Quantum Detection Theory for the Free-Space Channel

V. Vilnrotter[1] and C.-W. Lau[1]

The fundamental performance limits of optical communications over the free-space channel are developed using quantum theory, and presented in terms of concepts familiar to communications engineers. The compact Dirac notation generally employed in quantum mechanics is defined, and key concepts necessary for understanding quantum projection measurements are reviewed. A derivation that provides significant insights into the quantum measurement performed by the optimum receiver is developed by interpreting the familiar technique of photon counting in terms of quantum projection operators. The performance of the optimum quantum receiver for on–off keying and optical binary phase-shift-keying (BPSK) modulation is treated first as a noise-free (or pure-state) problem, then extended to include the effects of background radiation. The performance of the optimum quantum receiver is compared to that of classical optical receivers employing photon-counting and coherent detection techniques, and it is shown to be exponentially better in most cases.

I. Introduction

It is generally believed that the performance of communications systems is ultimately limited by thermal noise entering the receiver along with the signal. This idea can be illustrated with Shannon's capacity formula for bandlimited classical channels, according to which error-free communication is possible at rates less than $W \log_2(1 + [S/N])$, where W is the bandwidth, S is the average signal power, and N is the power of additive thermal noise. The above expression shows that, if there is no thermal noise, then error-free communication should be possible at arbitrarily high rates, even with bandlimited channels. This conclusion seems intuitively correct at first; however, there is a hidden assumption about the accuracy with which the signal can be measured. The Shannon formulation implicitly assumes that arbitrarily precise measurements are possible since, if this were not the case, there would have to be an effective noise term associated with the noisy signal measurements in the denominator of the capacity expression in addition to the thermal noise, and hence the denominator would never actually approach zero [1]. It turns out that indeed the signal cannot be measured with arbitrary precision even in the absence of thermal noise, due to inherent limitations imposed by the uncertainty principle; hence, channel capacity is ultimately limited by quantum effects that are not included in the classical model.

---

[1] Communications Systems and Research Section.

The research described in this publication was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

The classical model assumes that deterministic signals are observed in the presence of additive Gaussian noise. This model is perfectly adequate for describing communications systems operating at radio frequencies, where quantum effects are not readily detectable. However, at optical frequencies, quantum effects tend to be the dominant source of error, and therefore must be taken into account in the communications system model. The approach most consistent with the principles of quantum mechanics starts out by quantizing the received electromagnetic field, and seeks to determine those measurements on the received field that achieve the best results, such as, for example, minimizing the average probability of detection error [2]. The best measurements may not be readily realizable with physically available devices; however, it is often possible to determine the performance of the "quantum optimum" receiver analytically. Therefore, if these measurements could somehow be made and incorporated into a communications receiver, the performance of the quantum optimum receiver would represent the achievable limit on communications system performance consistent with the principles of quantum mechanics.

Another approach for evaluating optical communications systems assumes a classical instead of a quantized received field, but models the response of physically realizable detectors using the same statistics that a quantum mechanical model would provide [3]. This "quantum mechanically correct" detector response is then used as the fundamental observable on which the decisions are based. Receivers using this approach are often called "semi-classical" and have the advantage of employing well-known detection techniques; however, such receivers generally cannot attain optimum performance.

In this article, we present a thorough review of the quantum theory needed for understanding the concepts of quantum detection, and apply it to several modulation formats of interest for deep-space and terrestrial communications. This is intended as a tutorial for those unfamiliar with the notation generally employed in quantum mechanics, with emphasis on defining terms and concepts in a framework that is familiar to communications engineers.

A brief introduction to the subject of quantum mechanics is presented in the Appendix. Section II provides a review of state vectors and their properties using compact Dirac notation. In Section III, the notion of operators used to transform quantum states is presented, and the properties of specialized operators called "projection operators" are developed in Section IV. These projection operators are then used to introduce and define quantum measurements in Section V, as applied to communications problems. Using the tools developed in the previous sections, quantum detection is introduced in Section VI, where the density operator for a pure coherent state is defined and used to develop the structure of the optimum quantum receiver for binary pure-state problems.

In Section VII, both classical and quantum mechanical derivations of well-known photon-counting detection are presented. The quantum mechanical derivation of photon counting then provides a means to define "measurement states" in a heuristic manner, which can be used to develop a "state–space" interpretation of the pure-state quantum communications problem. Examples of quantum optimum receivers with commonly employed modulation formats are presented in Section VIII, whereas the effects of thermal radiation on the performance of binary quantum receivers is examined in Section IX. Section X addresses the pure-state multiple hypothesis problem again using the concept of measurement states, and concludes with an example of quantum multiple hypothesis testing. Finally, the article is summarized in Section XI.

II. Review of State Vectors

At any instant of time, the state of a quantum system is completely specified by a state vector $|\psi\rangle$ in a Hilbert space over the field of complex numbers [4]. The state vector, or "ket" $|\psi\rangle$, can be thought of as a column vector of infinite dimension. An equivalent "row vector" representation of the state is denoted by $\langle\psi|$ in Dirac notation.

If $|\psi_1\rangle$ and $|\psi_2\rangle$ are states of a system, then so is $|\psi\rangle = a_1|\psi_1\rangle + a_2|\psi_2\rangle$, where a_1 and a_2 are complex numbers. The row-vector representation is of the form $\langle\psi| = a_1^*\langle\psi_1| + a_2^*\langle\psi_2|$. The overlap between two states $|\psi\rangle$ and $|\varphi\rangle$ is the complex number $\langle\psi|\varphi\rangle$ or its complex conjugate $\langle\varphi|\psi\rangle$. If the overlap is zero, the states are orthogonal. The state is normalized if $\langle\psi|\psi\rangle = 1$. If $|\psi_1\rangle$ and $|\psi_2\rangle$ are orthonormal and $|\psi\rangle$ is normalized, then the overlap is

$$\left.\begin{aligned} \langle\psi_1|\psi\rangle &= \langle\psi|\psi_1\rangle^* = a_1 \\ \langle\psi_2|\psi\rangle &= \langle\psi|\psi_2\rangle^* = a_2 \end{aligned}\right\} \tag{1}$$

where $|a_1|^2 + |a_2|^2 = 1$, and we can interpret $|a_1|^2$ and $|a_2|^2$ as the probabilities that the system is found to be in states $|\psi_1\rangle$ and $|\psi_2\rangle$, respectively, after a measurement. Generalization to a superposition of an arbitrary number of possible states follows as

$$|\psi\rangle = \sum_n a_n|\psi_n\rangle \tag{2}$$

$$\sum_n |a_n|^2 = 1 \tag{3}$$

with the interpretation that $|a_n|^2$ is the probability that the system is found in state $|\psi_n\rangle$, provided the states are orthonormal so that $\langle\psi_m|\psi_n\rangle = \delta_{mn}$.

III. Operators Relevant to the Communications Problem

State vectors (kets) are transformed by operators, which can also be viewed as matrices of arbitrarily large dimension. A linear operator $\mathbf{A}$ satisfies

$$\mathbf{A}\left(|\psi\rangle + |\varphi\rangle\right) = \mathbf{A}|\psi\rangle + \mathbf{A}|\varphi\rangle \tag{4}$$

The Hermitian conjugate $\mathbf{A}^+$ of an operator is defined by the relations [4]

$$\left.\begin{aligned} (\mathbf{A}^+)^+ &= \mathbf{A} \\ (\mathbf{A}+\mathbf{B})^+ &= \mathbf{A}^+ + \mathbf{B}^+ \\ (\mathbf{B}\mathbf{A})^+ &= A^+\mathbf{B}^+ \\ (\lambda\mathbf{A})^+ &= \lambda^*\mathbf{A}^+ \end{aligned}\right\} \tag{5}$$

where λ is in general a complex number.

An operator $\mathbf{A}$ acting on any state of the system produces another state, which may not be normalized. An "observable" $\mathbf{A}$ is associated with an Hermitian operator, that is, an operator that is its own Hermitian conjugate: $\mathbf{A} = \mathbf{A}^+$. The eigenvalues of the Hermitian operator $\mathbf{A}$ are real and satisfy the equations

$$\left.\begin{aligned} \mathbf{A}|\lambda_n\rangle &= \lambda_n|\lambda_n\rangle \\ \langle\lambda_n|A &= \lambda_n\langle\lambda_n| \end{aligned}\right\} \tag{6}$$

where $|\lambda_n\rangle$ is the nth eigenstate of the operator $\mathbf{A}$ with eigenvalue λ_n.

The eigenstates of an Hermitian operator are orthogonal; hence, they can be normalized to form an orthonormal set. The set of eigenvalues of an operator is called the "spectrum" of the operator. If any state can be expanded as a superposition of eigenstates as $|\psi\rangle = \sum_n a_n|\lambda_n\rangle$, then the eigenstates form a complete orthonormal basis. Clearly, since the eigenstates are orthonormal, the probability that a measurement of the observable A yields λ_m is

$$|\langle\lambda_m|\psi\rangle|^2 = |\sum_n a_n\langle\lambda_m|\lambda_n\rangle|^2 = |a_m|^2 \tag{7}$$

The outer product of two states, $|\psi_1\rangle\langle\psi_2|$, acting on a state $|\psi\rangle$ yields $|\psi_1\rangle\langle\psi_2|\psi\rangle$. The outer product operator is Hermitian if and only if $|\psi_1\rangle = |\psi_2\rangle$. The Hermitian operator $|\psi_1\rangle\langle\psi_1|$ acting on a state $|\psi\rangle$ yields $|\psi_1\rangle\langle\psi_1|\psi\rangle$, which is seen to be the state $|\psi_1\rangle$ multiplied by the complex number $\langle\psi_1|\psi\rangle$, and can be interpreted as the projection of $|\psi\rangle$ onto $|\psi_1\rangle$.

If the set of eigenstates $\{|\lambda_n\rangle\}$ forms a complete orthonormal basis, then we can construct an "identity operator" as $\sum_n |\lambda_n\rangle\langle\lambda_n| = \mathbf{1}$, also referred to as a "resolution of the identity." The fact that an operator constructed in this manner is an identity operator can be demonstrated by observing its action on an arbitrary state $|\psi\rangle = \sum_n a_n|\lambda_n\rangle$:

$$\begin{aligned} \sum_m |\lambda_m\rangle\langle\lambda_m|\psi\rangle &= \sum_m |\lambda_m\rangle\langle\lambda_m| \sum_n a_n|\lambda_n\rangle \\ &= \sum_n a_n \sum_m \langle\lambda_m|\lambda_n\rangle|\lambda_m\rangle \\ &= \sum_n a_n|\lambda_n\rangle = |\psi\rangle \end{aligned} \tag{8}$$

where we made use of the fact that $\langle\lambda_m|\lambda_n\rangle = \delta_{mn}$.

If $\{|\lambda_n\rangle\}$ is a complete orthonormal set of eigenvectors of an Hermitian operator $\mathbf{A}$, then the operator $\mathbf{A}$ can be expanded in that basis as

$$\begin{aligned} \mathbf{A} = \mathbf{A1} &= \mathbf{A}\sum_n |\lambda_n\rangle\langle\lambda_n| \\ &= \sum_n \lambda_n|\lambda_n\rangle\langle\lambda_n| \end{aligned} \tag{9}$$

where we used the relation $\mathbf{A}|\lambda_n\rangle = \lambda_n|\lambda_n\rangle$.

An operator $\mathbf{A}$ can also be represented as a matrix, possibly of infinite dimension [2], whose mnth element is of the form $A_{mn} = \langle\varphi_m|\mathbf{A}|\varphi_n\rangle$, where $\{|\varphi_n\rangle\}$ represents an arbitrary orthonormal basis. The sum of the diagonal elements of a matrix is called the trace, here denoted by $\mathrm{Tr}(\mathbf{A}) = \sum_n \langle\varphi_n|\mathbf{A}|\varphi_n\rangle$. Finding the eigenvalues and eigenvectors of an operator is equivalent to diagonalizing the corresponding matrix by a unitary transformation—in other words, multiplying by a matrix U that obeys the condition $U^{-1} = U^{+}$.

