






Agnostic Dolinar receiver for coherent-state classificationFabio Zoratti  and Nicola Dalla Pozza 
*Scuola Normale Superiore, I-56126 Pisa, Italy*Marco Fanizza  and Vittorio Giovannetti 
NEST, Scuola Normale Superiore and Istituto Nanoscienze-CNR, I-56126 Pisa, Italy (Received 22 June 2021; accepted 23 September 2021; published 12 October 2021)

We consider the problem of discriminating quantum states, where the task is to distinguish two different quantum states with a complete classical knowledge about them, and the problem of classifying quantum states, where the task is to distinguish two classes of quantum states where no prior classical information is available but a finite number of physical copies of each class is given. In the case the quantum states are represented by coherent states of light, we identify intermediate scenarios where partial prior information is available. We evaluate an analytical expression for the minimum error when the quantum states are opposite and a prior on the amplitudes is known. Such a threshold is attained by complex positive operator-valued measurements that involve a highly nonlinear optical procedure. A suboptimal procedure that can be implemented with current technology is presented that is based on a modification of the conventional Dolinar receiver. We study and compare the performance of the scheme under different assumptions on the prior information available.

DOI: [10.1103/PhysRevA.104.042606](https://doi.org/10.1103/PhysRevA.104.042606)**I. INTRODUCTION**

Since its early studies, the discrimination of quantum states has been a central problem in quantum information theory due to the impossibility of perfectly distinguishing non-orthogonal quantum states. Its implications reflect not only on quantum communication scenarios [1–3], but also in metrology [4–7], sensing [8,9], quantum key distribution, and cryptography [10–14].

In the typical *discrimination* scenario [15–18], two parties, the transmitter and the receiver, agree on a known shared communication protocol, which defines the (possibly finite) set of quantum states to transmit and discriminate in the best way possible [19,20]. A different perspective on the problem has been adopted in recent years, following the growing trend of machine learning studies [21–28]. In the context of supervised learning, *classification* problems aim at assigning a sample to one of the available classes, of which a description is not known but multiple training samples are provided. Training samples could be copies of the quantum states to classify or other members of the family defining the classes. Classification problems are more general since the description provided by the communication protocol allows the generation of the training samples, therefore expressing a discrimination problem as a classification one. The latter is also more difficult since the classifier has to learn a description or a strategy for the discrimination from the (possibly noisy) training samples, in addition to performing the distinction.

Historically, the discrimination scenario has been investigated the most. Minimum error discrimination has been considered initially for two quantum states [19,20], where the optimal solution for the measurement operators assigning

the estimate has been given in a closed form. Optimality conditions for a bigger set of quantum states have been found [29], but the evaluation of the measurement operators and of the performance usually requires numerical procedures such as semidefinite programming [30]. When the quantum states exhibit symmetries, such evaluation can be further simplified [31–33]. Despite the advances in the field, in the case of the discrimination of optical states, physical realizations of the optimal receiver end are still an open problem, with the only exception of the Dolinar receiver for the discrimination of two coherent states [34–39]. Along with this scheme, other practical realizations of suboptimal receivers have been proposed for other sets of coherent states [40–44].

Regarding the quantum classification problem, early works frame the same scenario under different names, such as quantum matching (see Ref. [25] and references within) or quantum state identification [26], and programmable discrimination [45]. In the minimum error setting, solutions for the two-class problem came first for pure states [26,27] and then for general qubit mixed states [23,46,47], in the asymptotic and limited-training-samples regimes. Following papers [23,48,49] have focused on the performance comparison between a joint (collective) measurement strategy involving both the training samples and the one to distinguish, versus an estimate-and-discriminate strategy, where the training copies are used to estimate the classes of states, and the classical information extracted is used to set up the discrimination. The latter strategy results to be suboptimal to the former. The unambiguous version of the classification problem considered in this paper has been addressed in Refs. [50,51], which provided a strategy based on interferometers and photodetectors, that has been demonstrated experimentally [39]. Other results on

programmable discriminators can be found in Refs. [52–61]. An approach to programmable discriminators for coherent states based on reinforcement learning using passive optical elements, photodetectors, and classical adaptive control has been investigated recently in Ref. [62].

In the field of quantum optics, discrimination and classification problems have been formulated for the reading of an optical memory. The advantage of using quantum states of light for the discrimination has been established in a series of papers [63–67], while in Ref. [68] the reading has been framed as a classification problem between the vacuum state and a coherent state with unknown parameter, later set up to be a Gaussian *a priori* distribution around a mean value. The asymptotic behavior of the collective measurement strategy and the estimate-and-discriminate one has been evaluated and compared, confirming that the former gives better performances. In Ref. [68] it is also conjectured and given some evidence that this holds for *non-Gaussian* estimate-and-discriminate strategies.

Our work further investigates this comparison. We consider the simplest scenario concerning the classification of an unknown state belonging to one of two classes of coherent states that are assigned by giving access to a certain number of training copies. A practical scenario where this task could be relevant is an optical link through a stochastic channel where the attenuation is so unpredictable and random that, over sufficiently long time intervals, the signal intensity can be assumed to be completely unknown, and the same holds for the added phase. In this context one may try to exploit the existence of stability periods in the perturbations induced by the noise, to set communication protocols that consist in sending samples of the two types of training signals followed by the quantum state to classify. To begin with we show that, via simple linear optics, the problem we are facing can always be reduced to the special symmetric case where the two classes of inputs differs only by the sign of the associated coherent amplitudes. We hence evaluate the optimal bound for the probability of success of the classification task, under the assumption that the protocol to use is phase invariant. Such threshold can be explicitly calculated for any number of input copies of the training states, and for any given prior distribution of the coherent amplitudes that characterize them. Unfortunately the positive operator-valued measurement (POVM) that ensures the attainability of the optimal bound relies on highly nonlinear optical processes that are not feasible with current technology. In the alternative we propose a modification of the conventional Dolinar receiver [34] that we dub an agnostic Dolinar receiver, whose implementation is instead at reach with conventional quantum optical procedures. While being suboptimal when employed with a finite number n of training copies, in the proposed scheme it is shown to saturate to the optimal bound in the asymptotic limit $n \rightarrow \infty$. Most importantly, for all n , it yields a clear advantage when compared with respect to simple estimate-and-discriminate strategies that involve estimations performed on a fraction of the training samples.

The paper is organized as follows. In Sec. II we introduce the problem and review the original Dolinar receiver. Section III is the main section of the paper and includes all the original results of our work. Here we reduce the problem

to a symmetric scenario, provide an optimal bound for the problem, and present our apparatus. We study its performance by comparing it with an estimate-and-discriminate strategy based on a miscalibrated Dolinar scheme, and also compute the optimal error probability of the problem under different assumptions on the prior information available.

II. DISCRIMINATION AND CLASSIFICATION OF COHERENT OPTICAL SIGNALS

This section is dedicated to set the problem, introduce the notation, and review some basic facts.

A. Discrimination versus classification

The discrimination and classification of quantum states are two distinct primitives of quantum information processing that find applications in a variety of different contexts. Relying on the error probability as cost function to evaluate their efficiencies [19], we schematize these procedures in terms of the following minimum error discrimination (MED) and minimum error classification (MEC) problems.

MED problem. Given a set of known quantum states $\{\hat{\rho}_k\}_{k=1}^K$ and probabilities $\{p_k\}_{k=1}^K$, $\sum_k p_k = 1$ and an unknown quantum state $\hat{\rho} \in \{\hat{\rho}_k\}$ drawn from the set with probability $p \in \{p_k\}$, find the POVM operators $\{\hat{\Pi}_k\}_{k=1}^K$ that allow to identify ρ with minimum probability of error, $P_e^{(\text{MED})}$, or equivalently, with maximum probability of correct decision, $P_c^{(\text{MED})} = 1 - P_e^{(\text{MED})}$,

$$P_c^{(\text{MED})} = \sum_{k=1}^K p_k \text{tr}[\hat{\Pi}_k \hat{\rho}_k]. \quad (1)$$

It is worth reminding that for the special case with $K = 2$ an explicit solution for the MED problem is provided by the Helstrom theorem according to which the maximum value of $P_c^{(\text{MED})}$ is achieved by a binary projective measurement associated with positive and negative parts of the operator $p_1 \hat{\rho}_1 - p_2 \hat{\rho}_2$, leading to the optimal expression

$$P_{c,\text{max}}^{(\text{MED})} = \frac{1}{2}(1 + \text{tr}|p_1 \hat{\rho}_1 - p_2 \hat{\rho}_2|), \quad (2)$$

which we employ in the following as a benchmark for the efficiency of our schemes.

MEC problem. Given a training set of quantum states $\{\hat{\rho}_k\}_{k=1}^n$ and a set of labels $\{y(k) \mid y(k) \in [1, \dots, K]\}_{k=1}^n$ that associate each sample to its class $y(k)$, and given an unknown testing set of quantum states $\{\hat{\rho}_r\}_{r=1}^s$ and labels $\{z(r) \mid z(r) \in [1, \dots, K]\}_{r=1}^s$, find the POVM operators $\{\hat{\Pi}_z\}_{z=1}^K$ that identify their classes with minimum error probability $P_e^{(\text{MEC})}$, or equivalently with maximum probability of correct decision, $P_c^{(\text{MEC})} = 1 - P_e^{(\text{MEC})}$,

$$P_c^{(\text{MEC})} = \frac{1}{s} \sum_{r=1}^s \text{tr}[\hat{\Pi}_{z(r)} \hat{\rho}_r]. \quad (3)$$

Notice that, at variance with the MED problem, for the MEC problem a (complete classical) description of the quantum states of the training set $\{\hat{\rho}_k\}_{k=1}^n$ is not known, nor is the *a priori* probability of each class which is often inferred from the relative amount of the labels. Notice also that, even though in the subsequent sections of our paper we do not adopt

such options, in the general MEC setting (i) the samples of the training set are not necessarily organized into cluster of identical copies, and (ii) the testing samples $\{\hat{\rho}_r\}_{r=1}^s$ are not included in the training set $\{\hat{\rho}_k\}_{k=1}^n$.

