Quantum Detection and Channel Capacity Using State–Space Optimization

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

The fundamental performance limits and channel capacity of optical communications systems operating over the free-space channel will be examined using quantum detection theory. The performance of the optimum quantum receiver for on–off keying and optical binary phase-shift keying is first examined as a pure-state (no-noise) problem. The classical capacity of the binary symmetric channel for these two modulation schemes will be evaluated for the optimum quantum receiver by making use of the concept of quantum measurement states. The performance of M-ary pulse-position modulation, which requires a product state representation, will be evaluated along with the performance of certain dense signal sets. Performance comparisons with classical techniques show an improvement of over 5 dB in some cases when quantum detection is employed. As a further application of the quantum detection theory, the capacity of the binary channel with on–off keyed modulation and quantum detection is evaluated and shown to exceed the capacity obtained with classical photon counting.

I. Introduction

At present, modulated optical fields are generally detected by means of energy detectors either directly or by the use of phase-sensitive coherent detection techniques. At the extremely high frequencies of optical signals, energy detection becomes a viable option that can be used even to discriminate between individual photons, due to the high energy of photons in the optical regime. It is well-known that photon counting can overcome thermal noise in the detection electronics, leading to shot-noise-limited performance, where the only uncertainty is the inherent quantum-mechanical randomness in the weak optical fields. Coherent detection relies on the addition of a strong local optical field to generate a large cross-term between the received and local fields which, when detected using a suitable optical energy detector, can also overcome thermal noise and achieve shot-noise-limited performance. While these detection techniques are very sensitive, they cannot realize the full advantages of optimum quantum detection, which typically performs much better in terms of the signal energy required to achieve a given detection performance.

The principles of quantum detection have been developed by Liu, Helstrom, Kennedy, Yuen, and others during the last decades, and reported in numerous journal articles and books [1–4]. Their results on the detection of coherent-state signals by means of optimum quantum measurements have been summarized

[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.

in a recent review article [5], where quantum detection techniques have been applied to well-known optical modulation schemes such as on–off keying (OOK), binary phase-shift keying (BPSK), and pulse-position modulation (PPM), and compared to classical techniques such as photon-counting and coherent detection. While classical schemes rely on photon-counting or coherent measurements, optimum and near-optimum quantum measurements typically require both photon-counting and coherent techniques, together with real-time signal processing to control the amplitude and phase of the local field. In this way, improvements of a factor of two (3 dB) can often be achieved over classical detection in the absence of background radiation. Here we extend these results to a new class of signals called "dense" signal sets, where a large number of signals are packed into a given classical signal–space dimension, and compare the performance of optimum quantum detection with those of more familiar classical techniques. Examples of dense modulations include ternary and quadrature phase-shift keying (QPSK) modulations, as well as compound signals formed by combinations of different modulations, such as binary phase-shift keying–pulse-position modulation (BPSK-PPM). We shall show that quantum detection of these dense signal sets often provides very significant (5 dB or more) improvement over classical detection techniques.

II. Quantum Optical Communications

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

A. Definition of Quantum States

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. The state is normalized if $\langle\psi|\psi\rangle = 1$. Suppose $|\phi_1\rangle$ and $|\phi_2\rangle$ are orthonormal, that is $\langle\phi_1|\phi_2\rangle = \delta_{12}$, and $|\psi\rangle$ is normalized. If a state $|\psi\rangle$ can be expressed as the superposition of orthonormal states, $|\psi\rangle = a_1|\phi_1\rangle + a_2|\phi_2\rangle$, then their overlaps are

$$\left.\begin{aligned} \langle\phi_1|\psi\rangle &= \langle\psi|\phi_1\rangle^* = a_1 \\ \langle\phi_2|\psi\rangle &= \langle\psi|\phi_2\rangle^* = a_2 \end{aligned}\right\} \tag{1}$$

where $|a_1|^2 + |a_2|^2 = 1$ and $|a_1|^2$ and $|a_2|^2$ can be interpreted as the probabilities that the system is found to be in states $|\phi_1\rangle$ and $|\phi_2\rangle$, respectively, after a measurement. The overlap between two normalized states, $|\psi_1\rangle$ and $|\psi_2\rangle$, can also be interpreted geometrically as the cosine of the angle, θ_{12}, between the vectors representing the states in Hilbert space: $\langle\psi_1|\psi_2\rangle = \cos(\theta_{12})$. Generalization to a superposition of states follows as

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

where

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

with the interpretation that $|a_n|^2$ is the probability that the system is found in state $|\psi_n\rangle$ after a measurement.

There are situations where a single Hilbert space is not sufficient to describe the signal. An example of this is M-ary pulse-position modulation, or PPM, where an optical pulse is placed into one of M consecutive slots. This kind of modulation requires a "product-state" description of the form

$$|\psi\rangle = \prod_{j=1}^{M} |\psi_j\rangle = |\psi_1\rangle|\psi_2\rangle \cdots |\psi_M\rangle \tag{3}$$

where each of the $|\psi_j\rangle$ is a coherent state associated with the individual modes. For product-states, the overlap is computed by the rule

$$\begin{aligned}\langle\psi_m|\psi_k\rangle &= \prod_{j=1}^{M} \langle\psi_{m,j}|\psi_{k,j}\rangle \\ &= \langle\psi_{m,1}|\psi_{k,1}\rangle\langle\psi_{m,2}|\psi_{k,2}\rangle \cdots \langle\psi_{m,M}|\psi_{k,M}\rangle\end{aligned} \tag{4}$$

These concepts will now be applied to the quantum description of coherent optical fields, represented as a superposition of "number states" familiar from the quantum mechanical solution of the harmonic oscillator. This model will first be applied to problems involving the detection of single-mode optical fields, followed by more complex signal models requiring a product-state description of the signal set.