Given a state $|\psi\rangle = \sum_n a_n|\lambda_n\rangle$ and an Hermitian operator $\mathbf{A}$ with eigenstates $\{|\lambda_n\rangle\}$, the probability that a measurement of the observable A yields the eigenvalue λ_m is $|a_m|^2$. The mean value of the observable A is the average over all possible outcomes:

$$\bar{A} = \sum_n \lambda_n |a_n|^2 = \sum_n \lambda_n \langle\psi|\lambda_n\rangle\langle\lambda_n|\psi\rangle$$

$$= \langle\psi| \sum_n |\lambda_n\rangle\lambda_n\langle\lambda_n|\psi\rangle = \langle\psi|\mathbf{A}|\psi\rangle \tag{10}$$

directly from Eq. (9). For Hermitian operators, the "bracket" $\langle\psi|\mathbf{A}|\psi\rangle$ is always real.

If we do not know the state of the system, but know only the probabilities P_n that the system is in state $|\psi_n\rangle$, then the system is said to be in a mixture of states described by a density operator ρ, which is an Hermitian operator of the form

$$\rho = \sum_n P_n |\psi_n\rangle\langle\psi_n| \tag{11}$$

A density operator can also be defined for a pure quantum state $|\psi\rangle$ as $\rho = |\psi\rangle\langle\psi|$. The density operator ρ is Hermitian, its eigenvalues P_n are all non-negative, its eigenvectors form a complete orthonormal set, and $\mathrm{Tr}(\rho) = 1$.

For a mixture of states described by a density operator ρ, a further averaging over the statistics of the states is needed to describe the average value of the observable A, yielding

$$\bar{A} = \sum_n P_n \langle\psi_n|\mathbf{A}|\psi_n\rangle = \mathrm{Tr}(\rho\mathbf{A}) \tag{12}$$

where trace now refers to the sum of the diagonal elements of the new operator $\rho\mathbf{A}$; hence, each diagonal element is of the form $\langle\lambda_m|\rho\mathbf{A}|\lambda_m\rangle$. Equation (12) can be derived by writing the trace as the sum of the diagonal elements in the matrix representation of the operator $\rho\mathbf{A}$:

$$\begin{aligned}
\mathrm{Tr}(\rho\mathbf{A}) &= \sum_m \langle\lambda_m| \left(\sum_n P_n |\psi_n\rangle\langle\psi_n| \mathbf{A} \right) |\lambda_m\rangle \\
&= \sum_n P_n \sum_m \langle\lambda_m|\psi_n\rangle\langle\psi_n|\mathbf{A}|\lambda_m\rangle \\
&= \sum_n P_n \langle\psi_n|\mathbf{A} \sum_m |\lambda_m\rangle\langle\lambda_m|\psi_n\rangle \\
&= \sum_n P_n \langle\psi_n|\mathbf{A}\mathbf{1}|\psi_n\rangle \\
&= \sum_n P_n \langle\psi_n|\mathbf{A}|\psi_n\rangle \qquad (13)
\end{aligned}$$

This demonstrates the assertion made in Eq. (12). Note that for a pure quantum state represented by $\rho = |\psi\rangle\langle\psi|$, Eq. (13) reduces to $\mathrm{Tr}(\rho\mathbf{A}) = \langle\psi|\mathbf{A}|\psi\rangle$. The trace is invariant under unitary transformations and is equal to the sum of the eigenvalues of the operator. For example, defining the operator $\mathbf{A}$ as $\mathbf{A} = \sum_n \lambda_n |\lambda_n\rangle\langle\lambda_n|$, the trace of this operator is the sum of its eigenvalues: $\mathrm{Tr}(\mathbf{A}) = \sum_n \lambda_n$. Note that for any two Hermitian operators $\mathbf{A}$ and $\mathbf{B}$, the eigenvalues are real, and $\mathrm{Tr}(\mathbf{AB}) = \mathrm{Tr}(\mathbf{BA})$. This can be shown by observing that we can always write the trace of the product of two matrices [5] as $\mathrm{Tr}(\mathbf{AB}) = \sum_m \sum_n \langle\varphi_n|\mathbf{A}|\varphi_m\rangle\langle\varphi_m|\mathbf{B}|\varphi_n\rangle = \mathrm{Tr}(\mathbf{BA})$, where $\{|\varphi_n\rangle\}$ is an arbitrary orthonormal basis.

IV. Projection Operators

The outcome of the measurement of an observable B answers the question, "What is the value of B?" If we are interested in answering only the weaker question, "Does the value of B lie within a specified region of the outcome space?" then it suffices to measure a projection operator [5], constructed from the eigenstates of the operator $\mathbf{B}$.

Consider the spectrum of eigenvalues $\{b_n\}$ of an Hermitian operator $\mathbf{B} = \sum_n b_n |b_n\rangle\langle b_n|$ and define an elemental projection operator as the outer product $|b_k\rangle\langle b_k|$, where $|b_k\rangle$is an eigenstate of $\mathbf{B}$. When this elemental projection operator is applied to $\mathbf{B}$, the result is

$$\begin{aligned}
|b_k\rangle\langle b_k|\mathbf{B} &= |b_k\rangle\langle b_k| \sum_n b_n |b_n\rangle\langle b_n| \\
&= \sum_n b_n |b_k\rangle\langle b_k|b_n\rangle\langle b_n| \\
&= b_k |b_k\rangle\langle b_k| \qquad (14)
\end{aligned}$$

which is seen to be a projection of the operator $\mathbf{B}$ onto the one-dimensional subspace of all multiples of $|b_k\rangle$. Extending this to a multiple-dimensional subspace spanned by any subset of the spectrum of eigenvalues, say I_1, and denoting the corresponding projection operator by $\Pi(I_1)$ yields

$$\Pi(I_1) = \sum_{n:b_n \in I_1} |b_n\rangle\langle b_n| \qquad (15)$$

This projection operator has only two eigenvalues, 0 and 1, as can be seen by noting that, if $b_m \in I_1$, then $\Pi(I_1)|b_m\rangle = \sum_{n:b_n \in I_1} |b_n\rangle\langle b_n|b_m\rangle = |b_m\rangle$; however, if $b_m \notin I_1$, then the overlap is zero. In other words, the outcome of the measurement of $\Pi(I_1)$ is 1 if the value of the observable B falls within I_1, and 0 otherwise. If the state of the system before a measurement is $|\psi\rangle$, then the expected value of the outcome of a measurement of $\Pi(I_1)$ is the probability that the value b_n of the observable B lies within I_1:

$$\overline{\Pi}(I_1) = \langle\psi|\Pi(I_1)|\psi\rangle = \sum_{n:b_n \in I_1} |\langle b_n|\psi\rangle|^2 = \Pr(b_n \in I_1) \tag{16}$$

Note that the expected value of a projection operator is a probability, namely the probability that the result of the measurement falls within the domain of the operator.

For an exhaustive union of disjoint subsets $I_1, I_2, I_3, \cdots$, the resolution of the identity

$$\sum_k \Pi(I_k) = \mathbf{1} \tag{17}$$

holds, where $\Pi(I_k) = \sum_{n:b_n \in I_k} |b_n\rangle\langle b_n|$. Note that $\Pi^2 = \Pi$; hence, applying the same projection operator repeatedly does not change the outcome.

V. Quantum Measurements for Communications Applications

Consider a quantum system described by a density matrix ρ and an Hermitian operator $\mathbf{B}$ whose eigenvectors $|b_n\rangle$ are used to construct projection operators via the elemental projectors $|b_n\rangle\langle b_n|$. The probability of obtaining the number b_n as an outcome of the measurement of the operator $\mathbf{B}$ is

$$\Pr(b_n) = \langle b_n|\rho|b_n\rangle = \mathrm{Tr}\big[\rho|b_n\rangle\langle b_n|\big] \tag{18}$$

Partition the outcome space into an exhaustive set of disjoint regions, $I_1, I_2, \cdots$, and define the projection operator associated with the jth region as $\Pi_j = \sum_{n:b_n \in I_j} |b_n\rangle\langle b_n|$. The probability of obtaining an outcome in the jth region is

$$\Pr(b_n \in I_j) = \sum_{n:b_n \in I_j} \langle b_n|\rho|b_n\rangle = \mathrm{Tr}\left[\rho\Pi_j\right] \tag{19}$$

When more than one density operator is of interest, the conditional probability of obtaining an outcome in the jth region, given that the quantum system is in a mixture of states defined by the kth density operator ρ_k, can be expressed as

$$\Pr(b_n \in I_j|\rho_k) = \mathrm{Tr}\left[\rho_k\Pi_j\right] \tag{20}$$

This conditional probability plays a key role in quantum hypothesis testing, where typically the kth proposition corresponds to the quantum system described by the density operator ρ_k. In what follows, we examine some examples of the application of projection operators to problems in quantum detection.

A. Quantum Detection with Binary Decisions

Consider the problem of trying to determine which of two possible quantum states is present, modeling for example the reception of electromagnetic fields in a communications receiver. Under hypothesis H_0, the received electromagnetic field is in a mixture of states governed by the density operator ρ_0, and, under hypothesis H_1, it is governed by ρ_1. Assume equal a priori probabilities, so that $P(H_0) = P(H_1) = 1/2$. Suppose a receiver measures an operator $\mathbf{B}$ by applying the "detection operators" [5]:

$$\left.\begin{aligned} \Pi_0 &= \sum_{n:b_n \in I_0} |b_n\rangle\langle b_n| \\ \Pi_1 &= \sum_{n:b_n \in I_1} |b_n\rangle\langle b_n| \end{aligned}\right\} \tag{21}$$

which effectively partition the outcome space R into two disjoint regions, I_0 and I_1, such that $R = I_0 \cup I_1$ and $I_0 \cap I_1 = \emptyset$. In other words, detection operators are projection operators that partition the outcome space into an exhaustive set of disjoint decision regions. Since for binary detection the outcome space contains only two regions, $\Pi_0 + \Pi_1 = \mathbf{1}$, only one of the detection operators needs to be applied. If the outcome of the measurement falls within the region I_1, H_1 is selected; otherwise, the receiver chooses H_0. The detection and false-alarm probabilities, Q_d and Q_0, respectively, can be expressed in terms of either detection operator. Selecting Π_1, the detection and false-alarm probabilities can be expressed as

$$\left.\begin{aligned} Q_d &= \Pr(\text{choosing } H_1|\rho_1) = \sum_{n:b_n \in I_1} \langle b_n|\rho_1|b_n\rangle = \mathrm{Tr}\left[\rho_1\Pi_1\right] \\ Q_0 &= \Pr(\text{choosing } H_1|\rho_0) = \sum_{n:b_n \in I_1} \langle b_n|\rho_0|b_n\rangle = \mathrm{Tr}\left[\rho_0\Pi_1\right] \end{aligned}\right\} \tag{22}$$

and the conditional probabilities of a correct decision are

$$\left.\begin{aligned} P(C|H_1) &= Q_d = \mathrm{Tr}\left[\rho_1\Pi_1\right] \\ P(C|H_0) &= 1 - Q_0 = 1 - \mathrm{Tr}\left[\rho_0\Pi_1\right] \end{aligned}\right\} \tag{23}$$

It follows that, for equally probable signals, the average probability of correct decision is

$$\begin{aligned} P(C) &= \frac{1}{2}\left\{\mathrm{Tr}\left[\rho_1\Pi_1\right] + 1 - \mathrm{Tr}\left[\rho_0\Pi_1\right]\right\} \\ &= \frac{1}{2}\Big\{1 + \mathrm{Tr}\big[(\rho_1 - \rho_0)\Pi_1\big]\Big\} \end{aligned} \tag{24}$$

and the average probability of error can be expressed as

$$P(E) = 1 - P(C) = \frac{1}{2}\Big\{1 - \mathrm{Tr}\big[(\rho_1 - \rho_0)\Pi_1\big]\Big\} \tag{25}$$

It is clear from the form of Eq. (25) that the probability of error is minimized by choosing the projection operator Π_1 to maximize the quantity $\mathrm{Tr}\left[(\rho_1 - \rho_0)\Pi_1\right]$.

