COMMUNICATION SCIENCES

AND

ENGINEERING

XIV. PROCESSING AND TRANSMISSION OF INFORMATION

Academic and Research Staff

Prof. Peter Elias
Prof. Robert G. Gallager
Prof. Robert S. Kennedy
Prof. Adrian Segall
Prof. Jeffrey H. Shapiro
Prof. John M. Wozencraft
Dr. Horace P. H. Yuen

Graduate Students

Roger J. Camrass
Vincent Chan
Samuel J. Dolinar, Jr.
John G. Himes
Pierre A. Humblet
Jeffry R. Long
Thomas L. Marzetta
Gerard P. Massa
Nam-soo Myung
Howard F. Okrent
Woo H. Paik
John J. Poklemba
Stanley R. Robinson
Mahmoud Tebyani
Alexander S. Theodoru

RESEARCH OBJECTIVES AND SUMMARY OF RESEARCH

1. Optical Communication

The broad objectives of these investigations are (i) to formulate communication models for important optical channels from the underlying physical processes, (ii) to determine the fundamental limits on the communication performance that can be realized with these channels, (iii) to develop communication techniques that achieve, or approach, these limits, and (iv) to establish the validity and guide the evolution of the theoretical results through experiment.

a. Quantum Communication Theory

National Aeronautics and Space Administration (Grant NGL 22-009-013)

Robert S. Kennedy, Jeffrey H. Shapiro, Horace P. H. Yuen

We are investigating the fundamental limits imposed upon optical communications systems by quantum effects. Vincent Chan has completed a doctoral dissertation[1] and presented a paper on the implementation of quantum measurements.[2] During the past year we have also defined the conditions that are necessary and sufficient for a receiver to be optimum[3, 4] and for the feedback in an (optical) closed-loop receiver to be optimum. We have determined the structure and performance of quantum receivers for the space communication channel[5, 6] and made a study (see Sec. XIV-A) of possible means of generating a class of quantum states to yield improved system performance.[7]

During the coming year, a major objective is to determine the extent to which the performance of optimized quantum receivers can be made substantially superior to that of conventional receivers. This can only occur in systems that operate at high information rates or employ novel quantum states. We now seek to determine whether it does occur in the first instance, and how to generate the states in the second instance.

A second objective is to establish the performance advantage of optical feedback receivers. These receivers have already been shown to offer significant improvement over conventional receivers in some isolated situations.[8, 9] We do not know yet whether this is an indication of a universal advantage or only an isolated benefit.

References

1. V. W. Chan, "Characterization of Measurements in Quantum Communications," Ph.D. Thesis, Department of Electrical Engineering, M. I. T., August 1974.

2. V. W. Chan, "Two Realizations of Quantum Measurements Characterized by Generalized Operator-Valued Measures," 1974 IEEE International Symposium on Information Theory, Notre Dame, Indiana, October 28-31.

3. H. P. Yuen, R. S. Kennedy, and M. Lax, "Optimum Testing of Multiple Hypotheses in Quantum Detection Theory," 1974 IEEE International Symposium on Information Theory, Notre Dame, Indiana, October 28-31.

4. H. P. Yuen, R. S. Kennedy, and M. Lax, "Optimum Testing of Multiple Hypotheses in Quantum Detection Theory" (to appear in IEEE Trans. on Information Theory).

5. R. S. Kennedy, "The M-ary Pure State Quantum Detection Problem," 1974 IEEE International Symposium on Information Theory," Notre Dame, Indiana, October 28-31.

6. R. S. Kennedy, "Uniqueness of the Optimum Receiver for the M-ary Pure State Problem," Quarterly Progress Report No. 113, Research Laboratory of Electronics, M. I. T., April 15, 1974, pp. 129-130.

7. H. P. Yuen, "Performance Improvement of Optical Communication Systems by a Quantum State Generating Receiver," 1974 IEEE International Symposium on Information Theory, Notre Dame, Indiana, October 28-31.

8. S. J. Dolinar, "An Optimum Quantum Receiver for the Binary Coherent State Channel," Quarterly Progress Report No. 111, Research Laboratory of Electronics, M.I.T., October 15, 1973, pp. 115-120.

9. S. J. Dolinar, "An Optimum Receiver for the Binary Coherent State Quantum Channel," 1974 IEEE International Symposium on Information Theory, Notre Dame, Indiana, October 28-31.

b. Improved Low-Visibility Communication

National Science Foundation (Grant GK-41464)

Robert S. Kennedy, Jeffrey H. Shapiro

This investigation, which is carried out jointly with the M. I. T. Center for Materials Science and Engineering, is concerned with the performance of terrestrial line-of-sight communication systems under conditions of low visibility. Its objective is to determine the extent to which performance can be improved through appropriate system design, and to develop the devices for achieving this improvement. The potential for improvement resides in the energy and information contained in the scattered component of the received field.

Preliminary experimental results obtained during the past year indicate that, in fact, substantial performance improvements can be realized, but the variability of those results dictates that a more exhaustive set of propagation measurements be undertaken to determine the magnitude of the energy in the scattered component of the received field as a function of meteorological conditions. Accordingly, the major objective in 1975 will be to obtain the required experimental data for a line-of-sight path in the Boston area that is approximately 10 miles long. This will entail measurements of the scattered and unscattered energy, the angular spectrum of the scattered field, the time dispersion of the scattered field (multipath spread), and the frequency dispersion of the scattered field (fading rate). Measurements will be made at 0.63, 0.69, and 2.06 μm wavelength. The data will be correlated with meteorological conditions and used to predict the improvement in system performance that can be expected with systems that are specifically designed for low-visibility operation.