B. The Coherent-State Representation of Optical Signals

Coherent states, representing electromagnetic radiation produced by physical devices such as lasers, are an important class 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 the number eigenstates:

$$|\alpha\rangle = e^{-(1/2)|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{(n!)^{1/2}} |n\rangle \tag{5}$$

Each number eigenstate $|n\rangle$ contains n photons, and hence the probability of obtaining exactly n photons as the outcome of an experiment can be computed as

$$|\langle\alpha|n\rangle|^2 = e^{-|\alpha|^2} \frac{|\alpha|^{2n}}{n!} \tag{6}$$

For any n, these are recognized as Poisson probabilities for the number of photons, with the average number of photons equal to $|\alpha|^2$. Coherent states are not orthogonal, as can be seen by considering the overlap between two arbitrary coherent states, $|\alpha_1\rangle$ and $|\alpha_2\rangle$. Orthogonality requires that the overlap vanish altogether; however, for coherent states, the squared magnitude of the overlap is not zero but instead is given by

$$|\langle\alpha_1|\alpha_2\rangle|^2 = \left| e^{-(|\alpha_1|^2+|\alpha_2|^2)/2} \sum_n \sum_m \frac{\alpha_1^n}{\sqrt{n!}} \frac{(\alpha_2^*)^m}{\sqrt{m!}} \langle n|m\rangle \right|^2 = \left| e^{-(|\alpha_1|^2+|\alpha_2|^2-2\alpha_1\alpha_2^*)/2} \right|^2 = e^{-|\alpha_1-\alpha_2|^2} \tag{7}$$

where we made use of the orthogonality of the number states to simplify the expression. Equation (7) demonstrates that there is always some overlap between coherent states, regardless of how great the difference between average photon counts may be.

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

The concept of "measurement states," used extensively in the quantum mechanical derivation, can be illustrated by the following example employing on–off keying (OOK) modulation. This modulation can be described classically as "one of the signals has zero amplitude, while the other signal has complex amplitude α." Suppose there are two hypotheses, H_0 and H_1, denoting absence and presence of signal, respectively. If the background radiation can be neglected, then either no photons or an average of $|\alpha|^2$ photons are received. The received field is assumed to be from a coherent laser; hence, the photons are Poisson distributed with conditional densities

$$\left.\begin{aligned} P(n|H_0) &= \begin{cases} 1, & n=0 \\ 0, & n \geq 1 \end{cases} \\ P(n|H_1) &= \frac{|\alpha|^{2n}}{n!} e^{-|\alpha|^2} \end{aligned}\right\} \tag{8}$$

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. 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: $P(n \geq 1|H_1) = 1 - e^{-|\alpha|^2}$. With equal a priori probabilities, $P(H_0) = P(H_1) = 1/2$, the probability of correct detection becomes

$$P(C) = \sum_{i=1}^{2} P(C|H_i)P(H_i) = 1 - \frac{1}{2} e^{-|\alpha|^2} \tag{9}$$

yielding the average probability of error $P(E) = 1 - P(C) = (1/2)e^{-|\alpha|^2}$.

In the quantum mechanical formulation, the received field is in one of two states, $|\psi_0\rangle = |0\rangle$ or $|\psi_1\rangle = |\alpha\rangle$, corresponding to 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

$$|\psi_1\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \tag{10}$$

where α is a complex number, the set $\{|n\rangle\}$ represents the number eigenstates, and $|\alpha|^2$ again represents the average number of photons in the signal. 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 described in [5], defined as

$$\left.\begin{aligned} \Pi_0 &= |0\rangle\langle 0| \\ \Pi_1 &= \sum_{n=1}^{\infty} |n\rangle\langle n| = \mathbf{1} - |0\rangle\langle 0| \end{aligned}\right\} \tag{11}$$

where **1** is the identity operator with a number eigenstate basis. When the projection operator Π_0 is applied to the ground or "vacuum" state (the state of the received field under hypothesis H_0), the ground state is recovered:

$$\Pi_0|\psi_0\rangle = |0\rangle\langle 0|0\rangle = |0\rangle \equiv |w_0\rangle \tag{12}$$

Here we define $|w_0\rangle$ as the measurement state corresponding to $|\psi_0\rangle$. The probability of observing zero photons, given that H_0 is true, can be expressed as [5]

$$\Pr(0|H_0) = |\langle w_0|\psi_0\rangle|^2 = \langle 0|\Pi_0|0\rangle = 1 \tag{13}$$

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

$$\begin{aligned}\Pi_1|\psi_1\rangle &= \sum_{n=1}^{\infty} |n\rangle\langle n| \left(e^{-|\alpha|^2/2} \sum_{m=0}^{\infty} \frac{\alpha^m}{\sqrt{m!}}|m\rangle \right) \\ &= e^{-|\alpha|^2/2} \sum_{n=1}^{\infty} \frac{\alpha^n}{\sqrt{n!}}|n\rangle \equiv |W_1\rangle \end{aligned} \tag{14}$$

Note that $|W_1\rangle$ is not normalized. Denote the normalized version of $|W_1\rangle$ as $|w_1\rangle$, with the interpretation that it is a measurement state for $|\psi_1\rangle$, that is, $\Pi_1|\psi_1\rangle = \langle w_1|\psi_1|w_1\rangle$. Since $\langle w_0|W_1\rangle = e^{-|\alpha|^2/2} \sum_{n=1}^{\infty} \left(\alpha^n/\sqrt{n!}\right)\langle 0|n\rangle = 0$, it follows that the measurement states $|w_0\rangle$ and $|w_1\rangle$ are orthonormal. Similar to Eq. (13), the probability of obtaining a count greater than 0 when observing the received coherent state is

$$\begin{aligned}\Pr(n \geq 1|H_1) &= |\langle w_1|\psi_1\rangle|^2 = \langle \psi_1|\Pi_1|\psi_1\rangle \\ &= \left(e^{-|\alpha|^2/2} \sum_{m=0}^{\infty} \frac{(\alpha^*)^m}{\sqrt{m!}}\langle m| \right) \sum_{n=1}^{\infty} |n\rangle\langle n| \left(e^{-|\alpha|^2/2} \sum_{i=0}^{\infty} \frac{\alpha^i}{\sqrt{i!}}|i\rangle \right) \\ &= e^{-|\alpha|^2} \sum_{n=1}^{\infty} \frac{|\alpha|^{2n}}{n!} = e^{-|\alpha|^2} \left(e^{|\alpha|^2} - 1 \right) = 1 - e^{-|\alpha|^2} \end{aligned} \tag{15}$$

Again interpreting this as the squared magnitude of the projection of $|\psi_1\rangle$ onto $|w_1\rangle$, it follows that $P(C) = \sum_{i=1}^{2} P(C|H_i)P(H_i) = 1 - (1/2)e^{-|\alpha|^2}$, yielding the probability of error $P(E) = 1 - P(C) = (1/2)e^{-|\alpha|^2}$, exactly as with the classical derivation. We can see, therefore, that the detection operation can be interpreted as the projection of the signal states onto a properly chosen set of orthonormal measurement states.