B. The Optimum Measurement for Binary Detection

Helstrom demonstrated that the detection operator that maximizes the quantity $\mathrm{Tr}\,[(\rho_1 - \rho_0)\Pi_1]$ is of the form [7]

$$\Pi^* = \sum_{k:\eta_k \geq 0} |\eta_k\rangle\langle\eta_k| \tag{26}$$

where $|\eta_k\rangle$ are the eigenvectors and η_k the eigenvalues of the operator $(\rho_1 - \rho_0)$, that is,

$$(\rho_1 - \rho_0)|\eta_k\rangle = \eta_k|\eta_k\rangle \tag{27}$$

The probability of error for this optimum projector is

$$\begin{aligned} P^*(E) &= \frac{1}{2}\Big\{1 - \mathrm{Tr}\big[(\rho_1 - \rho_0)\Pi^*\big]\Big\} \\ &= \frac{1}{2}\left\{1 - \sum_k \langle\eta_k|(\rho_1 - \rho_0)\Pi^*|\eta_k\rangle\right\} \\ &= \frac{1}{2}\left\{1 - \sum_k \eta_k\langle\eta_k|\Pi^*|\eta_k\rangle\right\} \\ &= \frac{1}{2}\left\{1 - \sum_k \eta_k \sum_{m:\eta_m>0} \langle\eta_k|\eta_m\rangle\langle\eta_m|\eta_k\rangle\right\} \\ &= \frac{1}{2}\left\{1 - \sum_{k:\eta_k>0} \eta_k\right\} \end{aligned} \tag{28}$$

as shown in [2].

For "pure states" represented by $|\psi_0\rangle$ and $|\psi_1\rangle$, the density operators are $\rho_0 = |\psi_0\rangle\langle\psi_0|$ and $\rho_1 = |\psi_1\rangle\langle\psi_1|$. The eigenvectors of the operator $(\rho_1 - \rho_0)$ are now linear combinations of $|\psi_0\rangle$ and $|\psi_1\rangle$, yielding an expression of the form

$$|\eta_k\rangle = z_{k0}|\psi_0\rangle + z_{k1}|\psi_1\rangle, \quad k = 0, 1 \tag{29}$$

where the coefficients are determined by substituting into Eq. (27) and equating coefficients. For the binary pure-state problem, there is one positive and one negative eigenvalue. The optimum projection operator is simply $\Pi^* = |\eta_1\rangle\langle\eta_1|$, and its application yields the detection and false-alarm probabilities [8]

$$Q_d = \mathrm{Tr}\left[\rho_1 \Pi^*\right] = \langle \eta_1 | \rho_1 | \eta_1 \rangle = \langle \eta_1 | \psi_1 \rangle \langle \psi_1 | \eta_1 \rangle$$

$$= |\langle \psi_1 | \eta_1 \rangle|^2 = |z_{10} \langle \psi_1 | \psi_0 \rangle + z_{11}|^2$$

$$= \frac{1}{2}\left[1 + \sqrt{1 - |\langle \psi_1 | \psi_0 \rangle|^2}\right] \tag{30}$$

$$Q_0 = \mathrm{Tr}\left[\rho_0 \Pi^*\right] = \frac{1}{2}\left[1 - \sqrt{1 - |\langle \psi_1 | \psi_0 \rangle|^2}\right] \tag{31}$$

The probability of correct detection becomes

$$P^*(C) = \frac{1}{2}\left[Q_d + (1 - Q_0)\right]$$

$$= \frac{1}{2}\left[\frac{1}{2} + \frac{1}{2}\sqrt{1 - |\langle \psi_1 | \psi_0 \rangle|^2} + 1 - \frac{1}{2} + \frac{1}{2}\sqrt{1 - |\langle \psi_1 | \psi_0 \rangle|^2}\right]$$

$$= \frac{1}{2}\left[1 + \sqrt{1 - |\langle \psi_1 | \psi_0 \rangle|^2}\right] \tag{32}$$

while the probability of error is

$$P^*(E) = 1 - P^*(C) = \frac{1}{2}\left[1 - \sqrt{1 - |\langle \psi_1 | \psi_0 \rangle|^2}\right] \tag{33}$$

Thus, for the case of equally likely binary decisions, the error probability associated with the optimum projection measurement depends on the overlap between the two quantum states.

The above derivation implicitly assumes that quantum states corresponding to different hypotheses are linearly independent. It has been shown [5] that, when this is the case, and when minimization of the average error probability is the goal, the optimum quantum measurement corresponds to finding the eigenvalues of projection operators. However, projection operators are a subset of a larger class of probability operator measures (p.o.m.'s) that describe the outcome of arbitrary quantum measurements, assigning a probability $\Pr(\Delta) = \mathrm{Tr}\left[\rho \Pi(\Delta)\right]$ to the outcome that the measurement yields a value within a region Δ of the spectrum of eigenvalues. We shall not consider these generalized measurements here, but focus instead on the minimization of the average error probability among linearly independent states, which can be described in terms of projection operators.

VI. Quantum Optical Communications

We begin by describing the quantum mechanical representation of a single mode of a coherent optical field, which may be modulated in various ways in order to carry information from the transmitter to the receiver in an optical communications system.

A. The Coherent-State Representation

Coherent states, representing coherent radiation produced by lasers, are an important set of states for optical communications. It has been shown [6] that the coherent states of a single mode of radiation

$|\alpha\rangle$ can be expressed in the form of a superposition of orthonormal eigenstates $|n\rangle$, known as number eigenstates

$$|\alpha\rangle = e^{-(1/2)|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{(n!)^{1/2}} |n\rangle \tag{34}$$

Each number eigenstate $|n\rangle$ contains exactly n photons (see Appendix). The density operator associated with a coherent state is of the form

$$\rho = |\alpha\rangle\langle\alpha| = e^{-|\alpha|^2} \sum_n \sum_m \frac{\alpha^n (\alpha^*)^m}{(n!m!)^{1/2}} |n\rangle\langle m| \tag{35}$$

When ρ is viewed as an infinite dimensional matrix, the diagonal terms are

$$\langle n|\rho|n\rangle = e^{-|\alpha|^2} \frac{|\alpha|^{2n}}{n!} \tag{36}$$

which are recognized as Poisson probabilities for n, with an average number of photons equal to $|\alpha|^2$. The off-diagonal terms are

$$\langle m|\rho|n\rangle = e^{-|\alpha|^2} \frac{\alpha^n (\alpha^*)^m}{(n!m!)^{1/2}} \tag{37}$$

Coherent states are not orthogonal, as can be seen by considering the overlap between two arbitrary coherent states, $|\alpha\rangle$ and $|\beta\rangle$. Orthogonality of states requires that their overlap vanish; however, for coherent states, the squared magnitude of the overlap is

$$\begin{aligned} |\langle\alpha|\beta\rangle|^2 &= \left| e^{-(|\alpha|^2+|\beta|^2)/2} \sum_n \sum_m \frac{\alpha^n}{\sqrt{n!}} \frac{(\beta^*)^m}{\sqrt{m!}} \langle n|m\rangle \right|^2 \\ &= \left| e^{-(|\alpha|^2+|\beta|^2)/2} \sum_n \frac{(\alpha\beta^*)^n}{n!} \right|^2 = \left| e^{-(|\alpha|^2+|\beta|^2-2\alpha\beta^*)/2} \right|^2 = e^{-|\alpha-\beta|^2} \end{aligned} \tag{38}$$

by virtue of the orthogonality of the number states. Equation (38) demonstrates that there is always some overlap between coherent states, regardless of how great the average photon count in each state may be.

B. Classical Derivation of Photon-Counting Receiver (On–Off Keying)

Suppose there are two hypotheses, H_0 and H_1, denoting absence and presence of signal, respectively. In the absence of background radiation, either no photons or an average of $\lambda = |\alpha|^2 > 0$ photons in the signal pulse is received. The received field is assumed to be from a coherent laser, hence the number of photons is Poisson distributed with conditional densities

$$\left.\begin{aligned} P(n|H_0) &= \begin{cases} 1, & n = 0 \\ 0, & n \geq 1 \end{cases} \\ P(n|H_1) &= \frac{\lambda^n}{n!} e^{-\lambda} \end{aligned}\right\} \tag{39}$$

At the end of each signaling interval, the receiver records the total number of detected photons and decides which hypothesis is true by computing the two likelihood functions, $\Lambda_i \equiv P(n|H_i), i = 0, 1$, and selecting the hypothesis corresponding to the larger of the two. The likelihood functions under the two hypotheses are given by Eq. (39), where it can be seen that, if no photons are detected, $n = 0$, then $\Lambda_0 = 1$, and $\Lambda_1 = e^{-\lambda} < 1$ for $\lambda > 0$. Therefore, $\Lambda_0 > \Lambda_1$ and H_0 is selected whenever the photon count is zero. If one or more photons are observed, $n \geq 1$ and $\Lambda_1 > \Lambda_0$, in which case H_1 is selected.

In the absence of noise, H_0 is always decoded correctly, so $P(C|H_0) = P(0|H_0) = 1$. If at least one photon is detected, H_1 is decoded correctly: $P(n \geq 1|H_1) = 1 - e^{-\lambda}$. With equal a priori probabilities, $P(H_0) = P(H_1) = 1/2$, the probability of correct detection becomes

$$P(C) = \sum_{i=1}^{2} P(C|H_i)P(H_i) = 1 - \frac{1}{2}e^{-\lambda} \tag{40}$$

Finally, the average probability of error is $P(E) = 1 - P(C) = (1/2)e^{-\lambda}$.

C. Quantum Mechanical Derivation of Photon-Counting Receiver (On–Off Keying)

In the quantum mechanical formulation, the received field is in one of two states, $|\psi_0\rangle = |0\rangle$ or $|\psi_1\rangle$, corresponding to the hypotheses H_0 and H_1. The signal field is assumed to be in a pure coherent state, which can be expressed in the number representation as [6]

$$|\psi_1\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \tag{41}$$

where α is a complex number, the set $\{|n\rangle\}$ is the number eigenstates, and $|\alpha|^2$ again represents the average number of photons in the signal. The density operators for these two states are of the form

$$\left.\begin{aligned} \rho_0 &= |\psi_0\rangle\langle\psi_0| = |0\rangle\langle 0| \\ \rho_1 &= |\psi_1\rangle\langle\psi_1| \end{aligned}\right\} \tag{42}$$

The likelihood functions defined in Eq. (39) are based on the number of observed photons. Since counting photons represents a measurement of energy, the eigenvalues of the photon-counting operator yield measurements of energy. As shown in the Appendix, the number eigenstates of the harmonic oscillator form a complete set of energy eigenstates whose eigenvalues represent the exact number of photons in a given eigenstate. A measurement that determines whether or not the received state is the ground state corresponds to an application of the detection operators Π_0 and Π_1, defined as

$$\left.\begin{aligned} \Pi_0 &= |0\rangle\langle 0| \\ \Pi_1 &= \sum_{n=1}^{\infty} |n\rangle\langle n| = \mathbf{1} - |0\rangle\langle 0| \end{aligned}\right\} \tag{43}$$

where $\mathbf{1}$ is the identity operator in the number state basis. When the projection operator Π_0 is applied to the ground state or "vacuum state" (the state of the received field under hypothesis H_0), the ground state is recovered,

$$\Pi_0|\psi_0\rangle = |0\rangle\langle 0|0\rangle = |0\rangle \equiv |w_0\rangle \tag{44}$$

where we define $|w_0\rangle$ as a "measurement state" corresponding to $|\psi_0\rangle$. The probability of observing the eigenvalue 0, given that H_0 is true, is

$$\Pr(0|H_0) = \mathrm{Tr}(\rho_0\Pi_0) = \langle 0|\Pi_0|0\rangle = |\langle w_0|\psi_0\rangle|^2 = 1 \tag{45}$$

which can be interpreted as the squared magnitude of the projection of $|\psi_0\rangle$ onto $|w_0\rangle$.