In parallel with this experimental program we shall continue our determination of the structure and performance of receivers that best utilize the scattered component of the received field (see Sec. XIV-B).

c. Propagation and Communication through Atmospheric Turbulence

National Science Foundation (Grant GK-41098)

Jeffrey H. Shapiro

Because the primary effects of atmospheric turbulence on optical wave propagation are of a spatial nature, accurate spatial channel models are required to assess the limitations imposed by turbulence on spatial or spatiotemporal modulation systems. Thus a major thrust of our research has been to develop propagation models for the turbulent atmosphere that preserve the relevant spatial features of the propagation process and offer insight into the design and performance of optical communication systems. The second emphasis in our research is on the study of adaptive communication systems, i.e., systems in which channel estimates are made and used to compensate for the effects of turbulence.

During the past year, we have completed a study of the normal-mode decomposition of the turbulent atmosphere.[1] We found that the atmosphere exhibits far-field and near-field propagation regimes that are similar to those of the free-space channel.

Previously we had used the far-field propagation results to show how earth-to-space power transfer through atmospheric turbulence may be maximized by tracking (reciprocity-pointing) a pilot tone received on the ground from a laser aboard the spacecraft.[2] In the past year we have evaluated the reduction in reciprocity-tracking performance that occurs because the ground transmitter must "point ahead" to compensate for spacecraft motion.[3] Although the gain reduction may be substantial, the average power received at the spacecraft from the adaptive system increases continually in proportion to the transmitter area on the ground, unlike the performance of a nonadaptive transmitter which saturates when the transmitter diameter on the ground exceeds the atmospheric phase-coherence length. Work continues on the performance of phase-compensation transmitters and the bit error rate of reciprocity-tracking communication systems.

We have used the near-field propagation results to study coherent imaging through turbulence. We have found that a channel matched-filter (CF) receiver is essentially the optimum imager when the atmospheric Green's function is known a priori, and we showed that this receiver achieves diffraction-limited resolution.[4] We are investigating transmitted-reference techniques for implementation of the CF receiver.

References

1. J. H. Shapiro, "Normal-Mode Approach to Wave Propagation in the Turbulent Atmosphere," Appl. Opt. 13, 2614-2619 (1974).

2. J. H. Shapiro, "Optimum Power Transfer through Atmospheric Turbulence Using State Knowledge," IEEE Trans., Vol. COM-19, No. 4, pp. 410-414, August 1971.

3. J. H. Shapiro, "Point-Ahead Limitation on Reciprocity Tracking," J. Opt. Soc. Am. 65, 65 (1975).

4. J. H. Shapiro, "Optimum Adaptive Imaging through Atmospheric Turbulence," Appl. Opt. 13, 2609-2613 (1974).

2. Complexity of Networks and Algorithms

U. S. Army Research Office – Durham (Contract DAHC04-74-C-0027)

Peter Elias

During the past year Donna J. Brown completed a Master's thesis on the complexity of decoders for optimal variable-length codes.[1] A paper on this topic is in preparation. A paper on the complementary topic of codes that can be easily decoded and are uniformly pretty good for a large class of sources is scheduled for publication.[2] At present, Howard F. Okrent is exploring the relation between the retrieval models that we have been using and the more dynamic model of a data algebra, in which a data base is represented by the sequence of operations which built it up.

References

1. Donna J. Brown, "Complexity of Acceptors for Prefix Codes," S.M. Thesis, Department of Electrical Engineering, M.I.T., May 1974.
2. Peter Elias, "Universal Codeword Sets and Representations of the Integers" (to appear in IEEE Trans. on Information Theory).

JSEP

3. Information Theory of Data Processing Systems

Joint Services Electronics Program (Contract DAAB07-74-C-0630)

National Science Foundation (Grant GK-37582)

Peter Elias

The complexity problems summarized in Section XIV-2 arose from consideration[1] of the minimal complexity of computations performed by Turing machines and other halting automata when the freedom of representation of input and output which is usual in communications problems is allowed. The most natural application of such results is to problems of information storage and retrieval. Richard A. Flower[2] has completed a doctoral thesis giving informational lower bounds to the complexity of updating a data base. A paper that applies earlier results on the complexity of static retrieval and Flower's results on updating to several simple examples has been accepted for publication.[3] Future plans include a systematic treatment of the general static retrieval problem.

References

1. Peter Elias, "Minimum Times and Memories Needed to Compute the Values of a Function," J. Comput. & Syst. Sci. 9, 196-212 (1974).
2. Richard A. Flower, "An Analysis of Optimal Retrieval Systems with Updates," Ph.D. Thesis, Department of Electrical Engineering, M. I. T., August 1974; to appear as Technical Report 487, Research Laboratory of Electronics, M. I. T.
3. Peter Elias and Richard A. Flower, "The Complexity of Some Simple Retrieval Problems" (to appear in the Journal of the Association for Computing Machinery).