D. State–Space Derivation of Quantum Receiver Performance with Binary Signals

In the quantum formulation, the two signal states characterizing OOK modulation, $|\psi_0\rangle = |0\rangle$ and $|\psi_1\rangle = |\alpha\rangle$, define a plane in Hilbert space. We have shown that the states produced by the application of the photon-counting projection operators to the signal states can be referred to as measurement states that span the two-dimensional subspace of Hilbert space defined by the two signal states. The measurement

states are orthonormal and, for photon counting, one of the measurement states is lined up with the ground state representing the null hypothesis, while the other measurement state is orthogonal to it in the plane defined by the two signal states. Note that the ground state is always detected correctly, since with photon counting the ground state is colinear with one of the measurement states and therefore its projection onto the corresponding measurement state is one. However, since the signal state is not orthogonal to the ground state, but instead is at an angle $\cos(\theta) = \langle 0|\alpha\rangle$, there is a nonzero projection onto both measurement states; this is the reason for the occurrence of detection errors in the reception of optical OOK signals even in the complete absence of background light.

Operating entirely in the two-dimensional signal subspace of Hilbert space, it is possible to find the minimum average probability of error by rotating the measurement states within this "signal plane" and calculating the error probability for each rotation until a minimum is reached. Representing the plane defined by two signal states and also containing the two measurement states $|w_0\rangle$ and $|w_1\rangle$ as in Fig. 1, the angle θ is defined as $\theta = \cos^{-1}(|\langle\psi_1|\psi_0\rangle)$ and corresponds to the overlap between the two signal states. We start with $\varphi_0 = 0$, representing photon-counting detection, and rigidly rotate the measurement states.

The conditional probabilities of correct detection are given by

$$\left.\begin{aligned} P(C|H_0) &= |\langle\psi_0|w_0\rangle|^2 \equiv \cos^2(\varphi_0) \\ P(C|H_1) &= |\langle\psi_1|w_1\rangle|^2 \equiv \cos^2(\varphi_1) \end{aligned}\right\} \tag{16}$$

The maximum value of the probability of correct detection can be found as a function of the rotation angle φ_1, by differentiating $P(C)$ with respect to φ_1 and equating to zero:

$$\frac{\partial P(C)}{\partial \varphi_1} = \frac{1}{2}\big\{ -2\cos(a-\varphi_1)\sin(a-\varphi_1) + 2\cos(\varphi_1)\sin(\varphi_1)\big\} = 0 \tag{17}$$

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

$$\begin{aligned} P^*(C) &= \frac{1}{2}\left[\cos^2\left(\frac{a}{2}\right) + \cos^2\left(\frac{a}{2}\right)\right] = \frac{1}{2}\big[1+\cos(a)\big] = \frac{1}{2}\big[1+\sin(\theta)\big] \\ &= \frac{1}{2}\left[1+\sqrt{1-|\langle\psi_0|\psi_1\rangle|^2}\right] \end{aligned} \tag{18}$$

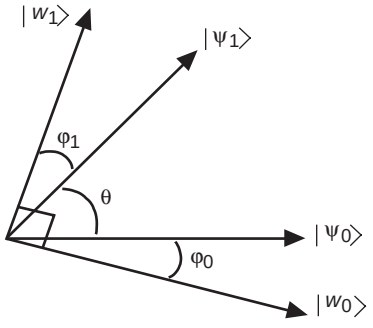

Fig. 1. The signal and measurement states for the binary OOK problem.

where we let $\sin(\theta) = \sqrt{1-\cos^2(\theta)} = \sqrt{1-|\langle\psi_0|\psi_1\rangle|^2}$ in the last step. Using the maximum value of $P(C)$, the minimum probability of error can be expressed as

$$P^*(E) = 1 - P^*(C) = \frac{1}{2}\left[1 - \sqrt{1-|\langle\psi_0|\psi_1\rangle|^2}\right] \tag{19}$$

which agrees exactly with the independently derived performance in [4]. Therefore, the operation performed by the optimum quantum receiver can be viewed as a rotation of the measurement states in the plane defined by the signal states. Using Eq. (7), the error probability of the optimum quantum receiver for OOK can be further expressed in terms of the average number of photons in the signal averaged over both hypotheses, $K_s = (1/2)|\alpha|^2$, as

$$P^*(E) = \frac{1}{2}\left[1 - \sqrt{1-e^{-|\alpha|^2}}\right] = \frac{1}{2}\left[1 - \sqrt{1-e^{-2K_s}}\right] \tag{20}$$

Another binary modulation format of interest for deep-space optical communications is "optical BBSK," which can be described quantum mechanically as two coherent states with the same average photon energy, but with π radians out of phase. For optical BPSK, the signal states are defined as $|\psi_0\rangle = |\alpha\rangle$ and $|\psi_1\rangle = |-\alpha\rangle$, where the average number of signal photons (averaged over both symbols) is $K_s = |\alpha|^2$. Therefore, $|\langle\psi_0|\psi_1\rangle|^2 = e^{-4K_s}$ for optical BPSK, and the error probability is given by $P^*(E) = (1/2)\left[1 - \sqrt{1-e^{-4K_s}}\right]$. Signals that employ phase modulation, such as BPSK, require phase-sensitive coherent measurements to distinguish the symbols; such measurements can be implemented by adding a strong local field that is in phase with the optical field to the received signal, and detecting the resulting sum field using classical energy detection. For optical BPSK, the error probability for this coherent receiver is given by the following expression [4]: $P(E) = Q\left(\sqrt{4|\alpha|^2}\right) = Q\left(\sqrt{4K_s}\right)$, where $Q(x) \equiv \left(1/\sqrt{2\pi}\right)\int_x^\infty e^{-y^2/2}dy$.

Performance curves for binary OOK and BPSK formats using both quantum and classical detection are shown in Fig. 2. Note that photon counting exhibits the same exponential behavior as optimum quantum detection (both curves have the same slope), implying that photon counting is nearly optimal for the detection of OOK signals. However, quantum detection of optical BPSK signals is exponentially 3 dB better than coherent detection, practically achieving a 2.6-dB reduction in the required signal energy at an error probability of 10^{-5}.

