rangle\langle \alpha|] = 1 - e^{-|\alpha|^2}. \end{aligned} \quad (19)$$

In the corresponding qubit framework, the measurement operators are written as

$$|\mu_{k,0}\rangle\langle \mu_{k,0}| = |\xi_0\rangle\langle \xi_0|, \quad |\mu_{k,1}\rangle\langle \mu_{k,1}| = I - |\xi_0\rangle\langle \xi_0|, \quad (20)$$

which give the correct conditional probabilities,

$$\begin{aligned} p_{0|0} &= |\langle \mu_0|\xi_0\rangle|^2 = 1, \\ p_{1|1} &= |\langle \mu_1|\xi_1\rangle|^2 = 1 - \chi = 1 - e^{-|\alpha|^2}. \end{aligned} \quad (21)$$

The receiver estimates the transmitted symbol corresponding to the measurement where the outcome  $z_k = 1$  has been observed, that is where one or more photons have been detected. By the PPM definition, multiple outcomes equal to 1 are impossible to detect, while it could be possible to observe only 0. When all the outcomes are  $z_k = 0$ , the receiver chooses at random. The error probability is thus

$$P_e^{DD} = \frac{M-1}{M}\chi = \frac{M-1}{M}e^{-|\alpha|^2}. \quad (22)$$

It is natural to ask whether it is possible to improve this receiver structure by adapting the subsequent slot

measurements on the base of previous measurement outcomes. This idea was initially pointed out by Dolinar [6] who proposed an adaptive nulling of the received signal.

When the *nulling* operation is performed in the slot, a displacement  $D(-\alpha)$  is applied to the current coherent state: in the case of the ground state it is displaced to  $|\alpha\rangle$ , while in the case of the coherent state  $|\alpha\rangle$  it is displaced to  $|0\rangle$ . After the nulling operation, a photon counting is performed with operators (17). The result of the nulling operation is that the conditional probabilities (19) change in

$$\begin{aligned} q_{0|0} &= |\langle 0|D(-\alpha)|0\rangle|^2 = e^{-|\alpha|^2}, \\ q_{1|\alpha} &= \text{Tr}[(I - |0\rangle\langle 0|)D(-\alpha)|\alpha\rangle\langle \alpha|D(-\alpha)^\dagger] = 0. \end{aligned} \quad (23)$$

In our qubit framework, this measurement is described by the operators

$$|\mu_{k,0}\rangle\langle \mu_{k,0}| = |\xi_1\rangle\langle \xi_1|, \quad |\mu_{k,1}\rangle\langle \mu_{k,1}| = I - |\xi_1\rangle\langle \xi_1|. \quad (24)$$

with consequent conditional probabilities

$$\begin{aligned} p_{0|0} &= |\langle \mu_{k,0}|\xi_0\rangle|^2 = \chi = e^{-|\alpha|^2}, \\ p_{1|1} &= |\langle \mu_{k,1}|\xi_1\rangle|^2 = 0. \end{aligned} \quad (25)$$

The Dolinar receiver for PPM starts with an initial hypothesis  $y = 1$ . It applies the nulling to the (unknown) coherent state in the first slot, and performs the photon counting.

If the outcome is  $z_1 = 0$ , i.e. no photons are detected, the receiver believe in the temporary hypothesis  $y$  and simply direct detect the remaining slots. During this stage, the receiver keeps the temporary hypothesis unless an outcome  $z_k = 1$ , i.e. some photons, are observed, and in that case the hypothesis is changed in  $y = k$ .

On the contrary, if the measured outcome in the nulled slot is  $z_1 = 1$ , i.e. at least one photon is recorded, the receiver completely neglects the current hypothesis, replacing it with the sequent symbol,  $y = 2$ . The nulling algorithm is then repeated recursively from the next slot, reducing the original problem to its  $M - 1$  version.

It is easy to shown that the receiver fails in the correct discrimination of symbol  $x = i$  estimating  $y = k$  when two wrong outcomes are observed, in the nulled slot  $k$  and in the slot  $i$  during direct detection [6, 7]. The error probability of this scheme is

$$P_e^{CN} = \frac{1}{M} [(1 - \chi)^M - 1 + M\chi]. \quad (26)$$

Two architectures have been proposed [7] 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 *Generalized Kennedy* uses the same concept to improve the Kennedy receiver [15].

The first architecture, denoted in [7] as *Type I*, uses the same algorithm as the PPM Dolinar receiver to choose consecutive measurements on the basis of previous outcomes, but applying a constant displacement  $D(\beta)$ , with

$\beta \neq -\alpha$ , in place of the exact nulling. The second architecture, called *Type II*, also adds a phase-sensitive amplifier with gain  $G$  to squeeze the partially-nulled coherent state, further improving the performance. The error probability for this *Improved Conditional Nulling Type II* scheme, in our ideal assumptions of unit efficiency and no dark current, becomes

$$P_e^{ICN} = \frac{q_{1|\alpha}(1-(1-q_{0|0})^{M-1}) + e^{-|\alpha|^2}(Mq_{0|0}-1+(1-q_{0|0})^M)}{Mq_{0|0}} \quad (27)$$

with

$$q_{0|0} = \frac{\exp\left[-\frac{(\sqrt{G} + \sqrt{G-1})^2|\beta|^2}{1 + \sqrt{G-1}(\sqrt{G} + \sqrt{G-1})}\right]}{\sqrt{G}}, \quad (28)$$

$$q_{1|\alpha} = 1 - \frac{\exp\left[-\frac{(\sqrt{G} + \sqrt{G-1})^2|\alpha - \beta|^2}{1 + \sqrt{G-1}(\sqrt{G} + \sqrt{G-1})}\right]}{\sqrt{G}}.$$

Substituting  $G = 1$ , we obtain the expression for the Type I architecture. From the expression (27), the displacement  $\beta$  and the phase-sensitive amplifier  $G$  have been numerically optimized to reach the maximum performance of this receiver scheme.

In the qubit framework, the direct detection is performed again with operators (20), while when the displacement and squeezing operations are performed, the operators are described by POVM as in (10). The POVM parameters  $a_k$ ,  $b_k$  depend upon the values of  $\beta$  and  $G$  employed, and can be obtained by inversion of (13) with the transition probabilities (28).

The conditional nulling scheme and its improved versions share a common algorithm that creates a decision tree of the possible sequence of measurements. Proceeding from the root to the leaves, in each node it is decided which branch to take depending on the last outcome. However, it was pointed out in [7] that further performance improvements can be obtained by considering different displacements  $\beta_k$  for each slot  $k = 1 \dots M$ . Moreover, further generalization leads to time varying displacements  $\beta_k(t)$ .

In addition, the decision tree of these architectures is not symmetrical, and the direct detection of all the slots after the outcome  $z_k = 1$  has been observed in a nulled slot may not be the best strategy.

