

Quantum Receiver for Phase-Shift Keying at the Single-Photon Level

Jasminder S. Sidhu^{1,2,*}, Shuro Izumi³, Jonas S. Neergaard-Nielsen³, Cosmo Lupo^{2,†} and Ulrik L. Andersen³

¹*SUPA Department of Physics, The University of Strathclyde, Glasgow G4 0NG, United Kingdom*

²*Department of Physics and Astronomy, The University of Sheffield, Sheffield S3 7RH, United Kingdom*

³*Center for Macroscopic Quantum States (bigQ), Department of Physics, Technical University of Denmark, Fysikvej, Lyngby 2800 Kgs., Denmark*



(Received 11 September 2020; revised 4 January 2021; accepted 21 January 2021; published 26 February 2021)

Quantum enhanced receivers are endowed with resources to achieve higher sensitivities than conventional technologies. For application in optical communications, they provide improved discriminatory capabilities for multiple nonorthogonal quantum states. In this work, we propose and experimentally demonstrate a new decoding scheme for quadrature phase-shift encoded signals. Our receiver surpasses the standard quantum limit and outperforms all previously known nonadaptive detectors at low input powers. Unlike existing approaches, the receiver only exploits linear optical elements and on-off photodetection. This circumvents the requirement for challenging feed-forward operations that limit communication transmission rates and can be readily implemented with current technology.

DOI: [10.1103/PRXQuantum.2.010332](https://doi.org/10.1103/PRXQuantum.2.010332)

I. INTRODUCTION

Quantum mechanics places strict fundamental limits on our ability to discriminate nonorthogonal quantum states [1,2]. This is a deep-rooted property of quantum mechanics that, on one hand, fuels numerous applications in quantum information science such as quantum computing and quantum key distribution [3–5], and on the other hand, limits the performance of other protocols such as sensing, metrology [6–9], and communication. The mathematical framework around state discrimination is based on the theory of quantum detection [1], and it has been applied to study the discrimination of various quantum states [10–12].

Of particular importance is the efficient discrimination of weak coherent states. Coherent states are endowed with an intrinsic resilience to loss and, given their immediate availability, have become indispensable information carriers in the optical realization of classical [13] and quantum information protocols [14,15]. An alphabet of coherent states with very small amplitudes (down to the single-photon level) possesses large state overlaps, and thus exhibits strong quantum features. Such a small-amplitude

alphabet occurs often in quantum communication protocols and in classical communication schemes that aim to enhance channel capacities [16]. More specifically, the optimal discrimination of weak coherent states can be used to enhance the secure key rate in quantum key distribution, improve the success rate in entanglement distillation, and increase the distance of deep-space communication [17,18].

This work focuses on the discrimination of four weak coherent states with equal amplitude and equidistant phase separations, $\{|\alpha\rangle, |i\alpha\rangle, |-\alpha\rangle, |-i\alpha\rangle\}$, chosen with equal prior probabilities, where the amplitude α is real valued and positive. This ensemble is referred to as quadrature phase-shift keying (QPSK) and is commonplace in fibre networks [19]. It offers efficient encoding of two bits of information in one mode of the electromagnetic field. Efficient readout of the encoded information can be accomplished by measuring conjugate quadratures via a heterodyne detection [13,20]. However, the optimal bound on the discrimination error (that is, the minimum average error in discriminating QPSK coherent states), known as the Helstrom bound, is significantly lower than that attainable through heterodyne detection [21,22]. A practical setup for discriminating the QPSK coherent states at the exact Helstrom limit is unknown. However, it is possible to surpass the heterodyne limit and approach the Helstrom bound using different decoding strategies. These schemes generally use a combination of linear optics, photodetection, and globally optimized displacement operations to distinguish coherent states through conditional signal nulling.

*jsmdrsidhu@gmail.com

†c.lupo@sheffield.ac.uk

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

Specifically, the average error probability can be decreased by adaptively updating the displacement phase [23–29] as well as the amplitude [30,31]. Alternative suboptimal receivers use hybrid strategies that combine homodyne detection with adaptive displacements followed by photodetection to beat the heterodyne detection limit for all signal amplitudes [32,33].

While receivers based on adaptive feedback indeed exhibit superior performance, e.g., outperforming the heterodyne detection limit for all amplitudes, they are technically challenging to implement and may impose practical limitations on the optical communication. Indeed, the bandwidth of the optical communication will be intrinsically limited by the feedback mechanism since the feedback delay significantly degrades the receiver performance [28]. Alternatively, adaptive receivers can be realized by spatially dividing a signal state into multiple modes [34,35], but this significantly increases the complexity of the receiver. It is therefore essential to devise a detection system that beats the heterodyne detection limit without the use of feedback techniques [36]. It has been shown that, for large coherent state amplitudes ($\alpha \gtrsim 2$), this is possible by solely using linear optics and photodetection without the adoption of feedback [24,37]. However, obtaining the same advantage for weak coherent states was, up to now, an open question.

In this paper, we introduce, characterize, and experimentally demonstrate a new decoding strategy for QPSK states, comprised of linear optics and on-off photodetection. Notably, we do not make use of adaptive measurements, feed forward, or photon-number resolution. We show that adaptive feedback is not necessary to beat the conventional heterodyne decoding limit in the fully quantum, weak coherent amplitude regime ($\alpha \lesssim 0.5$). We experimentally realize the receiver and evidence strong agreement with theoretical predictions that account for the system efficiency. This work demonstrates a fundamental advance towards suboptimal optical receivers, and provides an immediate, practical strategy to surpass the heterodyne detection limit with currently available technology. Our strategy is compatible with photon-number-resolving detection that can increase the robustness of the receiver against noise and extend the performance of our scheme to higher input intensities [25,27,38].

The theoretical framework of this paper is presented in Secs. II and III. In Sec. IV, we present our new receiver for QPSK decoding. We demonstrate that our receiver outperforms previous decoding strategies in the weak amplitude regime, and present an experimental demonstration of this in Sec. V. Conclusions are summarized in Sec. VI.

II. THEORETICAL FRAMEWORK

Consider the problem of identifying a quantum state ρ_x drawn from a known finite set $\{\rho_1, \rho_2, \dots, \rho_n\}$ with prior

probabilities $\{p_1, p_2, \dots, p_n\}$ [4,39]. We focus on a single-shot scenario where only a single instance of the state is available. When the states ρ_x are mutually orthogonal, detectors placed along the orthogonal directions will be able to perform perfect state discrimination. However, perfect discrimination of nonorthogonal states is not possible from a single-shot experiment and finding an optimal optical receiver is generally a difficult task. We consider this problem within the framework of ambiguous state discrimination, i.e., we allow for a finite probability of error that we aim to minimize.

Consider a general scheme for structured detection where the unknown state ρ is mixed with a known ancillary state σ through a unitary transformation U . The two output systems are then measured by applying a given measurement M , which is characterized by the positive operator-valued measure elements M_y , with $y = \{1, \dots, m\}$. This is shown schematically in the inset of Fig. 1. While the measurement is fixed, the ancillary state σ and the unitary U can be chosen within given sets, respectively denoted by \mathcal{S} and \mathcal{U} .

We now determine the optimal unitary U and ancilla σ that maximize the average probability of successful discrimination. For given ρ_x , U , and σ , the Born rule gives the probability of obtaining the measurement outcome y as

$$p_{U,\sigma}(y|x) = \text{Tr}[(U\rho_x \otimes \sigma U^\dagger)M_y]. \quad (1)$$

The Bayes rule allows us to compute the probability of input x given output y :

$$p_{U,\sigma}(x|y) = \frac{p_{U,\sigma}(y|x)p(x)}{p_{U,\sigma}(y)}. \quad (2)$$

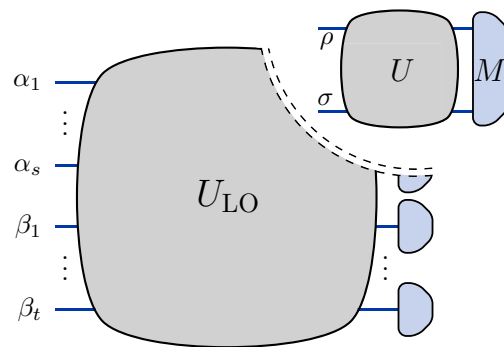


FIG. 1. A scheme for discrimination of multiple coherent states using linear optics and photon detection. The input ensemble is made of s -mode coherent states of unknown amplitudes. The scheme uses t ancillary modes prepared in coherent states of known amplitudes. The final measurement is realized as $s + t$ independent photon detectors. Inset: a general ancilla-assisted discrimination scheme, with ρ an unknown state, σ a known ancillary state, U_{LO} is a unitary transformation constructed from linear optics (LO) and M a measurement.