JSEP

A. GENERALIZED COHERENT STATES: STATISTICS OF TWO-PHOTON LASERS AND ELIMINATION OF QUANTUM NOISE

National Aeronautics and Space Administration (Grant NGL 22-009-013)

Horace P. H. Yuen

For a radiation mode of frequency ω with photon annihilation operator a, the coherent states $|\alpha\rangle(a|\alpha\rangle = \alpha|\alpha\rangle)$ have relatively large quantum fluctuations[1,2] when $\omega/2\pi \geq 10^{12}$ Hz. With $a = a_1 + ia_2$, $\alpha = \alpha_1 + i\alpha_2$ for self-adjoint a_1, a_2 and real α_1, α_2, a coherent state $|\alpha\rangle$ gives $\langle \Delta a_1^2\rangle = \langle \Delta a_2^2\rangle = 1/4$. A noise energy $\hbar\omega/4$ (actually a power spectral density, that is, power per unit frequency) is obtained if either of the quadrature components a_1 or a_2 is measured. This quantum noise is frequently dominant over other noise sources in a communication situation; for example, an equivalent noise temperature of ~3400°K is obtained for $\hbar\omega/4$ at the YAG laser frequency. States with $\langle \Delta a_1^2\rangle \ll 1/4$ and correspondingly larger $\langle \Delta a_2^2\rangle$, however, are permitted by the uncertainty principle $\langle \Delta a_1^2\rangle \langle \Delta a_2^2\rangle \geq 1/16$. (The roles of a_1 and a_2 are completely symmetrical here.) If such small $\langle \Delta a_1^2\rangle$ states can be generated by an explicit physical process, they can be used profitably as local oscillators in signal reception. These states may also offer significant improvement in the signal-to-noise ratio $(S/N)_1 = \langle a_1\rangle^2/\langle \Delta a_1^2\rangle$. We shall show that stimulated two-photon emission puts out "generalized coherent states" possessing the desired quantum noise behavior and forming a natural generalization of the usual coherent states.

An ordinary coherent state is a special degenerate case of a generalized coherent state, which is basically a minimum uncertainty state. In general, a generalized coherent state differs from a coherent state in several ways: it is generated by different photon processes, or equivalently it describes the quantum states of radiation obtained from different photon processes, and it has different quantum statistical properties. These and other differences, which are interrelated, will now be described.

The generalized coherent states $|\beta\rangle_g$ are the eigenstates of b with eigenvalues β. For complex numbers μ, ν,

$$b = \mu a + \nu a^{\dagger}; \qquad |\mu|^2 - |\nu|^2 = 1 \tag{1}$$

so that $[b, b^{\dagger}] = I$. From this commutator it follows at once[3] that b has exactly the same properties as a. In particular, it acts as the lowering operator for eigenstates of $b^{\dagger}b$, and itself possesses an overcomplete set of eigenstates.[2] Note that b is an operator on the Hilbert space of radiation states with frequency ω, and not on radiation states at a different frequency. We are considering only a single frequency.

The general coherent-state wave function of $|\beta\rangle_g$, from the differential operator representation of a and b, is

$$\langle \alpha|\beta\rangle_g = \mu^{-1/2} \quad \exp\left\{-\frac{1}{2}|\alpha|^2 - \frac{1}{2}|\beta|^2 - (\nu/2\mu)\alpha^{*2} + (\nu^*/2\mu)\beta^2 + (1/\mu)\alpha^*\beta + i\theta_o\right\}, \tag{2}$$

where θ_o is a real constant. The density operator ($\rho =$) $|\beta\rangle_g\, {}_g\langle\beta|$ cannot be represented as a positive superposition of coherent states. In fact, the P-representation of $|\beta\rangle_g\, {}_g\langle\beta|$ does not exist (or is very singular),[1,2] except when $\nu = 0$ where it reduces to a two-dimensional δ-function as an ordinary coherent state.

For the state $|\beta\rangle_g$, from (2), we have

$$\langle a\rangle = {}_g\langle\beta|a|\beta\rangle_g = \beta' = \mu^*\beta - \nu\beta^*; \qquad \langle b\rangle = \beta \tag{3}$$

$$\langle \Delta a_1^2\rangle = \frac{1}{4}|\mu-\nu|^2, \qquad \langle \Delta a_2^2\rangle = \frac{1}{4}|\mu+\nu|^2 \tag{4}$$

$$\langle \Delta a_1 \Delta a_2\rangle = \langle (a_1-\beta_1')(a_2-\beta_2')\rangle = i(1+\mu^*\nu-\nu^*\mu)/4. \tag{5}$$

Define

$$a' = a \exp\left\{i \tan^{-1} \frac{i(\mu^*\nu-\nu^*\mu)}{2|\mu||\nu| + \mu^*\nu + \nu^*\mu}\right\}. \tag{6}$$

Then it is not difficult to show that (2) represents the minimum uncertainty states for the product $\langle \Delta a_1'^2\rangle \langle \Delta a_2'^2\rangle \geq 1/16$, with

$$\langle \Delta a_1'^2\rangle = \frac{1}{4}(|\mu|-|\nu|)^2; \qquad \langle \Delta a_2'^2\rangle = \frac{1}{4}(|\mu|+|\nu|)^2. \tag{7}$$

When $\mu = \delta\nu$ for real δ, (2) becomes the usual minimum uncertainty states[3] for $\langle \Delta a_1^2\rangle \langle \Delta a_2^2\rangle$. The eigenstates of a_1, a_2 are also included in the limit $\mu, \nu \to \infty$. Thus generalized coherent states are also "generalized minimum uncertainty states." In particular, it follows from (4) and (7) that the quantum fluctuations in a_1 and a_2 (or a_1' and a_2') can be exchanged in a generalized coherent state, while $\langle \Delta a_1^2\rangle$ and $\langle \Delta a_2^2\rangle$ are always larger than 1/4 in coherent states and their classical superpositions.

The possibility of having an absolutely small $\langle \Delta a_1^2\rangle \ll 1/4$ in $|\beta\rangle_g$ is an important advantage in its utilization as a local-oscillator state. For example, in the reception of a coherent-state signal with a size-limited detector, it is possible, by employing a local oscillator that generates radiation in state $|\beta\rangle_g$, to attenuate the quantum noise by an amount that compensates the large diffraction loss incurred in free-space or unguided propagation. We shall not go into the details of this problem here.