B. Quantum optical setting

In the rest of the paper we consider the case where the system of interest is a single optimal mode of the electromagnetic field described by the annihilation and creation operators \hat{a} , \hat{a}^\dagger fulfilling canonical commutation rules, $[\hat{a}, \hat{a}^\dagger] = 1$ [69,70]. In this context we focus on a MEC problem where the training set is formed by identical copies of $K = 2$ coherent states,

$$|\alpha_k\rangle = \hat{D}(\alpha_k)|\emptyset\rangle, \quad k \in \{1, 2\}, \quad (4)$$

whose complex amplitudes α_1 and α_2 are unknown (in this expression $|\emptyset\rangle$ and $\hat{D}(\alpha) = \exp[\alpha\hat{a}^\dagger - \alpha^*\hat{a}]$ are respectively the vacuum state and the displacement operator of the model). Specifically we work under the assumption of having n copies of each training state, and that the density matrix we need to classify is guaranteed to be a coherent state $|\delta\rangle$ that coincides either with $|\alpha_1\rangle$ or with $|\alpha_2\rangle$, with flat prior probabilities (i.e., $p_1 = p_2 = 1/2$). The resulting global input state we operate can hence be expressed in the following multi-mode compact form:

$$|\alpha_1^{\otimes n}, \alpha_2^{\otimes n}, \delta\rangle = |\alpha_1\rangle^{\otimes n} \otimes |\alpha_2\rangle^{\otimes n} \otimes |\delta\rangle, \quad (5)$$

which in principle is characterized by four unknown real parameters (the complex numbers α_1 and α_2), and by one quantum binary variable (the testing state $|\delta\rangle \in \{|\alpha_1\rangle, |\alpha_2\rangle\}$). In the MED version of the problem, i.e., when the complex quantities α_1, α_2 are assigned or equivalently when the number n of the training copies of the MEC problem are infinitely many so that the values of the amplitudes can be recovered through quantum process tomography, the optimal success probability can be computed as in Eq. (2), leading to the value

$$\begin{aligned} P_{c,\max}^{(\text{MED})} &= \frac{1}{2}(1 + \sqrt{1 - 4p_1p_2|\langle\alpha_1|\alpha_2\rangle|^2}) \\ &= \frac{1}{2}(1 + \sqrt{1 - 4p_1p_2 e^{-|\alpha_1 - \alpha_2|^2}}), \end{aligned} \quad (6)$$

which can be attained via the Dolinar detection scheme which we review in the next section. The procedure we have in mind to solve the MEC problem for finite n is a variation of the scheme that relies on basic linear optical manipulations of the state $|\alpha_1^{\otimes n}, \alpha_2^{\otimes n}, \delta\rangle$ to compensate for the absence of classical information on the values of the amplitudes α_1 and α_2 . We call such a procedure an agnostic Dolinar receiver and we present it in Sec. III.

C. Dolinar receiver

As anticipated the Dolinar receiver is an experimental technique that allows one to practically attain the optimal threshold limit (6) for a binary MED problem aimed to discriminate between two assigned coherent input states $|\alpha_1\rangle$ and $|\alpha_2\rangle$ which are produced with prior probabilities p_1 and p_2 . It is worth noticing that in this special context, due to the fact that the values of the complex amplitudes α_1 and α_2 are known, one can always reduce the problem to the case of a symmetric configuration in which $|\alpha_1\rangle, |\alpha_2\rangle$ are traded with

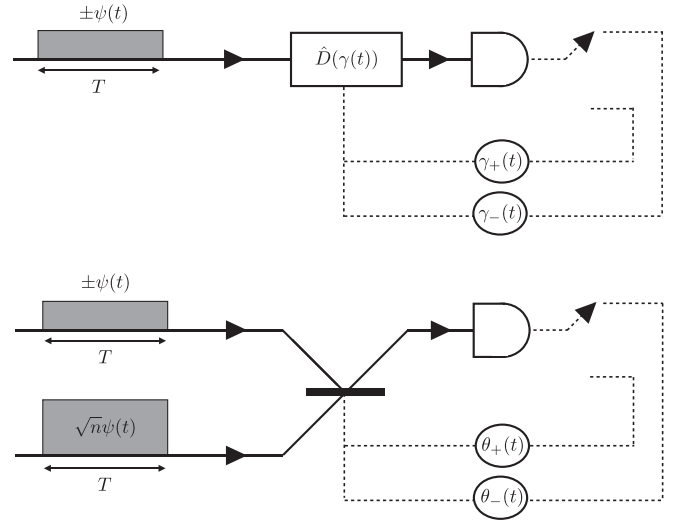


FIG. 1. Time-continuous description of the conventional Dolinar receiver (upper panel) and the agnostic Dolinar receiver (lower panel): solid black lines represent optical signals, and dashed lines represent instead classical control lines. In the upper panel, the rectangle with the $\hat{D}(\gamma(t))$ symbol represents a time-dependent displacement gate, which is followed by a photon counter that switches between two classically controlled quantities represented in the circles. In the lower panel, the crossing with an extra horizontal rectangle is a beam splitter with time-dependent reflectivity $\theta_k(t) \in \{\theta_+(t), \theta_-(t)\}$, which is classically controlled and selected by the photon counter.

the couple $|\pm\bar{\alpha}\rangle$ with $\bar{\alpha} = (\alpha_1 - \alpha_2)/2$, or equivalently to a maximally energetically unbalanced setting where instead $|\alpha_1\rangle$ and $|\alpha_2\rangle$ get replaced by $|2\bar{\alpha}\rangle$ and $|\emptyset\rangle$, respectively. Such mappings in fact simply rely on acting on the input $|\delta\rangle$ via optical displacements, i.e., transformations that can be physically implemented by mixing the signal with an intense coherent ancillary state through a beam splitter of high transmissivity. Specifically in the cases we are considering this accounts in replacing $|\delta\rangle$ with

$$\hat{D}(-\beta)|\delta\rangle = e^{(\alpha\beta^* - \alpha^*\beta)/2}|\delta - \beta\rangle, \quad (7)$$

with $\beta = (\alpha_1 + \alpha_2)/2$ for the symmetric configuration, and $\beta = -\alpha_2$ for the maximally energetically unbalanced setting. In view of these facts in the following paragraphs, without loss of generality we assume the symmetric setting posing $\alpha_1 = -\alpha_2 = \alpha$. We also point out that in our presentation of the Dolinar scheme we rely on the continuous-time formulation of the problem discussed in Ref. [37] (see, however, Appendix A for a description based on the sequence of beam splitters and photon detectors).

The Dolinar receiver works by continuously applying a displacement $\hat{D}(\gamma_k(t))$ on the input state and performing a photon counting on the displaced signal (see upper panel of Fig. 1). The displacement is optimized such that the parity of the counting at the end of the signal gives the final estimate of the input coherent state. The rationale of the scheme follows the same idea of another suboptimal scheme, the Kennedy receiver [71]. The key idea behind the Kennedy receiver is to rigidly shift the two states by α to obtain the

pair $\{|2\alpha\rangle, |0\rangle\}$ (i.e., to map the symmetric setting into the maximally energetically unbalanced one), and then perform photon counting. The vacuum always counts zero photons, and thus the Kennedy receiver uses $\hat{\Pi}_+ = \mathbb{I} - |\emptyset\rangle\langle\emptyset|$, $\hat{\Pi}_- = |\emptyset\rangle\langle\emptyset|$ as the POVM for the discrimination. This apparatus does not reach the optimal Helstrom bound (6), and this gap is filled exactly by the Dolinar receiver [34,36,37], where the optimal shift is defined from an optimization procedure. More precisely, consider the input field for our system $\psi_k(t)$, $0 < t < T$, corresponding to the coherent state $|\pm\alpha\rangle$, being represented by

$$\psi_k(t) = \pm\psi e^{i\omega_0 t} \quad (k = \pm), \quad (8)$$

with ω_0 the optical pulsation frequency and T the pulse duration, that are linked through the normalization condition

$$|\alpha|^2 = \int_0^T |\psi(t)|^2 dt = |\psi_k|^2 T. \quad (9)$$

It can be shown that all the following results do not depend on our choice of T and thus, for clarity reasons, we set $T = 1$ and $\psi = \alpha$. This input signal is displaced by a value $\gamma_k(t)$, which is classically controlled via feedback, and the resulting sum signal is monitored with a photon counter. Every time our counter “clicks,” the added signal is discontinuously changed from $\gamma_+(t)$ to $\gamma_-(t)$ and vice versa, which is the displacement employed when the provisional estimation for the quantum state is $z(t) = +$ or $z(t) = -$ corresponding to the sign of $|\pm\alpha\rangle$. Equivalently, the parity of the total number of photons counted until the current time t gives the estimation for the quantum state. The final discrimination result is declared at the end of the signal, $t = T = 1$.

We can use the following argument to find the optimal choice for $\gamma_k(t)$. The provisional correct decision probability at time t can be written as

$$P_c(t) = P[z(t) = +|k = +]p_+ + P[z(t) = -|k = -]p_-. \quad (10)$$

Let us now assume that $k = +$, that is, the actual coherent state to discriminate is represented by the input field $\psi_+(t)$ of Eq. (8). Then, the process $z(t)$ can be interpreted as a telegraph process driven by nonhomogeneous Poisson processes [37] with rates

$$\begin{aligned} \lambda(t) &= |\psi_+(t) - \gamma_+(t)|^2, \\ \mu(t) &= |\psi_+(t) - \gamma_-(t)|^2, \end{aligned} \quad (11)$$

that allows the evaluation of the differential equations for the conditional probabilities of correct decision $q_+(t) = P[z(t) = +|k = +]$ and $q_-(t) = P[z(t) = -|k = -]$ as

$$\begin{aligned} \frac{dq_+(t)}{dt} &= \mu(t) - [\lambda(t) + \mu(t)]q_+(t), \\ \frac{dq_-(t)}{dt} &= \mu(t) - [\lambda(t) + \mu(t)]q_-(t). \end{aligned} \quad (12)$$

Hence, the differential equation for the correct detection probability results is

$$\frac{dP_c(t)}{dt} = \frac{q'_+ + q'_-}{2} = \mu(t) - [\lambda(t) + \mu(t)]P_c(t). \quad (13)$$

We can now extremize with respect to $\gamma_+(t) = -\gamma_-(t)$ at each fixed time to find the optimal displacement, which leads to the differential equation

$$\begin{aligned} \frac{dP_c(t)}{dt} &= -4|\alpha|^2 \frac{P_c(t)(1 - P_c(t))}{1 - 2P_c(t)} \\ &= |\alpha|^2 \left(1 - 2P_c(t) - \frac{1}{1 - 2P_c(t)} \right) \end{aligned} \quad (14)$$

with solution

$$P_{c,\text{Dol}}^{(\text{MED})}(t) = \frac{1}{2} \left(1 + \sqrt{1 - 4p_+p_-e^{-4|\alpha|^2 t}} \right), \quad (15)$$

which for $t = 1$ reaches the maximum value (6) dictated by the Helstrom bound [19]. An experimental realization of this apparatus has been realized in Ref. [38].