The best guess for x , given the measurement output y , is the one that maximizes the conditional probability

$$\begin{aligned} p_{U,\sigma}(\hat{x}|y) &= \max_x p_{U,\sigma}(x|y) \\ &= \frac{1}{p_{U,\sigma}(y)} \max_x p_{U,\sigma}(y|x)p(x), \end{aligned} \quad (3)$$

where $p_{U,\sigma}(\hat{x}|y)$ is the probability of successfully identifying the input state given the output measurement y .

The average success probability is then given by

$$\begin{aligned} p_{U,\sigma} &= \sum_y p_{U,\sigma}(y) p_{U,\sigma}(\hat{x}|y) \\ &= \sum_y \max_x p_{U,\sigma}(y|x)p(x). \end{aligned} \quad (4)$$

The optimization routine consists in finding the ancillary state $\sigma \in \mathcal{S}$ and the unitary $U \in \mathcal{U}$ that maximize $p_{U,\sigma}$. This yields the optimized success probability

$$p_s = \sup_{U \in \mathcal{U}, \sigma \in \mathcal{S}} p_{U,\sigma}. \quad (5)$$

Note that this quantity is a function of the sets \mathcal{U} and \mathcal{S} only, in addition to the input ensemble and measurement M . In the following section, we apply this approach to the problem of discriminating a quaternary coherent state alphabet using the linear optics toolbox.

III. COHERENT STATE DISCRIMINATION WITH THE LINEAR OPTICS TOOLBOX

Coherent state discrimination represents a concrete example of quantum state discrimination and has many applications in quantum optics. Here we focus on the problem of ambiguous discrimination of coherent states using linear optics and on-off photon detection. For this, the unknown state ρ is a coherent state over s optical modes, σ is a known coherent state over t modes, and the measurement M is modewise on-off photon detection. The unitary U_{LO} is chosen from the set of linear optical (LO) transformations over $N = s + t$ modes. While the following framework can be extended to nonclassical ancillary states, we focus on coherent states given their availability and widespread use in quantum information.

It is known that linear N -mode unitaries can be constructed from passive linear optical transformations followed by modewise phase-displacement operations [40,41]. In turn, passive linear optical transformations are realized using specific arrangements of beam splitters and phase shifters [42,43].

Consider an unknown coherent state of amplitude α_x ($s = 1$) and $N - 1$ auxiliary coherent states of amplitudes $\beta_j \in \mathbb{C}$ for $j \in \{2, \dots, N\}$. Mixing these coherent states

through an N -mode passive, linear optical unitary U , followed by modewise displacements δ_j , yields as output on mode j a coherent state with amplitude

$$\gamma_j = U_{j1}\alpha_x + \sum_{k=2}^N U_{jk}\beta_{k-1} + \delta_j. \quad (6)$$

The ancillary state amplitudes, passive linear unitary, and the displacements must be chosen to maximize the average success probability of state discrimination. The complexity of this optimization scales quadratically with the number of modes N due to the decomposition of the unitary [42,43]. We significantly reduce this complexity by noting that the amplitudes in Eq. (6) are also attained if the signal α_x is instead mixed with $N - 1$ vacuum modes at the same unitary, with displacements $\epsilon_j = \sum_{k=2}^N U_{jk}\beta_{k-1} + \delta_j$ on the j th mode. The original optimization is then equivalent to requiring a general displacement $\epsilon_j \in \mathbb{C}$ such that $\gamma_j = U_{j1}\alpha_x + \epsilon_j$. Note that our use of ancillary vacuum modes renders only the first column of the unitary important, which amounts to a quadratic speedup of the optimization process. Hence, we write

$$\gamma_j = u_j\alpha_x + \epsilon_j, \quad (7)$$

where $u = (u_1, u_2, \dots, u_N)$ is a unit vector. The objective is then to determine the optimal choice of u and ϵ , collectively referred to as optimization parameters, to maximize the state discrimination.

We now define the optimized success probability for coherent state discrimination. Each mode is subject to an on-off photon detection. Denote a photon detection on the j th mode by $y_j = 1$, and a no detection event by $y_j = 0$. These mutually exclusive events occur with probabilities

$$p_{u_j,\epsilon_j}(y_j = 1|x) = 1 - e^{-|\gamma_j|^2}, \quad (8)$$

$$p_{u_j,\epsilon_j}(y_j = 0|x) = e^{-|\gamma_j|^2}, \quad (9)$$

respectively. The overall output of the N -mode measurement is represented by the binary vector $y = (y_1, y_2, \dots, y_N)$, and the average success probability in Eq. (4) reads

$$\begin{aligned} p_{u,\epsilon} &= \sum_y \max_x p_{u,\epsilon}(y|x)p(x) \\ &= \sum_y \max_x \prod_{j=1}^N p_{u_j,\epsilon_j}(y_j|x)p(x). \end{aligned} \quad (10)$$

The latter is the objective function that we maximize by finding the optimal choice for the parameters u and ϵ .

Note that, without loss of generality, u can be assumed real given the objective function is a function of the modulus of γ_j alone. The objective function in Eq. (10) is then optimized over the parameters $\epsilon_j \in \mathbb{C}$ and $u_j \in \mathbb{R}$ for all $j \in \{1, 2, \dots, N\}$ with $\sum_j |u_j|^2 = 1$. This corresponds to an optimization over a total of $3N - 1$ real parameters, which scales linearly with the number of modes. We implement a nonlinear constrained global optimization of the success probability. The constraints ensure that the vector u has unit norm. The numerical optimizer takes advantage of primitive implementations of gradient-based and direct search algorithms for finding constrained local maximum. We have implemented these routines in *Mathematica* [46].

IV. QUADRATURE PHASE-SHIFT KEYING

In QPSK, the unknown states are coherent states, $\rho_x \equiv |\alpha_x\rangle = |i^x \alpha\rangle$ with $x \in \{0, 1, 2, 3\}$, given with equal probability, and $\alpha > 0$ [19], where $|\alpha|^2$ is the mean photon number. A practical measurement scheme to distinguish these states is heterodyne detection, which, on average, is successful with probability

$$p_{\text{het}} = \frac{1}{4} \left(1 + \operatorname{erf} \left[\frac{\alpha}{\sqrt{2}} \right] \right)^2. \quad (11)$$

To apply our optimization routine in Sec. III to QPSK, we first fix the number of ancillary modes. We illustrate the achievable average error probability of distinguishing the QPSK alphabet with different numbers of ancillary modes in Fig. 2. The number of modes required to minimize the error probability depends on the amplitude of the coherent states. Specifically, our numerical optimization suggests that a decoder equipped with one ancillary coherent mode is optimal in the weak amplitude regime $\alpha \lesssim 0.5$. For larger amplitudes $\alpha \gtrsim 0.75$, two ancillary modes is

the optimal choice. Additional ancillary modes increases the complexity of implementation, while the decoding improvements over two ancillary modes are negligible. In the following, we concentrate on the weak amplitude regime and will hence consider only one ancillary mode (i.e., $N = 2$).