In principle, it is possible to have a state with $\langle \Delta a_1^2\rangle \to 0$, that is, an eigenstate of a_1. But for a fixed total radiative energy tr $\rho a^\dagger a \leq S$, a decrease of $\langle \Delta a_1^2 \rangle$ in a state $|\beta\rangle_g$ can be obtained only at the expense of spending a portion of available energy S in the form of added quantum noise energy tr $\rho(\Delta a)^\dagger (\Delta a) = |\nu|^2$. Thus $\langle \Delta a_1^2 \rangle \to 0$ requires $S \to \infty$. This consideration is not important in the context of local oscillators in a receiver, where a large amount of power is usually available to make $\langle \Delta a_1^2 \rangle$ sufficiently small. Thus a four order of magnitude reduction from the noise 1/4 of a_1 ($|\mu| \sim 10^2$) for a GHz bandwidth optical signal at $\omega/2\pi \sim 10^{15}$ Hz requires only a $|\nu|^2$ corresponding to ~1 microwatt. This will usually bring it down to the level of other extraneous noises. Furthermore, even a ten order reduction of the quantum noise requires ~1 W under the same condition.

On the other hand, the radiative power constraint places a major limitation on the signal-to-noise ratio $(S/N)_1$ when a transmitter generates information-carrying radiation in a state $|\beta\rangle_g$. (Note that it is not meaningful to talk about the (S/N) of a local oscillator, since the signal comes from the transmitter.) By spending a fraction $S/(2S+1)$ of S as quantum noise energy $|\nu|^2$, it can be shown that $(S/N)_1$ of a transmitter state $|\beta\rangle_g$ becomes $4S^2 + 4S$, compared with 4S in a coherent state. This leads to a higher information capacity even when the "channel" a_2 is not used. In the presence of other extraneous noises quantum-noise reduction in a transmitter state $|\beta\rangle_g$ can also be obtained with only a relatively small $|\nu|^2$ as we have shown.

These advantages are not available in an ordinary coherent state. While a coherent state is generated by a one-photon laser, a generalized coherent state may be obtained from a two-photon laser. For this purpose, consider the general quadratic Hamiltonian

$$H = \hbar\omega\left[f_1 a^\dagger a + f_2^* a^2 + f_2 a^{\dagger 2} + f_3^* a + f_3 a^\dagger\right], \tag{8}$$

where the c-numbers f_i may be time-dependent. The unitary time-development operator $U(t, t_o)$ corresponding to (8) transforms the vacuum state $|0\rangle$ at $t_o = 0$ to a state $|\beta(t)\rangle_g$ at time t with $\mu(t)$, $\nu(t)$ independent of f_3. Any μ, ν, β can be obtained by a proper choice of the f. This $U(t, t_o)$ is readily calculated explicitly by normal ordering techniques.[3]

Under this $U(t, t_o)$, a generalized coherent state will remain a generalized coherent state for all time, but with μ, ν time-variant. For example, when $f_1 = 1$, f_2 = constant, $f_3 = 0$ with $\lambda = (1-4|f_2|^2)^{1/2} \geq 0$, a $|\beta\rangle$ at t = 0 is changed to (2) with $\beta(t) = \beta$ and $\mu(t) = \cos \lambda\omega t + (i/\lambda) \sin \lambda\omega t$, $\nu(t) = i(2f_2/\lambda) \sin \lambda\omega t$. The quantum noise is oscillatory in this case, but monotone behavior can also be obtained. Thus when $f_1 = 1$, $f_2 = \frac{1}{2} \tanh (4\omega c_o t) - ic_o$, $f_3 = 0$, we get $\langle \Delta a_1^2(t) \rangle = \frac{1}{4} \exp(-4\omega c_o t)$, and so forth. For the resonant case $f_1 = 1$, $\omega f_2 = c_o \exp(-2i\omega t)$, $f_3 = 0$, we have $\mu(t) = e^{i\omega t} \cosh 2c_o t$, $\nu(t) = i\, e^{-i\omega t} \sinh 2c_o t$, and $|\beta(t)\rangle_g$ becomes the minimum uncertainty state for $a' = a \exp[i(\omega t - \pi/4)]$ with $\langle \Delta a_1'^2(t) \rangle = \frac{1}{4} \exp(-4c_o t)$.

An interpretation of the radiation Hamiltonian (8) can be given as the atom-field interaction in stimulated two-photon emission. (A somewhat more complicated interpretation as the photon interaction in a degenerate parametric process can also be given.) Four possible configurations (Fig. XIV-1) can be distinguished; population inversion is required in all cases. In Fig. XIV-1, cases (a) and (c) correspond to an

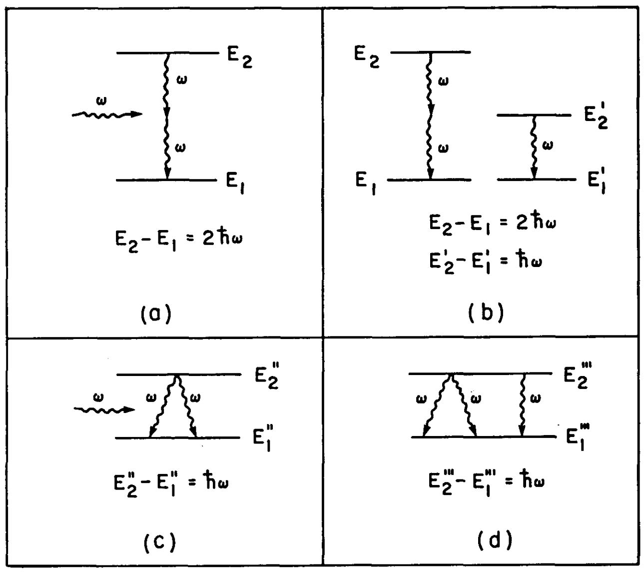

Fig. XIV-1. Transition processes for stimulated two-photon emission. In cases (a) and (c) the gain is obtained through a large external pump field at frequency ω; in cases (b) and (d) the gain is initially obtained from the usual one-photon lasing mechanism.