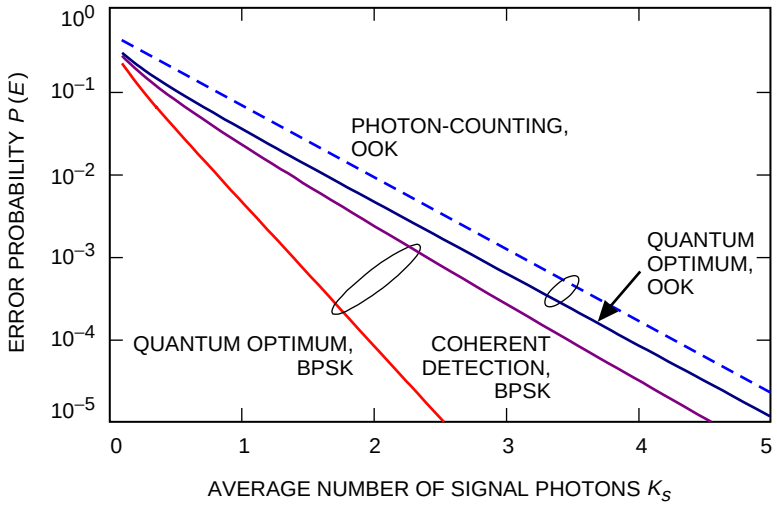

Fig. 2. Performance of optical OOK and BPSK signals.

The binary signals considered above spanned a one-dimensional classical signal space, since both OOK and BPSK can be represented classically as points on the real line: thus, OOK is represented by the points $(0, \alpha)$, and BPSK by the points $(-\alpha, \alpha)$. If the number of signals in this one-dimensional classical signal space is increased by adding another signal with amplitude $(-\alpha)$, for example, then we obtain a "dense signal set" characterized by the classical amplitudes $(-\alpha, 0, \alpha)$. We define a signal set as dense if its classical representation contains more signals than dimensions. Since distinct coherent states are linearly independent, the corresponding quantum signals $|-\alpha\rangle, |0\rangle, |\alpha\rangle$ span a three-dimensional subspace of Hilbert space. Therefore, three measurement states are needed to decode these signals optimally. A graphical representation of the subspace spanned by the three signals is shown in Fig. 3, along with the orthonormal measurement states that must be rotated to achieve optimum detection.

The probability of error for this ternary problem has been calculated by Helstrom in [4] and shown to be of the form $P(E) = 1 - (1/3)(a^2 + 2c^2) \approx (1/3)\exp(-N_s)$, where a and c are complicated algebraic expressions. For comparison, the performance of the classical coherent detection receiver is given by $P(E) = (4/3)\,\mathrm{erfc}\left(\sqrt{K_s}\right)$, which is asymptotically proportional to $\exp(-K_s/2)$. The error performance of these two receivers is shown in Fig. 4, where it can be seen that at low error probabilities quantum detection enjoys a significant, nearly 3-dB, advantage over classical detection for ternary signals.

The state–space solution to the ternary problem is similar to the binary problem discussed above, except now three orthonormal measurement states must be rotated until the maximum average detection probability, or minimum average error probability, is found.

III. Quantum Decoding of Higher Dimensional Signal Sets

Higher dimensional optimization is the process of finding the orientation of the orthogonal measurement states that yield the minimum probability of error. The solution to the M hypotheses problem can be found by rotating the measurement states in an M dimensional space. The solution is iterative, starting with the known optimum solution for the first two signals selected at random. The following example demonstrates the solution for $M = 4$ but can easily be generalized to M dimensions. Using the signal state overlap angles defined earlier as $\theta_{ij} = \cos^{-1}\left(|\langle\psi_i|\psi_j\rangle|\right)$, the signal states are plotted in a four-dimensional Hilbert space with the axes defined by x–y–z–w. Figure 5 illustrates the algorithm, showing the first three signal states, $|\psi_1\rangle$, $|\psi_2\rangle$, and $|\psi_3\rangle$, with optimum placement of the measurement states around the first two signal states, in the plane defined by this pair.

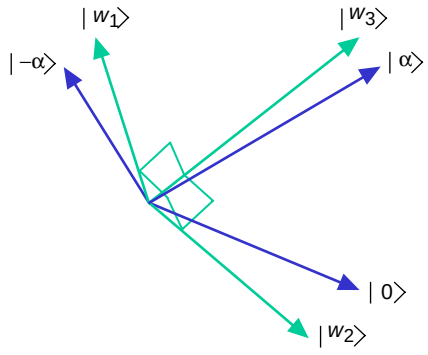

Fig. 3. Signal and measurement states for the ternary problem.

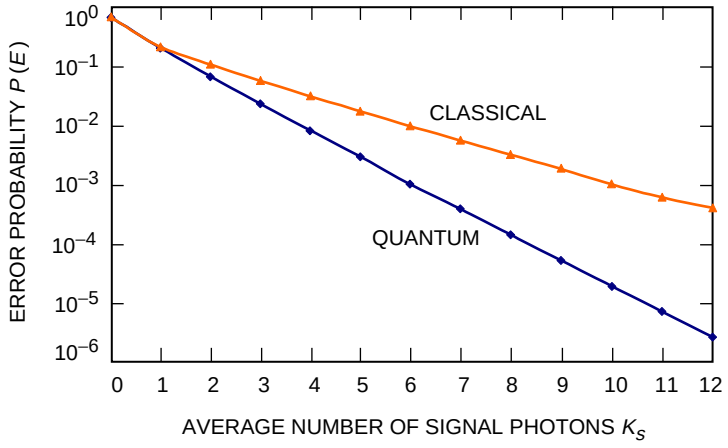

Fig. 4. Performance of ternary signals with optimum quantum and classical detection.