When the received field is in a coherent state, application of the projection operator Π_1 yields

$$\begin{aligned} \Pi_1|\psi_1\rangle &= \sum_{n=1}^{\infty} |n\rangle\langle n| \left(e^{-|\alpha|^2/2} \sum_{m=0}^{\infty} \frac{\alpha^m}{\sqrt{m!}} |m\rangle \right) \\ &= e^{-|\alpha|^2/2} \sum_{m=0}^{\infty} \sum_{n=1}^{\infty} \frac{\alpha^m}{\sqrt{m!}} |n\rangle\langle n|m\rangle \\ &= e^{-|\alpha|^2/2} \sum_{n=1}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \equiv |W_1\rangle \end{aligned} \tag{46}$$

We now define the normalized version of $|W_1\rangle$ as $|w_1\rangle$, and interpret it as a measurement state for $|\psi_1\rangle$. Since $\langle w_0|W_1\rangle = e^{-|\alpha|^2/2} \sum_{n=1}^{\infty} (\alpha^n/\sqrt{n!})\langle 0|n\rangle = 0$, it follows that the measurement states $|w_0\rangle$ and $|w_1\rangle$ are orthonormal. The probability of obtaining an eigenvalue greater than 0 when observing the received coherent state is

$$\Pr(n \geq 1|H_1) = \mathrm{Tr}(\rho_1 \Pi_1) = \langle \psi_1|\Pi_1|\psi_1\rangle = |\langle w_1|\psi_1\rangle|^2$$

$$= \left(e^{-|\alpha|^2/2} \sum_{m=0}^{\infty} \frac{(\alpha^*)^m}{\sqrt{m!}} \langle m| \right) \sum_{n=1}^{\infty} |n\rangle\langle n| \left(e^{-|\alpha|^2/2} \sum_{i=0}^{\infty} \frac{\alpha^i}{\sqrt{i!}} |i\rangle \right)$$

$$= e^{-|\alpha|^2} \sum_{m=0}^{\infty} \frac{(\alpha^*)^m}{\sqrt{m!}} \langle m| \sum_{n=1}^{\infty} |n\rangle\langle n| \sum_{i=0}^{\infty} \frac{\alpha^i}{\sqrt{i!}} |i\rangle$$

$$= e^{-|\alpha|^2} \sum_{m=0}^{\infty} \sum_{n=1}^{\infty} \sum_{i=0}^{\infty} \frac{(\alpha^*)^m}{\sqrt{m!}} \frac{\alpha^i}{\sqrt{i!}} \langle m|n\rangle\langle n|i\rangle$$

$$= e^{-|\alpha|^2} \sum_{n=1}^{\infty} \frac{|\alpha|^{2n}}{n!} = e^{-|\alpha|^2} \left(e^{|\alpha|^2} - 1 \right) = 1 - e^{-|\alpha|^2} \tag{47}$$

Again interpreting this as the squared magnitude of the projection of $|\psi_1\rangle$ onto $|w_1\rangle$, it follows that $P(C) = \sum_{i=1}^{2} P(C|H_i)P(H_i) = 1 - (1/2)e^{-\lambda}$, yielding an average probability of error $P(E) = 1 - P(C) = (1/2)e^{-\lambda}$, the same as that computed from the classical derivation.

VII. State–Space Interpretation of the Quantum Optimum Receiver for the Binary Pure-State Detection Problem

The two signal states characterizing the binary two-state problem define a plane in Hilbert space. A graphical interpretation of the signal states in the subspace defined by the first three number states is shown in Fig. 1. Although the following arguments apply in a more general context [5], here we restrict our attention to the example of optical on–off keying worked out in the previous section because of the physical insights it provides. Continuing with the interpretation of measurement states defined above, we express the kets produced by the detection operators Π_0 and Π_1 acting on the signal states $|\psi_0\rangle$ and $|\psi_1\rangle$ as $|W_0\rangle = \Pi_0|\psi_0\rangle$ and $|W_1\rangle = \Pi_1|\psi_1\rangle$. After normalization, the states produced by the application of the photon-counting projection operators to the signal states may be referred to as measurement

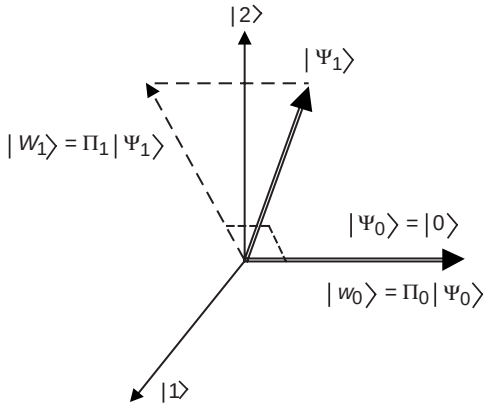

Fig. 1. Projection of signal states onto three-dimensional subspace defined by the first three number states: on–off keying.

states spanning the two-dimensional subspace defined by the signal states. These measurement states are orthonormal, and, for optimum photon counting, one of the detection operators lines up with the ground state representing the null hypothesis: therefore, the ground state is always detected correctly. However, since the signal state is not orthogonal to the ground state, it projects a nonzero component onto both detection operators: its projection onto the ground state is the mechanism by which detection errors occur in the reception of optical on–off keyed signals.

Operating entirely in the plane determined by the signal states, it is possible to find the minimum average probability of error by rotating the measurement states within the signal plane and calculating the error probability for each rotation until a minimum is reached. Again considering the example of optical on–off signaling, we represent the plane defined by two signal states, and containing two orthogonal measurement states, as in Fig. 2. The angle θ is defined as $\theta = \cos^{-1}(|\langle\psi_1|\psi_0\rangle)$ and represents the overlap between the two signal states.

It is clear that the conditional probabilities of correct detection are given by

$$\left.\begin{aligned} P(C|H_0) &= |\langle\psi_0|w_0\rangle|^2 \equiv \cos^2(\phi_0) \\ P(C|H_1) &= |\langle\psi_1|w_1\rangle|^2 \equiv \cos^2(\phi_1) \end{aligned}\right\} \tag{48}$$

where $0 \leq \varphi_0$, $\varphi_1 \leq \pi/2$, and we interpret $\cos(\varphi_i) = |\langle\psi_i|w_i\rangle|$, $i = 0, 1$, as the overlap between the signal and corresponding measurement states. Note from Fig. 2 that $\varphi_0 = \pi/2 - \theta - \varphi_1$. With equally probable signals, the average probability of correct detection is

$$\begin{aligned} P(C) &= \frac{1}{2}\left[|\langle\psi_0|w_0\rangle|^2 + |\langle\psi_1|w_1\rangle|^2\right] \\ &= \frac{1}{2}\left[\cos^2(a - \varphi_1) + \cos^2(\varphi_1)\right] \end{aligned} \tag{49}$$

where we define $a \equiv \pi/2 - \theta$ for convenience. The maximum value of the probability of correct detection, as a function of the rotation angle φ_1, can be easily found by differentiating $P(C)$ with respect to φ_1 and equating to zero:

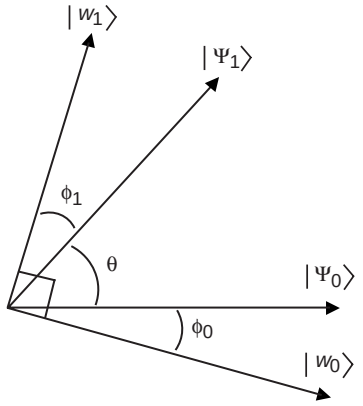

Fig. 2. Signal and measurement states for the binary pure-state problem.

$$\frac{\partial P(C)}{\partial \varphi_1} = -2\cos(a-\varphi_1)\sin(a-\varphi_1) + 2\cos(\varphi_1)\sin(\varphi_1) = 0 \tag{50}$$

This yields the optimum rotation angle $\varphi_1^* = a/2 = 1/2([\pi/2] - \theta)$. Substituting φ_1^* into the expression for $P(C)$ yields the maximum value of the probability of correct detection as

$$\begin{aligned} P^*(C) &= \frac{1}{2}\left[\cos^2\left(\frac{a}{2}\right) + \cos^2\left(\frac{a}{2}\right)\right] \\ &= \frac{1}{2}\big[1+\cos(a)\big] = \frac{1}{2}\big[1+\sin(\theta)\big] \\ &= \frac{1}{2}\left[1+\sqrt{1-|\langle\psi_0|\psi_1\rangle|^2}\right] \end{aligned} \tag{51}$$

where we let $\sin(\theta) = \sqrt{1-\cos^2(\theta)} = \sqrt{1-|\langle\psi_0|\psi_1\rangle|^2}$ in the last step.

Using this maximum value of $P(C)$, the minimum probability of error becomes

$$P^*(E) = 1 - P^*(C) = \frac{1}{2}\left[1-\sqrt{1-|\langle\psi_0|\psi_1\rangle|^2}\right] \tag{52}$$

which is seen to be identical to the error probability for the optimum quantum receiver in Eq. (33). Therefore, the optimum receiver can be viewed as a simple rotation of the measurement states in the plane defined by the signal states.

A. Numerical Performance Analysis for Optimizing the Quantum Measurement

The behavior of the error probability with rotation within the subspace defined by the signal states can also be addressed using numerical techniques. A Matlab script was written to generate the error-probability curves as a function of the angle φ_1 defined above. The controllable parameters are the number of photons and the number of terms to use in the vector. For on–off keying, the zero state can be approximated as a 1×10 vector of the form

$$|0\rangle_{10} = (\,1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0\,)^T$$

To generate a finite vector approximation to the coherent state vector, the following equation can be used, as in Eq. (22):

$$|\alpha\rangle_N = e^{-(1/2)|\alpha|^2} \sum_0^{N-1} \frac{\alpha^n}{(n!)^{1/2}} |n\rangle \tag{53}$$

For a 1×10 coherent state with 1 photon, this approximation yields the vector

$$|\alpha\rangle_{10} = (\,0.6065 \quad 0.6065 \quad 0.4289 \quad 0.2476 \quad 0.1238 \quad 0.0554 \quad 0.0226 \quad 0.0085 \quad 0.0030 \quad 0.0010\,)^T$$

which is not normalized and hence does not strictly represent a state vector.

An approximation to the magnitudes of the two states was calculated along with the angle between them. A new set of axes was defined, where one axis was lined up with the zero vector and the other orthogonal to it, but in the plane defined by the two original signal states. The projections of the state vectors on the new axes were measured and the probability of error calculated. Next, the new axis was rotated by one degree and the probability of error recalculated for the new axis orientation. This process was repeated until the second axis lined up with the coherent state vector, and the results are displayed in Fig. 3. Note that for a given state the error probability is symmetric around the optimum angle φ^*, in agreement with the analysis.

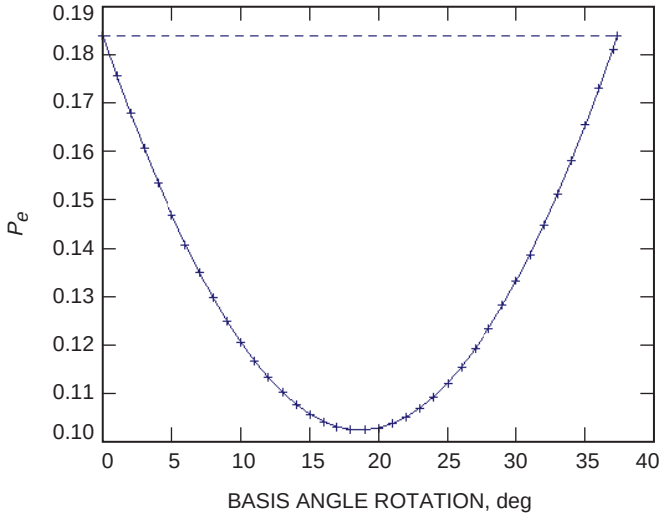

Fig. 3. Error probability as a function of measurement state rotation.

VIII. Examples of Optical Modulation Formats

The following are some specific examples of optical modulation formats that have been considered in the literature and are of potential interest for deep-space applications.

A. Orthogonal Signal States

With orthogonal signal states, $\langle\psi_0|\psi_1\rangle = 0$. Substituting into Eq. (51), we find that $P^*(E) = 0$. Therefore, truly error-free communication could be achieved if practical orthogonal signal states could be generated—for example, if the number states $|\psi_0\rangle = |0\rangle$ and $|\psi_1\rangle = |1\rangle$ could somehow be prepared.