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 3). In order to focus on the transition probabilities, we use the qubit representation to describe the measurement in the  $k+1$ -th slot, i.e. after  $k$  outcomes has been observed, with  $\phi_k$  the parameters that specify the local measurement. For example, in the

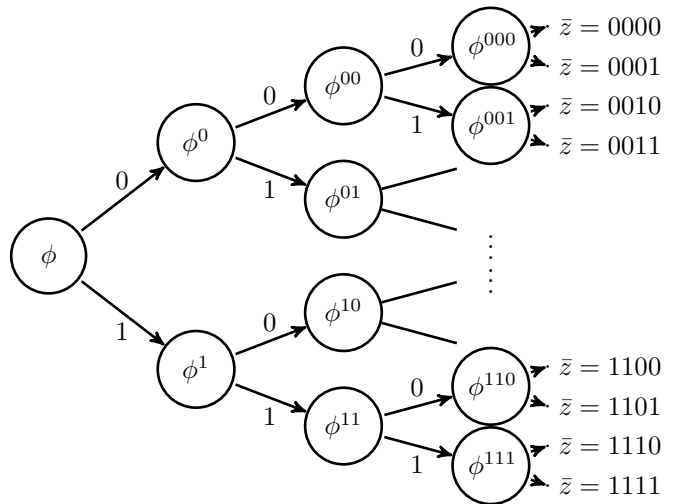


FIG. 3: Strategy tree for an adaptive receiver algorithm, in the case of a 4-PPM. Each node represents a measurements, identified by a parameter, and each branch a possible outcome. The parameter  $\phi_k = \pi_k(\bar{z}_k)$  employed at the measurement  $k+1$  after observing the outcomes sequence  $\bar{z}_k$  is indicated as  $\phi^{\bar{z}_k}$  to shorten the notation. The parameter defining the measurement in the first slot depends only upon the a priori information, and is denoted with  $\phi$ . The algorithm proceeds from the root on the left to the leaves on the right following the path indicated by the outcomes.

case of the projective measurements,  $\phi_k$  coincides with the angle in the definition (12). We specify the function  $\phi_k = \pi_k(\bar{z}_k)$  to define the adaptive strategy [19].

The receiver starts with the first measurement specified by  $\phi_0$ . Then, depending on the outcome  $z_1 = 0$  or  $z_1 = 1$ , it proceeds with a measurement in the second slot defined by  $\phi_1 = \pi_1(0)$  or  $\phi_1 = \pi_1(1)$  respectively. In general it results  $\pi_1(0) \neq \pi_1(1)$ . 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 parameters  $\phi_k = \pi_k(\bar{z}_k)$ , for all  $k$  and  $\bar{z}_k$  is necessary. Since the number of parameters 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 simplify the optimization problem by applying the dynamic programming algorithm, that is the topic of the next Section.

#### IV. 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 [16].

Consider a discrete time dynamic system described by

the update equation

$$s_{k+1} = f_k(s_k, u_k, w_k), \quad k = 0, \dots, N-1, \quad (29)$$

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 information 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 (29) 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 [20] is associated with the system evolution, that in our case we can write as

$$g(s_N), \quad (30)$$

and depends upon the final system state  $s_N$ . Since the evolution (29) is influenced by the random variables  $w_0, \dots, w_{N-1}$ , the final system state  $s_N$  is a random variable and the expected reward we want to maximize is

$$\mathbb{E}_{s_N} [g(s_N)] = \int_{S_N} d\sigma g(\sigma) \wp_{s_N}(\sigma), \quad (31)$$

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 (29) for  $k = N-1$ , the expected reward can be rewritten as

$$\begin{aligned} \mathbb{E}_{s_{N-1}, w_{N-1}} [g(f_{N-1}(s_{N-1}, u_{N-1}, w_{N-1}))] &= \\ &= \int_{S_{N-1}} d\rho \int_{W_{N-1}} d\omega g(f_{N-1}(\rho, u_{N-1}, \omega)) \\ &\quad \times \wp_{s_{N-1}, w_{N-1}}(\rho, \omega). \end{aligned} \quad (32)$$

By (29) we can iterate the substitutions backward in the index  $k$  to obtain a reformulation of the expected reward (31) in terms of the controls and the initial system state  $s_0$ ,

$$\begin{aligned} \int_{W_0} d\omega_0 \cdots \int_{W_{N-1}} d\omega_{N-1} \wp_{w_0, \dots, w_{N-1}}(\omega_0, \dots, \omega_{N-1}) \\ \times g(f_{N-1}(\dots f_0(s_0, u_0, \omega_0), \dots, u_{N-1}, \omega_{N-1})) \end{aligned} \quad (33)$$

where we use the fact that the initial state  $s_0$  is known, and we have explicitly indicated the composition of the update functions  $f_{N-1}, f_{N-2}, \dots, f_0$  from  $s_0$  with the variables  $u_0, \dots, u_{N-1}, w_0 = \omega_0, \dots, w_{N-1} = \omega_{N-1}$ .

In seeking the maximization of (33), we can employ different strategies. For example, the values of the control  $u_0, \dots, u_{N-1}$  can be determined before the 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  $s_k$ ,

$$u_k = \pi_k(s_k), \quad s_k \in S_k, \quad u_k \in U_k, \quad (34)$$

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 *reward-to-go* function at time  $k$  from the current system state  $s_k$  as the function

$$J_k : S_k \times U_k \times \Phi_{k+1} \times \dots \times \Phi_{N-1} \mapsto \mathbb{R} \quad (35)$$

specified as

$$\begin{aligned} J_k(\rho, \nu, \pi_{k+1}, \dots, \pi_{N-1}) &= \\ &\int_{W_k} d w_k \cdots \int_{W_{N-1}} d w_{N-1} \wp_{w_k, \dots, w_{N-1} | s_k}(\omega_k, \dots, \omega_{N-1} | \rho) \\ &\quad \times g(f_{N-1}(\dots f_k(\rho, \nu, \omega_k), \dots, u_{N-1}, \omega_{N-1})) \end{aligned} \quad (36)$$

with  $\Phi_k = U_k^{S_k}$  the set of all possible functions  $\pi_k : S_k \mapsto U_k$ .

Define  $\bar{\pi}^* = (\pi_0^*, \pi_1^*, \dots, \pi_{N-1}^*)$  the optimal policy, that is the one that maximize  $J_0$

$$\bar{\pi}^* := \operatorname{argmax}_{\bar{\pi}} J_0(s_0, \bar{\pi}(s_0)), \quad (37)$$

and define

$$J_0^*(s_0) := J_0(s_0, \bar{\pi}^*(s_0)) \quad (38)$$

the optimal reward from  $s_0$ .

The dynamic programming algorithm relies on the following idea.

### Principle of Optimality [16]

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  $k$  with  $s_k = \sigma$ ,

$$\max_{\pi_k, \dots, \pi_{N-1}} J_k(\sigma, \pi_k(\sigma), \dots, \pi_{N-1}) =: J_k^*(\sigma). \quad (39)$$

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  $\pi_k$ . Instead, the dynamic programming algorithm decomposes the main problem into a sequence of subproblems.

### Dynamic Programming Algorithm [16]

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

$$J_N^*(\sigma) = g(\sigma), \quad \sigma \in S_N \quad (40)$$

2. for  $k = N-1, \dots, 0$ , for all  $\rho \in S_k$  evaluate the optimal control and the optimal reward-to-go function at time  $k$ , namely

$$\pi_k^*(\rho) = \operatorname{argmax}_{\nu \in U_k} J_k(\rho, \nu, \pi_{k+1}^*, \dots, \pi_{N-1}^*) \quad (41)$$

$$= \operatorname{argmax}_{\nu \in U_k} E_{w_k} [J_{k+1}^*(f_k(\rho, \nu, w_k))], \quad (42)$$

$$J_k^*(\rho) = J_k(\rho, \pi_k^*(\rho), \pi_{k+1}^*, \dots, \pi_{N-1}^*). \quad (43)$$

3. the optimal reward and the optimal policy are

$$J_0^*(s_0) = J_0(s_0, \pi_0^*(s_0), \dots, \pi_{N-1}^*), \quad (44)$$

$$\bar{\pi}^* = (\pi_0^*, \pi_1^*, \dots, \pi_{N-1}^*). \quad (45)$$

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  $s_k = \rho$  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.