We also find an analytical solution that is an excellent approximation of the numerical optimal in the regime of weak amplitudes, with $N = 2$. Furthermore, this solution does not require each parameter to be tuned to specific values of α . This near optimal receiver is attained through

$$u = \frac{1}{\sqrt{2}}(1, 1) \quad \text{and} \quad \epsilon = \frac{1}{2}(i + 1, i - 1). \quad (12)$$

Physically, this is realized by mixing the input and ancilla modes on a 50% beam splitter. The two modes are then displaced by $(i + 1)/2$ and $(i - 1)/2$, respectively. This is shown in Fig. 3, where $D(\epsilon)$ denotes the phase-space displacement of amplitude ϵ . With this, the computational overheads are greatly reduced, and an experimental implementation to discriminate QPSK states below the standard quantum limit can be easily performed for weak signal amplitudes. For this near optimal choice of parameters, we obtain the following analytical expression for the average success probability (see Appendix A):

$$p_s = \frac{1}{4} \left(1 + 2 \exp \left[-\frac{1 + \alpha^2}{2} \right] \sinh \left[\frac{\alpha}{\sqrt{2}} \right] \right)^2. \quad (13)$$

This is close to the numerically optimized success probability. Note that the independence of the success probability, and hence the optimal parameters, on $|\alpha|$ is only valid in the weak amplitude regime. Intuitively, this is due to a fundamental scale in phase space that is given by the shot noise, which in our units is equal to 1. The optimal

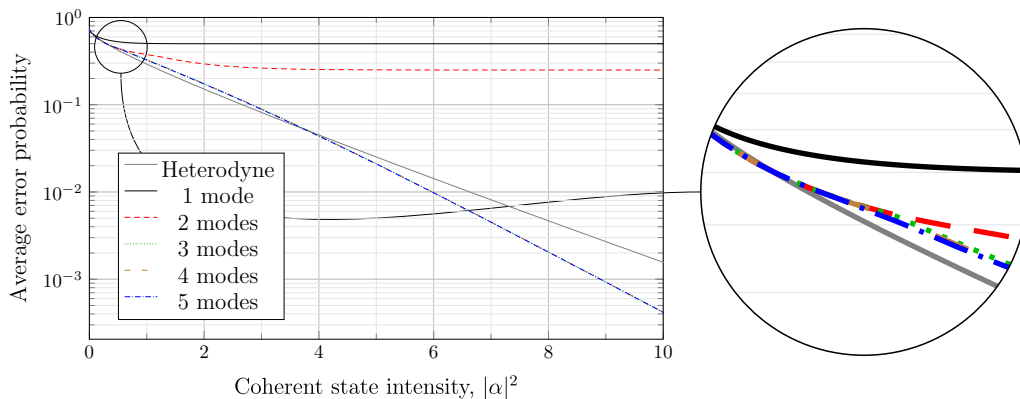


FIG. 2. Average error rate for QPSK signal discrimination with varying signal intensities. Here the number of modes N corresponds to the dimensions of the unitary transformation of the input signal mode and $N - 1$ ancillary modes. At weak photon numbers ($|\alpha|^2 \lesssim 0.5$), a decoder implementing two modes (single vacuum mode) attains the smallest average error probability. For larger photon numbers ($|\alpha|^2 \gtrsim 1$), three modes are optimal and sufficient to outperform heterodyne measurements; using more than three modes does not make any noticeable difference. The magnification on the right-hand side illustrates this clearly.

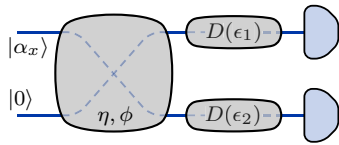


FIG. 3. An optimal receiver for QPSK discrimination. The unknown coherent state is first mixed with a vacuum state at a beam splitter (BS) with transmissivity η and phase ϕ . Each mode is then independently displaced in phase space by ϵ_1, ϵ_2 , before being detected using bucket detectors. This scheme is easy to implement given its independence of ancillary states and adaptive strategies.

parameters change substantially when $\alpha \gtrsim 1$. In the weak amplitude regime, where $\alpha < 1$, these parameters remains fairly constant.

We benchmark the success probability of our optimal receiver scheme with the Helstrom bound and heterodyne detection in Fig. 4. Our scheme outperforms both heterodyne detection and the hybrid scheme of Müller *et al.* [33] in the small amplitude regime. Unlike the receiver of Müller *et al.*, our scheme is implemented using only linear optics and photon counting and does not rely on adaptive conditioning of the optimization parameters.

While we have considered QPSK alphabets, our optimized routine can be applied to an arbitrary constellation

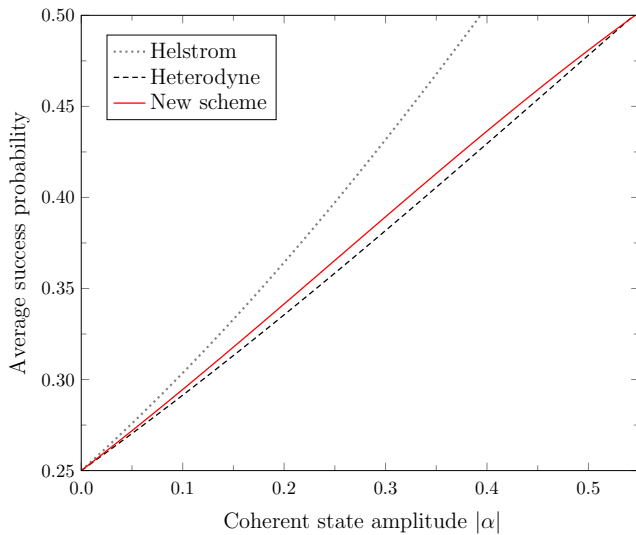


FIG. 4. Optimised average success probability of distinguishing QPSK coherent states as a function of the signal amplitude α is shown with a solid red line. For comparison, the success probabilities attainable with heterodyne measurements is indicated with a dashed black line and the Helstrom bound [22] with a dotted gray line. Note that our decoder beats the SQL for weak amplitudes.

of coherent states. In Appendix B, we consider the application of our detection scheme for an M -ary phase-shift keying code.

V. EXPERIMENTAL DEMONSTRATION

We experimentally demonstrate our QPSK decoder in a temporal mode representation [34,35,47]. In the spatial mode representation used thus far, beam splitters divide the signal coherent state into multiple modes while maintaining the encoded phase information. This multimode splitting can equivalently be accomplished in the temporal domain by splitting the signal coherent state in multiple time bins. The splitting ratio of the beam splitters corresponds to the ratio of the widths of each time bin. The displacement operations performed individually on each mode in the spatial mode version can be implemented in the temporal mode version by instantly updating the displacement operation in time. The multiple single-photon detectors in the spatial version are replaced with a *single* single-photon detector, but it is now required that this detector has sufficiently high time resolution to detect photons in each time bin.

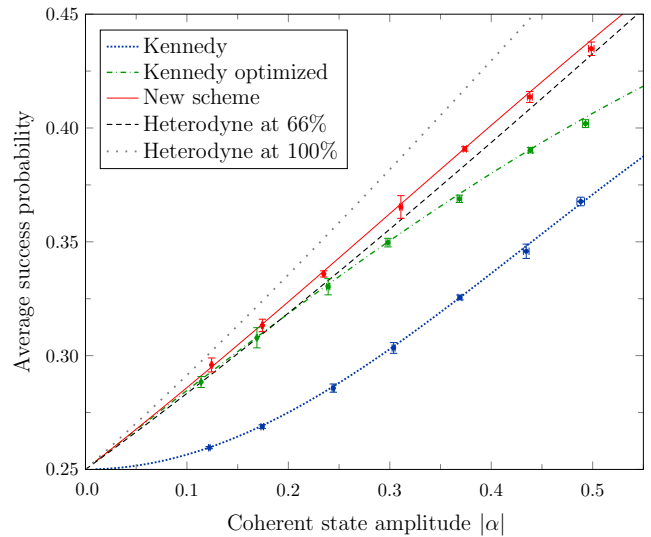


FIG. 5. Average success probability of distinguishing QPSK coherent states as a function of the signal amplitude α for different receivers. The theoretical success probability of our scheme in the experimental condition is shown with a solid red line, together with experimental data points. The success probabilities from heterodyne [from Eq. (11)] measurements accounting for 66% and 100% detection efficiencies are shown with dashed black and dotted gray lines, respectively. The blue dotted line is for a conventional nulling Kennedy receiver [44], shown together with experimental data points. The green dash-dot line is for the Kennedy receiver with optimized displacement amplitudes [45], also shown with experimental data. Error bars denote one standard deviation from five realizations.