initial radiation state ρ having finite tr $\rho a^{\dagger} a$ with $f_3 = 0$; cases (b) and (d) correspond to an initial $|0\rangle$ with $f_3 \neq 0$. In cases (a) and (b), f_2 is proportional to an atomic polarization, whereas in (c) and (d) it is proportional to the square of the polarization; f_3 is always proportional to a polarization. Note that f_1 must always be 1 in order that H include the free radiation energy at frequency ω.

When such a quantum oscillator operates far above threshold, the fluctuations in atomic polarizations are small physically so that the polarizations may be treated as c-numbers in a first approximation. Mathematically the amplitude stabilization and a phase linewidth inversely proportional to the average photon number are general features of self-sustained oscillators, quantum or classical.[4,5] Therefore, a device of this type operating far above threshold can be expected to produce a state $|\beta\rangle_g$ with further small fluctuations in β, in a way exactly analogous to ordinary lasers ($f_1 = 1$, $f_2 = 0$) where $|\alpha\rangle$ is produced with small fluctuations in α.[4,1] An investigation of the detailed average and statistical behavior of such a two-photon laser with a full quantum treatment of the atoms can be carried out with existing techniques.[4, 3]

There have been many experimental observations of two-photon processes, including spontaneous[6] and enhanced[7] two-photon emission, although two photons of different frequencies are usually involved. Stimulated two-photon emission was first proposed for giant pulse generation.[8] There is no apparent reason why an oscillator such as that in Fig. XIV-1 cannot actually be achieved.

When the processes illustrated by Fig. XIV-1, cases (a) and (c), are used for amplification of the incoming radiation with $f_3 = 0$ in (8), an initial $|\beta\rangle$ becomes a $|\beta\rangle_g$ with the same β; that is, $U(t, t_o)|\beta\rangle = |\beta\rangle_g$. According to (3) the average field then changes from β to $\mu^*\beta - \nu\beta^*$. In general, this process may be interpreted as the action of one linear amplifier and one linear attenuator acting separately on the two quadrature components β_1 and β_2. For example, when μ, ν are real, we obtain ${}_g\langle\beta|a_1|\beta\rangle_g = (\mu-\nu)\beta_1$ and ${}_g\langle\beta|a_2|\beta\rangle_g = (\mu+\nu)\beta_2$. This phase-dependent behavior of the two-photon amplifier can be physically understood by observing its similarity to a degenerate parametric amplifier. Note that the quantum noises in a_1 and a_2 will be correspondingly attenuated and amplified so that the signal-to-noise ratios $(S/N)_1$ and $(S/N)_2$ remain invariant for all time whenever $f_3 = 0$. Such invariance, of course, is required by the uncertainty principle. This process should not be confused with the possibility of obtaining through a different process a higher $(S/N)_1$ in a transmitter state $|\beta\rangle_g$ under the meaningful constraint of a fixed radiative power. Since the two-photon process operates as an amplifier in the region above threshold, it provides a mechanism for the practical realization of an ideal linear amplifier which we shall not discuss here.

References

1. R. J. Glauber, in C. Dewitt et al. (Eds.), Quantum Optics and Electronics (Gordon and Breach, New York, 1965), p. 65.
2. J. R. Klauder and E. C. G. Sudarshan, Fundamentals of Quantum Optics (W. Benjamin, New York, 1968).
3. W. H. Louisell, Quantum Statistical Properties of Radiation (John Wiley and Sons, Inc., New York, 1973).
4. M. Lax, in M. Chretien et al. (Eds.), Statistical Physics, Vol. 2 (Gordon and Breach, New York, 1968), p. 271.
5. M. Lax, Phys. Rev. 160, 290 (1967).
6. M. Lipeles, R. Novick, and N. Tolk, Phys. Rev. Letters 15, 690 (1965).
7. S. Yatsiv, M. Rokni, and S. Barak, Phys. Rev. Letters 20, 1282 (1968).
8. P. P. Sorokin and N. Braslau, IBM J. Res. Develop. 8, 117 (1964).

B. ESTIMATION WITH FEEDBACK FOR DOUBLY STOCHASTIC POISSON PROCESSES

National Science Foundation (Grant GK-41464)

Stanley R. Robinson

1. Introduction

In this report we consider the structure and performance of a class of causal, minimum mean-square error (MMSE) estimators of a Gauss-Markov process observed through a conditional Poisson process. Results are derived for the case of a scalar observation and presented without proof for the vector observation case. Applications to closed-loop phase estimation in optical communication receivers are discussed.

2. Estimator Structure with Feedback

Assume that the Gauss-Markov process of interest has a finite-state representation described by the Ito equation

$$d\underline{x}_t = \underline{F}_t\underline{x}_t dt + \underline{G}_t d\underline{w}_t$$

$$\underline{y}_t = \underline{H}_t\underline{x}_t; \quad \underline{x}_{t_o} = \underline{x}_o,$$

where $\underline{F}$, $\underline{G}$, and $\underline{H}$ may be time-variant matrices, and $\{\underline{w}_t; t \geq t_o\}$ is a standardized vector Wiener process.