B. On–Off Keying

For on–off keying (OOK), we let

$$|\psi_0\rangle = |0\rangle$$

$$|\psi_1\rangle = e^{-(1/2)|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{(n!)^{1/2}} |n\rangle$$

where the average number of photons is $|\alpha|^2$ for the optical pulse and zero for the ground state: the average number of received photons, averaged over the signal distribution with this modulation format, is therefore

$K_s = (1/2)|\alpha|^2$. The squared magnitude of the overlap between the two states is $|\langle\psi_0|\psi_1\rangle|^2 = e^{-|\alpha|^2}$, yielding $P^*(E) = (1/2)[1 - \sqrt{1 - e^{-|\alpha|^2}}] = (1/2)[1 - \sqrt{1 - e^{-2K_s}}]$ for optimum quantum detection. In comparison, the error probability for photon-counting detection was shown to be $P(E) = (1/2)e^{-|\alpha|^2} = (1/2)e^{-2K_s}$.

C. Optical Binary Phase-Shift Keying (BPSK)

In this case, we have two coherent states with the same average photon energy but π radians out of phase. The signal states are defined as $|\psi_0\rangle = |\alpha\rangle$ and $|\psi_1\rangle = |\beta\rangle$, where the complex amplitudes are related as $\beta = -\alpha$ and the average number of signal photons is $K_s = |\alpha|^2$. Therefore, $|\langle\psi_0|\psi_1\rangle|^2 = e^{-4K_s}$, and the error probability becomes $P^*(E) = (1/2)[1 - \sqrt{1 - e^{-4K_s}}]$. A physically realizable receiver structure that achieves this error probability has been devised and analyzed by Dolinar [9].

1. Near-Optimum Optical BPSK. Exponentially optimum (or "near-optimum") performance can be obtained by adding a local field of the same amplitude, in phase, to the received field followed by photon counting, as shown by Kennedy [10]. With this technique, the field amplitude under one hypothesis is shifted to the ground state, but doubled under the other. The error probability for this near-optimum detection scheme is $P(E) = (1/2)e^{-4|\alpha|^2} = (1/2)e^{-4K_s}$.

2. Coherent Detection: the Classical Limit. If we add a local field of great amplitude, in phase, to the received field and detect the resulting sum field using classical energy detection, then the performance of the classical coherent optical detector is obtained. For optical BPSK modulation, the error probability for this receiver is given by [5]

$$P(E) = Q\left(\sqrt{4|\alpha|^2}\right) = Q\left(\sqrt{4K_s}\right)$$

where

$$Q(x) \equiv \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-y^2/2} dy$$

Performance curves for the various modulation formats and detection options are shown in Fig. 4.

IX. Optimum Binary Detection in the Presence of Background Radiation

When background radiation is present, the received field is in a mixture of states described by an appropriate density operator. As described in Subsection 5.B, the probability of error can be calculated by finding the significant eigenvalues of the difference operator $(\rho_1 - \rho_0)$ and applying the formula for the error probability given by Eq. (28). Equivalently, the matrix representation of the difference operator can be diagonalized, the projection operator that selects only the positive eigenvalues applied, and the trace of the resulting diagonal matrix determined [2].

The following analysis refers to on–off keying with noise. For this scenario, a density matrix ρ_1 can be defined for the signal and noise case. The controllable values are the size of the matrix, average number of noise photons, and average number of signal photons. From [2], the elements of this density matrix in the number state basis are

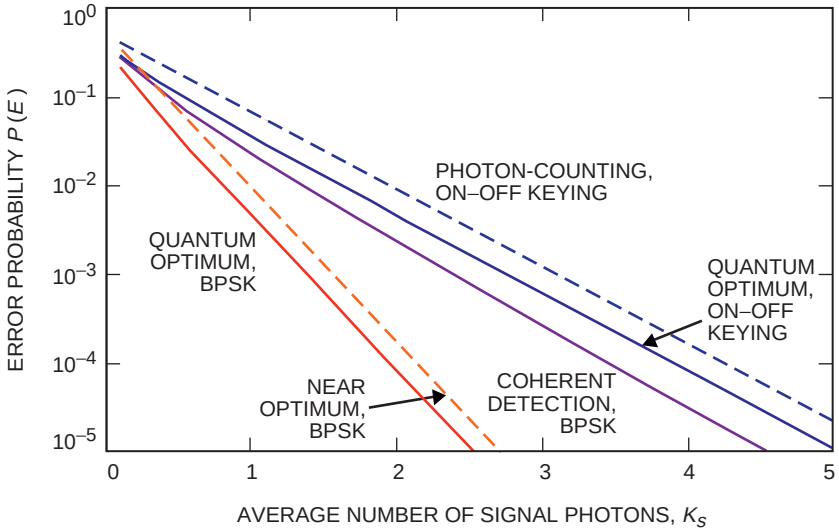

Fig. 4. Performance of binary optical receivers: no background radiation.

$$\left.\begin{aligned} \langle n|\rho_1|m\rangle &= (1-v)\left(\frac{n!}{m!}\right)^{1/2} v^m \left(\frac{\mu^*}{N}\right)^{m-n} \exp^{(-(1-v)|\mu|^2)} L_n^{m-n}[-(1-v)\frac{|\mu|^2}{v}], \quad m \geq n \\ \langle n\,|\rho_1|\,m\rangle &= \langle m\,|\rho_1|\,n\rangle^*, \quad m < n \\ v &= \frac{N}{(N+1)} \end{aligned}\right\} \tag{54}$$

where N represents the average number of noise photons, μ is the complex envelope of the signal, $|\mu|^2$ is the average number of signal photons, and $L_n^{m-n}(x)$ is a Laguerre polynomial. Although matrices as large as 20×20 were used in the calculations, an example of a density matrix with the first 4×4 elements filled in is

$$\left[\langle m|\rho_1|n\rangle\right] = \begin{pmatrix} 0.6921 & 0.3446 & 0.1213 & 0.0349 & \\ 0.3446 & 0.2345 & 0.1047 & 0.0365 & \cdots \\ 0.1213 & 0.1047 & 0.0582 & 0.0245 & \\ 0.0349 & 0.0365 & 0.0245 & 0.0123 & \\ & & \vdots & & \ddots \end{pmatrix}$$

with $|\mu|^2 = 1$ and $N = 0.1$ photons on the average.

A density operator ρ_0 for the null hypothesis, or noise-only case, can be generated using the corresponding equation from [2],

$$\rho_0 = \sum_{n=0}^{\infty} (1-v)v^n \,|n\rangle\,\langle n| \tag{55}$$

The corresponding matrix example for $N = 0.1$ is

$$\left[\langle m|\rho_0|n\rangle\right] = \begin{pmatrix} 0.9091 & 0 & 0 & 0 & \\ 0 & 0.0826 & 0 & 0 & \cdots \\ 0 & 0 & 0.0075 & 0 & \\ 0 & 0 & 0 & 0.0007 & \\ & & \vdots & & \ddots \end{pmatrix}$$

The noise matrix is subtracted from the signal-plus-noise matrix to yield

$$\left[\langle m|\rho_1 - \rho_0|n\rangle\right] = \begin{pmatrix} -0.2170 & 0.3446 & 0.1213 & 0.0349 & \\ 0.3446 & 0.1519 & 0.1047 & 0.0365 & \cdots \\ 0.1213 & 0.1047 & 0.0507 & 0.0245 & \\ 0.0349 & 0.0365 & 0.0245 & 0.0117 & \\ & & \vdots & & \ddots \end{pmatrix}$$

Next the resulting difference matrix is diagonalized, yielding

$$\left[\langle \eta_k|\rho_1 - \rho_0|\eta_k\rangle\right] = \begin{pmatrix} -0.0085 & 0 & 0 & 0 & \\ 0 & 0.0064 & 0 & 0 & \cdots \\ 0 & 0 & -0.4292 & 0 & \\ 0 & 0 & 0 & 0.4286 & \\ & & \vdots & & \ddots \end{pmatrix}$$

The probability of error is found from this diagonalized matrix. The probability of correct detection is found by adding all the positive diagonal terms, yielding the probability of error,

$$P(E) = 1 - P(C) = \frac{1}{2}\left\{1 - \sum_{k:\eta_k>0} \eta_k\right\} \tag{56}$$

The probability of error was computed for several values of average noise photons and is shown in Fig. 5.

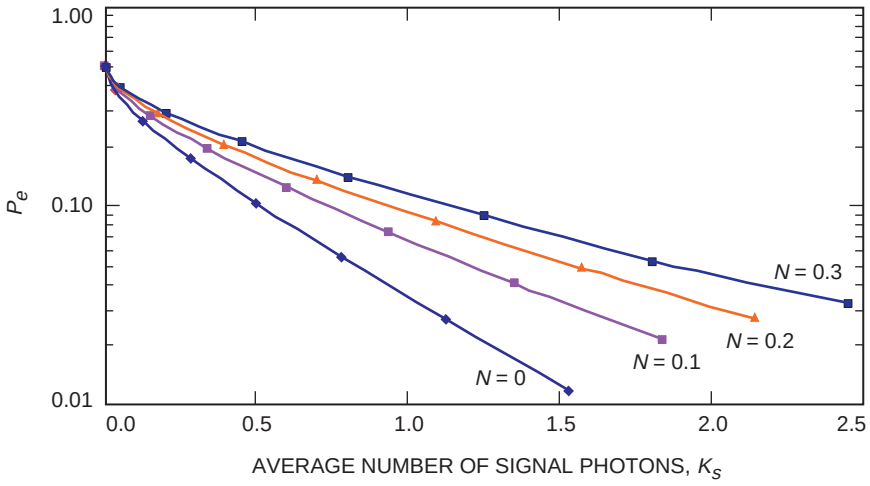

Fig. 5. Performance of quantum optimum receiver in the presence of background radiation: on–off keying.

The algorithm for computing the error probability of BPSK-modulated signals in the presence of noise is similar to the on–off keying algorithm discussed above. The difference is that now both density operators, ρ_1 and ρ_0, contain signal plus noise and are given by the right side of Eq. (54) with $\mu_1 = +\alpha$ and $\mu_0 = -\alpha$, respectively. The rest of the calculations are exactly the same as for the on–off keying case, yielding the results shown in Fig. 6.

The performance of classical coherently detected BPSK signals observed in the presence of noise is also shown in Fig. 6 for comparison. This technique is called "threshold detection" in [5], with error probability $P(E) = Q(\sqrt{4K_s/(2N+1)})$, which is seen to be similar to the expression for the noiseless case, but with K_s replaced by $K_s/(2N+1)$. Note that, for high background levels, the performance of the classical threshold detector is nearly as good as that of the optimum quantum detector; hence, the physically realizable classical detector is a good approximation to the optimum quantum detector under conditions of high background radiation.

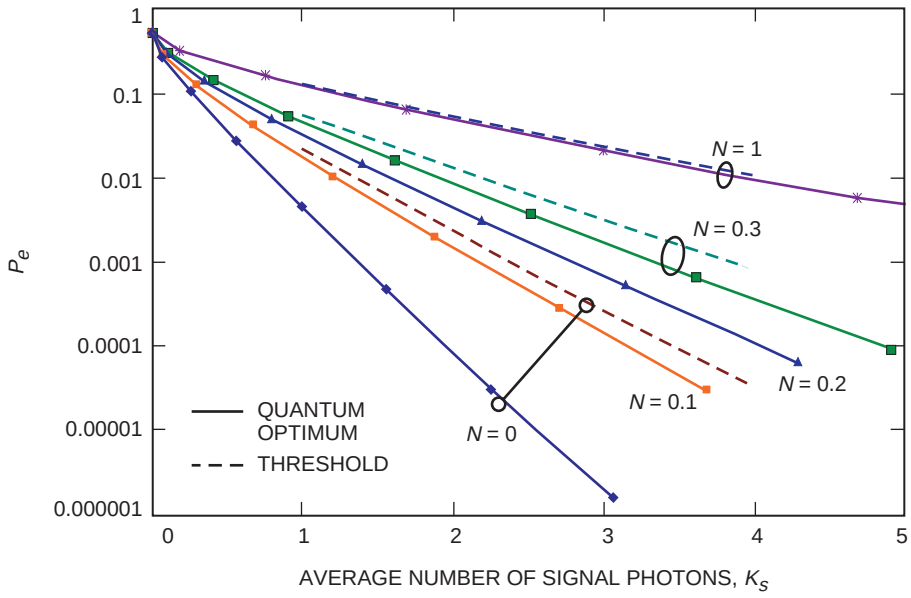

Fig. 6. Performance of quantum optimum and threshold receivers in the presence of background radiation: BPSK signaling.