III. BUILDING AN AGNOSTIC DOLINAR RECEIVER

In this section we present the results of our work, which lead to a scheme for solving the MEC problem associated with a binary set of coherent inputs introduced in Sec. II B. As a preliminary step, following the discussion at the beginning of Sec. II C, in Sec. III A we show that via some trivial physical manipulations of the input data we can always restrict the analysis to the special case of a binary MEC problem where $\alpha_2 = -\alpha_1$, hence reducing from four to two the number of unknown real parameters associated with the input state, Eq. (5). In Sec. III B we find the equivalent of the Helstrom bound in Sec. II B for the MEC scenario: as we will see the attainability of such an optimal threshold relies on the possibility of implementing highly nonlinear optical processes which represent an impressive challenge for current technology. In Sec. III C we hence focus on a more realistic procedure based on an adaptive scheme where the displacements $\hat{D}(\gamma_k(t))$ of the original Dolinar receiver are replaced by partial coherent mixing with a fraction of the copies of the training set. In this context we show that if the value $|\alpha_1 - \alpha_2|$ is known (an assumption which would be trivially granted in the MED version of the problem but not in the MEC scenario where it allows to reduce the number of the unknown real parameters needed from two to one), the new setting attains a probability of success that, already for medium size values of n , approaches the one of the nonlinear optimal bound of Sec. III B. In Sec. III D we fix the issue associated with the lack of knowledge of the parameter $|\alpha_1 - \alpha_2|$ by exploiting part of the copies of the training set to obtain a preliminary estimation of such a term: the performance of the resulting scheme is hence studied and compared with those one would obtain by using a miscalibrated Dolinar scheme.

A. Mapping the MEC problem into a symmetric scenario

The aim of this section is to show that when studying the MEC problem introduced in Sec. II B, we can safely assume the amplitudes of the unknown coherent states of the training set to have opposite phases and equal absolute values (i.e., $\alpha_2 = -\alpha_1$). This simplification is analogous to the reduction of the general MED problem to a symmetric configuration; in the present case, however, this formal passage is slightly

more subtle due to the fact that we do not have prior classical knowledge of α_1 and α_2 .

A key ingredient of the analysis is represented by what we may call an m -mode *concentrator gate* $\hat{U}_C^{(m)}$ [51,68,72,73], i.e., an m -mode unitary transformation implementable via an array of properly concatenated beam splitters that, acting on a collection of m copies of a generic (possibly unknown) coherent state $|\alpha\rangle$, manages to move all their photons in a single output mode via the mapping

$$|\alpha\rangle^{\otimes m} \mapsto \hat{U}_C^{(m)} |\alpha\rangle^{\otimes m} = |\sqrt{m}\alpha\rangle \otimes |\emptyset\rangle^{\otimes m-1}. \quad (16)$$

Applying this to the $(2n+1)$ -mode input state (5) that formally defines the MEC problem we are facing, we can map it into an equivalent form where all photons are concentrated into the following three-mode coherent state:

$$|\sqrt{n}\alpha_1, \sqrt{n}\alpha_2, \delta\rangle = |\sqrt{n}\alpha_1\rangle \otimes |\sqrt{n}\alpha_2\rangle \otimes |\delta\rangle \quad (17)$$

(the net operation involves a collection of extra $2(n-1)$ irrelevant vacuum states $|\emptyset\rangle$). Notice also that with a three-port beam splitter [69] defined by the 3×3 scattering matrix

$$S_n = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{4n+2}} & \frac{1}{\sqrt{4n+2}} & -\sqrt{\frac{2n}{2n+1}} \\ \sqrt{\frac{n}{2n+1}} & \sqrt{\frac{n}{2n+1}} & \sqrt{\frac{1}{2n+1}} \end{bmatrix}, \quad (18)$$

we can then unitarily transform (17) [and hence (5)] into the further equivalent form

$$|\sqrt{2n+1}\alpha'\rangle \otimes |\delta'\rangle \otimes \left| \frac{\delta + n(\alpha_1 + \alpha_2)}{\sqrt{2n+1}} \right\rangle, \quad (19)$$

where now

$$\alpha' = \sqrt{\frac{2n}{2n+1}} \frac{\alpha_1 - \alpha_2}{2}, \quad (20)$$

and where

$$\delta' = \sqrt{\frac{2n}{2n+1}} \left(\frac{\alpha_1 + \alpha_2}{2} - \delta \right) \quad (21)$$

is a variable that for $\delta = \alpha_1, \alpha_2$ assumes the values $\pm\alpha'$. Therefore, since the coherent state $|\sqrt{2n+1}\alpha'\rangle$ can be mapped into $|\alpha'\rangle^{\otimes(2n+1)}$ via the action of the inverse of a concentrator gate $\hat{U}_C^{(2n+1)}$, our original MEC problem associated with the input (5) can be casted into the new MEC problem where starting from a collection of $2n+1$ copies of the coherent state $|\alpha'\rangle$ we are asked to decide whether the state $|\delta'\rangle$ is equal to either $|\alpha'\rangle$ or $|\alpha'\rangle$. Notice that in doing so we are implicitly neglecting the last coherent state component of Eq. (19); this, however, does not represent a huge loss of information since the residual dependence that such a term bears upon δ vanishes in the asymptotic limit of $n \rightarrow \infty$, and in any case the analysis shows that a scheme that is capable to efficiently solve a MEC in the symmetric scenario can also be applied to the generic one. Notice further that, via the action of a phase shifter gate aimed to flip the sign of the amplitude of an incoming input state, we can also convert the $2n+1$ copies of $|\alpha'\rangle$ into a state of the form $|\alpha'\rangle^{\otimes m} \otimes |\alpha'\rangle^{\otimes 2n+1-m}$ with $0 \leq m \leq 2n+1$. Relying on these observations in the remainder of the paper we will thus focus on the symmetric version of our MEC problem where starting from the beginning it is assumed $\alpha_1 = -\alpha_2 = \alpha$, hence replacing the input

state (5) with the vector

$$|\alpha^{\otimes n}, \delta\rangle = |\alpha\rangle^{\otimes n} \otimes |\delta\rangle, \quad (22)$$

characterized by two unknown real parameters (the phase and the absolute values of the complex number α), and by the quantum binary variable $|\delta\rangle \in \{|\pm\alpha\rangle\}$. [N.B., formally speaking in the above expression the total number of copies of $|\alpha\rangle$ we can extract from (5) requiring $\alpha_1 = -\alpha_2$ would be $2n$; hereafter, however, we reparametrize this with n just to allow for the possibility of having an odd number of input copies.]

B. Optimal bound for the problem

As already anticipated in the asymptotic limit $n \gg 1$, the optimal upper bound for the success probability of a generic apparatus aimed to solve the MEC problem associated with the input (22) reduces to the Helstrom limit (6) attainable via the Dolinar scheme, i.e., the quantity

$$P_{c,\max}^{(\text{MED})} = \frac{1}{2} (1 + \sqrt{1 - e^{-4|\alpha|^2}}). \quad (23)$$

Estimating the optimal MED performance in the finite-copy case can be useful to compare our results with a fundamental bound. In this section we will find this bound under the assumption that the protocol we use is phase invariant, i.e., insensitive to the phase value of the amplitude α that enters in Eq. (22), a constraint which is reasonable to impose in the MED scenario where no prior information on α is granted. For this purpose, first of all we invoke once more the action of a concentrator gate (16), to replace $|\alpha^{\otimes n}, \delta\rangle$ with a two-mode input state $|\sqrt{n}\alpha, \delta\rangle$. Then we focus on the two-element POVM $\{\hat{E}_+, \hat{E}_- = \hat{I} - \hat{E}_+\}$ which, acting globally on the two modes of the model, aims to discriminate the density matrix $\hat{\rho}_+ = |\sqrt{n}\alpha, \alpha\rangle\langle\sqrt{n}\alpha, \alpha|$ from $\hat{\rho}_- = |\sqrt{n}\alpha, -\alpha\rangle\langle\sqrt{n}\alpha, -\alpha|$ under phase-invariant assumptions. It is worth stressing that a similar calculation was performed in Ref. [68] for a slightly different setting where the two states under scrutiny were $|\sqrt{n}\alpha, \alpha\rangle$ and $|\sqrt{n}\alpha, 0\rangle$ and where the analysis was confined in large- n limit; as we will see in the following, at variance with those results, due to the symmetric structure of the inputs we employ, our analysis allows us to present closed analytical expressions also for the finite- n limit. Specifically we associate \hat{E}_+ to $\hat{\rho}_+$ and \hat{E}_- to $\hat{\rho}_-$, and we enforce the phase-invariant constraint by requiring them to commute with the global phase operator $e^{i\phi(\hat{n}_1 + \hat{n}_2)}$ with $\hat{n}_1 = \hat{a}_1^\dagger \hat{a}_1$ and $\hat{n}_2 = \hat{a}_2^\dagger \hat{a}_2$ being the number operators of the two modes of the model. By Schur's lemma [74] it then follows that the POVM elements must satisfy the identities $\hat{E}_\pm = \sum_{m=0}^{\infty} \hat{\Pi}_m \hat{E}_\pm \hat{\Pi}_m = \sum_{m=0}^{\infty} \hat{E}_{\pm,m}$, where $\hat{\Pi}_m$ is the projector on the subspace total photon number $n_1 + n_2 = m$ and $\{\hat{E}_{+,m}, \hat{E}_{-,m} = \hat{I}_m - \hat{E}_{+,m}\}$ is a binary POVM in this subspace. For any two states $\hat{\rho}_+$ and $\hat{\rho}_-$, POVMs $\{\hat{E}_\pm\}$ with this property, the probability of error satisfies