We wish to estimate $\underline{y}_t$ from the observation record of a doubly stochastic Poisson process (DSPP) $N_{t_o,t} = \{N_\sigma ; t_o < \sigma \leq t\}$ whose rate parameter, because of optical feedback introduced at the receiver, has the particular form

$$\lambda_t = \underline{c}^T(\underline{y}_t - \hat{\underline{y}}_t) + \lambda_o$$

$$= \underline{c}^T\underline{H}_t(\underline{x}_t - \hat{\underline{x}}_t) + \lambda_o.$$

Here the elements of the vector $\underline{c}$ and λ_o may be time-variant (but known) and independent of the process $\underline{y}_t$, the superscript T indicates transpose, and $\hat{\underline{y}}_t$ is the causal MMSE estimate of $\underline{y}_t$: $\hat{\underline{y}}_t = E[\underline{y}_t | N_{t_o,t}]$.

The exact stochastic differential equation for the causal MMSE estimate can be shown[1,2] to be

$$d\hat{\underline{x}}_t = \underline{F}_t\hat{\underline{x}}_t dt + \underline{\Sigma}_t\underline{H}_t^T\underline{c}_t(\lambda_o)^{-1}\left[dN_t - \lambda_o dt\right] \tag{1}$$

$$\hat{\underline{y}}_t = \underline{H}_t\hat{\underline{x}}_t; \qquad \hat{\underline{x}}_{t_o} = E[\underline{x}_o],$$

where $\underline{\Sigma}_t$ is the conditional error covariance $\underline{\Sigma}_t = E\left[(\underline{x}_t-\hat{\underline{x}}_t)(\underline{x}_t-\hat{\underline{x}}_t)^T \mid N_{t_o,t}\right]$. The exact conditional covariance is described by the Ito equation:

$$d\underline{\Sigma}_t = \left(\underline{F}_t\underline{\Sigma}_t + \underline{\Sigma}_t\underline{F}_t^T + \underline{G}_t\underline{G}_t^T\right)dt - \underline{\Sigma}_t\underline{M}\,\underline{\Sigma}_t(\lambda_o)^{-2}\,dN_t$$

$$+ E\left[(\underline{x}_t-\hat{\underline{x}}_t)(\underline{x}_t-\hat{\underline{x}}_t)^T\,(x_t-\hat{\underline{x}}_t)^T\,\underline{H}_t^T\underline{c} \mid N_{t_o,t}\right]\lambda_o^{-1}[dN_t-\lambda_o dt],$$

with $\underline{\Sigma}_{t_o} = \underline{\Sigma}_o = \text{cov}(\underline{x}_o)$ (a priori covariance) and $\underline{M} = \underline{H}_t^T\underline{c}\,\underline{c}^T\underline{H}_t$.

Following Snyder,[2] we seek to approximate the conditional covariance equation by dropping the last term on the right. In contrast to Snyder's approximation, it is sufficient to assume here that the conditional probability density of the errors is symmetrical about the origin. This appears intuitively to be a very reasonable assumption. The resulting approximate conditional covariance equation is given by

$$d\underline{\Sigma}_t = \left(\underline{F}_t\underline{\Sigma}_t + \underline{\Sigma}_t\underline{F}_t^T + \underline{G}_t\underline{G}_t^T\right)dt - \underline{\Sigma}_t\underline{M}\,\underline{\Sigma}_t\lambda_o^{-2}dN_t \tag{2}$$

and $\underline{\Sigma}_{t_o} = \underline{\Sigma}_o$.

Equations 1 and 2 describe the structure of the causal MMSE estimator, consistently with the assumptions made thus far. We observe that (2) is in the familiar matrix Riccati form that appears in many other estimation problems,[3-5] except that it is coupled to the observations so that conditional performance cannot be precomputed. Moreover, because of the nature of the data (dN_t/dt is impulsive at event times), implementation of (2) is particularly simple. Given the set of event times: $I = \{t_i;\ t_i$ are event times $\epsilon(t_o, t)$, and $i = 1, 2, \dots\}$, we can express the solution for $\underline{\Sigma}_t$ by the recursion relation

$$\underline{\Sigma}_t = \underline{\phi}_{v_2v_1}(t,t_o)\,\underline{\phi}^T_{v_2v_2}(t,t_o) + \sum_{i=0}^{p}\underline{\phi}_{v_2v_2}(t,t_i)\,\underline{N}_{t_i}\,\phi^T_{v_2v_2}(t,t_i),$$

where p is the largest index in I,

$$\underline{\Phi}(t,t_o) = \left[\begin{array}{c|c} \underline{\phi}_{v_1v_1} & \underline{0} \\ \hline \underline{\phi}_{v_2v_1} & \underline{\phi}_{v_2v_2} \end{array}\right]$$

is the transition matrix of the linear system

$$\begin{bmatrix} \dot{\underline{v}}_1 \\ \hline \dot{\underline{v}}_2 \end{bmatrix} = \left[\begin{array}{c|c} -\underline{F}_t^T & \underline{0} \\ \hline -\underline{G}_t\underline{G}_t^T & \underline{F} \end{array}\right] \begin{bmatrix} \underline{v}_1 \\ \hline \underline{v}_2 \end{bmatrix}$$

and

$$\underline{N}_{t_i} = \begin{cases} \underline{\Sigma}_o & ; \quad i = 0 \\ -\underline{\Sigma}_{t_i} \underline{M}\, \underline{\Sigma}_{t_i} \lambda_o^{-2}; & i \neq 0. \end{cases}$$

Observe that all the $\underline{\Phi}$ are precomputable and by the nature of the transition matrix (for a stable system), the conditional covariance will evolve in time toward the a priori covariance in a "long enough time interval" where there are no events.[3-6]

3. Estimator Performance

Since the conditional covariance is data-dependent, other methods for computation of performance are of interest. We shall determine first the Cramer-Rao lower bound for any causal unbiased estimate of $\underline{x}_t$ based on the data $N_{t_o, t}$ and then show that the conditional covariance averaged over the statistics of the data (i.e., the actual MS error covariance) is precisely that given by the lower-bound equation. Although estimation of $\underline{y}_t$ is the desired result, we shall investigate the performance of estimation of $\underline{x}_t$, since for the stated linear observation knowing the estimation performance of $\underline{x}_t$ is sufficient to determine the performance of the estimates of $\underline{y}_t$.