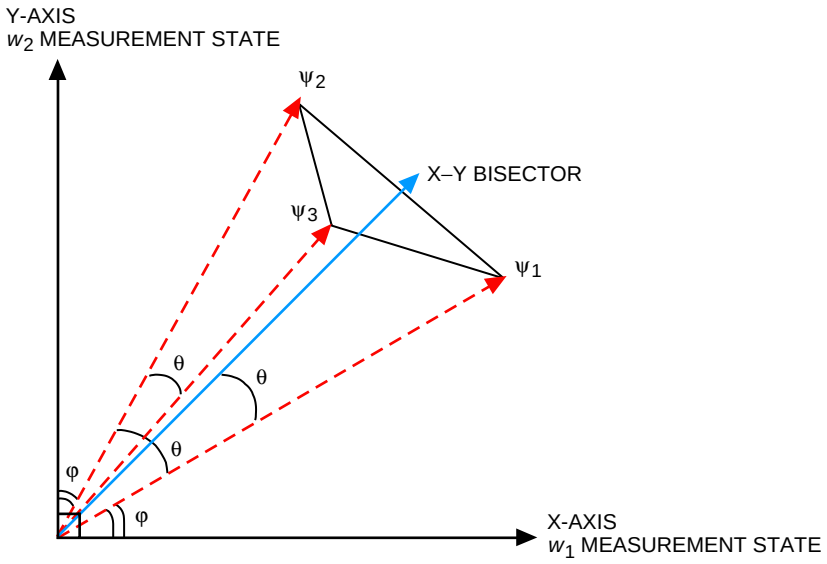

Fig. 5. Two-dimensional optimal solution.

Next, the Cartesian coordinates for the state $|\psi_4\rangle$ relative to $|\psi_1\rangle$, $|\psi_2\rangle$, and $|\psi_3\rangle$ are determined using the signal-state overlap angle relations defined above. Once all the signal-state coordinates have been determined, the measurement states $|w_1\rangle$, $|w_2\rangle$, $|w_3\rangle$, and $|w_4\rangle$ can be rotated to obtain the minimum probability of error.

The problem can be solved iteratively, starting with two dimensions, since it is possible to rotate the projections of a vector in any lower dimension without affecting the projections of that vector in the higher dimensions. This principle is illustrated in Fig. 6, where the projection of vector V_1 in the x–y plane is given by $V_1 \cos(\psi)$. The vector V_1 can be rotated such that the projection onto the z-axis is unchanged while the x- and y-projection components vary. Generalizing this result to any M-dimensional vector, the projections onto the $(M-1)$ coordinate axes can be varied by rotating the $(M-1)$ measurement coordinates without affecting the projection onto the next higher dimension. Therefore, the optimum measurement-state orientation for each dimension can be determined from the lower-dimensional solution, and this procedure can be continued until the complete M-dimensional solution is obtained.

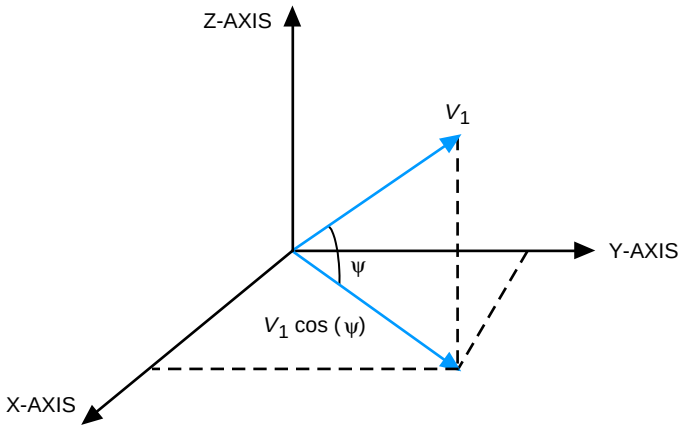

Fig. 6. Vector rotation.

Next we illustrate a detailed solution for the four-dimensional case. Initially, the measurement states are aligned with the Cartesian coordinate axes: $|w_1\rangle$ with the x-axis, $|w_2\rangle$ with the y-axis, $|w_3\rangle$ with the z-axis, and $|w_4\rangle$ with the w-axis. We start with the solution for $M = 2$, which defines a plane containing two signal states and two measurement states. The three-dimensional optimal rotation will be in the plane defined by the z-axis and x–y bisector. As the states are rotated into the third dimension, the projections of $|\psi_1\rangle$ and $|\psi_2\rangle$ onto $|w_1\rangle$ and $|w_2\rangle$ must be kept equal to each other in order to keep their projections onto the fixed $|w_1\rangle$ and $|w_2\rangle$ measurement states maximum for any rotation. The reason for this is that the minimum probability of error is achieved when the average projection of all signal states onto their corresponding measurement states is maximum. From Fig. 1, it can be seen that in order to keep the angles between $|\psi_1\rangle, |\psi_2\rangle$ and $|w_1\rangle$, $|w_2\rangle$ equal, the rotation must be in the plane defined by the z-axis and x–y bisector. The rotation of the measurement states away from the z-axis towards the x–y bisector is accomplished with the use of the following rotation matrix:

$$\begin{pmatrix} \cos\left(\frac{\pi}{4}\right) & -\sin\left(\frac{\pi}{4}\right) & 0 & 0 \\ \sin\left(\frac{\pi}{4}\right) & \cos\left(\frac{\pi}{4}\right) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} \cos\theta & 0 & \sin\theta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\theta & 0 & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} \cos\left(-\frac{\pi}{4}\right) & -\sin\left(-\frac{\pi}{4}\right) & 0 & 0 \\ \sin\left(-\frac{\pi}{4}\right) & \cos\left(-\frac{\pi}{4}\right) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (21)$$

where θ is the rotation angle away from the z-axis towards the x–y bisector. As the measurements states are rotated, the probability of error is found for each incremental θ by computing the following quantities:

$$\left.\begin{aligned} P_c &= \frac{1}{3}\left(\langle w_1|\psi_1\rangle^2 + \langle w_2|\psi_2\rangle^2 + \langle w_3|\psi_3\rangle^2\right) \\ P_e &= 1 - P_c \end{aligned}\right\} \quad (22)$$

Thus, the measurement states are incrementally rotated and the error probability calculated until the angle that yields the minimum probability of error is determined: this is the optimum rotation angle for this dimension. Next, this three-dimensional solution is used to obtain the four-dimensional solution corresponding to four signal states. As before, the goal is to rotate the $|w_4\rangle$ measurement state away from the w-axis while keeping the projections of $|\psi_1\rangle$, $|\psi_2\rangle$, and $|\psi_3\rangle$ onto $|w_1\rangle$, $|w_2\rangle$, and $|w_3\rangle$ maximal. The rotation that satisfies these conditions is the rotation within the plane defined by the w-axis and the x–y–z trisector. Defining θ' as the angle between the x–y–z trisector and the x–y bisector, and γ as the rotation angle away from the w-axis towards the x–y–z trisector, γ is incremented once again until the minimum probability of error is found using the four-dimensional extension of Eq. (22). This approach can be easily generalized to any higher dimensional signal sets by iteratively applying the above algorithm.