X. Multiple Hypotheses: Orthogonal Envelopes

Following Helstrom [5], the performance of the optimum quantum receiver for the case of equal-energy, equally probable signals is considered. The signals are assumed to have orthogonal classical envelopes, which means that the classical complex envelopes $S_k(t)$ obey the condition

$$\int_0^T S_k^*(t) S_m(t) dt = 0, \quad k \neq m \tag{57}$$

In this case, there are assumed to be M hypotheses, represented by M orthogonal classical envelopes modulating electromagnetic plane waves normally incident on the receiving aperture, and with temporal variation proportional to $S_k(t)$. An example of this modulation format is M-ary pulse-position modulation (PPM), where a single pulse is placed in one of M consecutive slots. The kth aperture-field mode is assumed to be matched to the kth signal, such that when the kth signal is present, the state of the aperture field is in a coherent state described by the "product-state"

$$|\alpha_k\rangle = \prod_{j=1}^{M} |\alpha_{k,j}\rangle$$

$$= |\alpha_{k,1}\rangle|\alpha_{k,2}\rangle \cdots |\alpha_{k,M}\rangle \tag{58}$$

where each of the $|\alpha_{k,j}\rangle$ are coherent states associated with the individual modes, with $|\alpha_{k,j}| = |\alpha|\delta_{k,j}$, and where $K_s = |\alpha|^2$ is the average number of signal photons in each signal. For product states, the magnitude squared of the overlap is computed according to the rule

$$|\langle\alpha_m|\alpha_k\rangle|^2 = \prod_{j=1}^{M} |\langle\alpha_{m,j}|\alpha_{k,j}\rangle|^2$$

$$= |\langle\alpha_{m,1}|\alpha_{k,1}\rangle|^2 |\langle\alpha_{m,2}|\alpha_{k,2}\rangle|^2 \cdots |\langle\alpha_{m,M}|\alpha_{k,M}\rangle|^2 \tag{59}$$

Using the expression for the magnitude squared of the overlap between two coherent states in Eq. (38), this quantity can be expressed as

$$|\langle\alpha_m|\alpha_k\rangle|^2 = \exp\left(\sum_{j=1}^{M} \left(2\alpha_{mj}^*\alpha_{kj} - \left|\,\alpha_{mj}\,\right|^2 - |\alpha_{kj}|^2\right)\right)$$

$$= \begin{cases} 1, & m = k \\ e^{-2K_s}, & m \neq k \end{cases} \tag{60}$$

The M signals define an M-dimensional subspace. Since the M product states are linearly independent, the optimum strategy for minimizing the average probability of error is to project the received signal state onto M orthonormal measurement states spanning the same subspace [5], and to select the signal corresponding to the measurement state with the greatest projection. The state–space interpretation is similar to the binary case: because of symmetry, the M orthogonal measurement states are aligned with the M signal states in such a way as to maximize the projection of each state onto the corresponding measurement state, thus minimizing the probability of error. It has been shown [11] that for equally likely signals, the minimum symbol-error probability for the optimum quantum receiver is

$$P(SE) = \frac{M-1}{M^2}\left[\sqrt{1+(M-1)e^{-K_s}} - \sqrt{1-e^{-K_s}}\right]^2 \tag{61}$$

For $M = 2$, which may represent binary PPM, the overlap can again be interpreted as the cosine of the angle θ between two states in state–space. For this case, $\cos^2(\theta) = e^{-2K_s}$ from Eq. (60). The probability of correct decision is maximized by rotating the two orthogonal measurement states to be symmetrically placed around the signal states, as indicated in Fig. 7.

The error probability is given by $P(E) = (1/2)[1 - \sqrt{1 - |\langle\psi_1|\psi_2\rangle|^2}]$ as before, but now $|\langle\psi_1|\psi_2\rangle|^2 = e^{-2K_s}$, as compared to e^{-K_s} for the corresponding overlap with on–off keying, and e^{-4K_s} for BPSK signals.

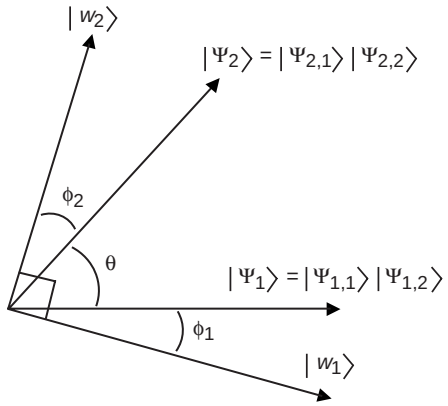

Fig. 7. Signal and measurement states for binary orthogonal modulation.

XI. Numerical Performance Analysis: *M*-ary Decisions

This part of the article describes the simulation of M-ary orthogonal-envelope signals in the absence of noise. The solution for M-ary decisions can be obtained by using an iteration on the lower dimensional optimization. Due to the symmetry of states and measurement axes, the minimum probability of error occurs when the angles between the signal states and the corresponding measurement states are equal. This minimizing angle can be found by first solving for the minimizing angle between signal and measurement states for $M-1$ states. Once the $M-1$ dimension solution is known, the next dimension can be found by rotating away from the Mth dimensional axis while keeping the angles between the $M-1$ signal and measurement states equal to each other. The problem reduces to finding the orientation in the Mth dimension, which yields equal angles for each signal and measurement state pair. Starting with $M=2$, the solution for any Mth dimensional problem can be obtained using this technique.

To illustrate the solution principle, consider the case $M=3$. The angle between each of the signal states can be calculated using the equation

$$\theta = \cos^{-1}\left(\exp(-K_s)\right) \tag{62}$$

where K_s represents the average number of signal photons. Once θ is determined, the signal states and the measurement states are known. The minimum probability of error would occur when the signal states each projected equally onto the corresponding measurement axis [5]. The symmetry of the system makes it possible to start with the optimal solution for the $M=2$ case and apply a third state to the system. The two signal states in the (x,y) plane were initially aligned so the angles between the signal states and the corresponding measurement states were equal, similar to the OOK case discussed before. Initially the measurement states were aligned so the $|w_1\rangle$ measurement state aligned with the x-axis, the $|w_2\rangle$ measurement state aligned with the y-axis, and the $|w_3\rangle$ measurement state aligned with the z-axis. The projection of each signal state onto the proper measurement state was then calculated and the probability of error found using

$$\left.\begin{aligned} P_e &= 1 - P_c \\ P_c &= \frac{1}{3}\left(\langle w_1|\psi_1\rangle^2 + \langle w_2|\psi_2\rangle^2 + \langle w_3|\psi_3\rangle^2\right) \end{aligned}\right\} \tag{63}$$

The measurement axes were then rotated around the line $y = -x$. The reason for using $y = -x$ is to keep the two measurement states that started out in the x–y plane an equal distance from their corresponding measurement states. The probability of error is calculated for each step, and the process is repeated until the $|w_3\rangle$ measurement state aligns with $|\psi_3\rangle$. Figure 8 depicts the states and the rotations.

The dotted black lines with triangle terminators represent the x–y–z axes. The solid blue lines with no terminators represent the rotated measurement states, and the dashed–dotted magenta lines with square terminators represent the signal states. The probability of error for each rotation value was calculated and is shown in Fig. 9 for the case $K_s = 1$. As shown in Fig. 9, the rotation angle that achieves the minimum probability of error is 12.16 degrees. Similar to the binary case discussed earlier, the minimum error probability is achieved when the angles between the signal states and the corresponding measurement states are equal. The minimum probability of error was found for a number of different values of K_s and plotted along with the theoretical results in Fig. 10. The theoretical results were calculated using Eq. (60). The theoretical curve is denoted by the blue line with circles, whereas the numerically generated curve is denoted by the magenta line with X's. The two curves are virtually indistinguishable, verifying that rotation of the measurement axes achieves the performance of the optimum quantum receiver. This technique can be extended to higher dimensions by starting with the solution for a given dimension and effectively rotating it into the next higher dimension.

For comparison, the symbol-error probability for the photon-counting receiver in the absence of background radiation is

$$P(SE) = \frac{M-1}{M} e^{-K_s} \tag{64}$$

Performance comparisons of photon-counting and optimum quantum receivers are shown in Figs. 11(a) through 11(c) for $M = 2, 16$, and 256. A good example of an M-ary modulation format is PPM, where one pulse is placed into one of M consecutive slots leading to $log_2 M$ bits of information. Since each

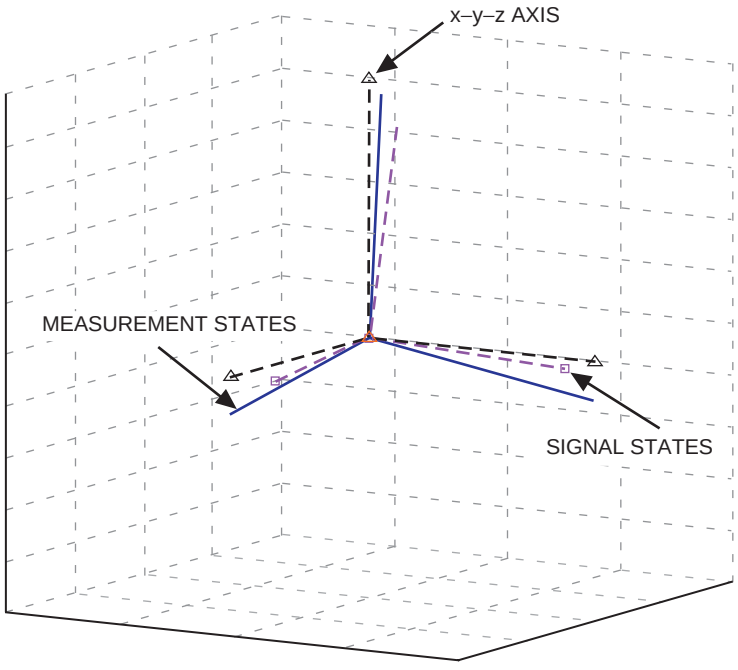

Fig. 8. Three-dimensional representation of signal and measurement states.

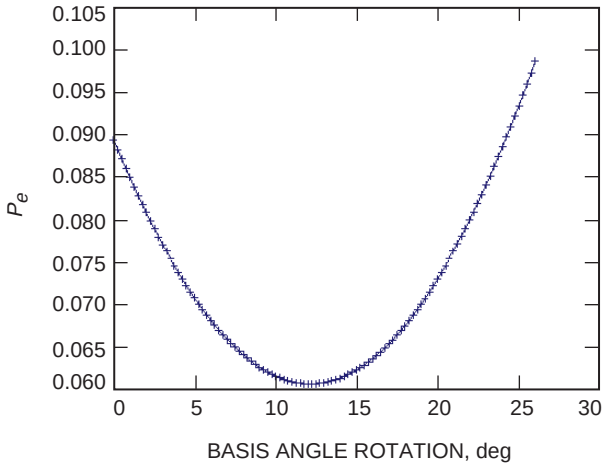

Fig. 9. Error probability as a function of measurement state rotation for the case *M* = 3.

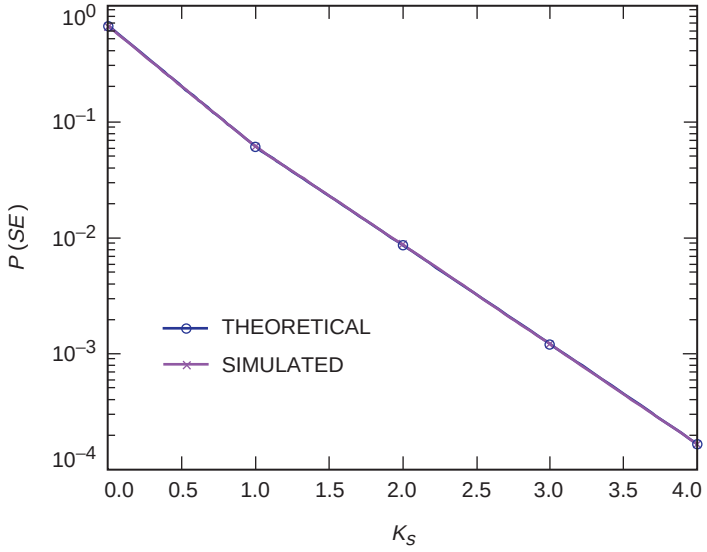

Fig. 10. Theoretical and numerically calculated performance of *M*-ary quantum receiver with *M* = 3.