#### A. Reformulation of the Dynamic Programming Algorithm

In some optimization problems, it could be of interest to include the probability of the actual system state  $s_k$  among the information to pass from an iteration to another, as a component of the system state. In this case, the probability of the system state  $\wp_{s_k}(\cdot)$  refers to the other components. As we will see later, this allows us to reduce the size of the system state thus speeding up the computation of the algorithm.

A consequence of this new definition is that in the steps 1. and 2. of the dynamic programming algorithm must be evaluated assuming a possible value for the probability of the state. This is not a problem since even in the original formulation the system state probability  $\wp_{s_k}(\cdot)$

is not known, but it is obtained by the backward steps of the algorithm. In the same manner, non-admissible values of the system state probabilities are excluded by the backward procedure.

Including the system state probability allows us to define a reward-to-go function which uses its value, with joint probabilities rather than conditional probabilities, as in

$$\begin{aligned} \tilde{J}_k(\rho, \nu, \pi_{k+1}, \dots, \pi_{N-1}) &= \\ &\int_{W_k} dw_k \cdots \int_{W_{N-1}} dw_{N-1} \wp_{s_k, w_k, \dots, w_{N-1}}(\rho, \omega_k, \dots, \omega_{N-1}) \\ &\quad \times g(f_{N-1}(\dots f_k(\rho, u_k, \omega_k), \dots, u_{N-1}, \omega_{N-1})) \end{aligned} \quad (46)$$

$$\begin{aligned} &= \wp_{s_k}(\rho) \int_{W_k} dw_k \cdots \int_{W_{N-1}} dw_{N-1} \\ &\quad \times g(f_{N-1}(\dots f_k(\rho, u_k, \omega_k), \dots, u_{N-1}, \omega_{N-1})) \\ &\quad \times \wp_{w_k, \dots, w_{N-1}|s_k}(\omega_k, \dots, \omega_{N-1}|\rho), \end{aligned} \quad (47)$$

$$= \wp_{s_k}(\rho) J_k(\rho, \nu, \pi_{k+1}, \dots, \pi_{N-1}) \quad (48)$$

In addition, the dynamic programming algorithm can be reformulated with integrals rather than expectations, as in the following procedure.

### Dynamic Programming Algorithm (Reformulation)

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

$$\tilde{J}_N^*(\sigma) = \tilde{g}(\sigma) := g(\sigma)\wp_{s_N}(\sigma), \quad \sigma \in S_N \quad (49)$$

2. for  $k = N-1, \dots, 0$ , for all  $\rho \in S_k$  evaluate the optimal control and the optimal reward-to-go function at time  $k$ , namely

$$\pi_k^*(\rho) = \operatorname{argmax}_{\nu \in U_k} \tilde{J}_k(\rho, \nu, \pi_{k+1}^*, \dots, \pi_{N-1}^*) \quad (50)$$

$$= \operatorname{argmax}_{\nu \in U_k} \int_{W_k} d\omega_k \tilde{J}_{k+1}^*(f_k(\rho, \nu, \omega_k)), \quad (51)$$

$$\tilde{J}_k^*(\rho) = \tilde{J}_k(\rho, \pi_k^*(\rho), \pi_{k+1}^*, \dots, \pi_{N-1}^*). \quad (52)$$

3. the optimal reward and the optimal policy are

$$J_0^*(s_0) = \tilde{J}_0(s_0, \pi_0^*(s_0), \dots, \pi_{N-1}^*), \quad (53)$$

$$\bar{\pi}^* = (\pi_0^*, \pi_1^*, \dots, \pi_{N-1}^*). \quad (54)$$

Note that since  $J_k$  and  $\tilde{J}_k$  differs by only the factor  $\wp_{s_k}$ , the values of  $\pi_k^*(\rho)$  maximizing the two definitions of the reward-to-go function are the same.

## V. OPTIMIZATION ALGORITHM

In this Section we follow the dynamic programming algorithm to optimize the parameters tree depicted in Figure 3. More details are reported in Appendix A, where some preliminary lemmas are explained in order to better understand the algorithm.

### A. System 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  $y$ . By extension, we can define  $k = 0$  as the time step before the measurement in the first slot.

In order to define the system state  $s_k$ , we need to evaluate the information that is necessary to express the probability of correct decision. In general, a receiver strategy considers first a measurement stage, where information about the system under investigation are acquired performing measurement on it, and later an estimation stage, where the information are processed in order to formulate the answer to our detection problem.

The estimation stage is a (possibly probabilistic) map  $h$  that assigns to each measurement outcome sequences  $\bar{z}_M$  an estimate  $y$  of the transmitted symbol  $x$ . The final probability of correct decision is therefore rewritten as

$$\begin{aligned} P_c &= \text{P}[y = x] = \sum_{i=1}^M \text{P}[y = i, x = i] \\ &= \sum_{i=1}^M \sum_{z \in \mathcal{Z}_M} \text{P}[y = i, x = i, \bar{z}_M = z] \\ &= \sum_{i=1}^M \sum_{z \in \mathcal{Z}_M} \text{P}[y = i | x = i, \bar{z}_M = z] p_{i,z}. \end{aligned} \quad (55)$$

where  $\text{P}[x = i, \bar{z}_M = z] = p_{i,z}$ .

The performance index (55) is maximized by a deterministic map  $h$  according to the maximum a posteriori (MAP) criterion,

$$h(z) = \underset{i \in \{1, \dots, M\}}{\operatorname{argmax}} p_{i,z}, \quad (56)$$

such that the probability of correct decision reads

$$P_c = \sum_{z \in \mathcal{Z}_M} \text{P}[x = h(z), \bar{z}_M = z] = \sum_{z \in \mathcal{Z}_M} \max_{i \leq M} p_{i,z} \quad (57)$$

The probability of correct decision (57) corresponds to the expected reward function (33), namely

$$\mathbb{E}_{w_0, \dots, w_{N-1}} [g(s_M)] = \sum_{z \in \mathcal{Z}_M} \left( \max_{i \leq M} p_{i|z} \right) p_z, \quad (58)$$

and therefore we can identify  $w_k = z_{k+1}$  and

$$g(s_M(z)) = \max_{i \leq M} p_{i|z}. \quad (59)$$

While the estimation stage is completely optimized by (56), in order to optimize the measurement stage, it seems natural to consider the information described by the vector of the conditional probabilities given the outcome  $\bar{z}_k$ ,

$$\bar{p}_{\bar{z}_k} = \begin{bmatrix} p_{1|\bar{z}_k} \\ p_{2|\bar{z}_k} \\ \vdots \\ p_{M|\bar{z}_k} \end{bmatrix}. \quad (60)$$

However, not all the entries in (60) are necessary in order to solve (56) and therefore define the *system state* of the receiver algorithm. In fact, given the outcome sequence  $\bar{z}_k$ , due to the local binary discrimination in each slot it turns out that only two conditional probabilities are necessary,  $p_{m|\bar{z}_k}$  and  $p_{m|\bar{z}_k}$ , with  $m(\bar{z}_k)$  the maximum a posteriori (MAP) estimate of  $x$  among the symbols  $1, \dots, k$ , i.e.

$$m(\bar{z}_k) = \underset{i \leq k}{\operatorname{argmax}} p_{i|\bar{z}_k}. \quad (61)$$

These considerations lead to a definition of the state as in

$$s_k(\bar{z}_k) = (p_{m|\bar{z}_k}, p_{M|\bar{z}_k}, m(\bar{z}_k), p_{\bar{z}_k}). \quad (62)$$

where we include in the last component the probability of the state, as explained in the previous Section IV A. However, due to the reformulation of the dynamic programming algorithm, it suffices to define as a system state  $s_k \in S_k$  the triple

$$s_k(\bar{z}_k) = (p_{m, \bar{z}_k}, p_{M, \bar{z}_k}, m(\bar{z}_k)), \quad (63)$$

$$S_k = \{(u, v, i) : 0 \leq u \leq 1, 0 \leq v \leq 1, u + v \leq 1, i \in \{1, \dots, k\}\}, \quad (64)$$

with joint probabilities rather than conditional ones, and use the definition of probability of correct decision (57). This definition is not equivalent to (62), since we cannot recover (62) from (63), but still it suffices to perform the optimization. Note that  $m(\bar{z}_k)$  can be defined in the same manner of (61) with joint probabilities.