IV. Error Performance of PPM and Dense Signal Sets

First we examine PPM signals, which require a product-state representation in the quantum model, as described in [4,5]. As an example of product-state signals, the description of binary PPM with complex amplitude α is of the form $|\psi_1\rangle = |\alpha\rangle|0\rangle, |\psi_2\rangle = |0\rangle|\alpha\rangle$. This signal set spans a two-dimensional subspace of the product space, as illustrated in Fig. 7. Using the expression for the magnitude squared of the overlap between two coherent product states, as in Eq. (4), the overlaps are $\langle\psi_1|\psi_1\rangle = \langle\psi_2|\psi_2\rangle = 1, \langle\psi_1|\psi_2\rangle = \langle\psi_2|\psi_1\rangle = e^{-K_s}$, where K_s is the average number of photons in a signal set. It has been shown [3,4] that, for equally likely signals, the minimum symbol-error probability for the optimum quantum receiver is

$$P(SE) = \frac{M-1}{M^2}\left[\sqrt{1+(M-1)e^{-K_s}} - \sqrt{1-e^{-K_s}}\right]^2 \tag{23}$$

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.

The probability of a correct decision is maximized by rotating the two orthogonal measurement states so that they are symmetrically placed around the signal states in the plane defined by the signals. The error probability is then given by $P(E) = (1/2)\left[1 - \sqrt{1 - |\langle\psi_1|\psi_2\rangle|^2}\right]$ 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. The performance of quantum and classical receivers is compared in Fig. 8 for $M = 2$ and $M = 8$, where it can be seen that optimum quantum detection is approximately 3 dB better than classical detection in both cases.

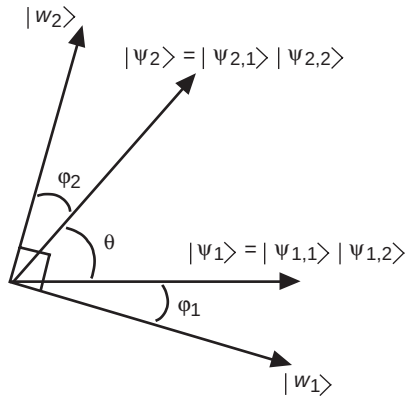

Fig. 7. Signal and measurement states for binary orthogonal modulation.

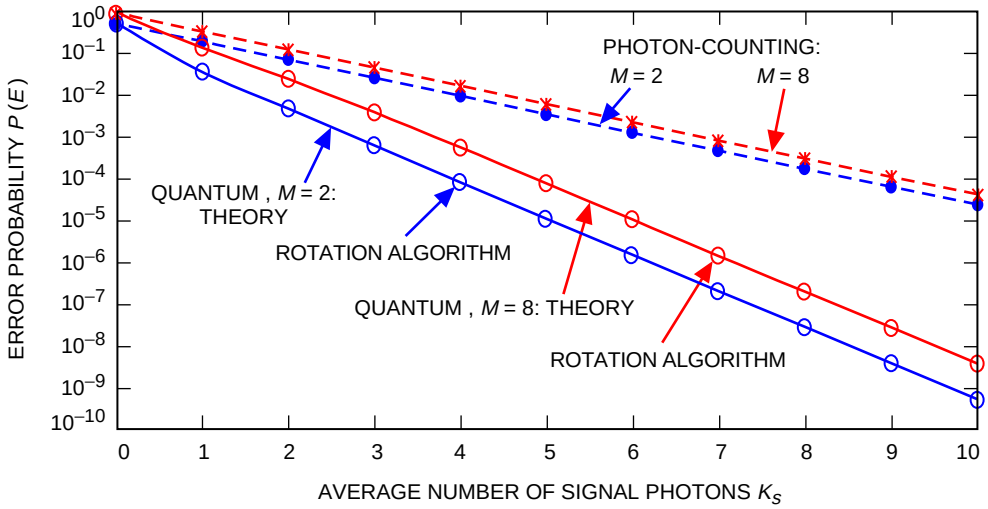

Fig. 8. Performance of PPM signals with optimum quantum and classical detection: $M = 2, 8$.

A. QPSK Signals

The classical description of QPSK signals includes a pair of orthogonal waveforms together with their negatives, such as, for example, the following set of four sinusoidal waveforms that also uses the negative of each orthogonal component: $\{s_1(t) = \sin(\omega t), s_2(t) = -\sin(\omega t), s_3(t) = \cos(\omega t), s_4(t) = -\cos(\omega t)\}$. Since there are four signals in two dimensions, this is an example of a dense signal set by the above definition. The quantum description of coherent-state QPSK is of the form $\{|\psi_1\rangle = |\alpha\rangle, |\psi_2\rangle = |-\alpha\rangle, |\psi_3\rangle = |i\alpha\rangle, |\psi_4\rangle = |-i\alpha\rangle\}$, with the pairwise overlaps $\langle\psi_1|\psi_2\rangle = \langle\alpha|-\alpha\rangle = e^{-(1/2)|\alpha+\alpha|^2} = e^{-2|\alpha|^2}$ and $\langle\psi_3|\psi_4\rangle = e^{-(1/2)|i\alpha+i\alpha|^2} = e^{-2|\alpha|^2}$; the four remaining overlaps are $\langle\psi_1|\psi_3\rangle = \langle\psi_1|\psi_4\rangle = \langle\psi_2|\psi_3\rangle = \langle\psi_2|\psi_4\rangle = e^{-(1/2)|\alpha\pm i\alpha|^2} = e^{-|\alpha|^2}$. Using the rotation algorithm described above for optimally aligning the four orthogonal measurement states with the QPSK signal states yields the error-probability performance shown in Fig. 9 (denoted by large circles). The exact error probability for QPSK signals has been derived in [4]. For the case of large signal energy, the error probability for quantum detection approaches $P(E)_{\text{quantum}} \cong (1/2)e^{-2K_s}, K_s >> 1$, whereas in the same limit the classical receiver performs as $P(E)_{\text{classical}} \cong \sqrt{(2/\pi K_s)}e^{-K_s/2}, K_s >> 1$. The quantum receiver is exponentially better by a factor of 4, or 6 dB, than the coherent classical receiver; it can be seen in Fig. 9 that, for an error probability of 10^{-3}, the optimum quantum receiver outperforms the classical receiver by 5 dB.