$$P_e = \frac{1}{2} \left(1 - \sum_{m=0}^{\infty} \frac{\text{tr}[\hat{E}_- \hat{\Pi}_m \hat{\rho}_+ \hat{\Pi}_m + \hat{E}_+ \hat{\Pi}_m \hat{\rho}_- \hat{\Pi}_m]}{2} \right) \geq \frac{1}{2} \left(1 - \sum_{m=0}^{\infty} \frac{\|\hat{\Pi}_m(\hat{\rho}_+ - \hat{\rho}_-)\hat{\Pi}_m\|_1}{2} \right), \quad (24)$$

where the inequality comes from the Helstrom bound. Remembering that in our case $\hat{\rho}_\pm$ are the coherent states $|\sqrt{n}\alpha\rangle\langle\sqrt{n}\alpha| \otimes |\pm\alpha\rangle\langle\pm\alpha|$ it follows that the optimal choice for \hat{E}_\pm is given by

$$\hat{E}_\pm := \sum_{m=0}^{\infty} |m; \pm\rangle\langle m; \pm|, \quad (25)$$

with

$$|m, \pm\rangle := \sum_{n_1+n_2=m} \sqrt{\binom{m}{n_1}} \frac{\sqrt{n}^{n_1} (\pm 1)^{n_2}}{\sqrt{n+1}^m} |n_1, n_2\rangle. \quad (26)$$

Accordingly as shown in Appendix B, the minimum error probability [Eq. (24)] reduces to

$$P_{e,\min}^{(\text{MEC},n)} = \frac{1}{2} \left(1 - \frac{1}{2} \sum_{m=0}^{\infty} \mathfrak{p}(m; \mu) \sqrt{1 - \left(\frac{N-1}{N+1}\right)^{2m}} \right), \quad (27)$$

where $\mu = \sqrt{n+1} \alpha$ and $\mathfrak{p}(m; \mu)$ is the probability of drawing m from a Poissonian of mean $|\mu|^2$, namely,

$$\mathfrak{p}(m; \mu) = \frac{|\mu|^{2m}}{m!} \exp[-|\mu|^2]. \quad (28)$$

Notice in particular that, for fixed $|\alpha|^2$, the above expression admits the following asymptotic expansion at large n :

$$P_{e,\min}^{(\text{MEC},n)} \simeq \frac{1}{2} \left[1 - \frac{1}{2} \left(\sqrt{1 - e^{-4|\alpha|^2}} - \frac{1}{n} \frac{2|\alpha|^2 e^{-4|\alpha|^2}}{(1 - e^{-4|\alpha|^2})^{3/2}} \right) \right] + O\left(\frac{1}{n^2}\right), \quad (29)$$

with the leading term corresponding to value dictated by the Helstrom limit (23). The same analysis can be repeated for input states of our problem with a given prior $p(\alpha)$ by simply replacing $\hat{\rho}_\pm$ with

$$\hat{\rho}_\pm^{(\text{ave})} = \int_{\mathbb{C}^2} d\alpha p(\alpha) |\sqrt{n}\alpha\rangle\langle\sqrt{n}\alpha| \otimes |\pm\alpha\rangle\langle\pm\alpha|, \quad (30)$$

obtaining in this case the following optimal minimum error probability

$$\bar{P}_{e,\min}^{(\text{MEC},n)} = \frac{1}{2} \left(1 - \frac{1}{2} \sum_{m=0}^{\infty} \bar{\mathfrak{p}}(m) \sqrt{1 - \left(\frac{n-1}{n+1}\right)^{2m}} \right), \quad (31)$$

where now

$$\bar{\mathfrak{p}}(m) = \int_{\mathbb{C}^2} d\alpha p(\alpha) \mathfrak{p}(m; \sqrt{n+1}\alpha) \quad (32)$$

is the average photon number distribution.

C. Agnostic Dolinar receiver with prior information on the input mean photon number

The implementation of the optimal covariant measure (25) is highly nontrivial as it requires to discriminate between nonorthogonal states $|m, \pm\rangle$ that involve complex superposition of two-mode Fock states [see Eq. (26)]. To compensate for this here we introduce a preliminary version of the agnostic Dolinar scheme that assumes that only the phase of the

complex parameter α of Eq. (22) is unknown, but grants full knowledge about the mean photon number $|\alpha|^2$ of the inputs. In other words we interpolate between the symmetric version of the MEC problem defined at the end of the previous section (two unknown real terms and one quantum binary variable), and the corresponding MED problem (zero unknown real terms and one quantum binary variable). As schematically shown in the lower panel of Fig. 1, the basic idea is to replace the displacement operations of the original Dolinar configuration, whose values γ_k assume full knowledge of α , with coherent mixing of the testing input with the single-mode concentrated version $|\sqrt{n}\alpha\rangle$ of the training copies, via a beam-splitter operation characterized by a time-dependent reflectivity $\theta_k(t)$, where k can be either $+$ or $-$. Because of the symmetry of the problem, we will call $\theta(t) = \theta_+(t) = \theta_-(t) + \pi$. Note that in the alternative formulation of the Dolinar receiver [36] discussed in Appendix A where the input coherent state is sliced by the sequence of the beam splitter and on each slice a displacement and photon counting process is applied, the scheme of the current section substitutes the displacement operations with additional beam splitters, which mix slices of the n training copies and the sliced input field in an optimized way to give output rates in Eq. (33). It is finally worth stressing that, since we do not rely on reference signals, by construction, the proposed detection strategy is explicitly phase insensitive; accordingly the optimal bound (27) constitutes a proper reference for testing the efficiency of the scheme.

In order to evaluate the optimal $\theta(t)$ and maximize the correct decision probability we follow the same procedure of Sec. II C, that we will now discuss with more detail. First of all, following the same procedure, we define our new $\lambda(t), \mu(t)$ as

$$\begin{aligned} \lambda(t) &= |\alpha|^2 |\cos \theta(t) - \sqrt{n} \sin \theta(t)|^2, \\ \mu(t) &= |\alpha|^2 |\cos \theta(t) + \sqrt{n} \sin \theta(t)|^2. \end{aligned} \quad (33)$$

Let us now assume that $k = +$. The number of photons counted in an interval $(t, t + \Delta t]$ is a Poisson variable $N(t, t + \Delta t)$ with parameter $\lambda(t)\Delta t$ or $\mu(t)\Delta t$ depending whether the provisional hypothesis is respectively $z(t) = +$ or $z(t) = -$. These rates allows us to evaluate the conditional probabilities of correct decision $q_+(t) = P[z(t) = + | k = +]$ and $q_-(t) = P[z(t) = - | k = -]$ following the Dolinar receiver strategy, which changes the provisional hypothesis when a photon is detected [37]. From the difference equation

$$\begin{aligned} q_+(t + \Delta t) &= P[N(t, t + \Delta t) = 0, z(t) = + | k = +] \\ &\quad + P[N(t, t + \Delta t) = 1, z(t) = - | k = +] + o(\Delta t) \\ &= P[N(t, t + \Delta t) = 0 | z(t) = +, k = +] q_+(t) \\ &\quad + P[N(t, t + \Delta t) = 1 | z(t) = -, k = +] (1 - q_+(t)) \\ &\quad + o(\Delta t) \\ &= (1 - \lambda(t)\Delta t) q_+(t) + \mu(t)\Delta t (1 - q_+(t)) + o(\Delta t) \end{aligned} \quad (34)$$

follows the differential equation

$$\frac{dq_+(t)}{dt} = \mu(t) - [\lambda(t) + \mu(t)]q_+(t).$$

In a similar fashion, from the expression for $q_-(t + \Delta t)$ and employing symmetric arguments for the displacements, we get the other differential equation of Eq. (12). With this same procedure, it is obvious that the equation for $P_c(t)$ maintains the same form of Eq. (13), i.e.,

$$\frac{dP_c(t)}{dt} = \mu(t) - [\lambda(t) + \mu(t)]P_c(t), \quad (35)$$

where now, however, the terms $\mu(t)$ and $\lambda(t)$ are defined in Eq. (33) instead of Eq. (11). We can now extremize Eq. (35) to obtain an equation for the optimal control function $\theta^*(t) = \theta_{\text{opt}}^{|\alpha|^2, (n)}(t)$ given $|\alpha|^2$ and n : this yields the solution

$$\tan(2\theta^*(t)) = \frac{\sqrt{n}}{n-1} \frac{1}{P_c^*(t) - \frac{1}{2}}, \quad (36)$$

where $P_c^*(t) = P_{c, \text{opt}}^{|\alpha|^2, (n)}(t)$ is the the associated optimal probability of success that can be obtained by solving a differential equation which is more concisely expressed with the change of variable $\xi(t) = P_c^*(t) - 1/2$, i.e.,

$$\begin{aligned} \frac{d\xi(t)}{dt} &= |\alpha|^2(-\xi(t)(n+1) + \sqrt{(n-1)^2\xi^2(t) + n}) \\ &= |\alpha|^2 \left[-2\xi(t) - (n-1) \left(\xi(t) - \sqrt{\frac{n^2}{(n-1)^2} + \xi(t)^2} \right) \right] \end{aligned} \quad (37)$$

(notice the explicit dependence upon $|\alpha|^2$), which for $n \rightarrow \infty$ converges to Eq. (14). With the separation of variables we can formally integrate Eq. (37), obtaining

$$\begin{aligned} \frac{1}{2}|\alpha|^2 t &= -\frac{n-1}{4n} \tanh^{-1} \left(\frac{(n-1)\xi(t)}{\sqrt{(n-1)^2\xi^2(t) + n}} \right) \\ &+ \frac{n+1}{8n} \left[\tanh^{-1} \left(\frac{2\xi(t)(n+1)\sqrt{(n-1)^2\xi^2(t) + n}}{2(n^2+2)\xi^2(t) + n} \right) \right. \\ &\left. - \log(1 - 4\xi^2(t)) \right]. \end{aligned} \quad (38)$$

This expression cannot be inverted explicitly but it can be evaluated numerically. It turns out that the resulting $P_c^*(t)$ does not coincide with the optimal bound of Eq. (27). Still it remains close to such function being increasing in n and asymptotically reaching the performance of the Dolinar scheme given in Eq. (23). This is explicitly shown in Fig. 2, where we plot the resulting associated probability of error $P_e^* = 1 - P_c^*(t = 1)$ as a function of the training set n for a known value of $|\alpha|^2$; notice that the asymptotic limit (23) is reached quickly for $n \approx 20$ even in the full quantum limit ($|\alpha| < 1$).