PPM symbol contains a pulse, the average signal energy and average pulse energy are both equal to K_s, unlike for OOK, where the average symbol energy is half of the average pulse energy. Note that, for high-dimensional signaling and modest error probabilities on the order of $P(SE) \cong 10^{-3}$, conventional photon counting performs approximately 1.3 dB worse than the optimum quantum receiver in the absence of background radiation, but this performance gap increases at the lower error probabilities.

The solution to the M-ary coherent state detection problem in the presence of noise is not known. However, upper bounds to the performance of the optimum quantum receiver have been obtained by evaluating various suboptimum receiver structures, many of which have been detailed in [5]. One such receiver is the M-ary quantum receiver optimum in the absence of noise, but its performance was found to degrade severely with the addition of noise photons, approaching a lower limit determined by the signal dimension and noise energy but essentially independent of the signal energy [5].

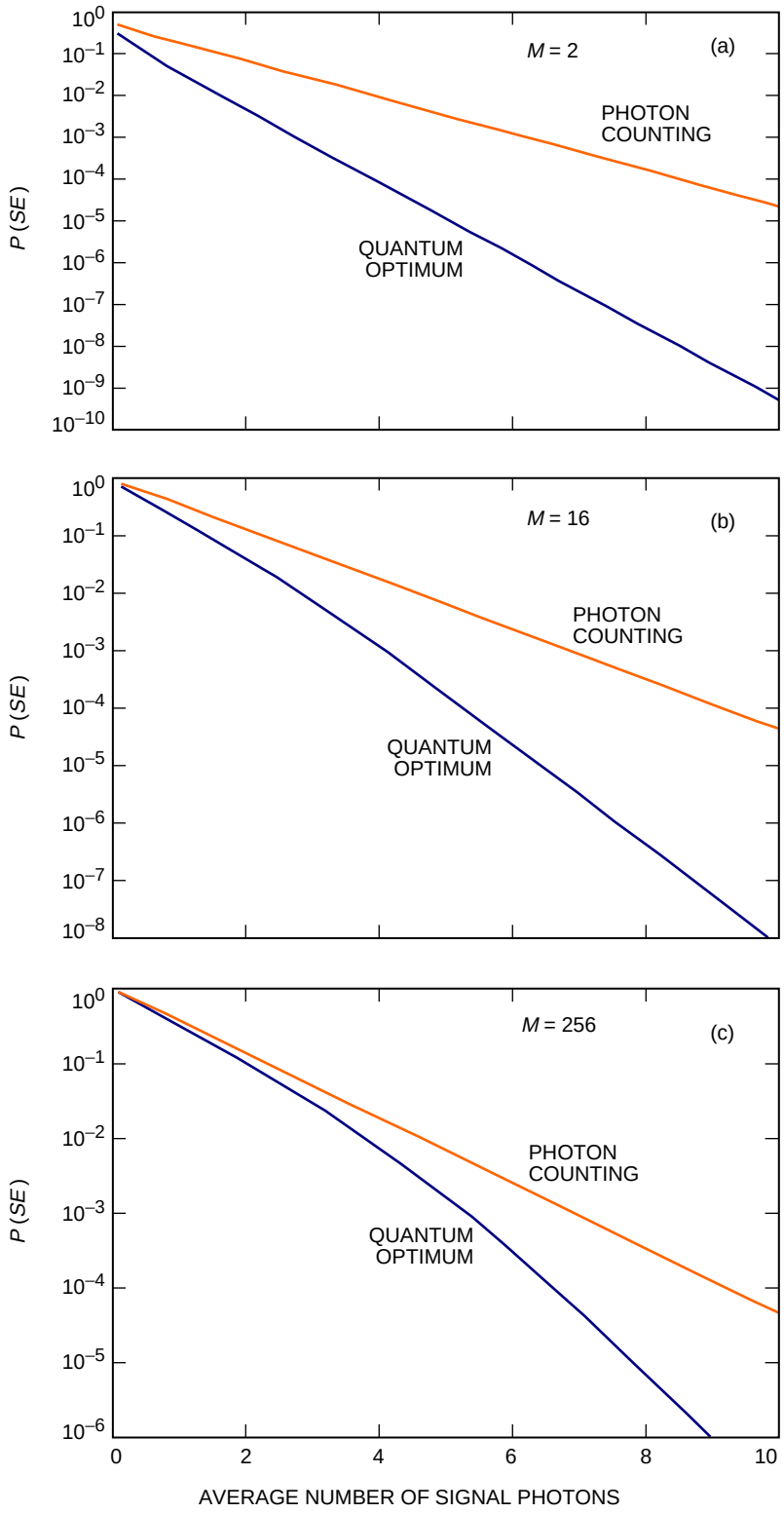

Fig. 11. Performance of optical receivers, no background: (a) *M* = 2, (b) *M* = 16, and (c) *M* = 256.

XII. Summary and Conclusions

In this article, we have reviewed the theoretical background needed to understand key concepts in quantum detection. This framework was then used to develop the structure of the optimum quantum receiver for binary hypotheses, both for noise-free, or "pure state" problems, and for detection in the presence of noise. The concept of "measurement states" was introduced, starting with a quantum mechanical formulation of photon-counting detection for simple on–off keyed modulation. It was demonstrated that, when operating on the signal states, the projection operators that correspond to photon counting created orthogonal bases that spanned the signal space and, therefore, could be interpreted as measurement states after appropriate normalization. A straightforward derivation of the performance of the quantum optimum receiver was obtained by analytically deriving the rotation of the measurement states needed to achieve optimum performance. This rotation procedure was verified numerically and then extended to higher dimensional multimode modulation schemes. It was shown that the same rotation procedure can be incremented to compute optimum performance for the next higher dimension, namely $M = 3$, by starting with the optimum solution for $M = 2$ and performing a single rotation into the third dimension. Extension to higher dimensions is conceptually straightforward but remains to be implemented. Finally, published analytic results were used to compute the symbol-error probabilities of 16- and 256-dimensional PPM in the absence of noise. It was shown that for high-dimensional signaling conventional photon counting performs within 1.3 dB of the optimum quantum receiver at error probabilities of approximately 0.001, which are of interest for deep-space communication.

References

[1] J. W. S. Liu, "Reliability of Quantum-Mechanical Communications Systems," Massachusetts Institute of Technology, Research Laboratory of Electronics, Technical Report 468, Cambridge, Massachusetts, December 31, 1968.

[2] C. W. Helstrom, J. W. S. Liu, and J. P. Gordon, "Quantum Mechanical Communications Theory," *Proceedings of the IEEE*, vol. 58, no. 10, pp. 1578–1598, October 1970.

[3] S. Karp, E. L. O'Neill, and R. M. Gagliardi, "Communication Theory for the Free-Space Optical Channel," *Proceedings of the IEEE*, vol. 58, no. 10, pp. 1611–1626, October 1970.

[4] S. M. Barnett and P. M. Radmore, *Methods in Theoretical Quantum Optics*, Oxford: Oxford Science Publications, Clarendon Press, 1997.

[5] C. W. Helstrom, *Quantum Detection and Estimation Theory, Mathematics in Science and Engineering*, vol. 123, New York: Academic Press, 1976.

[6] R. J. Glauber, "Coherent and Incoherent States of the Radiation Field," *The Physical Review*, vol. 131, no. 6, pp. 2766–2788, 1963.

[7] C. W. Helstrom, "Detection Theory and Quantum Mechanics," *Information and Control*, vol. 10, pp. 254–291, March 1968.

[8] C. W. Helstrom, "Detection and Quantum Mechanics (II)," *Information and Control*, vol. 13, pp. 156–171, August 1968.

[9] S. J. Dolinar, Jr. "An Optimum Receiver for the Binary Coherent State Quantum Channel," *MIT Research Laboratory Of Electronics Quarterly Progress Report 111*, Massachusetts Institute of Technology, Cambridge, Massachusetts, pp. 115–120, October 15, 1973.

[10] R. S. Kennedy, "A Near-Optimum Receiver for the Binary Coherent State Quantum Channel," *MIT Research Laboratory Of Electronics Quarterly Progress Report 108*, Massachusetts Institute of Technology, Cambridge, Massachusetts, pp. 219–225, January 15, 1973.

[11] H. P. Yuen, R. S. Kennedy, and M. Lax, "Optimum Testing of Multiple Hypotheses in Quantum Detection Theory," *IEEE Transactions on Information Theory*, vol. IT-21, no. 2, pp. 125–134, March 1975.

[12] B. H. Bransden and C. J. Joachain, *Introduction to Quantum Mechanics*, New York: John Wiley & Sons, 1989.

Appendix

Fundamentals of Quantum Mechanics

I. Wave Functions

The wave function forms the basis of quantum mechanics. The wave function or state function will be represented as $\Psi(x, y, z; t)$, which represents a complex probability amplitude that contains all the information about the quantum system [10]. A "system" can be electromagnetic radiation, a collection of subatomic particles, or a single particle. The probability of finding a particle at a particular point within a volume V at some time t is proportional to the squared magnitude of the wave function:

$$P(x, y, z, t) \propto |\Psi(x, y, z, t)|^2 \tag{A-1}$$

The wave function has a number of useful properties. Since the wave function $\Psi(x, y, z; t)$ is a probability amplitude, the probability of finding the particle somewhere must be unity:

$$\int_{-\infty}^{\infty} |\Psi(r, t)|^2 dr = \int_{-\infty}^{\infty} \Psi^*(r, t)\Psi(r, t) dr = 1 \tag{A-2}$$

where $dr \equiv dxdydz$ and $r \equiv (x, y, z)$.

In order to account for interference effects, the wave function obeys the principle of superposition. If a state is defined by a wave function Ψ_1 and another state is defined by another wave function Ψ_2, then any linear combination of the two yields a possible wave function:

$$\Psi = a\Psi_1 + b\Psi_2 \tag{A-3}$$

An important point is that, in general, $|\Psi|^2 \neq |a\Psi_1|^2 + |b\Psi_2|^2$. There is an interference term that must be accounted for. Let

$$\left.\begin{aligned} \Psi_1 &= |\Psi_1| \, e^{j\alpha_1} \\ \Psi_2 &= |\Psi_2| \, e^{j\alpha_2} \end{aligned}\right\} \tag{A-4}$$

The squared magnitude of the linear combination of wave functions in Eq. (A-3) yields

$$|\Psi|^2 = |a\Psi_1|^2 + |b\Psi_2|^2 + 2\mathrm{Re}\left\{ab^* \left|\Psi_1\right| \left|\Psi_2\right| \exp^{j(\alpha_1-\alpha_2)}\right\} \tag{A-5}$$

where the 2Re term accounts for interference.