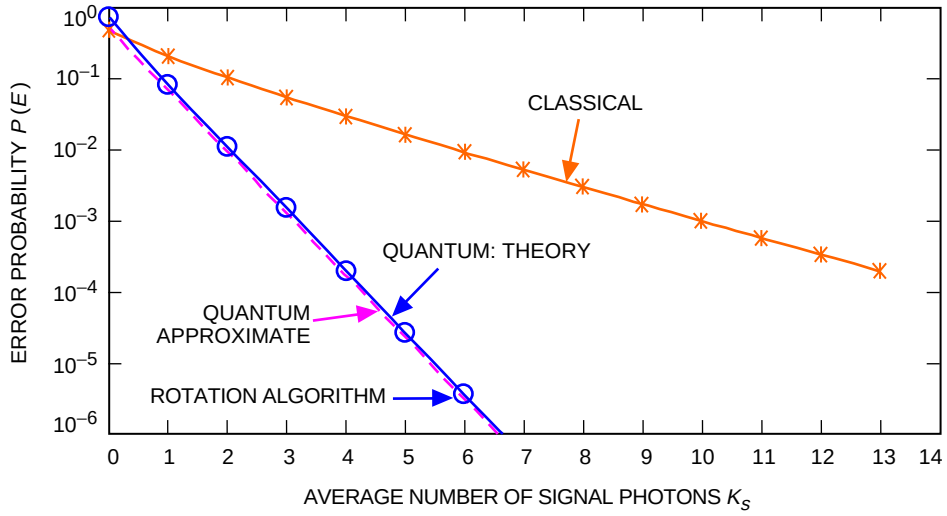

Fig. 9. Performance of QPSK and BPSK-2PPM signals with quantum and classical detection.

B. BPSK-2PPM Signals

This signal set is defined as $\{s_1(t) = (0, \alpha), s_2(t) = (0, -\alpha), s_3(t) = (\alpha, 0), s_4(t) = (-\alpha, 0)\}$, where the notation $(0, \alpha)$ refers to no signal in the first T-second slot and a signal with amplitude α in the second slot, etc. It is a dense signal set because two PPM symbols define two dimensions in signal space, but the signal set contains four symbols. The quantum states corresponding to this signal set are product states as defined in Eq. (3), with $|\psi_1\rangle = |0\rangle|\alpha\rangle, |\psi_2\rangle = |0\rangle|-\alpha\rangle, |\psi_3\rangle = |\alpha\rangle|0\rangle, |\psi_4\rangle = |-\alpha\rangle|0\rangle$. Pairwise overlaps must be computed using the tensor-product rule defined in Eq. (4), resulting in the following overlaps: $\langle\psi_3|\psi_4\rangle = \langle\alpha|-\alpha\rangle\langle 0|0\rangle = e^{-(1/2)|\alpha+\alpha|^2} = e^{-2|\alpha|^2}$ and $\langle\psi_1|\psi_3\rangle = \langle\psi_1|\psi_4\rangle = \langle\psi_2|\psi_3\rangle = \langle\psi_2|\psi_4\rangle = e^{-|\alpha|^2}$. Note that this overlap matrix is identical to that of QPSK signals; therefore, optimization of the product states yields the same performance as QPSK, despite the different state descriptions. The performance of the coherent receiver observing BPSK-2PPM signals is also exactly the same as for QPSK, since both signal sets can be represented as biorthogonal extensions of orthogonal signals in classical signal space, and hence the results of Fig. 9 apply.

V. Capacity of the Binary Channel with Quantum Detection

As an interesting application of state–space optimization, we compute the capacity of the binary OOK channel with quantum detection, and compare it to the capacity obtained with classical photon counting. Note that, for the noiseless quantum model, photon counting leads to a z-channel, whereas optimum quantum detection results in a binary symmetric channel (BSC). For an arbitrary rotation of the measurement states with respect to the signal states, the transition probabilities are not equal, and hence a generalized (asymmetric) binary channel model must be considered. Our approach for determining the capacity of the binary channel is to compute the mutual information between input and output for each rotation of the measurement states, starting with photon counting where one of the measurement states is aligned with the ground state, and compute the mutual information as a function of symbol input probability, β, for each rotation away from this configuration. For each rotation, the maximum of the mutual information as a function of β is recorded. The global maximum of the mutual information over all input probabilities and rotations is the capacity of the binary channel.

The input alphabet is denoted by A and the output alphabet by B. The input alphabet consists of the two symbols $a_1 = 0$ and $a_2 = 1$. Likewise, the output can take on one of two values, namely $b_1 = 0$ or $b_2 = 1$. The probability that a 0 is transmitted is β, whereas the probability of a transmitted 1 is $1 - \beta \equiv \overline{\beta}$. The probability that b_2 is received, given that a_1 was transmitted, is p, while the probability that b_1 is received, given that a_2 was transmitted, is q. These relationships are illustrated in Fig. 10.

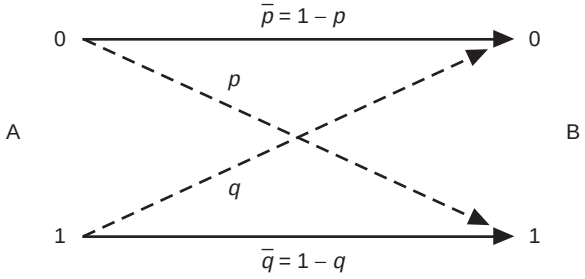

Fig. 10. The binary channel.