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 (63) 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  $z_{k+1} = 0$  or  $z_{k+1} = 1$ , employing the parameters  $\phi_k = \pi_k(\bar{z}_k)$  in the measurement. The joint probabilities  $p_{M, [\bar{z}_k 0]}$ ,  $p_{M, [\bar{z}_k 1]}$  can be easily obtained from the transition probabilities in equations (14), in the case  $k < M - 1$  by

$$\begin{aligned} p_{M, [\bar{z}_k 0]} &= \hat{P}_{0|0} p_{M, \bar{z}_k}, \\ p_{M, [\bar{z}_k 1]} &= \hat{P}_{1|0} p_{M, \bar{z}_k}, \end{aligned} \quad (65)$$

while if  $k = M - 1$  by

$$\begin{aligned} p_{M, [\bar{z}_k 0]} &= \hat{P}_{0|1} p_{M, \bar{z}_k}, \\ p_{M, [\bar{z}_k 1]} &= \hat{P}_{1|1} p_{M, \bar{z}_k}. \end{aligned} \quad (66)$$

Note that given the system state  $s_k(\bar{z}_k)$ , the probability  $p_{M, \bar{z}_k}$  equals the joint probability

$$p_{i, \bar{z}_k} = \text{P}[x = i, \bar{z}_k], \quad i = k + 1, k + 2, \dots, M$$

due to the equal a priori probability and the same quantum states  $|\xi_0\rangle$  in the first  $k$  slots for all symbols  $x = k + 1, \dots, M$  (see Lemma 3 in Appendix A).

In the case of  $p_{m(\bar{z}_k 0), [\bar{z}_k 0]}$  and  $p_{m(\bar{z}_k 1), [\bar{z}_k 1]}$ , in order to obtain the update equation we apply its definition,

$$\begin{aligned} p_{m(\bar{z}_k 0), [\bar{z}_k 0]} &= \max_{i \leq k+1} \{p_{i, [\bar{z}_k 0]}\} \\ &= \max \left\{ \hat{P}_{0|1} p_{k+1, \bar{z}_k}, \max_{i \leq k} \{ \hat{P}_{0|0} p_{i, \bar{z}_k} \} \right\} \\ &= \max \left\{ \hat{P}_{0|1} p_{M, \bar{z}_k}, \hat{P}_{0|0} p_{m(\bar{z}_k), \bar{z}_k} \right\}, \end{aligned} \quad (67)$$

$$\begin{aligned} p_{m(\bar{z}_k 1), [\bar{z}_k 1]} &= \max_{i \leq k+1} \{p_{i, [\bar{z}_k 1]}\} \\ &= \max \left\{ \hat{P}_{1|1} p_{k+1, \bar{z}_k}, \max_{i \leq k} \{ \hat{P}_{1|0} p_{i, \bar{z}_k} \} \right\} \\ &= \max \left\{ \hat{P}_{1|1} p_{M, \bar{z}_k}, \hat{P}_{1|0} p_{m(\bar{z}_k), \bar{z}_k} \right\}. \end{aligned} \quad (68)$$

Thereby, the symbol  $m(\bar{z}_k)$  is updated into

$$m([\bar{z}_k z_{k+1}]) \in \{m(\bar{z}_k), k + 1\} \quad (69)$$

accordingly with the term maximizing (67) and (68). 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 information, and due to the equal a priori distribution for the symbols, we can define

$$s_0 = \left( \frac{1}{M}, \frac{1}{M}, \dots \right), \quad (70)$$

where it is unnecessary to specify  $m(\emptyset)$ , since the first update of the system state leads to the trivial maximization

$$m(z_1) = \operatorname{argmax}_{i \leq 1} p_{i, z_1} = 1. \quad (71)$$

The parameters  $\phi_k = \pi_k(\bar{z}_k)$  appear (implicitly) in the transition probabilities of update equations (65)-(69), highlighting their role of control variables. In the next Section we formulate the corresponding expected reward function, and following the dynamic programming we define  $\phi_k$  as function of the state  $s_k$ ,

$$\phi_k = \pi_k(s_k(\bar{z}_k)) \quad (72)$$

optimizing for each  $\bar{z}_k$ .

## B. Reward-to-go 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.

Consider now the system state  $s_k(z)$  before the  $k + 1$ -th measurement, and define the reward-to-go function

$$\begin{aligned} \tilde{J}_k(s_k(z), \phi_k, \pi_{k+1}, \dots, \pi_{M-1}) &= \\ &= \sum_{z' \in \mathcal{Z}_{M-k}} p_{m([z z']), [z z']} \end{aligned} \quad (73)$$

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 (73) is implicit in the transition probabilities of the outcomes.

It is trivial to see that

$$\tilde{J}_0(s_0, \pi_0(s_0), \dots, \pi_{M-1}) = P_c \quad (74)$$

and

$$\begin{aligned} \tilde{J}_{M-1}(s_{M-1}(z), \phi_{M-1}) &= \sum_{z_M=0,1} p_{m([z z_M]), [z z_M]}(\phi_{M-1}) \\ &= p_{m([z 0]), [z 0]}(\phi_{M-1}) + p_{m([z 1]), [z 1]}(\phi_{M-1}) \\ &= \text{P}[z_M = 0 | x = i] p_{i, z} + \text{P}[z_M = 1 | x = j] p_{j, z}, \end{aligned} \quad (75)$$

with  $i = h([z 0])$  and  $j = h([z 1])$ .

We consider for the moment the possibility to perform arbitrary local projective measurements, as explained in Section II B. Therefore, the transition probabilities  $\text{P}[z_M = z | x = i] = \hat{p}_{z|i}$  can be evaluated as in (14).

As explained in the Lemma 1 in Appendix A, expression (75) 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 the analytical maximization

$$\tilde{J}_{M-1}^*(s_{M-1}) = \frac{1}{2} \left[ p_m + p_M + \sqrt{(p_m + p_M)^2 - 4p_m p_M \chi} \right] \quad (76)$$

obtained employing the angle in the  $M$ -th measurement

$$\pi_{M-1}^*(s_{M-1}) = \frac{1}{2} \frac{\cos 2\theta(p_m - p_M) + j \sin 2\theta(p_m + p_M)}{\quad} \quad (77)$$

where in both (76) and (77) we drop the dependency from  $\bar{z}_{M-1}$ .