D. Agnostic Dolinar receiver with no prior information on the input mean photon number

The scheme of the previous section requires the exact knowledge of $|\alpha|^2$, which is not granted in the original MEC problem. Here, we compensate for such lack of information by splitting the n training copies of $|\alpha\rangle$ into two sets, one of size

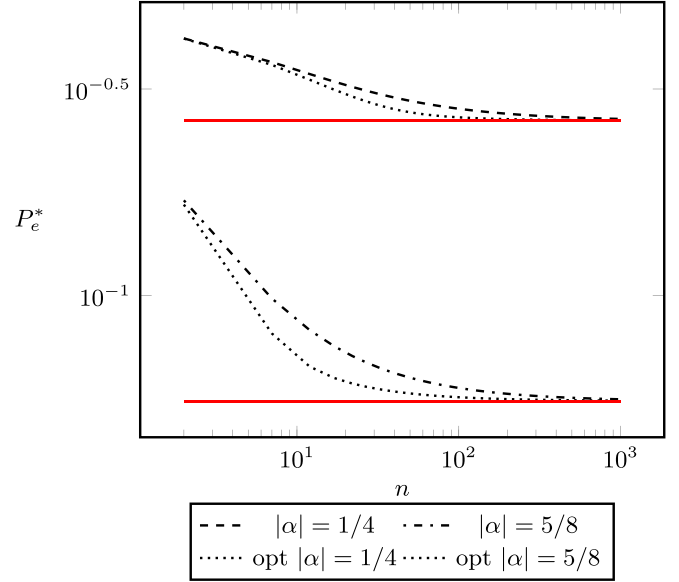


FIG. 2. Error probability $P_e^* = 1 - P_c^*(t = 1)$ of an agnostic Dolinar receiver (dashed and dash-dotted lines) as a function of the number n of training copies, for different (known) values of the mean photon number $|\alpha|^2$. The horizontal solid red lines correspond to the Helstrom values (23) attained by the conventional Dolinar scheme for the associated MED problem. The dotted lines correspond to the optimal error probability obtainable using the optimal bound for the phase-invariant scheme of Eq. (27)

m used to obtain an estimate of the value of $|\alpha|^2$, and the other set of size $n - m$ copies to realize the apparatus described in the former section. Obviously, *a priori* there is no optimal choice for the size of the two sets, as there is a trade-off between a good parameter estimate and the performance of the apparatus. Studying the optimal way to split our sets will be the aim of this section. To estimate the classical value of $|\alpha|^2$ we examine two strategies, photon counting and heterodyne detection. In the former case, the outcomes of the measurement are discrete values $k \in \mathbb{N}$ associated with the count of photons of the state $|\alpha\rangle$ which get distributed according to the Poissonian probability $p(k; \alpha)$ defined as in Eq. (28). When applied to a coherent state $|\sqrt{m}\alpha\rangle$, we can obtain an estimate $\frac{k}{m}$ for $|\alpha|^2$. Note that due to the discrete nature of the outcomes, such an estimate comes in discrete steps. Heterodyne detection is obtained by mixing the coherent state $|\alpha\rangle$ with a strong local oscillator with higher optical frequency [69], and then measuring both quadratures. The measurement outcomes are continuous and can be represented with a complex value $\beta \in \mathbb{C}$ obtained with probability

$$P[\beta; \alpha] = \frac{e^{-|\beta-\alpha|^2}}{\pi}. \quad (39)$$

The estimate for $|\alpha|^2$ can be obtained from the absolute value $|\beta|^2$ of the complex outcome, which is obtained with probability

$$\begin{aligned} P[|\beta|^2; \alpha] &= \int_0^{2\pi} \frac{e^{-(|\alpha|^2 + |\beta|^2 - 2|\alpha\beta| \cos \phi)}}{\pi} |\beta| d\phi \\ &= 2|\beta| e^{-|\alpha|^2 - |\beta|^2} I_0(2|\alpha\beta|), \end{aligned} \quad (40)$$

with $I_0(\cdot)$ the modified Bessel function of the first kind.

To obtain the real performance of the apparatus, we take the expectation value over this probability distribution of the performance of the apparatus using the *wrong* estimate for $|\alpha|^2$. Namely, we use Eqs. (36) and (38) with the amplitude $|\tilde{\alpha}|^2$ estimated from $|\sqrt{m}\alpha\rangle$ to obtain $\theta_{\text{opt}}^{|\tilde{\alpha}|^2, (n-m)}(t)$. Averaging the performance for all the estimates gives the probability of correct decision as a function of α , n , and m .

We compare the performance with the Helstrom bound (6), but also with an estimate-and-discriminate procedure which assumes using all the n copies to obtain a classical estimate of α ($m = n$) and then apply the original Dolinar receiver. The performance of this straightforward method (which we dub miscalibrated Dolinar) is studied in the next section, Sec. III D 1, while the performances of the agnostic Dolinar are studied in Secs. III D 2 and III D 3.

1. Estimate-and-discriminate scheme based on miscalibrated Dolinar receiver

In the original setting of the Dolinar receiver, the value of α that uniquely determines the coherent state $|\alpha\rangle$ was known with arbitrary precision. In the estimate-and-discriminate approach we analyze here, the idea is to use all the n copies of $|\alpha\rangle$ of the state (22) to get an estimate β of α and then use this to build up the Dolinar procedure.

To evaluate the performance of the scheme let us first consider what happens when a Dolinar receiver setup for the discrimination of $|\beta\rangle, |-\beta\rangle$, $\beta \neq \alpha$ is applied to the coherent states $|\alpha\rangle, |-\alpha\rangle$. For this purpose we can use Eq. (13), which is still valid, using the optimal displacement evaluated from β , obtaining a differential equation that can be solved analytically. As a result we get the following probability of success:

$$P_{c,\text{Dol}}^{(\beta;\alpha)} = \frac{1}{2} + \frac{\text{Re}[\alpha\beta^*](1 - e^{-2(|\alpha|^2+|\beta|^2)})}{(|\alpha|^2 + |\beta|^2)\sqrt{1 - e^{-4|\alpha|^2}}}, \quad (41)$$

which, by construction, is upper bounded by the optimal value $P_{c,\text{max}}^{(\text{MED})}$ of Eq. (23) (see Fig. 3). The average success probability of the estimate-and-discriminate approach can now be obtained by averaging Eq. (41) with respect to probability $P[\beta;\alpha, n]$ of getting a certain value of β from our n copies of $|\alpha\rangle$, i.e.,

$$P_{c,\text{E\&D}}^{(\text{MEC})} = \int_{\mathbb{C}^2} d\beta P[\beta;\alpha, n] P_{c,\text{Dol}}^{(\beta;\alpha)}. \quad (42)$$

A plot of this quantity as a function of n for few values of α can be found in Fig. 4 under the assumption that β is recovered via heterodyne detection, so that

$$P[\beta;\alpha, n] = \frac{n}{\pi} \exp[-n|\alpha - \beta|^2]. \quad (43)$$

2. Miscalibrated agnostic Dolinar receiver

In this section, we study the performance of our classifier in the two-step procedure where we split our set of states into two different sets. The first one, of size m , is used to obtain an estimate of the value of $|\alpha|$, while the second one of size $n - m$ is used as input for the agnostic Dolinar re-

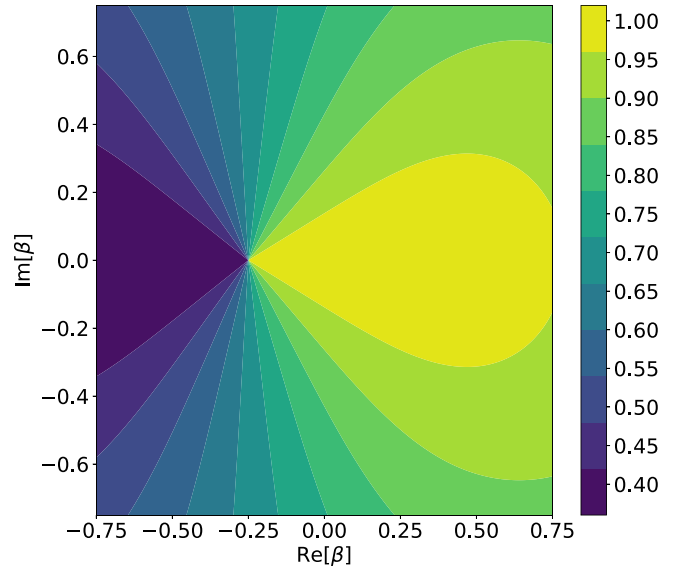


FIG. 3. Density plot of the ratio between the probability of correct decision of a miscalibrated Dolinar receiver as in Eq. (41) and the optimal threshold $P_{c,\text{max}}^{(\text{MED})}$ of Eq. (23) as a function of the complex estimate β for $\alpha = 0.25$.

ceiver described in Sec. III C. The choice of the optimal m is highly nontrivial and we resort to a numerical procedure, here illustrated for $n = 15$. In the upper plot of Fig. 5 it is studied the dependence of the probability of correct decision on the size of the estimating set m with a photon counting estimator. Other values of n show a similar trend. The same setting, but with heterodyne-detection estimation, is plotted in the same figure, in the lower panel. As can be seen from