According to Shannon's first theorem, an average of $H(A)$ bits of information is needed to specify one input symbol, where $H(A)$ is the entropy of the source defined as $H(A) = \sum_A P(a) \log \left[1/P(a)\right]$ [7]. However, if we are allowed to observe the output symbol produced by that input, then we need only $H(A|B) = \sum_A P(a,b) \log \left[1/P(a|b)\right]$ bits to specify an input symbol, on the average. It follows, therefore, that observation of a single output symbol provides us with $H(A) - H(A|B)$ bits of information, on the average. This difference is called the mutual information between input A and output B, denoted by $I(A;B)$. It is non-negative and symmetric so that $I(A;B) = I(B;A)$; hence, we can also write $I(A;B) = H(B) - H(B|A)$. Writing the conditional output entropy, $H(B|A)$, explicitly yields

$$H(B|A) = \sum_A P(a) \sum_B P(b|a) \log \left[\frac{1}{P(b|a)}\right]$$

$$= \beta \left(p \log \frac{1}{p} + \bar{p} \log \frac{1}{\bar{p}}\right) + \overline{\beta} \left(q \log \frac{1}{q} + \bar{q} \log \frac{1}{\bar{q}}\right) \tag{24a}$$

The output entropy can be expressed as

$$H(B) = \sum_B P(b) \log \left[\frac{1}{P(b)} \right]$$

$$= \left(\beta\bar{p} + \overline{\beta}q\right) \log \left(\frac{1}{\beta\bar{p} + \overline{\beta}q} \right) + \left(\beta p + \overline{\beta}\bar{q}\right) \log \left(\frac{1}{\beta p + \overline{\beta}\bar{q}} \right) \tag{24b}$$

where we made use of the fact that $P(b) = P(b|a_1)P(a_1) + P(b|a_2)P(a_2)$. Combining Eqs. (24a) and (24b) yields the mutual information for the binary channel as

$$I(A;B) = \left[\left(\beta\bar{p} + \overline{\beta}q\right) \log \left(\frac{1}{\beta\bar{p} + \overline{\beta}q} \right) + \left(\beta p + \overline{\beta}\bar{q}\right) \log \left(\frac{1}{\beta p + \overline{\beta}\bar{q}} \right) \right]$$

$$- \left[\beta \left(p \log \frac{1}{p} + \bar{p} \log \frac{1}{\bar{\beta}} \right) + \overline{\beta} \left(q \log \frac{1}{q} + \bar{q} \log \frac{1}{\bar{q}} \right) \right] \tag{25}$$

Note that the mutual information of the z-channel and the BSC can be obtained by setting $p = 0$ and $p = q$, respectively.

Our approach for determining the capacity of the quantum channel is to start with a rotation angle of zero between the ground state and its measurement state (corresponding to photon counting, as we have shown above), and to compute the mutual information defined in Eq. (25) as a function of $\beta, 0 \leq \beta \leq 1$ for each rotation in the signal plane, until the measurement state corresponds with the signal state. Since different rotations yield different projections onto the measurement states, the values of p and q change with each rotation.

Examples of mutual information and capacity for the binary channel with OOK modulation are shown in Fig. 11, as a function of the input probability β, for an average value of one photon per symbol (or two photons per signal pulse). Only the limiting cases of optimum quantum measurement and photon

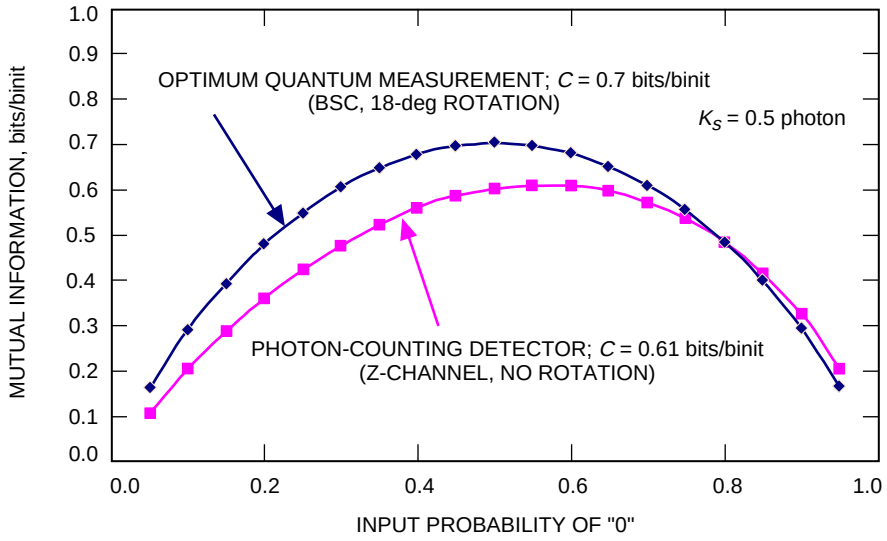

Fig. 11. Mutual information and capacity of the binary channel, with quantum and classical detection.

counting are included. From Fig. 2, the error probabilities for $K_s = 0.5$ are approximately 0.1025 and 0.18 for quantum and direct detection, respectively. The global maximum value of mutual information was found to occur with optimum quantum measurement, at an input probability of $\beta = 0.5$. With photon-counting detection, for which the asymmetric z-channel is the correct representation, the maximum mutual information occurs at a higher value of input probability, namely at $\beta = 0.55$. The value of the maximum mutual information was found to be 0.7 bits/binit for quantum detection and 0.61 bits/binit for photon counting, verifying that optimum quantum detection achieves higher capacity as well as better average error performance than photon-counting detection.

VI. Conclusions

In this article, we have examined the improvement that can be gained over classical detection techniques by the use of optimum quantum detection. The basics of quantum theory were reviewed and applied to determine the performances of several optical modulation schemes. The performance of an OOK optical receiver was evaluated both classically and quantum mechanically to introduce the concept of quantum-mechanical measurement states starting with familiar photon-counting detection. To evaluate the optimum quantum detector, a state–space solution was applied. The concept of dense signal states was introduced, and the performance of the optimum quantum measurement compared to the corresponding classical solution. In all cases evaluated, namely OOK, ternary signals, QPSK, and BPSK 2PPM, the optimum quantum receiver performed significantly better than its classical counterpart, with improvements of 5 dB demonstrated in some cases. Finally, the rotation algorithm was applied to determine the capacity of the binary channel when optimum quantum detection is used; it was found that the optimum quantum approach attained higher channel capacity than the classical receiver. We conclude, therefore, that the optimum quantum receiver performs better both in terms of average error probability and channel capacity than its classical counterpart.

Acknowledgments

The authors would like to thank Sam Dolinar and Jon Hamkins of the Information Processing Group for informative discussions during the course of this research.

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] 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.

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

[5] V. Vilnrotter and C.-W. Lau, "Quantum Detection Theory for the Free-Space Channel," *The InterPlanetary Network Progress Report 42-146, April–June 2001*, Jet Propulsion Laboratory, Pasadena, California, pp. 1–34, August 15, 2001.
http://ipnpr.jpl.nasa.gov/progress_report/42-146/146B.pdf

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

[7] N. Abramson, *Information Theory and Coding*, New York: McGraw Hill, 1963.