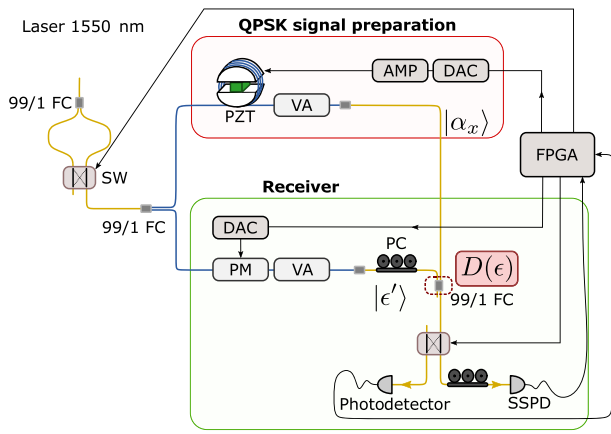


FIG. 6. Experimental setup. FC is the fiber coupler, SW is the optical switch, PM is the phase modulator, PZT is the piezo transducer, VA is the variable attenuator, PC is the polarization controller, DAC is the digital to analog converter, AMP is the amplifier, and SSPD is the superconducting nanowire single photon detector.

Our experimental setup is shown in Fig. 6. We use a continuous wave laser at 1550 nm, which is split into two optical paths in order to individually prepare the signal coherent state and the auxiliary coherent state for the displacement operation. A variable attenuator and a piezo transducer respectively control the amplitude and the phase of the signal state. A phase modulator on the auxiliary coherent state path controls the phase of the displacement operation with a maximum frequency of 1 MHz. Since the temporal width of the signal state is defined to be 100 μ s, the displacement phase can be changed to the desired condition with little adversary effect from the finite bandwidth of the phase modulation. The signal state is combined with

the auxiliary state $|\epsilon'\rangle$ at the 99/1 fiber coupler corresponding to the physical implementation of the displacement operation. Using an optical switch, the interfered beam is guided to either a photodetector for the purpose of stabilizing the relative phase between signal and auxiliary states, or a superconducting nanowire single-photon detector (SSPD) for data acquisition. Because the conventional photon detector cannot measure the laser power highly attenuated to photon level, the laser power is also switched between high and low by an optical switch after the laser source. A field-programmable gate array (FPGA) collects the electrical signals from the SSPD and generates the signal driving the phase modulator. We achieve a total system efficiency of about 66%, where the transmission efficiency from before the 99/1 fiber coupler to the SSPD is approximately 90%. The detection efficiency of the SSPD is measured to be approximately 73% and the dark count noise around 25 Hz, which corresponds to 2.5×10^{-3} counts per signal [28]. The dark count noise, as well as the nonunit visibility, are critical experimental imperfections that limit the performance of receivers based on displacement and photon detection [25,27]. Nevertheless, since we demonstrate our strategy with very weak coherent amplitude conditions, the error probability is so high that the contribution of the error induced by the dark count noise and the visibility imperfection is negligibly small.

We experimentally investigate the performance of three types of two-mode receivers based on displacement operations and photon detections. The three schemes are illustrated on phase-space diagrams in Fig. 7 and the experimentally obtained performances of the receivers are depicted in Fig. 5 [48]. (1) A conventional nulling Kennedy receiver that implements the displacement

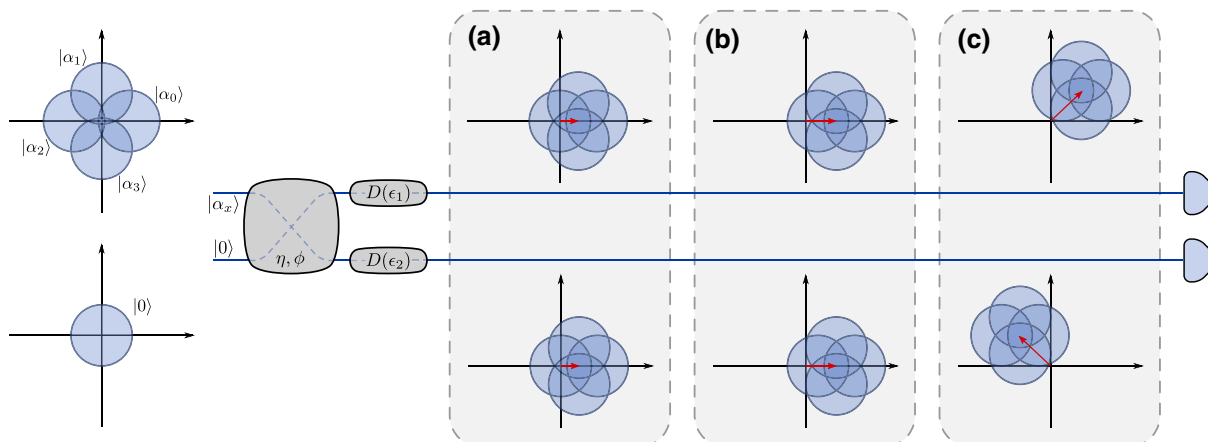


FIG. 7. Illustration of the three receiver schemes implemented in the experiment for discriminating between the four input QPSK states $|\alpha_0\rangle, |\alpha_1\rangle, |\alpha_2\rangle, |\alpha_3\rangle$. (a) In the nulling Kennedy receiver [44], both BS outputs are displaced such that the $|\alpha_2\rangle$ state is shifted to the phase-space origin. (b) The amplitude-optimized Kennedy receiver [45] is similar to the nulling receiver, but by displacing $|\alpha_2\rangle$ further past the origin, the success probabilities for weak amplitudes are significantly improved. (c) Our optimal receiver also optimizes the phases of the displacements on the two BS outputs, leading to further improvements.

operations such that one of the QPSK signals is displaced to the vacuum state, i.e., $|\epsilon_1| = |\epsilon_2| = |\alpha|/\sqrt{2}$ with $\eta = \frac{1}{2}$ [44] [Fig. 7(a)]. The mean and error bars of the success probabilities are evaluated from five independent procedures with 4×10^4 data points for each procedure. The signal amplitude is calibrated from the observed photon count rate by blocking the auxiliary state path and the error bars are evaluated from ten independent procedures. (2) The Kennedy receiver with optimized displacement amplitude [45] [Fig. 7(b)]. Our numerical analysis indicates that, for both nulling and displacement amplitude-optimized receivers, the displacement phases for modes 1 and 2 should be set to the same value to maximize the average success probability in the very weak amplitude case. For the displacement amplitude-optimized receiver, the near optimal performance for the weak coherent signal amplitude is obtained with $|\epsilon_1| = |\epsilon_2| = \frac{1}{2}$ and $\eta = \frac{1}{2}$. The conventional approaches are unable to beat the heterodyne limit, given by the black dash-dot line, in the weak coherent amplitude range. (3) Our strategy of optimizing the phase of the displacement operations [Fig. 7(c)], implemented with the near optimal parameters, provides an improved performance that overcomes the heterodyne limit. Since our system has a finite detection efficiency of 66%, the success probability is degraded and the performances in the actual experimental condition are plotted for the optimal receiver by a dashed curve and open circles for theory and experiment, respectively.