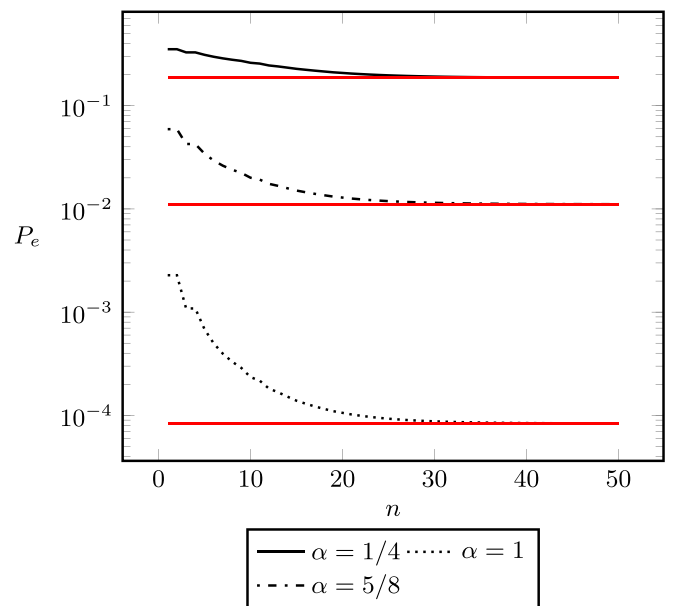


FIG. 4. Error probability $1 - P_{c,\text{E\&D}}^{(\text{MEC})}$ of the estimate-and-discriminate scheme based on a miscalibrated Dolinar receiver where we use the n training copies of $|\alpha\rangle$ of the input (22) to estimate the value of α via heterodyne measurements. The red solid line is the Helstrom bound $1 - P_{c,\text{max}}^{(\text{MED})}$ from (23).

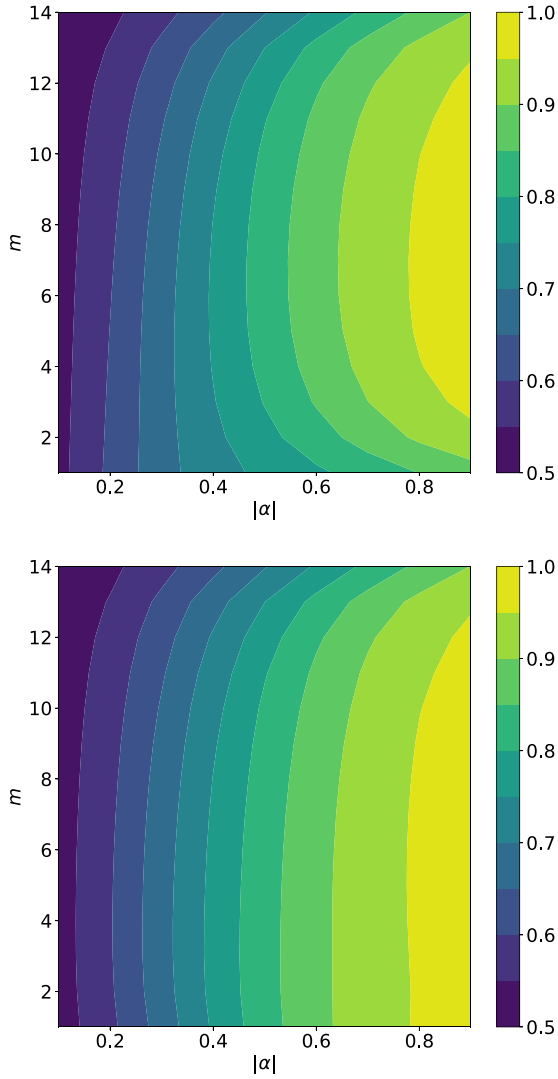


FIG. 5. Probability of correct decision computed as indicated in Sec. III D as a function of α and the size m of the copies used to estimate it, for a total number of copies $n = 15$. The estimation of $|\alpha|$ is performed via photon counting in the upper plot and with heterodyne detection in the lower one.

the figure, the optimal choice of m depends on the value of $|\alpha|$, but the optimal value belongs to a big plateau that allows us to ignore this dependence without losing too much performance. For this reason, we can choose *a priori* the value of m for each n , independent from α , looking at the plateau in the former figures. With this choice, we can finally compare our results with the estimate-and-discriminate strategy of Sec. III D 1 where all the training copies were used to estimate α with heterodyne detection and then a miscalibrated original Dolinar receiver was employed. These results are summarized in Fig. 6. The red solid line is the Helstrom bound, while the black lines are the estimate-and-discriminate performance for $n = 4$ and $n = 8$. The orange and blue lines correspond to photon counting and heterodyne detection, respectively. We can clearly see a divergence in the optimal performance and the estimate-and-discriminate procedure, due to the difference

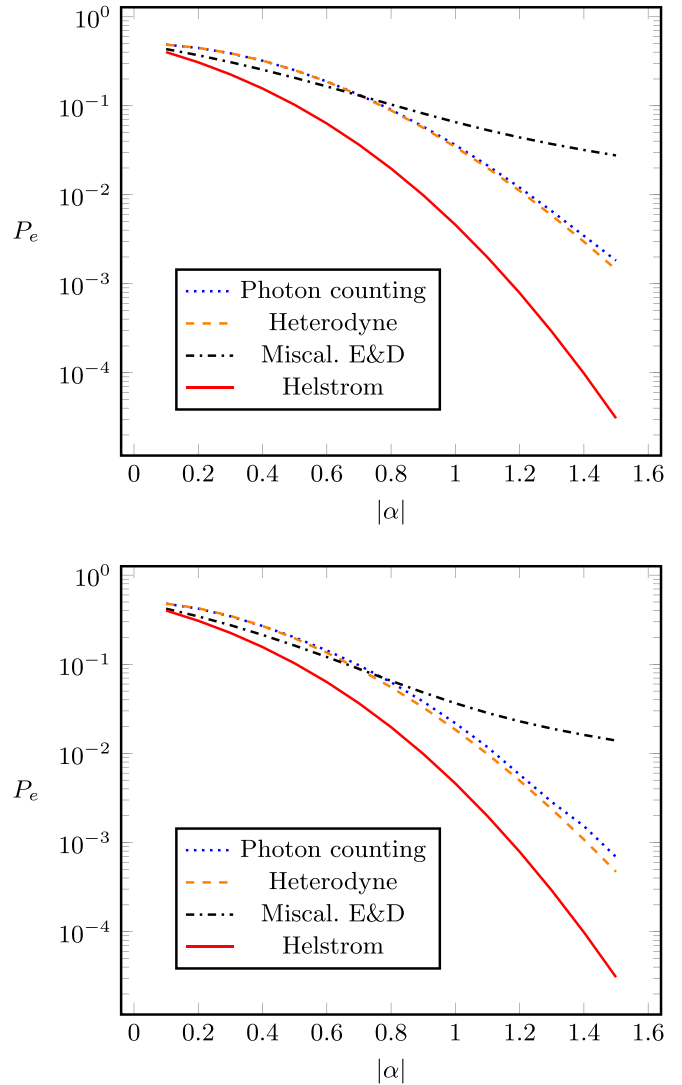


FIG. 6. Error probability as a function of α for different classifiers, for size $n = 4$ (upper part) and $n = 8$ (lower part) of the training set. The red solid line is the Helstrom bound (23) while the agnostic Dolinar receivers employing photon counting and heterodyne measurements are depicted with a blue dotted line and orange dashed line, respectively. The miscalibrated E&D line is relative to the estimate-and-discriminate procedure based on full estimation with heterodyne detection and the use of a conventional (miscalibrated) Dolinar receiver.

in the concavity of the two plots. This does not happen with our strategy that remains close to the optimal bound. For low values of the distance between the states, the performance of the estimate-and-discriminate procedure is slightly better than ours, and this is due to our *a priori* choice of m , namely, ($n = 4 \rightarrow m = 2, n = 8 \rightarrow m = 3$), that is near the plateau for high (greater than 0.3) values of $|\alpha|$ but is not optimal for low values. If we chose the best m for each value of the distance, our performance would be better than the estimate-and-discriminate procedure, but this cannot be done for the reasons discussed above.

3. Performances in the presence of prior on $|\alpha|$

In this section we analyze the agnostic Dolinar scheme when we have a prior on the value of $|\alpha|$ but no information on the value of the phase $\arg \alpha$. In this case we can average the performance of Fig. 5 to obtain an expected probability of error for each size of the estimating set m , and choose the best m for the given prior. As an example we consider the case in which the prior distribution for $|\alpha|$ is given by the Rice distribution

$$\begin{aligned}
 p(|\alpha|; \sigma, x_c) &= \int_0^{2\pi} \frac{\exp\left[-\frac{(|\alpha|^2 + x_c^2 - 2x_c|\alpha|\cos\theta)}{2\sigma^2}\right]}{2\pi\sigma^2} |\alpha| d\theta \\
 &= \frac{|\alpha|}{\sigma^2} \exp\left[-\frac{|\alpha|^2 + x_c^2}{2\sigma^2}\right] I_0\left(\frac{|\alpha|x_c}{\sigma^2}\right), \quad (44)
 \end{aligned}$$

where I_0 is the modified Bessel function of the first kind. The obtained results are reported in Fig. 7, where the optimal error probability bound given in Eq. (31) is compared with the performance of our scheme, employing both heterodyne detection (blue lines) and photon counting (orange lines) to estimate $|\alpha|$, for $n = 4$ and $n = 8$. In this plot, which is evaluated with $\sigma = 0.1$ as a function of x_c , we can observe that with both measurements the performances remain close to the optimal ones as x_c increases, maintaining the ordering with respect to n (greater gives lower error probability).