The lower bound for the covariance of any causal unbiased estimate $\underline{x}_t^*$, computed from the data $N_{t_o, t}$, is given[7] by

$$E\left[\left(\underline{x}_t-\underline{x}_t^*\right)\left(\underline{x}_t-\underline{x}_t^*\right)^T\right] \geq E\left[(\underline{x}_t-\hat{\underline{x}}_t)(\underline{x}_t-\hat{\underline{x}}_t)^T\right] \geq \underline{V}_t, \tag{3}$$

where

$$\dot{\underline{V}}_t = \underline{F}_t\underline{V}_t + \underline{V}_t\underline{F}_t^T + \underline{G}_t\underline{G}_t^T - \underline{V}_t\underline{Q}_t\underline{V}_t; \qquad \underline{V}_{t_o} = \underline{\Sigma}_o \tag{4}$$

and

$$\underline{Q}_t = E\left[\left(\underline{c}^T\underline{H}_t(\underline{x}-\hat{\underline{x}}_t)+\lambda_o\right)^{-1}\right]\underline{M}$$

$$= \left[\lambda_o^{-1}+\lambda_o^{-3}E\left[\left(\underline{c}^T\underline{H}_t(\underline{x}_t-\hat{\underline{x}}_t)\right)^2\right]+\dots\right]\underline{M}$$

$$\approx \lambda_o^{-1}\underline{M}. \tag{5}$$

Therefore, under the assumption that higher moments of the error are small relative to inverse powers of λ_o, we see that the lower bound is given by (3) with $\underline{Q}_t$ given by (5).

In contrast, now consider the actual MS performance of the estimator in question. We denote $E[\underline{\Sigma}_t] = \bar{\underline{\Sigma}}_t$ and take the expected value of (2) to obtain[1]

$$\dot{\bar{\underline{\Sigma}}} = \underline{F}_t\bar{\underline{\Sigma}}_t + \bar{\underline{\Sigma}}_t\underline{F}_t^T + \underline{G}_t\underline{G}_t^T - E[\underline{\Sigma}_t\underline{M}\,\underline{\Sigma}_t]\,\lambda_o^{-1}; \qquad \bar{\underline{\Sigma}}_{t_o} = \underline{\Sigma}_o.$$

The last term on the right involves moments that we are unable to calculate, so we seek to solve the problem indirectly by rewriting the equation as

$$\dot{\bar{\underline{\Sigma}}}_t = \underline{F}_t\bar{\underline{\Sigma}}_t + \bar{\underline{\Sigma}}_t\underline{F}_t^T + \underline{G}_t\underline{G}_t^T - \Delta\underline{G} - \bar{\underline{\Sigma}}_t\underline{M}\,\bar{\underline{\Sigma}}_t\lambda_o^{-1} \tag{6}$$

and

$$\Delta\underline{G} = E[(\underline{\Sigma}_t-\bar{\underline{\Sigma}}_t)\,\underline{M}(\underline{\Sigma}_t-\bar{\underline{\Sigma}}_t)]\,\lambda_o^{-1}.$$

We can easily verify that $\Delta\underline{G}$ is a positive semidefinite matrix. Therefore, if we assume that solutions exist to (4, with 5) and (6) and denote their solutions respectively by $\underline{P}_1 = \underline{P}_2 + \Delta\underline{P}$ and $\underline{P}_2$, we can determine an equation for $\Delta\underline{P}$:

$$\Delta\dot{\underline{P}} = (\underline{F}_t-\underline{P}_2\underline{M})\Delta\underline{P} + \Delta\underline{P}(\underline{F}_t-\underline{P}_2\underline{M})^T + \Delta\underline{G} - \Delta\underline{P}\,\underline{M}\Delta\underline{P}; \qquad \Delta\underline{P}_{t_o} = \underline{0}.$$

Again, the equation for $\Delta\underline{P}$ is in the form of the widely studied matrix Riccati equation. One property of the solution is that $\Delta\underline{P}$ is a positive semidefinite matrix.[5,6,8,9] That is, $\underline{P}_1 - \underline{P}_2 \geq 0$. $\underline{P}_1$ is the solution to the lower-bound equation, however, so that $\underline{P}_2 - \underline{P}_1 \geq 0$. These conditions can only be met simultaneously if $\underline{P}_1 = \underline{P}_2$; that is, the actual error covariance is precisely equal to the lower-bound covariance. (An ancillary result is $\Delta\underline{P} = \Delta\underline{G} = 0$.) Therefore the estimator is efficient in the sense that its error covariance is equal to that computed by using the Cramer-Rao lower bound. The fact that an estimator for a DSPP with a rate parameter that depends only on a linear function of the estimation error is efficient has a striking parallel with a linear observation of a Gaussian signal process in additive Gaussian noise.[3,7]