VI. CONCLUSIONS AND DISCUSSIONS

Nonorthogonal quantum states are the building blocks of quantum communication protocols. Weak coherent states are commonly used in these applications given their nonorthogonality, relative ease of generation in laboratories, and resilience to loss. Motivated by this, we have looked at designing practical receivers to discriminate coherent states. In particular, we have focused on the optimal receivers that can be obtained by combining linear optics (including passive linear optics and phase-space displacements) and on-off photodetectors.

The natural decoders for discriminating coherent states are homodyne and heterodyne detections. These detectors have the advantage of being already commonly employed in standard telecommunications. However, they are limited by the shot noise. Several works have focused on a design of the structured receiver that could beat the shot noise limit. In particular, for the problem of decoding QPSK, the only known sub-shot-noise strategies in the low amplitude regime exploit feed forward [30,31,33]. Although feasible in principle, approaches based on feed forward remain technically demanding.

Furthermore, the delay due to practical factors associated with the feedback, such as the signal propagation

time, the speed of signal processing after photon detection and the response time of the single-photon detector, and the delay of updating the displacement operation, may significantly degrade the error probability [28]. The delay becomes more critical if the temporal extent of the signal carrier becomes shorter and therefore the bandwidth of the feedback would limit the possible bandwidth of the optical communication. On the other hand, since static strategies without feedback control do not require real-time signal processing, our receiver is compatible with the high-speed optical communication where the temporal width of the signal carrier is intrinsically short.

Here, we develop a novel sub-shot-noise QPSK decoding. Our scheme employs linear optics and on-off photodetectors, without feed-forward operations. We demonstrate that it outperforms heterodyne detection, as well as all previous nonadaptive detectors in the weak pulse regime. Experimental implementation of our receiver demonstrates results consistent with our theoretical analysis. Going beyond QPSK, we have shown that our scheme allows us to beat heterodyne detection for 3-PSK decoding, but not for 5-PSK and 6-PSK decoding (Appendix B). This suggests that our scheme can beat heterodyne in M -PSK detection only for $M \leq 4$.

An interesting experimental platform to test some of this work is the hybrid spatiotemporal architecture for universal linear optics [49]. This scheme would be useful to implement the optimized unitary receivers that we construct based on the design by Reck *et al.* [42]. It would be interesting to see how this proposal compares with the experimental minimum error measurements proposed by Solís-Prosser *et al.* [2].

Our results pave the way for a number of research questions. First, photon-number-resolving detectors may further improve our decoding strategy, especially in the region of higher signal amplitudes. Second, the effects of nonclassical ancillary states are also expected to deliver improvements to the attainable success probability [50]. The extent of improvement to our scheme is an interesting line of future work. Third, though we have focused on ambiguous state discrimination, our approach may also be useful for unambiguous discrimination of coherent states [51]. Finally, it may be combined with error-correcting codes and exploited to demonstrate the phenomenon of superadditivity in quantum communication [52].

ACKNOWLEDGMENTS

J.S.S. and C.L. acknowledge the EPSRC for funding via the Quantum Communications Hub (Grant No. EP/M013472/1). S.I., J.S.N.-N., and U.L.A. acknowledge the Danish National Research Foundation through the Center for Macroscopic Quantum States (DNRF142). We would like to thank Masahiro Takeoka and Saikat Guha for useful discussions on an earlier version of the manuscript

and T. Yamashita, S. Miki, and H. Terai for providing and installing the superconducting nanowire single-photon detector.

J.S.S. performed the numerical optimization and wrote the initial version of the manuscript. C.L. designed the project and steered the direction of research. S.I., J.S.N.-N., and U.L.A. designed the experiment and wrote about this, while S.I. performed the experiment and analyzed the data. All authors contributed equally in selecting relevant literature and checking the manuscript for accuracy.

APPENDIX A: ALMOST OPTIMAL DISCRIMINATION OF QPSK

Consider the receiver in Fig. 3, where the BS has 50% transmissivity and zero phase, and the complex displacements on both modes have amplitudes $\epsilon_1 = (i + 1)/2$, $\epsilon_2 = (i - 1)/2$.

For a given x , the action of the BS maps the input and ancillary coherent states to two coherent states of amplitudes

$$\begin{aligned} \gamma_+ &= \frac{1}{\sqrt{2}} e^{ix\pi/2} \alpha + \frac{i+1}{2}, \\ \gamma_- &= \frac{1}{\sqrt{2}} e^{ix\pi/2} \alpha + \frac{i-1}{2}. \end{aligned} \quad (\text{A1})$$

Defining $p(y|x)$ as the probability of obtaining the measurement outcome y given the input state α_x , then the probability of individual detectors to click is given by

$$\begin{aligned} p_{\pm}(1|x) &= 1 - e^{-|\gamma_{\pm}|^2} = 1 - \exp\left[-\frac{1 + \alpha^2}{2} - \frac{\alpha}{\sqrt{2}}\right. \\ &\quad \left. \times \left\{ \sin\left(\frac{x\pi}{2}\right) \pm \cos\left(\frac{x\pi}{2}\right) \right\}\right]. \end{aligned} \quad (\text{A2})$$

In conclusion, we table the detection probability for each input state:

$$x = 0 \rightarrow \begin{cases} p_+(1|0) = 1 - \exp\left[-\frac{1 + \alpha^2 + \sqrt{2}\alpha}{2}\right], \\ p_-(1|0) = 1 - \exp\left[-\frac{1 + \alpha^2 - \sqrt{2}\alpha}{2}\right], \end{cases} \quad (\text{A3})$$

$$x = 1 \rightarrow \begin{cases} p_+(1|1) = 1 - \exp\left[-\frac{1 + \alpha^2 + \sqrt{2}\alpha}{2}\right], \\ p_-(1|1) = 1 - \exp\left[-\frac{1 + \alpha^2 - \sqrt{2}\alpha}{2}\right], \end{cases} \quad (\text{A4})$$

$$x = 2 \rightarrow \begin{cases} p_+(1|2) = 1 - \exp\left[-\frac{1 + \alpha^2 - \sqrt{2}\alpha}{2}\right], \\ p_-(1|2) = 1 - \exp\left[-\frac{1 + \alpha^2 + \sqrt{2}\alpha}{2}\right], \end{cases} \quad (\text{A5})$$

$$x = 3 \rightarrow \begin{cases} p_+(1|3) = 1 - \exp\left[-\frac{1 + \alpha^2 - \sqrt{2}\alpha}{2}\right], \\ p_-(1|3) = 1 - \exp\left[-\frac{1 + \alpha^2 + \sqrt{2}\alpha}{2}\right]. \end{cases} \quad (\text{A6})$$

From this table, we compute the maximum likelihood estimation for each combination of detection events, and the associated average probability of successful state discrimination:

$$p_s = \frac{1}{4} \left(1 + 2 \exp\left[-\frac{1 + \alpha^2}{2}\right] \sinh\left(\frac{\alpha}{\sqrt{2}}\right) \right)^2. \quad (\text{A7})$$

Hence, we have found a fixed receiver setup, where the displacements are independent of $|\alpha|$. Note that the independence of the success probability, and hence the optimal parameters, on $|\alpha|$ is only valid in the weak amplitude regime.

In Fig. 8 we illustrate the difference between the numerically optimized success probability and the approximate analytic solution in Eq. (A7). Note that the difference is negligible and increases as the signal amplitudes approach the classical regime ($\alpha \gtrsim 1$), where the receiver is no longer independent of α . Specifically, in this regime, vacuum fluctuations have minimal effect and any change to signal amplitudes effect large changes in the attainable

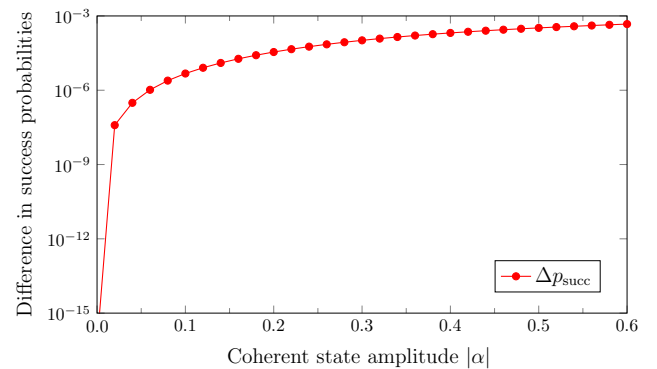


FIG. 8. Difference between the numerically optimized success probability of discriminating QPSK constellations, with the corresponding approximate solution in Eq. (A7). We note that in the weak signal regime, our analytic solution is a good approximate to the optimized solution.

success probability. This highlights that it is possible to implement a signal amplitude independent receiver and still obtain a good approximate to the fully optimized receiver, provided $|\alpha| \lesssim 1$.