IV. CONCLUSIONS

To summarize, here we have introduced the minimum error discrimination and minimum error classification problems of quantum states. The former, a central problem in quantum information theory, assumes classical knowledge of the quantum states to discriminate. The latter, risen with the recent studies on machine learning, trades the classical description with the availability of multiple training copies assigned to the classes of quantum states to distinguish. In quantum optics, in the case of the binary discrimination of coherent states, an apparatus realizing the optimal discrimination is known (Dolinar receiver), while the corresponding one for the classification problem is still missing. Between these two scenarios, we identify some intermediate setups with increasing level of classical knowledge: for instance, we can assume that the coherent states have opposite phases, or in addition to that, a prior distribution on the amplitude of the quantum states. We evaluate an optimal bound for this later problem, leveraging on the fact that the POVM associated with the optimal classifier must be phase invariant on the quantum states defined by the training copies and the state to distinguish. This bound asymptotically approaches the Helstrom limit in the limit of infinite training copies, which is the optimal bound for the discrimination problem.

We extend the Dolinar receiver with an agnostic formulation with and without prior information on the input mean photon number. In the case the prior is unknown, a fraction of the training copies is measured to estimate the mean photon number, either with a heterodyne measurement or via photon counting. The remaining training copies are employed in the classification device. We compare the performances of these schemes with the optimal bound previously evaluated and with a miscalibrated estimate-and-discriminate

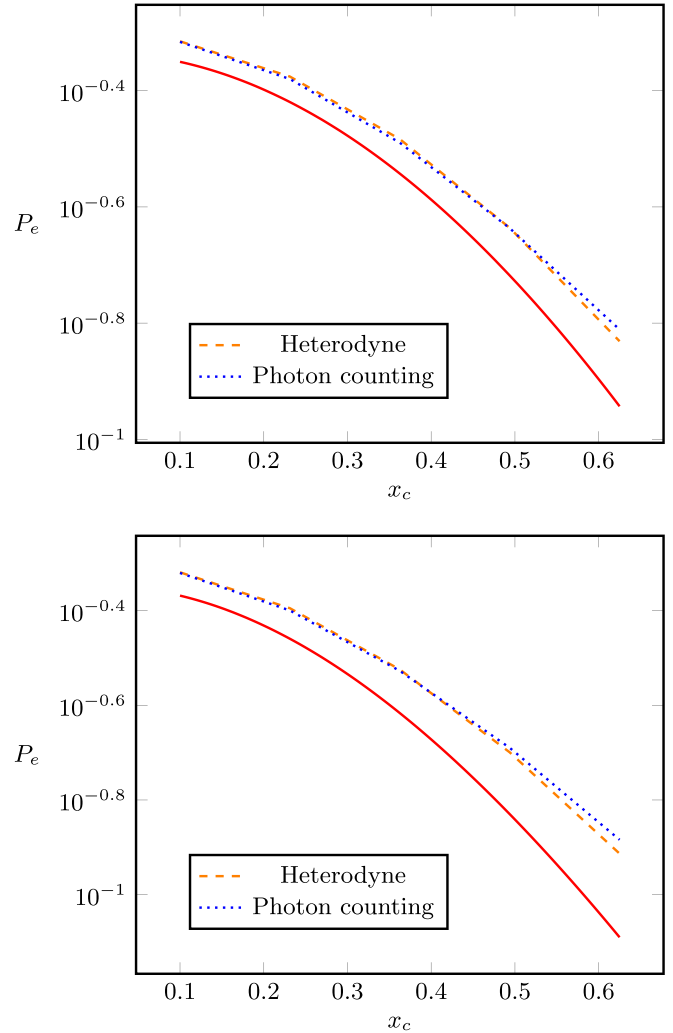


FIG. 7. Error probability comparison given a prior Rice distribution [Eq. (44)] on $|\alpha|$ with $\sigma = 0.1$ as a function of x_c , for $n = 4$ (upper part) and $n = 8$ (lower part). The red solid lines represent the associated optimal bound for phase-insensitive schemes in Eq. (31). The performance of the agnostic Dolinar receiver is plotted for heterodyne (dashed orange) and photon-counting (blue dotted) measurements.

apparatus, where all the training copies are employed in the amplitude estimation, used in a later stage by a Dolinar receiver. The trend of the schemes employing both heterodyne and photon-counting measurements follows the optimal bound with a clear gap, but outperforms the estimate-and-discriminate strategy. This confirms and extends the results of Ref. [68], where the behavior of the estimate-and-discriminate strategy was evaluated asymptotically in the number of training copies.

As future outlooks, one can narrow the gap with the optimal bound with adaptive strategies that estimate the coherent-state amplitudes and perform a partial discrimination at the same time. On the experimental side, the proposed classifiers can in principle be already implemented as they require state-of-the-art devices (beam splitters, phase shifters, photon counters, and local laser sources) commonly present in current laboratories.

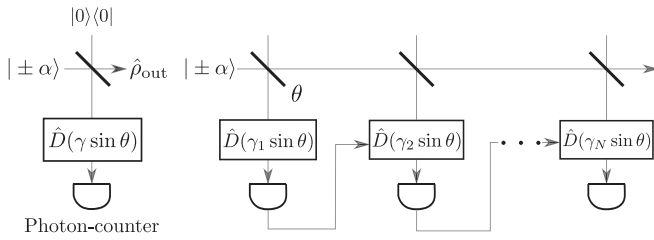


FIG. 8. Discretized description of the Dolinar receiver. The rectangles represent a displacement gate, which is followed by photon counters, while the diagonal lines represent beam splitters.

ACKNOWLEDGMENTS

This work has received support from MIUR via PRIN 2017 (Progetto di Ricerca di Interesse Nazionale): project QUSHIP (2017SRNBRK). The authors would like to thank Matteo Rosati for the useful discussions.

APPENDIX A: AN EQUIVALENT DESCRIPTION OF THE DOLINAR RECEIVER

Here we review the alternative formulation of the Dolinar receiver presented in Ref. [36], which is depicted in Fig. 8 and comes from the equivalence between a continuous photon-counting process and a sequence of beam splitters and photon detectors [75]. The input state comes in the apparatus from the left and goes through a sequence of very similar steps. Each of the diagonal rectangles is a beam splitter of very small reflectivity $\theta \ll 1$. The input state is mixed with the vacuum

$|0\rangle$ via this beam splitter, displaced with the displacement gate $\hat{D}(\gamma_k \sin \theta)$, and then undergoes photon counting. The measurement result is used, in addition to the known value of α , to decide the next displacement parameter γ_{k+1} . Then, the discrimination result will simply be the parity of the total number of photons counted. It can be shown that, with the correct choice of γ_k , and for the number of steps going to infinity, this apparatus tends to the Helstrom bound.

APPENDIX B: DERIVATION OF EQ. (28)

Observe that the following identities hold:

$$\begin{aligned} \sum_{m=0}^{\infty} \hat{\Pi}_m | \sqrt{n}\alpha \rangle \langle \sqrt{n}\alpha | \otimes | \pm\alpha \rangle \langle \pm\alpha | \hat{\Pi}_m \\ = \sum_{m=0}^{\infty} p(m; \mu) |m, \pm\rangle \langle \pm, m|, \end{aligned} \quad (B1)$$

$$\begin{aligned} \sum_{m=0}^{\infty} \| \Pi_m(\hat{\rho}_+ - \hat{\rho}_-) \Pi_m \|_1 \\ = \sum_{m=0}^{\infty} p(m; \mu) \sqrt{1 - |\langle m, + | m, - \rangle|^2}, \end{aligned} \quad (B2)$$

which imply Eq. (28) by noticing that

$$\langle m, + | m, - \rangle = \left(\frac{n-1}{n+1} \right)^m. \quad (B3)$$

-
- [1] N. Gisin and R. Thew, Quantum communication, *Nat. Photonics* **1**, 165 (2007).
- [2] L. Gyongyosi, S. Imre, and H. V. Nguyen, A survey on quantum channel capacities, *IEEE Commun. Surv. Tutorials* **20**, 1149 (2018).
- [3] A. Manzalini, Quantum communications in future networks and services, *Quantum Rep.* **2**, 221 (2020).
- [4] V. Giovannetti, Quantum-enhanced measurements: Beating the standard quantum limit, *Science* **306**, 1330 (2004).
- [5] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum Metrology, *Phys. Rev. Lett.* **96**, 010401 (2006).
- [6] V. Giovannetti, S. Lloyd, and L. Maccone, Advances in quantum metrology, *Nat. Photonics* **5**, 222 (2011).
- [7] S. Pirandola, R. Laurenza, C. Lupo, and J. L. Pereira, Fundamental limits to quantum channel discrimination, *npj Quantum Inf.* **5**, 50 (2019).
- [8] C. L. Degen, F. Reinhard, and P. Cappellaro, Quantum sensing, *Rev. Mod. Phys.* **89**, 035002 (2017).
- [9] T. Gefen, A. Rotem, and A. Retzker, Overcoming resolution limits with quantum sensing, *Nat. Commun.* **10**, 4992 (2019).
- [10] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Quantum cryptography, *Rev. Mod. Phys.* **74**, 145 (2002).
- [11] H.-K. Lo, M. Curty, and K. Tamaki, Secure quantum key distribution, *Nat. Photonics* **8**, 595 (2014).
- [12] K. Banaszek, Optimal receiver for quantum cryptography with two coherent states, *Phys. Lett. A* **253**, 12 (1999).
- [13] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani, J. L. Pereira, M. Razavi, J. S. Shaari, M. Tomamichel, V. C. Usenko, G. Vallone, P. Villoresi, and P. Wallden, Advances in quantum cryptography, *Adv. Opt. Photonics* **12**, 1012 (2020).
- [14] F. Cavaliere, E. Prati, L. Poti, I. Muhammad, and T. Catuogno, Secure quantum communication technologies and systems: From labs to markets, *Quantum Rep.* **2**, 80 (2020).
- [15] A. Chefles, Quantum state discrimination, *Contemp. Phys.* **41**, 401 (2000).
- [16] J. A. Bergou, Quantum state discrimination and selected applications, *J. Phys.: Conf. Ser.* **84**, 012001 (2007).
- [17] J. A. Bergou, Discrimination of quantum states, *J. Mod. Opt.* **57**, 160 (2010).
- [18] S. M. Barnett and S. Croke, Quantum state discrimination, *Adv. Opt. Photonics* **1**, 238 (2009).
- [19] C. Helstrom, *Quantum Detection and Estimation Theory*, Mathematics in Science and Engineering: A Series of Monographs and Textbooks (Academic Press, New York, 1976).
- [20] A. S. Holevo, Statistical problems in quantum physics, in *Proceedings of the Second Japan-USSR Symposium on Probability Theory*, edited by G. Maruyama and Y. V. Prokhorov (Springer, Berlin, 1973), pp. 104–119.
- [21] P. Wittek, Quantum classification, in *Quantum Machine Learning*, edited by P. Wittek (Academic Press, Boston, 2014), pp. 119–123.