Moreover, we can easily write down the update equation for the reward-to-go

$$\begin{aligned} \tilde{J}_k(s_k(z), \phi_k, \pi_{k+1}, \dots, \pi_{M-1}) &= \sum_{\substack{z' \in \\ \mathcal{Z}_{M-k}}} p_m([z \ z'], [z \ z']) \\ &= \sum_{z_{k+1}} \sum_{\substack{z'' \in \\ \mathcal{Z}_{M-k-1}}} p_m([z \ z_{k+1} \ z''], [z \ z_{k+1} \ z'']) \\ &= \tilde{J}_{k+1}(s_{k+1}([z \ 0]), \pi_{k+1}, \pi_{k+2}, \dots, \pi_{M-1}) \\ &\quad + \tilde{J}_{k+1}(s_{k+1}([z \ 1]), \pi_{k+1}, \pi_{k+2}, \dots, \pi_{M-1}) \end{aligned} \quad (78)$$

In equation (78), the role of  $\phi_k$  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, which depend on parameters  $\phi = \phi_k$  that appear in the transition probabilities of (67) and (68). Therefore, four possible evolutions must be considered in evaluating (78), indicated with A, B, C and D in equations (79)-(82). Since we want to maximize the probability of correct decision,  $\tilde{J}_k^*(s_k)$  is the maximum between these possibilities, as in (83).

Along with the reward-to-go function, we define the function  $\pi_k^*$  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.

### C. 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  $\tilde{J}_k$  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 used 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  $\tilde{J}_{M-1}^*$  and the function  $\pi_{M-1}^*$  for each  $(p_m, p_M)$ ,  $p_m + p_M \leq 1$  as in (76) and (77).
2. Repeat step 3. and 4. for  $k = M-1, \dots, 1$ .
3. From  $\tilde{J}_{k+1}^*$ , for each  $(p_m, p_M)$ ,  $p_m + p_M \leq 1$  evaluate  $\tilde{J}_k^*$  and  $\pi_k^*$  as in (83) and (84) respectively.
4. For each  $(p_m, p_M)$ ,  $p_m + p_M \leq 1$ , depending on the association A, B, C or D of (79)-(82) used in the

previous step, define the children nodes of  $s_k(z_k)$  generated with outcome  $z_{k+1} = 0$  and  $z_{k+1} = 1$

$$\text{children}(s_k) = \{s_{k+1}([\bar{z}_k 0]), s_{k+1}([\bar{z}_k 1])\} \quad (85)$$

Note that in  $s_{k+1}([\bar{z}_k 0])$  and in  $s_{k+1}([\bar{z}_k 1])$  we can define  $m([\bar{z}_k 0])$  and  $m([\bar{z}_k 1])$  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’’ indicating the value is  $m \leq k$ , that will be defined in later iterations of the optimization.

5. Evaluate the parameter in the first measurement and the performances of the adaptive algorithm from  $s_0$  as

$$\begin{aligned} \pi_0^* &= \operatorname{argmax}_{\phi} \tilde{J}_1^* \left( \frac{\hat{p}_{0|1}(\phi)}{M}, \frac{\hat{p}_{0|0}(\phi)}{M}, 1 \right) \\ &\quad + \tilde{J}_1^* \left( \frac{\hat{p}_{1|1}(\phi)}{M}, \frac{\hat{p}_{1|0}(\phi)}{M}, 1 \right) \end{aligned} \quad (86)$$

$$\begin{aligned} P_c &= \tilde{J}_1^* \left( \frac{\hat{p}_{0|1}(\pi_0^*)}{M}, \frac{\hat{p}_{0|0}(\pi_0^*)}{M}, 1 \right) \\ &\quad + \tilde{J}_1^* \left( \frac{\hat{p}_{1|1}(\pi_0^*)}{M}, \frac{\hat{p}_{1|0}(\pi_0^*)}{M}, 1 \right) \end{aligned} \quad (87)$$

In order to reconstruct the binary tree parameters 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 parameter  $\phi_k$ . 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  $z_k = 0$  or  $z_k = 1$ . 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 parameters.
8. Define the children nodes of the system state  $s_0$ , the one corresponding to the outcome  $z_1 = 0$ ,

$$s_1(0) = \left( \frac{\hat{p}_{0|1}(\pi_0^*)}{M}, \frac{\hat{p}_{0|0}(\pi_0^*)}{M}, 1 \right), \quad (88)$$

and the other corresponding to  $z_1 = 1$ ,

$$s_1(1) = \left( \frac{\hat{p}_{1|1}(\pi_0^*)}{M}, \frac{\hat{p}_{1|0}(\pi_0^*)}{M}, 1 \right). \quad (89)$$

$$\tilde{J}_{k,A}(s_k, \phi) = \tilde{J}_{k+1}^*(\overbrace{\hat{p}_{0|0}(\phi)p_m, \hat{p}_{0|0}(\phi)p_M, m}^{s_{k+1}([\bar{z}_k 0])}) + \tilde{J}_{k+1}^*(\overbrace{\hat{p}_{1|0}(\phi)p_m, \hat{p}_{1|0}(\phi)p_M, m}^{s_{k+1}([\bar{z}_k 1])}), \quad (79)$$

$$\tilde{J}_{k,B}(s_k, \phi) = \tilde{J}_{k+1}^*(\hat{p}_{0|0}(\phi)p_m, \hat{p}_{0|0}(\phi)p_M, m) + \tilde{J}_{k+1}^*(\hat{p}_{1|1}(\phi)p_M, \hat{p}_{1|0}(\phi)p_M, k+1), \quad (80)$$

$$\tilde{J}_{k,C}(s_k, \phi) = \tilde{J}_{k+1}^*(\hat{p}_{0|1}(\phi)p_M, \hat{p}_{0|0}(\phi)p_M, k+1) + \tilde{J}_{k+1}^*(\hat{p}_{1|0}(\phi)p_m, \hat{p}_{1|0}(\phi)p_M, m), \quad (81)$$

$$\tilde{J}_{k,D}(s_k, \phi) = \tilde{J}_{k+1}^*(\hat{p}_{0|1}(\phi)p_M, \hat{p}_{0|0}(\phi)p_M, k+1) + \tilde{J}_{k+1}^*(\hat{p}_{1|1}(\phi)p_M, \hat{p}_{1|0}(\phi)p_M, k+1) \quad (82)$$

$$\tilde{J}_k^*(s_k) = \max_{\phi} \{ \tilde{J}_{k,A}(s_k, \phi), \tilde{J}_{k,B}(s_k, \phi), \tilde{J}_{k,C}(s_k, \phi), \tilde{J}_{k,D}(s_k, \phi) \} \quad (83)$$

$$\pi_k^*(s_k) = \operatorname{argmax}_{\phi} \{ \tilde{J}_{k,AC}(\phi), \tilde{J}_{k,AD}(\phi), \tilde{J}_{k,BC}(\phi), \tilde{J}_{k,BD}(\phi) \} \quad (84)$$

9. Repeat step 10. for  $k = 2, \dots, M-1$ .
10. For each node  $s_k(\bar{z}_k)$  of the level  $k$  in the binary parameters tree, the parameter corresponding to the *next* measurement is

$$\pi_k^*(s_k(\bar{z}_k)) \quad (90)$$

and in the next level of the system state tree add  $s_{k+1}([\bar{z}_k 0])$  and  $s_{k+1}([\bar{z}_k 1])$ , replacing, if present, the label “previous” with the symbol  $m(s_k(\bar{z}_k))$ .

Once completed these trees, following the outcome sequence through the parameter tree we get the parameter  $\phi_k = \pi_k(s_k(\bar{z}_k))$  to employ in the  $k+1$ -th measurement. The region of estimation are defined in this way: if the sequence ends in  $z_M = 1$ , it is attributed to  $y = M$ , otherwise for  $z_M = 0$  it is assigned to  $y = m(s_{M-1}(\bar{z}_{M-1}))$ .

#### D. Adaptive Receiver Algorithm with Suboptimal Local Measurements

The adaptive receiver algorithm we proposed in Section III bases its performance on two key ingredients, the possibility of implementing arbitrary local projective measurements and the dynamic programming algorithm that optimize subsequent measurements on the base of previous results.