II. Dirac Notation

A notation defined by Dirac can be used to simplify quantum mechanical models. The state function $\Psi(x, y, z; t)$ is replaced with the symbol ket $|\Psi\rangle$ and $\Psi^*(x, y, z; t)$ is replaced with the symbol bra $\langle\Psi|$. The ket represents a vector in an infinite dimensional Hilbert space representing the state of a system. The overlap of the two, $\langle\Psi_2|\Psi_1\rangle$, gives the probability amplitude that a system in state $|\Psi_1\rangle$ is found in state $|\Psi_2\rangle$ or

$$\langle\Psi_2|\Psi_1\rangle \equiv \int \Psi_2^*(r)\Psi_1(r)dr \tag{A-6}$$

where $\langle\Psi_2|\Psi_1\rangle$ is also called the scalar product. The probability of finding $|\Psi_1\rangle$ in state $|\Psi_2\rangle$ is $|\langle\Psi_2|\Psi_1\rangle|^2$. If $\langle\Psi_2|\Psi_1\rangle = 0$, the two states are orthogonal, and the probability of finding $|\Psi_1\rangle$ in state $|\Psi_2\rangle$ is zero. The normalization condition can be expressed as $\langle\Psi|\Psi\rangle = 1$. Given a complex number c and some third-state Ψ_3, some useful relations are

$$\left.\begin{aligned} \langle\Psi_2|c\Psi_1\rangle &= c\langle\Psi_2|\Psi_1\rangle \\ \langle c\Psi_2|\Psi_1\rangle &= c^*\langle\Psi_2|\Psi_1\rangle \\ \langle\Psi_3|\Psi_1 + \Psi_2\rangle &= \langle\Psi_3|\Psi_1\rangle + \langle\Psi_3|\Psi_2\rangle \end{aligned}\right\} \tag{A-7}$$

III. Operators

In quantum mechanics, an observable is associated with an operator. An operator acts on a wave function and returns another wave function along with the desired quantity. The action of an operator on the wave function can be expressed as

$$Q\Psi_q(x,t) = q\Psi_q(x,t) \tag{A-8}$$

where Q is the operator and $\Psi_q(x,t)$ is the state function that returns the value q when the physical quantity Q is measured. Therefore, q is an eigenvalue of $\Psi_q(x,t)$. An example of this is the position operator. Since the wave function is already defined in the "position representation," the operator can be expressed as

$$X\Psi(x,t) = x\Psi(x,t) \tag{A-9}$$

where X is the position operator and the position is given by the real number x. Another example is the momentum operator. Since the position wave function is defined in the position basis, the momentum

operator in the position representation is not simply a number. Let the wave function be represented by a plane wave traveling in the x-direction:

$$\Psi(x,t) = A\exp\left\{i[kx-\omega t]\right\} \tag{A-10}$$

where A is the amplitude of the wave, k is the wave number, x is the position, and ω is the frequency. Using the ω, k, and Planck's constant, $\hbar = (h/2\pi)$, we can express the energy, E, and momentum, p, as

$$\left.\begin{array}{l} E = \hbar\omega \\ \\ p = \hbar k \end{array}\right\} \tag{A-11}$$

Substituting E and p into Eq. (A-10) yields

$$\Psi(x,t) = A\exp\left\{\frac{i[px-Et]}{\hbar}\right\} \tag{A-12}$$

We now have the wave function in terms of the momentum and energy. In order to extract the value of p, the following operation is used:

$$-i\hbar\frac{\partial}{\partial x}\Psi(x,t) = p\Psi(x,t) \tag{A-13}$$

This operation satisfies the operator condition in Eq. (A-8) and thus yields the momentum operator. The momentum operator in the position basis is

$$P = -i\hbar\frac{\partial}{\partial x} \tag{A-14}$$

The position operator can be used to find the average position of the particle or wave packet described by the wave function $\Psi(x,t)$. Let $|x\rangle$ be defined as a state with definite position value x, $X|x\rangle = x|x\rangle$. Now, the state $\Psi(x,t)$ has some position distribution. To calculate the average value of position, we average over the probability density $P(x)$:

$$\begin{aligned} \bar{x} &= \int dx P(x) x \\ &= \int dx \left|\Psi(x)\right|^2 x \\ &= \int \langle\Psi \mid x\rangle x \langle x \mid \Psi\rangle \, dx \\ &= \langle\Psi| \int |x\rangle x \langle x| \, dx \, |\Psi\rangle \end{aligned} \tag{A-15}$$

The term $\int |x\rangle x \langle x| \, dx$ is independent of the state. It can be used to determine the position of any wave function regardless of distribution. Since the position operator extracts position, we can see that the position operator X can be expressed as

$$X = \int |x\rangle x \langle x| \, dx = \int x|x\rangle\langle x| dx \tag{A-16}$$

This result can be generalized to any arbitrary operator. The operator can be expressed as an outer product in its eigenvector basis. Let Q be any operator and $|q\rangle$ be its eigenvectors. Any operator can be expressed as

$$Q = \int |q\rangle q \langle q| \, dq = \int q|q\rangle\langle q| dq \tag{A-17}$$

Another useful operator in quantum mechanics is the Hamiltonian. The Hamiltonian operator in terms of the potential energy V, mass m, and Planck's constant is

$$H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V \tag{A-18}$$

The Hamiltonian H represents the total energy of a particle. The two terms in the Hamiltonian represent the kinetic and potential energy of a particle. In essence, the Hamiltonian is an "energy" operator. The Hamiltonian measures the energy of a state. In quantum mechanics, the total energy (Hamiltonian) of a particle is given by Eq. (A-11). The Hamiltonian for a harmonic oscillator (such as a classical pendulum with small excursion) can be expressed as

$$H = \frac{p^2}{2m} + \frac{m\omega^2 x^2}{2} \tag{A-19}$$

These quantities are defined as before with m being the mass of the oscillator.

A. Non-Commuting Operators

The position and momentum operators in the position representation do not commute, as can be seen from the following example. Let $\Psi(x)$ be an arbitrary wave function. If the position and momentum operators are applied in one order and then in reverse, the results are

$$\left.\begin{aligned} XP\Psi(x) &= X\left(-i\hbar\frac{\partial}{\partial x}\Psi(x)\right) = -i\hbar X\frac{\partial\Psi(x)}{\partial x} \\ PX\Psi(x) &= -i\hbar\frac{\partial}{\partial x}\big(X\Psi(x)\big) = -i\hbar\left[X\frac{\partial\Psi(x)}{\partial x} + \Psi(x)\right] \end{aligned}\right\} \tag{A-20}$$

Taking the difference, $XP\Psi(x) - PX\Psi(x) = i\hbar\Psi(x)$. Clearly, the results are not the same when these operators are applied in different order. Defining the commutation bracket $[A, B]$ for two operators A and B as $[A, B] = AB - BA$, we can see from Eq. (A-20) that the position and momentum commutation rule is $[X, P] = i\hbar$. If two operators A and B commute, then they necessarily satisfy the relation $[A, B] = 0$. If $[A, B] \neq 0$, then the operators A and B do not commute. If two operators do not commute, then the corresponding observables *cannot* be measured simultaneously with complete precision. Conversely, if two observables commute, the corresponding observables *can* be measured simultaneously.

IV. The Harmonic Oscillator

The harmonic oscillator example demonstrates the simplicity of the Dirac notation. To begin, the harmonic oscillator problem will be solved classically. The Hamiltonian for the one-dimensional harmonic oscillator is defined as

$$H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}kx^2 \tag{A-21}$$

From the Hamiltonian, the Schroedinger differential equation can be derived as in [12]:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + \frac{1}{2}kx^2\psi(x) = E\psi(x) \tag{A-22}$$

It is convenient to rewrite Eq. (A-22) in terms of dimensionless quantities. To do so, let λ_w be the wavelength defined as

$$\lambda_w = \frac{2E}{\hbar\omega} \tag{A-23a}$$

where

$$\omega = \left(\frac{k}{m}\right)^{1/2} \tag{A-23b}$$

Also, let ξ be a dummy variable defined as

$$\xi = \alpha x \tag{A-24a}$$

where

$$\alpha = \left(\frac{m\omega}{\hbar}\right)^{1/2} \tag{A-24b}$$

In the equations above, k is the spring constant, m is the mass, E is the energy, and $\hbar$ is Planck's constant. Substituting Eqs. (A-23) and (A-24) into Eq. (A-22) yields the differential equation

$$\frac{d^2\psi(\xi)}{dx^2} + (\lambda_w - \xi^2)\psi(\xi) = 0 \tag{A-25}$$

The solution to this equation is of the form

$$\psi(\xi) = e^{-\xi^2/2}H(\xi) \tag{A-26}$$

where $H(\xi)$ are Hermite polynomials. Since the solution contains both even and odd states of $\psi(\xi)$, each case must be examined individually. For the even states, $\psi(-\xi) = \psi(\xi)$ and, thus, $H(-\xi) = H(\xi)$.

Solving the above equations, we can find that only discrete values of λ can exist. Specifically, for even states

$$\lambda_w = 4n + 1, \quad n = 0, 1, 2, \cdots \tag{A-27}$$

Similarly, for the odd states, $\psi(-\xi) = -\psi(\xi)$ and $H(-\xi) = -H(\xi)$. Using the above equations, the discrete values of λ_w for the odd states are

$$\lambda_w = 4n + 3, \quad n = 0, 1, 2, \cdots \tag{A-28}$$

Combining the even and odd solutions yields

$$\lambda_w = 2n + 1, \quad n = 0, 1, 2, \cdots \tag{A-29}$$

Using Eq. (A-23a), the discrete energy levels for the harmonic oscillator are

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega \tag{A-30}$$

To solve the harmonic oscillator problem using the Dirac notation, it is convenient to define two new operators:

$$\left.\begin{aligned} \sqrt{\hbar\omega}a^+ &= \sqrt{\frac{k}{2}}x - i\sqrt{\frac{1}{2m}}p \\ \sqrt{\hbar\omega}a &= \sqrt{\frac{k}{2}}x + i\sqrt{\frac{1}{2m}}p \end{aligned}\right\} \tag{A-31}$$

These two operators are called the raising and lowering (or creation and annihilation) operators, respectively. Since X and P are both operators, a^+ and a are also. The commutation rule for these operators is

$$[a, a^+] = 1 \tag{A-32a}$$

or

$$aa^+ = 1 + a^+a \tag{A-32b}$$

Thus, the Hamiltonian for the harmonic oscillator can be expressed as

$$H = \hbar\omega\left(a^+a + \frac{1}{2}\right) \tag{A-33}$$

Now, if $|\Psi\rangle$ is an eigenstate of H with eigenvalue E, then $a^+|\Psi\rangle$ is an eigenstate of H with eigenvalue $E + \hbar\omega$:

$$
\begin{aligned}
Ha^+|\Psi\rangle &= \hbar\omega a^+aa^+|\Psi\rangle + \frac{1}{2}\hbar\omega a^+|\Psi\rangle \\
&= \hbar\omega a^+(a^+a+1)|\Psi\rangle + \frac{1}{2}\hbar\omega a^+|\Psi\rangle \\
&= a^+(H+\hbar\omega)|\Psi\rangle = a^+(E+\hbar\omega)|\Psi\rangle \\
&= (E+\hbar\omega)a^+|\Psi\rangle
\end{aligned} \tag{A-34}
$$

Similarly, $a|\Psi\rangle$ is an eigenstate of H with eigenvalue $E-\hbar\omega$ with one exception. Since the Hamiltonian is the sum of positive definite terms, the eigenvalues must be positive. Thus, a base eigenstate that cannot be lowered is required. This base state is called the lowest energy eigenstate or zero state $|0\rangle$. When the lowering operator is applied to the zero state, the result is zero:

$$
a|0\rangle = 0 \tag{A-35}
$$

Note that the state itself is annihilated and nothing is left. Therefore, the zero state is truly the lowest energy eigenstate. The zero-state energy can be determined by

$$
\left.
\begin{aligned}
H|0\rangle &= \hbar\omega a^+a|0\rangle + \frac{1}{2}\hbar\omega|0\rangle \\
H|0\rangle &= \frac{1}{2}\hbar\omega|0\rangle
\end{aligned}
\right\} \tag{A-36}
$$

Therefore, the zero-state energy is $(1/2)\hbar\omega$. Now, starting with the zero state, an infinite ladder of eigenstates can be constructed by repeatedly applying the raising (creation) operator. Thus,

$$
H|n\rangle = \hbar\omega\left(n+\frac{1}{2}\right)|n\rangle \tag{A-37}
$$

Since the Hamiltonian is an energy operator, the discrete energy levels of the harmonic oscillator are

$$
E_n = \hbar\omega\left(n+\frac{1}{2}\right) \tag{A-38}
$$

Using operators and the Dirac notation yields a simpler derivation than the classical notation. Another point that should be made is that, since $|n\rangle$ is an eigenstate of H, it is also an eigenstate of $N = a^+a$, which is called the number operator. The eigenstates associated with the number operator are called the number states:

$$
N\,|n\rangle = n\,|n\rangle \tag{A-39}
$$

where n is an integer-valued eigenvalue. The number eigenstates are important in the representation of the states of electromagnetic radiation used to model communications systems.