- [22] S. Gambs, Quantum classification, [arXiv:0809.0444](https://arxiv.org/abs/0809.0444).
- [23] M. Guță and W. Kotłowski, Quantum learning: Asymptotically optimal classification of qubit states, *New J. Phys.* **12**, 123032 (2010).
- [24] M. Schuld, I. Sinayskiy, and F. Petruccione, Quantum computing for pattern classification, in *PRICAI 2014: Trends in Artificial Intelligence*, edited by D.-N. Pham and S.-B. Park (Springer International, Cham, 2014), pp. 208–220.
- [25] M. Sasaki and A. Carlini, Quantum learning and universal quantum matching machine, *Phys. Rev. A* **66**, 022303 (2002).
- [26] A. Hayashi, M. Horibe, and T. Hashimoto, Quantum pure-state identification, *Phys. Rev. A* **72**, 052306 (2005).
- [27] A. Hayashi, M. Horibe, and T. Hashimoto, Unambiguous pure-state identification without classical knowledge, *Phys. Rev. A* **73**, 012328 (2006).
- [28] C. Chen, D. Dong, B. Qi, I. R. Petersen, and H. Rabitz, Quantum ensemble classification: A sampling-based learning control approach, *IEEE Trans. Neural Networks Learning Syst.* **28**, 1345 (2017).
- [29] H. Yuen, R. Kennedy, and M. Lax, Optimum testing of multiple hypotheses in quantum detection theory, *IEEE Trans. Inf. Theory* **21**, 125 (1975).
- [30] Y. C. Eldar, A. Megretski, and G. C. Verghese, Designing optimal quantum detectors via semidefinite programming, *IEEE Trans. Inf. Theory* **49**, 1007 (2003).
- [31] Y. C. Eldar and G. D. Forney, On quantum detection and the square-root measurement, *IEEE Trans. Inf. Theory* **47**, 858 (2001).
- [32] Y. C. Eldar, A. Megretski, and G. C. Verghese, Optimal detection of symmetric mixed quantum states, *IEEE Trans. Inf. Theory* **50**, 1198 (2004).
- [33] N. Dalla Pozza and G. Pierobon, Optimality of square-root measurements in quantum state discrimination, *Phys. Rev. A* **91**, 042334 (2015).
- [34] S. J. Dolinar, Jr., Processing and transmission of information, Massachusetts Institute of Technology Research Laboratory of Electronics Quarterly Progress Report No. 111 (1973).
- [35] J. M. Geremia, Distinguishing between optical coherent states with imperfect detection, *Phys. Rev. A* **70**, 062303 (2004).
- [36] M. Takeoka, M. Sasaki, P. van Loock, and N. Lütkenhaus, Implementation of projective measurements with linear optics and continuous photon counting, *Phys. Rev. A* **71**, 022318 (2005).
- [37] A. Assalini, N. Dalla Pozza, and G. Pierobon, Revisiting the Dolinar receiver through multiple-copy state discrimination theory, *Phys. Rev. A* **84**, 022342 (2011).
- [38] R. L. Cook, P. J. Martin, and J. M. Geremia, Optical coherent state discrimination using a closed-loop quantum measurement, *Nature (London)* **446**, 774 (2007).
- [39] L. Bartůšková, A. Černoč, J. Soubusta, and M. Dušek, Programmable discriminator of coherent states: Experimental realization, *Phys. Rev. A* **77**, 034306 (2008).
- [40] K. Kato, M. Osaki, M. Sasaki, and O. Hirota, Quantum detection and mutual information for QAM and PSK signals, *IEEE Trans. Commun.* **47**, 248 (1999).
- [41] N. Dalla Pozza and N. Laurenti, Adaptive discrimination scheme for quantum pulse-position-modulation signals, *Phys. Rev. A* **89**, 012339 (2014).
- [42] C. R. Müller and C. Marquardt, A robust quantum receiver for phase shift keyed signals, *New J. Phys.* **17**, 032003 (2015).
- [43] M. P. da Silva, S. Guha, and Z. Dutton, *Optimal Discrimination of m Coherent States with a Small Quantum Computer* (AIP, New York, 2014).
- [44] M. Namkung and Y. Kwon, Sequential state discrimination of coherent states, *Sci. Rep.* **8**, 16915 (2018).
- [45] J. A. Bergou and M. Hillery, Universal Programmable Quantum State Discriminator that is Optimal for Unambiguously Distinguishing Between Unknown States, *Phys. Rev. Lett.* **94**, 160501 (2005).
- [46] G. Sentís, E. Bagan, J. Calsamiglia, and R. Muñoz-Tapia, Multicopy programmable discrimination of general qubit states, *Phys. Rev. A* **82**, 042312 (2010).
- [47] M. Fanizza, A. Mari, and V. Giovannetti, Optimal universal learning machines for quantum state discrimination, *IEEE Trans. Inf. Theory* **65**, 5931 (2019).
- [48] G. Sentís, J. Calsamiglia, R. Muñoz-Tapia, and E. Bagan, Quantum learning without quantum memory, *Sci. Rep.* **2**, 708 (2012).
- [49] G. Sentís, E. Bagan, J. Calsamiglia, and R. Muñoz-Tapia, Programmable discrimination with an error margin, *Phys. Rev. A* **88**, 052304 (2013).
- [50] M. Sedlák, M. Ziman, O. c. v. Příbyla, V. Bužek, and M. Hillery, Unambiguous identification of coherent states: Searching a quantum database, *Phys. Rev. A* **76**, 022326 (2007).
- [51] M. Sedlák, M. Ziman, V. Bužek, and M. Hillery, Unambiguous identification of coherent states. II. Multiple resources, *Phys. Rev. A* **79**, 062305 (2009).
- [52] J. Fiurášek, M. Dušek, and R. Filip, Universal Measurement Apparatus Controlled by Quantum Software, *Phys. Rev. Lett.* **89**, 190401 (2002).
- [53] J. Fiurášek and M. Dušek, Probabilistic quantum multimeters, *Phys. Rev. A* **69**, 032302 (2004).
- [54] M. Dušek and V. Bužek, Quantum-controlled measurement device for quantum-state discrimination, *Phys. Rev. A* **66**, 022112 (2002).
- [55] J. A. Bergou, E. Feldman, and M. Hillery, Optimal unambiguous discrimination of two subspaces as a case in mixed-state discrimination, *Phys. Rev. A* **73**, 032107 (2006).
- [56] C. Zhang, M. Ying, and B. Qiao, Universal programmable devices for unambiguous discrimination, *Phys. Rev. A* **74**, 042308 (2006).
- [57] B. He and J. A. Bergou, Programmable unknown quantum-state discriminators with multiple copies of program and data: A Jordan-basis approach, *Phys. Rev. A* **75**, 032316 (2007).
- [58] U. Herzog and J. A. Bergou, Optimum unambiguous identification of d unknown pure qudit states, *Phys. Rev. A* **78**, 032320 (2008).
- [59] D. Akimoto and M. Hayashi, Discrimination of the change point in a quantum setting, *Phys. Rev. A* **83**, 052328 (2011).
- [60] A. J. Colina, Programmed discrimination of multiple sets of qubits with added classical information, *Eur. Phys. J. D* **66**, 185 (2012).
- [61] T. Zhou, Success probabilities for universal unambiguous discriminators between unknown pure states, *Phys. Rev. A* **89**, 014301 (2014).
- [62] M. Bilkis, M. Rosati, R. M. Yepes, and J. Calsamiglia, Real-time calibration of coherent-state receivers: Learning by trial and error, *Phys. Rev. Research* **2**, 033295 (2020).
- [63] S.-H. Tan, B. I. Erkmen, V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, S. Pirandola, and J. H. Shapiro, Quantum

- Illumination with Gaussian States, *Phys. Rev. Lett.* **101**, 253601 (2008).
- [64] S. Lloyd, Enhanced sensitivity of photodetection via quantum illumination, *Science* **321**, 1463 (2008).
- [65] S. Pirandola, Quantum Reading of a Classical Digital Memory, *Phys. Rev. Lett.* **106**, 090504 (2011).
- [66] S. Pirandola, C. Lupo, V. Giovannetti, S. Mancini, and S. L. Braunstein, Quantum reading capacity, *New J. Phys.* **13**, 113012 (2011).
- [67] C. Lupo, S. Pirandola, V. Giovannetti, and S. Mancini, Quantum reading capacity under thermal and correlated noise, *Phys. Rev. A* **87**, 062310 (2013).
- [68] G. Sentís, M. Guță, and G. Adesso, Quantum learning of coherent states, *EPJ Quantum Technol.* **2**, 17 (2015).
- [69] A. Serafini, *Quantum Continuous Variables* (CRC Press, Boca Raton, FL, 2017).
- [70] D. Walls and G. J. Milburn, *Quantum Optics* (Springer, Berlin, 2008).
- [71] R. S. Kennedy, A near-optimum receiver for the binary coherent state quantum channel, MIT Research Laboratory of Electronics Quarterly Progress Report No. 108, 219 (1973).
- [72] S. Guha, Structured Optical Receivers to Attain Superadditive Capacity and the Holevo Limit, *Phys. Rev. Lett.* **106**, 240502 (2011).
- [73] M. Sedlák, M. Ziman, V. Bužek, and M. Hillery, Unambiguous comparison of ensembles of quantum states, *Phys. Rev. A* **77**, 042304 (2008).
- [74] B. Hall and B. Hall, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*, Graduate Texts in Mathematics (Springer, Berlin, 2003).
- [75] M. Ban, Quasicontinuous measurements of photon number, *Phys. Rev. A* **49**, 5078 (1994).