As already pointed out in Section II B, an arbitrary projective measurement can be implemented with the Dolinar receiver setup. However, it is well known that the implementation of this receiver is challenging due to the requirements of a fast feedback from the photon counter to the local oscillator during the time slot, and suboptimal discrimination schemes results to be more practical in experimental setup.

The dynamic programming algorithm can be set up with any local measurement scheme. In this section we consider the definition of an adaptive receiver scheme that uses a Generalized Kennedy receiver [15] in each

slot for the local binary discrimination, but leveraging on the dynamic programming to optimize the sequence of measurements. The resulting algorithm can again be represented with a tree as in Figure 3, but with the parameter  $\beta_k = \pi_k(\bar{z}_k)$  instead of  $\phi_k$ , representing the time constant displacement to apply in the slot.

The transition probabilities of a Generalized Kennedy scheme are reported in equations (28) with  $\beta = \beta_k$  and  $G = 1$ . In the framework setup, this transition probabilities are obtained from a pair of POVM, whose parameters can be obtained by inversion of (13).

Since the two expressions are not symmetric, we may consider two different associations between the outcomes  $\{0, 1\}$  and the qubits  $\{|\xi_0\rangle, |\xi_1\rangle\}$ . The association

$$\begin{cases} p_{0|0} = q_{0|0} \\ p_{1|1} = q_{1|\alpha} \end{cases} \quad (91)$$

comes directly from the definition of the Generalized Kennedy receiver. The other association,

$$\begin{cases} p_{0|0} = q_{1|\alpha} \\ p_{1|1} = q_{0|0} \end{cases} \quad (92)$$

can be obtained placing a displacement operation  $D(-\alpha)$  before the receiver, therefore inverting the role of the ground state and the non-nulled coherent state.

Replacing the local measurements leads to minor changes in the dynamic programming algorithm. The reward-to-go function  $\tilde{J}_{M-1}^*$  and the optimized parameter  $\pi_{M-1}^*$  for the last measurement are evaluated as

$$\tilde{J}_{M-1}^*(s_{M-1}) = \max_{\beta_{M-1}} \tilde{J}_{M-1}(s_{M-1}, \beta_{M-1}), \quad (93)$$

$$\pi_{M-1}^*(s_{M-1}) = \operatorname{argmax}_{\beta_{M-1}} \tilde{J}_{M-1}(s_{M-1}, \beta_{M-1}), \quad (94)$$

with

$$\begin{aligned} \tilde{J}_{M-1}(s_{M-1}, \beta_{M-1}) &= p_M + [(p_m + p_M) \sinh \alpha \beta_{M-1} \\ &+ (p_m - p_M) \cosh \alpha \beta_{M-1}] e^{-\beta_{M-1}^2 - \frac{|\alpha|^2}{4}}. \end{aligned} \quad (95)$$

The transition probabilities  $\{\hat{p}_{0|0}, \hat{p}_{1|1}\}$  of the projective measurements need to be replaced with  $\{p_{0|0}, p_{1|1}\}$ . Both the possible associations (91) and (92) must be considered in the update equations (79)-(82), as well as in the optimization of the first measurement in equations (86) and (87).

This adaptive receiver improves the performance of the Improved Conditional Nulling type I scheme due to the more generality obtained by the dynamic programming. On the other hand, the performance will be suboptimal since a suboptimal scheme is used for the local measurements.

## VI. RESULTS AND NUMERICAL ISSUES

In the previous Section we described the algorithm to optimize the sequences of parameters  $\phi_k$  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  $\chi$ . 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).

A first result is that in the case of  $M = 2$  the adaptive receiver with local projective measurements reaches the theoretical quantum limit. This is not surprising, because as already pointed out in [17] 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 the performances of the classical receiver, conditional nulling, type I and type II schemes are compared with respect to the theoretical limit. The performance of the adaptive receivers surpasses these schemes, and in the case of the adaptive receiver with projective measurements the performance overlaps with the theoretical one.

As the cardinality  $M$  increases, the performance of the adaptive receiver with local projective measurements slightly moves away from the theoretical optimum. In Figure 5, 6 and 7 the performance of the existing and adaptive receivers are compared for  $M = 3$ ,  $M = 4$  and  $M = 8$  respectively. The trend is the same in all the figures: the adaptive schemes outperform the existing conditional nulling, type I and type II receiver, placing the error probability curves between these and the theoretical limit. The adaptive schemes maintain 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 gets 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  $\tilde{J}_k^*$  for all possible system states  $s_k$  may require a non trivial amount of computational time and memory. In addition, this evaluation must be repeated from  $k = M - 1$

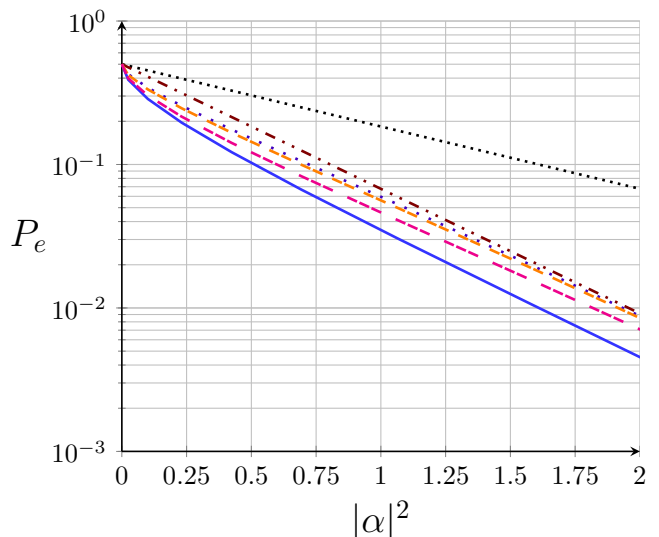


FIG. 4: (Color online) Performances comparison of different receiver schemes, for 2-PPM. The curves, from top to bottom, correspond to:  $\cdots$  classical direct detection (black),  $-\cdot-\cdot-$  conditional nulling receiver (red),  $\cdots$  type I improved conditional nulling scheme (violet),  $-\cdot-\cdot-$  type II improved conditional nulling scheme (orange),  $---$  adaptive scheme with local Generalized Kennedy measurements (magenta),  $---$  quantum theoretical limit and adaptive with projective measurements (overlapped, blue).

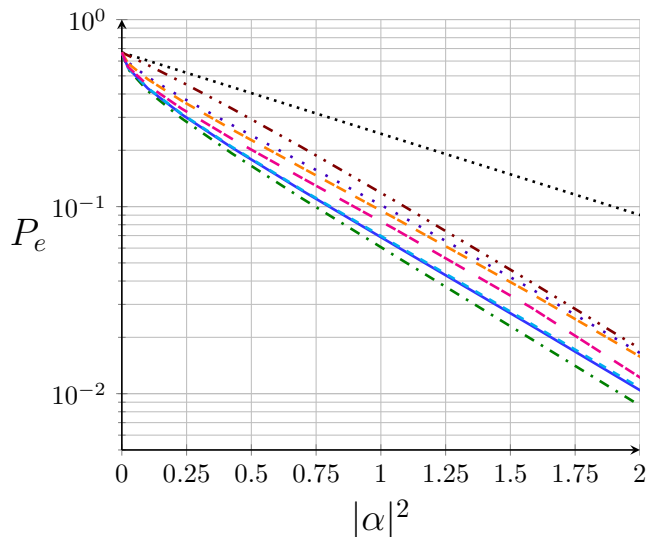


FIG. 5: (Color online) Performances comparison of different receiver schemes, for 3-PPM. The curves, from top to bottom, correspond respectively to:  $\cdots$  classical direct detection (black),  $-\cdot-\cdot-$  conditional nulling receiver (red),  $\cdots$  type I improved conditional nulling scheme (violet),  $-\cdot-\cdot-$  type II improved conditional nulling scheme (orange),  $---$  adaptive scheme with local Generalized Kennedy measurements (magenta),  $---$  retraced forward adaptive scheme with projective measurements (cyan),  $---$  adaptive with projective measurements (blue),  $---$  quantum theoretical limit (green).