APPENDIX B: M-PSK RECEIVERS

In this section, we explore the application of our detection scheme for an M -PSK (M -ary phase-shift keying) code. For arbitrary M , this code is defined by M equally likely coherent states with amplitudes $\alpha_k = \alpha e^{ik2\pi/M}$ for $k = \{1, 2, \dots, M\}$. We compare our scheme with heterodyne detection, which for M -PSK yields the average success probability (see Appendix C)

$$p_{\text{het}} = \frac{1}{M} \exp[-\alpha^2] + \frac{1}{2} \operatorname{erf} \left[\alpha \sin \left(\frac{\pi}{M} \right) \right] + \frac{1}{\sqrt{\pi}} \int_0^{\alpha \sin(\pi/M)} du \exp[-u^2] \operatorname{erf}[\sqrt{\alpha^2 - u^2}]. \tag{B1}$$

We have considered the cases in which $M = 3, 5, 6$, and optimized our detection scheme using a variable number of modes N . In Fig. 10 we illustrate the effect of increasing the number of ancillary modes on the average error probability of the optimal schemes. The average error probability of discriminating an M -PSK alphabet using heterodyne detection is also illustrated for comparison. The discrimination capability of heterodyne detection increases with increasing M .

In general, there is an advantage in using more ancillary modes. However, the average error probability quickly saturates, especially when the number of coherent states, M , is small. Specifically, for $M = 3$, there is no noticeable improvement in using more than one ancilla for all signal amplitudes. Similarly, for $M = 4$, there is no improvement in using more than two ancillas for all signal amplitudes (see Fig. 2). The use of more ancillas appears to be more beneficial as M increases. However, our scheme does not beat heterodyne detection for $M = 5, 6$. This suggests that

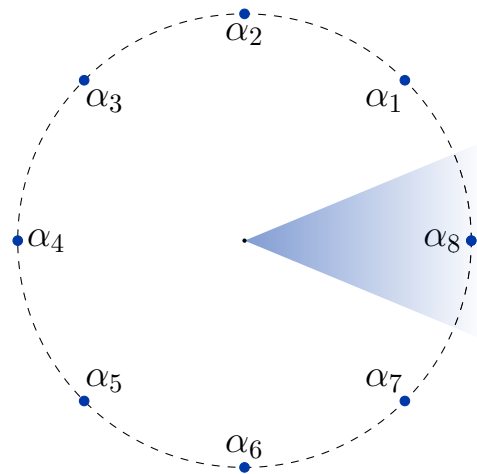


FIG. 9. A constellation of M coherent states distributed with phase separation $2\pi/M$ with $M = 8$.

the standard quantum limit can be exceeded only for $M = 3, 4$ when using linear passive optics, without feedback.

APPENDIX C: M-PSK AVERAGE SUCCESS PROBABILITY OF HETERODYNE DETECTION

In this section, we derive an expression for the success probability of distinguishing coherent states from the M -PSK alphabet. This is a set of M coherent states with amplitudes $\alpha_k = \alpha e^{j2\pi k/M}$. The M -PSK coherent states are represented in phase space in Fig. 9. Without loss of generality, we assume that α is real and set $\alpha = x_0/\sqrt{2}$. Measuring by heterodyne detection a single coherent state with amplitude $\alpha = x_0/\sqrt{2}$ leads to an outcome (x, y) with probability density distribution

$$P(x, y) = \mathcal{N} \exp \left[-\frac{(x - x_0)^2 + y^2}{2\sigma^2} \right], \tag{C1}$$

where $\mathcal{N} = (2\pi\sigma^2)^{-1}$, and $\sigma^2 = 1$ is the shot noise (variance of the vacuum). We need to express this in polar

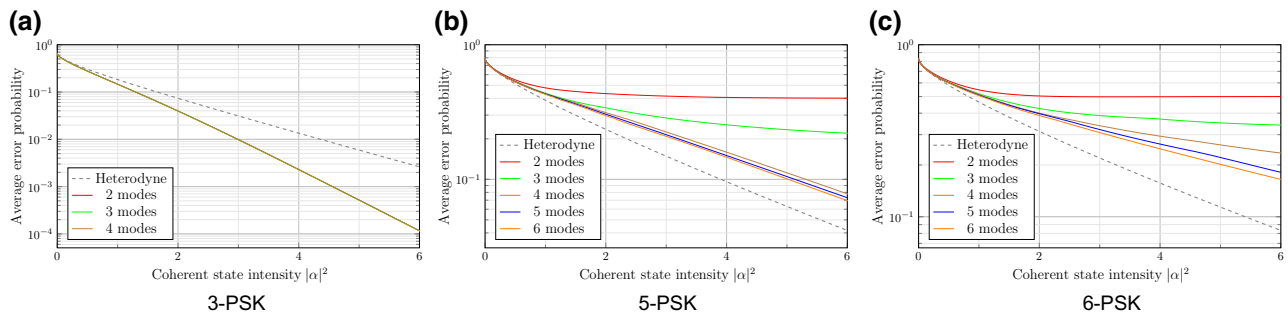


FIG. 10. Average error rate for M -PSK signal discrimination with $M = \{3, 5, 6\}$ and varying numbers of ancillary modes. We apply our nonadaptive linear optical optimization scheme and extend the discrimination beyond the regime of weak signals (where $|\alpha|^2 > 1$).

coordinates, i.e.,

$$P(r, \theta) = \mathcal{N} \exp \left[-\frac{r^2 + x_0^2 - 2rx_0 \cos \theta}{2\sigma^2} \right], \quad (\text{C2})$$

where

$$r = \sqrt{x^2 + y^2}, \quad \cos \theta = \frac{x}{r}. \quad (\text{C3})$$

Consider an ensemble of M coherent states with amplitudes $\alpha_k = (x_0/\sqrt{2}) e^{i2k\pi/M}$. If these states are equally likely then the probability of correct guessing from the output of heterodyne detection is

$$p_{\text{het}} = \int_{-\pi/M}^{\pi/M} d\theta \int_0^\infty r dr P(r, \theta) = \mathcal{N} e^{-x_0^2/(2\sigma^2)} \times \int_{-\pi/M}^{\pi/M} d\theta \int_0^\infty r dr \exp \left[-\frac{r^2 - 2rx_0 \cos \theta}{2\sigma^2} \right]. \quad (\text{C4})$$

This corresponds to the integral over the colored region in Fig. 9. We can simplify the integral with the change of variable

$$\rho := \frac{r^2 - 2rx_0 \cos \theta}{2\sigma^2}. \quad (\text{C5})$$

We have

$$p_{\text{het}} = \sigma^2 \mathcal{N} e^{-x_0^2/(2\sigma^2)} \int_{-\pi/M}^{\pi/M} d\theta \int_0^\infty d\rho e^{-\rho} + x_0 \mathcal{N} e^{-x_0^2/(2\sigma^2)} \int_{-\pi/M}^{\pi/M} d\theta \cos \theta \int_0^\infty dr \exp \left[-\frac{r^2 - 2rx_0 \cos \theta}{2\sigma^2} \right] \\ = \sigma^2 \mathcal{N} e^{-x_0^2/(2\sigma^2)} \frac{2\pi}{M} + x_0 \mathcal{N} \int_{-\pi/M}^{\pi/M} d\theta \cos \theta \exp \left[-\frac{x_0^2}{2\sigma^2} (1 - (\cos \theta)^2) \right] \int_0^\infty dr \exp \left[-\frac{(r - x_0 \cos \theta)^2}{2\sigma^2} \right]. \quad (\text{C6})$$