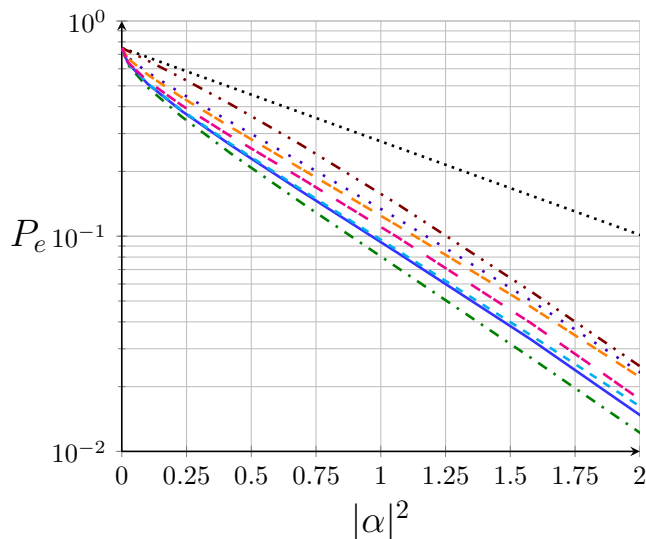


FIG. 6: (Color online) Performances comparison of different receiver schemes, for 4-PPM. The curves, from top to bottom, correspond respectively to:  $\cdots$  classical direct detection (black),  $-\cdot-$  conditional nulling receiver (red),  $\cdots$  type I improved conditional nulling scheme (violet),  $-\cdot-$  type II improved conditional nulling scheme (orange),  $---$  adaptive scheme with local Generalized Kennedy measurements (magenta),  $---$  retraced forward adaptive scheme with projective measurements (cyan),  $---$  adaptive with projective measurements (blue),  $-\cdot-$  quantum theoretical limit (green).

down to  $k = 1$ .

Since a numerical procedure is required to evaluate  $\tilde{J}_k^*$  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 parameters  $\pi_k^*$  in (84) and the evaluation of  $\tilde{J}_k^*$  in (83) requires to approximate the system state  $s_{k+1}$  in the grid when considering  $\tilde{J}_{k+1}^*(s_{k+1})$ . The issue of this approximation spread out in successive evaluations of  $\tilde{J}_k^*$ , 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 (63).

Some considerations can be done in order to lighten the computation. The first consideration is that for different cardinality  $M$ , the sequence of tables  $\tilde{J}_k^*$  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  $\tilde{J}_{\tilde{M}}^*, \tilde{J}_{\tilde{M}-1}^*, \dots, \tilde{J}_1^*$  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

$$s_0 = \left( \frac{1}{M_i}, \frac{1}{M_i}, \cdot \right) \quad (96)$$

end evaluate the probability of correct decision as

$$P_c = \max_{\phi} \tilde{J}_{M-M_i+1}^* \left( \frac{\hat{p}_{0|1}(\phi)}{M_i}, \frac{\hat{p}_{0|0}(\phi)}{M_i}, 1 \right) + \tilde{J}_{M-M_i+1}^* \left( \frac{\hat{p}_{1|1}(\phi)}{M_i}, \frac{\hat{p}_{1|0}(\phi)}{M_i}, 1 \right) \quad (97)$$

In addition, as already pointed out and proved in Lemma 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  $\tilde{J}_k^*$ , 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

$$\left\{ (u, v), u + v \leq 1, 0 \leq u \leq 1, 0 \leq v \leq \frac{1}{M-k+1} \right\}. \quad (98)$$

Furthermore, if we are interested in the performance for a single value of the cardinality  $M$ , we can take advantage of Lemma 4 in Appendix A. 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  $\tilde{J}_k^*$  to consider only the set

$$\left\{ (u, v), 0 \leq u \leq \frac{1}{M}, 0 \leq v \leq \frac{1}{M} \right\}. \quad (99)$$

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 parameters path for each measurement, evaluating the final joint probabilities and calculating the probability of correct decision as the sum in (55). Due to the discretization of  $S_k$  as a grid, the performance obtained retracing the parameters 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, 5, 6 and 7 in a dashed (cyan) 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 6).

In Figure 7, 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  $10^3 \times 10^3$  samples.

## VII. 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

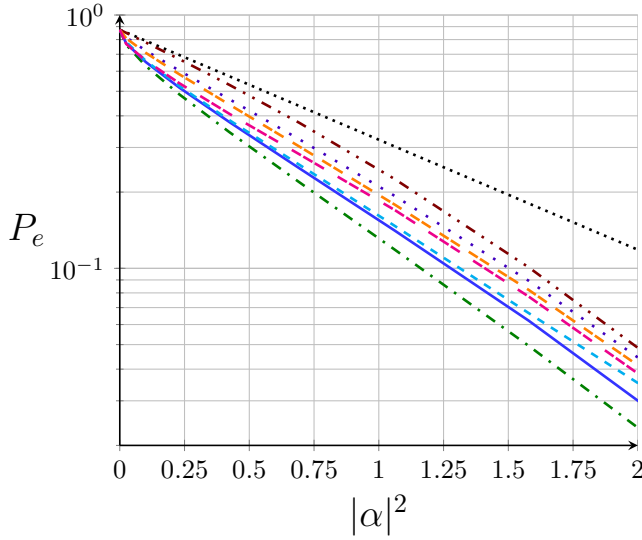


FIG. 7: (Color online) Performances comparison of different receiver schemes, for 8-PPM. The curves, from top to bottom, correspond respectively to:  $\cdots$  classical direct detection (black),  $-\cdot-$  conditional nulling receiver (red),  $\cdots$  type I improved conditional nulling scheme (violet),  $-\cdot-$  type II improved conditional nulling scheme (orange),  $---$  adaptive scheme with local Generalized Kennedy measurements (magenta),  $---$  retraced forward adaptive scheme with projective measurements (cyan),  $---$  adaptive with projective measurements (blue),  $-\cdot-$  quantum theoretical limit (green).

is then reformulated as a sequence of shorter measurements, one in each slot, that allows to design adaptive receiver 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. We consider both (optimal) projective and (suboptimal) Generalized Kennedy local measurements 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.

This work has been carried out within the Strategic-Research-Project QUINTET of the Department of Information Engineering, University of Padova and the Strategic-Research-Project QUANTUMFUTURE of the

University of Padova.

## Appendix A: Useful Lemmas

In this Appendix we give some useful Lemmas that helps to better understand the optimization algorithm.

**Lemma 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  $|\xi_0\rangle, |\xi_1\rangle$  respectively. The maximum probability of correct discrimination between  $i = 0$  and  $i = 1$  achievable with measurement operators (12) defined by  $\phi$  is*

$$\frac{1}{2} \left[ p_0 + p_1 + \sqrt{(p_0 + p_1)^2 - 4p_0p_1\chi} \right], \quad (\text{A1})$$

and the optimal angle is defined by

$$\phi^* = \frac{1}{2} \sqrt{\cos 2\theta(p_0 - p_1) + j \sin 2\theta(p_0 + p_1)}. \quad (\text{A2})$$

**Proof.** Consider the measurement operators (12), and without loss of generality associate the outcomes  $z = 0$  with the estimation of  $i = 0$  and  $z = 1$  with  $i = 1$ . The probability of correct discrimination can be written as

$$\begin{aligned} P_c &= P[z = 0|i = 0]p_0 + P[z = 1|i = 1]p_1 \\ &= \cos^2(\theta - \phi)p_0 + \sin^2(\theta + \phi)p_1 \\ &= \frac{1 + \cos(2\theta - 2\phi)}{2}p_0 + \frac{1 - \cos(2\theta + 2\phi)}{2}p_1 \end{aligned} \quad (\text{A3})$$

with transition probabilities defined by (14). The maximization of (A3) with respect to the angle  $\phi$  leads to the relation

$$\tan 2\phi = \tan 2\theta \frac{p_0 + p_1}{p_0 - p_1}, \quad (\text{A4})$$

solved by the angles verifying

$$\begin{aligned} \sin 2\phi &= \pm \frac{p_0 + p_1}{\sqrt{R}} \sin 2\theta, \\ \cos 2\phi &= \pm \frac{p_0 - p_1}{\sqrt{R}} \cos 2\theta, \end{aligned} \quad (\text{A5})$$

with  $R$  a normalization term,

$$R = \cos^2 2\theta(p_0 - p_1)^2 + \sin^2 2\theta(p_0 + p_1)^2. \quad (\text{A6})$$

Expression (A5) with the plus sign corresponds to the point of maximum, and the thesis (A2) follows. Substituting (A2) in (A3) gives

$$\begin{aligned} P_c &= \frac{1}{2} \left[ p_0 + p_1 + \sqrt{R} \right] \\ &= \frac{1}{2} \left[ p_0 + p_1 + \sqrt{(p_0 + p_1)^2 - 4p_0p_1 \cos^2 2\theta} \right]. \end{aligned} \quad (\text{A7})$$

□

**Corollary 1** *If the quantum states associated to the two symbols  $i = 0, 1$  is the same, i.e.  $|\xi_0\rangle = |\xi_1\rangle$ , the maximal probability of correct discrimination between  $i = 0$  and  $i = 1$  is the maximum of their a priori probability,*

$$P_c = \max\{p_0, p_1\}. \quad (\text{A8})$$

The Corollary follows from  $\chi = |\langle \xi_0 | \xi_1 \rangle|^2 = 1$ . 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 (A1) is always non lower than (A8), 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 (A1) 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 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  $|\xi_0\rangle$  in the  $k$ -th position of  $|\bar{\gamma}_i\rangle, |\bar{\gamma}_j\rangle$ . Hence, the transition probability is the same,

$$P[z_k|i] = |\langle \mu_{z_k} | \xi_0 \rangle|^2 = P[z_k|j] \quad (\text{A9})$$

and the joint probabilities are multiplied by the same factor

$$p_{i, z_k} = |\langle \mu_{z_k} | \xi_0 \rangle|^2 p_{i, z_{k-1}}, \quad p_{j, z_k} = |\langle \mu_{z_k} | \xi_0 \rangle|^2 p_{j, z_{k-1}}, \quad (\text{A10})$$

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  $|\xi_0\rangle$  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 with respect to a measurement in this slot. The  $k$ -th slot can give information only for the the discrimination of the  $k$ -th symbol, whether it is more likely or not respect the other.

**Corollary 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  $|\xi_0\rangle$ , so the joint probabilities after measurement  $l > k$  is

$$p_{i, [\bar{z}_k z_{k+1} \dots z_l]} = P[z_l | z_{l-1}, \dots, \bar{z}_k, |\xi_0\rangle] \cdot \dots \cdot P[z_{k+1} | \bar{z}_k, |\xi_0\rangle] p_{i, \bar{z}_k} \quad (\text{A11})$$

that is, all the joint probabilities are multiplied by the same factors.  $\square$

**Lemma 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

$$p_{x, \bar{z}_l} = P[z_l | \bar{z}_{l-1}, x] P[z_{l-1} | \bar{z}_{l-2}, x] \cdot \dots \cdot P[z_1 | x] \frac{1}{M}, \quad (\text{A12})$$

but since each symbol  $x = k + 1, \dots, M$  has a quantum state  $|\xi_0\rangle$  in position  $1, \dots, l$ ,  $l \leq k$ , the conditional probabilities give

$$p_{x, \bar{z}_l} = |\langle \mu_{\bar{z}_l} | \xi_0 \rangle|^2 |\langle \mu_{\bar{z}_{l-1}} | \xi_0 \rangle|^2 \cdot \dots \cdot |\langle \mu_{\bar{z}_1} | \xi_0 \rangle|^2 \frac{1}{M}, \quad (\text{A13})$$

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  $x = k + 1, k + 2, \dots, M$  have the quantum state  $|\xi_0\rangle$  in position  $1, \dots, k$ , 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** *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

$$p_{i, \bar{z}_k} = P[z_k | \bar{z}_{k-1}, i] P[z_{k-1} | \bar{z}_{k-2}, i] \cdot \dots \cdot P[z_1 | i] \frac{1}{M} \quad (\text{A14})$$

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  $p_{i, \bar{z}_k} \leq \frac{1}{M}$ .  $\square$

- 
- [1] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, *Reviews of Modern Physics* **84**, 621 (2012).
- [2] V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, J. H. Shapiro, and H. P. Yuen, *Physical Review Letters* **92**, 027902 (2004).
- [3] S. J. Dolinar, Research Laboratory of Electronics, MIT Quarterly Progress Report **111**, 115 (1973).
- [4] R. L. Cook, P. J. Martin, and J. M. Geremia, *Nature* **446**, 774 (2007).
- [5] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
- [6] S. J. Dolinar, *The Telecommunications and Data Acquisition Progress Report* **42-72**, 30 (1983).
- [7] S. Guha, J. L. Habif, and M. Takeoka, *Journal of Modern Optics* **58**, 257 (2011).
- [8] J. Chen, J. L. Habif, Z. Dutton, R. Lazarus, and S. Guha, *Nature Photonics* **6**, 374 (2012).
- [9] H. P. Yuen, R. S. Kennedy, and M. Lax, *IEEE Transactions on Information Theory* **21**, 125 (1975).
- [10] G. Cariolaro and G. Pierobon, *IEEE Transactions on Communications* **58**, 1213 (2010).
- [11] N. Tomassoni and M. G. Paris, *Physics Letters A* **373**, 61 (2008).
- [12] N. Dalla Pozza, N. Laurenti, and F. Ticozzi, p. 13 (2013), arXiv:quant-ph/1307.0014.
- [13] A. S. Holevo, *Soviet Math. (Iz. VUZ)* **26**, 1 (1982).
- [14] M. Takeoka, M. Sasaki, P. van Loock, and N. Lütkenhaus, *Physical Review A* **71**, 022318 (2005).
- [15] M. Takeoka and M. Sasaki, *Physical Review A* **78**, 022320 (2008).
- [16] D. P. Bertsekas, *Dynamic Programming and Optimal Control, Vol. I-II* (Athena Scientific, 2007), 3rd ed.
- [17] A. Acín, E. Bagan, M. Baig, L. Masanes, and R. Muñoz-Tapia, *Physical Review A* **71**, 032338 (2005).
- [18] The function  $\phi = \angle a + jb$  gives the argument of the complex number  $a + jb = \sqrt{a^2 + b^2} e^{j\phi}$ , with  $j = \sqrt{-1}$  the imaginary unit.
- [19] A different notation is used in Figure 3, as indicated in the caption.
- [20] Dynamic 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.