Now we substitute $\mathcal{N} = (2\pi\sigma^2)^{-1}$ and make the change of variables

$$Y := \sin \theta, \quad r' := \frac{r - x_0 \cos \theta}{\sqrt{2}\sigma}. \quad (\text{C7})$$

This yields

$$p_{\text{het}} = \frac{1}{M} \exp \left[-\frac{x_0^2}{2\sigma^2} \right] + \frac{x_0}{2\pi\sigma^2} \int_{-\pi/M}^{\pi/M} d\theta \cos \theta \exp \left[-\frac{x_0^2}{2\sigma^2} [1 - (\cos \theta)^2] \right] \int_0^\infty dr \exp \left[-\frac{(r - x_0 \cos \theta)^2}{2\sigma^2} \right] \\ = \frac{1}{M} \exp \left[-\frac{x_0^2}{2\sigma^2} \right] + \sqrt{\frac{\pi}{2}} \frac{x_0}{2\pi\sigma} \int_{-\sin(\pi/M)}^{\sin(\pi/M)} dY \exp \left[-\frac{x_0^2}{2\sigma^2} Y^2 \right] \frac{2}{\sqrt{\pi}} \int_{-x_0 \cos \theta/(\sqrt{2}\sigma)}^\infty dr' e^{-r'^2} \\ = \frac{1}{M} \exp \left[-\frac{x_0^2}{2\sigma^2} \right] + \sqrt{\frac{\pi}{2}} \frac{x_0}{2\pi\sigma} \int_{-\sin(\pi/M)}^{\sin(\pi/M)} dY \exp \left[-\frac{x_0^2}{2\sigma^2} Y^2 \right] \left(1 + \operatorname{erf} \left[\frac{x_0 \sqrt{1 - Y^2}}{\sqrt{2}\sigma} \right] \right), \quad (\text{C8})$$

where in the last equation we have applied the identities

$$\frac{2}{\sqrt{\pi}} \int_u^\infty dr' \exp[-r'^2] = 1 - \frac{2}{\sqrt{\pi}} \int_0^u dr' \exp[-r'^2] = 1 - \operatorname{erf}[u] = 1 + \operatorname{erf}[-u]. \quad (\text{C9})$$

We obtain

$$p_{\text{het}} = \frac{1}{M} \exp \left[-\frac{x_0^2}{2\sigma^2} \right] + \sqrt{\frac{\pi}{2}} \frac{x_0}{2\pi\sigma} \int_{-\sin(\pi/M)}^{\sin(\pi/M)} dY \exp \left[-\frac{x_0^2}{2\sigma^2} Y^2 \right] \\ + \sqrt{\frac{\pi}{2}} \frac{x_0}{2\pi\sigma} \int_{-\sin(\pi/M)}^{\sin(\pi/M)} dY \exp \left[-\frac{x_0^2}{2\sigma^2} Y^2 \right] \operatorname{erf} \left[\frac{x_0 \sqrt{1 - Y^2}}{\sqrt{2}\sigma} \right] \\ = \frac{1}{M} e^{-x_0^2/(2\sigma^2)} + \sqrt{\frac{\pi}{2}} \frac{x_0}{\pi\sigma} \int_0^{\sin(\pi/M)} dY e^{-x_0^2/(2\sigma^2)Y^2} + \sqrt{\frac{\pi}{2}} \frac{x_0}{\pi\sigma} \int_0^{\sin(\pi/M)} dY \exp \left[-\frac{x_0^2}{2\sigma^2} Y^2 \right] \operatorname{erf} \left[\frac{x_0 \sqrt{1 - Y^2}}{\sqrt{2}\sigma} \right]. \quad (\text{C10})$$

By applying the following change of variables

$$Y' := \frac{x_0}{\sqrt{2}\sigma} Y, \tag{C11}$$

we obtain

$$\begin{aligned} p_{\text{het}} &= \frac{1}{M} \exp\left[-\frac{x_0^2}{2\sigma^2}\right] + \frac{1}{2} \frac{2}{\sqrt{\pi}} \int_0^{x_0 \sin(\pi/M)/(\sqrt{2}\sigma)} dY' e^{-Y'^2} + \frac{1}{\sqrt{2\pi}} \frac{x_0}{\sigma} \int_0^{\sin(\pi/M)} dY \exp\left[-\frac{x_0^2}{2\sigma^2} Y^2\right] \operatorname{erf}\left[\frac{x_0 \sqrt{1-Y^2}}{\sqrt{2}\sigma}\right] \\ &= \frac{1}{M} \exp\left[-\frac{x_0^2}{2\sigma^2}\right] + \frac{1}{2} \operatorname{erf}\left[\frac{x_0}{\sqrt{2}\sigma} \sin \frac{\pi}{M}\right] + \frac{1}{\sqrt{\pi}} \int_0^{x_0 \sin(\pi/M)/(\sqrt{2}\sigma)} dY' \exp[-Y'^2] \operatorname{erf}\left[\sqrt{\frac{x_0^2}{2\sigma^2} - Y'^2}\right]. \end{aligned} \tag{C12}$$

Finally, setting $x_0 = \alpha\sqrt{2}$ and $\sigma^2 = 1$, we obtain

$$p_{\text{het}} = \frac{1}{M} e^{-\alpha^2} + \frac{1}{2} \operatorname{erf}\left[\alpha \sin \frac{\pi}{M}\right] + \frac{1}{\sqrt{\pi}} \int_0^{\alpha \sin(\pi/M)} dY' \exp[-Y'^2] \operatorname{erf}[\sqrt{\alpha^2 - Y'^2}], \tag{C13}$$

which concludes our proof for the average success probability of distinguishing an M -PSK alphabet using heterodyne measurements.

[1] C. W. Helstrom, Detection theory and quantum mechanics, *Inform. Control*, **10**, 254 (1967).
 [2] M. A. Solís-Prosser, M. F. Fernandes, O. Jiménez, A. Delgado, and L. Neves, Experimental Minimum-Error Quantum-State Discrimination in High Dimensions, *Phys. Rev. Lett.* **118**, 100501 (2017).
 [3] C. H. Bennett, Quantum Cryptography Using any two Nonorthogonal States, *Phys. Rev. Lett.* **68**, 3121 (1992).
 [4] S. M. Barnett, and S. Croke, Quantum state discrimination, *Adv. Opt. Photon.* **1**, 238 (2009).
 [5] C. H. Bennett, and G. Brassard, Quantum cryptography: Public key distribution and coin tossing, *Theor. Comput. Sci.* **560**, 7 (2014).
 [6] M. Ban, K. Kurokawa, R. Momose, and O. Hirota, Optimum measurements for discrimination among symmetric quantum states and parameter estimation, *Int. J. Theor. Phys.* **36**, 1269 (1997).
 [7] A. N. Jordan, J. Tollaksen, J. E. Troupe, J. Dressel, and Y. Aharonov, Heisenberg scaling with weak measurement: a quantum state discrimination point of view, *Quant. Stud. Math. Found.* **2**, 5 (2015).
 [8] J. S. Sidhu, and P. Kok, Geometric perspective on quantum parameter estimation, *AVS Quantum Sci.* **2**, 014701 (2020).
 [9] J. S. Sidhu, Y. Ouyang, E. T. Campbell, and P. Kok, Tight Bounds on the Simultaneous Estimation of Incompatible Parameters, *Phys. Rev. X* **11** (2021).
 [10] Y. C. Eldar, and G. D. Forney, On quantum detection and the square-root measurement, *IEEE Trans. Inf. Theory* **47**, 858 (2001).
 [11] C.-L. Chou, and L. Y. Hsu, Minimum-error discrimination between symmetric mixed quantum states, *Phys. Rev. A* **68**, 042305 (2003).

[12] Y. C. Eldar, A. Megretski, and G. C. Verghese, Optimal detection of symmetric mixed quantum states, *IEEE Trans. Inf. Theory* **50**, 1198 (2004).
 [13] K. Kikuchi, Fundamentals of coherent optical fiber communications, *J. Light. Technol.* **34**, 157 (2016).
 [14] T. C. Ralph, A. Gilchrist, G. J. Milburn, W. J. Munro, and S. Glancy, Quantum computation with optical coherent states, *Phys. Rev. A* **68**, 042319 (2003).
 [15] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, The security of practical quantum key distribution, *Rev. Mod. Phys.* **81**, 1301 (2009).
 [16] V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, J. H. Shapiro, and H. P. Yuen, Classical Capacity of the Lossy Bosonic Channel: The Exact Solution, *Phys. Rev. Lett.* **92**, 027902 (2004).
 [17] V. W. S. Chan, Free-space optical communications, *J. Light. Technol.* **24**, 4750 (2006).
 [18] H. Kaushal, and G. Kaddoum, Optical communication in space: Challenges and mitigation techniques, *IEEE Commun. Surv. Tutor.* **19**, 57 (2017).
 [19] V. K. Garg, and Y.-C. Wang, *Chapter 4 - Modulation and Demodulation Technologies* (Academic Press, 2005), 1st ed., p. 971.
 [20] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, Gaussian quantum information, *Rev. Mod. Phys.* **84**, 621 (2012).
 [21] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press Inc., 1976).
 [22] M. Osaki, M. Ban, and O. Hirota, Derivation and physical interpretation of the optimum detection operators for coherent-state signals, *Phys. Rev. A* **54**, 1691 (1996).
 [23] R. S. Bondurant, Near-quantum optimum receivers for the phase-quadrature coherent-state channel, *Opt. Lett.* **18**, 1896 (1993).

- [24] S. Izumi, M. Takeoka, M. Fujiwara, N. D. Pozza, A. Assalini, K. Ema, and M. Sasaki, Displacement receiver for phase-shift-keyed coherent states, *Phys. Rev. A* **86**, 042328 (2012).
- [25] S. Izumi, M. Takeoka, K. Ema, and M. Sasaki, Quantum receivers with squeezing and photon-number-resolving detectors for m -ary coherent state discrimination, *Phys. Rev. A* **87**, 042328 (2013).
- [26] F. E. Becerra, J. Fan, G. Baumgartner, J. Goldhar, J. T. Kosloski, and A. Migdall, Experimental demonstration of a receiver beating the standard quantum limit for multiple nonorthogonal state discrimination, *Nat. Photon.* **7**, 147 (2013).
- [27] F. E. Becerra, J. Fan, and A. Migdall, Photon number resolution enables quantum receiver for realistic coherent optical communications, *Nat. Photon.* **9**, 48 (2015).
- [28] S. Izumi, J. S. Neergaard-Nielsen, S. Miki, H. Terai, and U. L. Andersen, Experimental Demonstration of a Quantum Receiver Beating the Standard Quantum Limit at Telecom Wavelength, *Phys. Rev. Appl.* **13**, 054015 (2020).
- [29] I. Burenkov, M. Jabir, A. Battou, and S. Polyakov, Time-resolving quantum measurement enables energy-efficient, large-alphabet communication, *PRX Quantum* **1**, 010308 (2020).
- [30] C. R. Müller, and C. Marquardt, A robust quantum receiver for phase shift keyed signals, *New J. Phys.* **17**, 032003 (2015).
- [31] A. R. Ferdinand, M. T. DiMario, and F. E. Becerra, Multi-state discrimination below the quantum noise limit at the single-photon level, *NPJ Quantum Inf.* **3**, 43 (2017).
- [32] M. A. Usuga, C. Müller, C. Wittmann, M. Takeoka, U. L. Andersen, and G. Leuchs, in *CLEO/Europe and QECC 2009 Conference Digest* (Optical Society of America, 2009).
- [33] C. R. Müller, M. A. Usuga, C. Wittmann, M. Takeoka, C. Marquardt, U. L. Andersen, and G. Leuchs, Quadrature phase shift keying coherent state discrimination via a hybrid receiver, *New J. Phys.* **14**, 083009 (2012).
- [34] 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).
- [35] D. Sych, and G. Leuchs, Practical Receiver for Optimal Discrimination of Binary Coherent Signals, *Phys. Rev. Lett.* **117**, 200501 (2016).
- [36] F. E. Becerra, J. Fan, G. Baumgartner, S. V. Polyakov, J. Goldhar, J. T. Kosloski, and A. Migdall, m -ary-state phase-shift-keying discrimination below the homodyne limit, *Phys. Rev. A* **84**, 062324 (2011).
- [37] M. T. DiMario, E. Carrasco, R. A. Jackson, and F. E. Becerra, Implementation of a single-shot receiver for quaternary phase-shift keyed coherent states, *J. Opt. Soc. Am. B* **35**, 568 (2018).
- [38] K. Li, Y. Zuo, and B. Zhu, Suppressing the errors due to mode mismatch for m -ary psk quantum receivers using photon-number-resolving detector, *IEEE Photonics Technol. Lett.* **25**, 2182 (2013).
- [39] J. Bae, and L.-C. Kwek, Quantum state discrimination and its applications, *J. Phys. A: Math. Theoret.* **48**, 083001 (2015).
- [40] A. Ferraro, S. Olivares, and M. G. A. Paris, Gaussian states in continuous variable quantum information (to be published 2005).
- [41] P. Aniello, C. Lupo, and M. Napolitano, Exploring representation theory of unitary groups via linear optical passive devices, *Open Syst. Inf. Dyn.* **13**, 415 (2006).
- [42] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, Experimental Realization of any Discrete Unitary Operator, *Phys. Rev. Lett.* **73**, 58 (1994).
- [43] W. R. Clements, P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, and I. A. Walmsley, Optimal design for universal multiport interferometers, *Optica* **3**, 1460 (2016).
- [44] R. S. Kennedy, Research laboratory of electronics, (Quarterly Progress Report No. 108, MIT, 1973), p. 219.
- [45] C. Wittmann, M. Takeoka, K. N. Cassemiro, M. Sasaki, G. Leuchs, and U. L. Andersen, Demonstration of Near-Optimal Discrimination of Optical Coherent States, *Phys. Rev. Lett.* **101**, 210501 (2008).
- [46] See the Supplemental Material at <http://link.aps.org/supplemental/10.1103/PRXQuantum.2.010332>. Our *Mathematica* script that implements the numerical optimization for the success probability is available in supplementary materials.
- [47] S. Izumi, J. S. Neergaard-Nielsen, and U. L. Andersen, Tomography of a Feedback Measurement with Photon Detection, *Phys. Rev. Lett.* **124**, 070502 (2020).
- [48] See the Supplemental Material at <http://link.aps.org/supplemental/10.1103/PRXQuantum.2.010332>. The raw data for the experimental implementations of our new scheme, the nulling Kennedy receiver and the amplitude-optimised Kennedy receiver are available as supplementary materials. For each discrimination scheme, the data is available in csv and npy format (for direct import in Python via `numpy.load`).
- [49] D. Su, I. Dhand, L. G. Helt, Z. Vernon, and K. Brádler, Hybrid spatiotemporal architectures for universal linear optics, *Phys. Rev. A* **99**, 062301 (2019).
- [50] R. Nair, B. J. Yen, S. Guha, J. H. Shapiro, and S. Pirandola, Symmetric m -ary phase discrimination using quantum-optical probe states, *Phys. Rev. A* **86**, 022306 (2012).
- [51] A. Peres and D. R. Terno, Optimal distinction between non-orthogonal quantum states, *J. Phys. A: Math. General* **31**, 7105 (1998).
- [52] S. Guha, Structured Optical Receivers to Attain Superadditive Capacity and the Holevo Limit, *Phys. Rev. Lett.* **106**, 240502 (2011).