4. Extension to Detector Arrays

Our results can be extended to the use of detector arrays by using the data $\{\underline{N}_\sigma ; t_o < \sigma \leq t\}$, a doubly stochastic vector Poisson process with conditionally independent components. Assume that there are m elements in the array and that the vector DSPP has a vector rate parameter:

$$\underline{\lambda}_t = \underline{C}^T(\underline{y}_t - \hat{\underline{y}}_t) + \underline{\lambda}_o$$

$$= \underline{C}^T\underline{H}_t(\underline{x}_t - \hat{\underline{x}}_t) + \underline{\lambda}_o,$$

where $\underline{C} = [\underline{c}_1 \vdots \underline{c}_2 \vdots \dots \vdots \underline{c}_m]$ and $\underline{\lambda}_o = \left[\lambda_o^{(1)}\lambda_o^{(2)} \dots \lambda_o^{(m)}\right]^T$. By using the same assumptions as in the scalar case, the estimator structure can be shown[10] to be

$$d\hat{\underline{x}}_t = \underline{F}_t\hat{\underline{x}}_t dt + \underline{P}_t\underline{H}^T\underline{C}\,\underline{\Lambda}^{-1}[d\underline{N}_t - \underline{\lambda}_o dt],$$

where

$$\underline{\Lambda} = \text{diagonal matrix} = \left[\lambda_o^i\right]_{ii}$$

$$d\underline{N}_t = [dN_1 dN_2 \dots dN_m]^T$$

$$d\underline{P}_t = \underline{F}_t\underline{P}_t + \underline{P}_t\underline{F}_t^T + \underline{G}_t\underline{G}_t^T - \sum_{i=1}^{m} \underline{P}_t\underline{H}_t^T\underline{c}_i\underline{c}_i^T\underline{H}_t\underline{P}_t\left(\lambda_o^{(i)}\right)^{-2} dN_i$$

$$\underline{P}_{t_o} = \underline{\Sigma}_o.$$

Again, the estimator is efficient and if we denote $\underline{W}_t = E[\underline{P}_t]$, $\underline{W}_t$ is described by

$$d\underline{W}_t = \underline{F}_t\underline{W}_t + \underline{W}_t\underline{F}_t^T + \underline{G}_t\underline{G}_t^T - \sum_{i=1}^{m} \underline{W}_t\underline{H}_t^T\underline{c}_i\underline{c}_i^T\underline{H}_t\underline{W}_t\left(\lambda_o^i\right)^{-1}$$

$$\underline{W}_{t_o} = \underline{\Sigma}_o.$$

5. Applications: Phase Modulation in a Free-Space Channel

As an example assume that we receive a single plane wave with complex envelope: $U_s = A_t \exp[+j\phi_t]$, where A_t is a known amplitude and ϕ_t is the analog message process of interest with the scalar state representation $d\phi_t = -k\phi_t dt + \sqrt{2k}\, dw_t$. The closed-loop

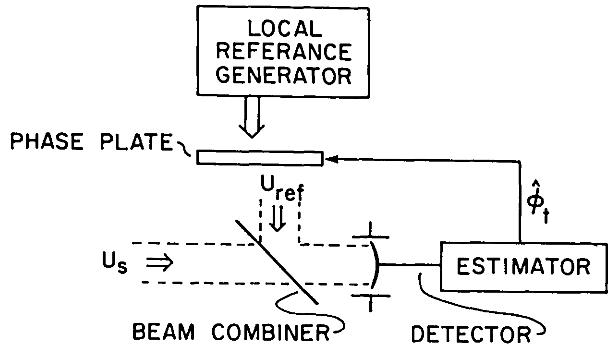

Fig. XIV-2. Closed-loop estimator.

estimator, shown in Fig. XIV-2, involves adding a locally generated estimate reference signal. Two types of reference will be considered:

"Heterodyne" $U_{ref} = B_t \exp[+j2\pi f_{1F}t + \hat{\phi}_t]$

"Homodyne" $U_{ref} = B_t \exp[+j(\hat{\phi}_t + \pi/2)]$.

Here the quotation marks indicate that the usual assumption of $B \gg A$ will not be made unless it is explicitly stated.[11,12] Under the assumption of small error (so that $\sin(\phi-\hat{\phi}) \approx \phi - \hat{\phi}$), the rate parameters are in the form $\lambda_t = c(\phi_t - \hat{\phi}_t) + \lambda_o$, where

$$\text{Homodyne} \begin{cases} c = 2AB\beta \\ \lambda_o = \beta[A^2+B^2] + \lambda d \end{cases}$$

$$\text{Heterodyne} \begin{cases} c = 2AB\beta \sin(2\pi f_{1F}t) \\ \lambda_o = \beta[A^2+B^2+2AB\cos(2\pi f_{1F}t)] + \lambda d \end{cases}$$

with $\beta = \eta Ad/h\nu$. Here η is the detector quantum efficiency, $h\nu$ is the energy of a photon, Ad is the detector area, and λd is the dark current rate parameter.

We note that the heterodyne rate parameter is similar in form to a phase subcarrier modulation considered by others,[13,14] for which it has been shown by simulation that the estimator is asymptotically efficient. The estimator structure is a special application of Eqs. 1 and 2 and will not be repeated here. We note, however, that the structure is quite similar to a phase-locked loop (PLL) where we have "closed the loop" optically.[15]

The mean-square error performance of the two systems is also of interest.

Using (4) and assuming that the system is in statistical steady state so that $\dot{V} = 0$, we can compute the performance from the algebraic equation. In the heterodyne case, we must also assume that f_{1F} is large enough that we need only consider the slowly varying components of c/λ_o.[16]

$$E\left[(\phi_t-\hat{\phi}_t)^2\right]\Bigg|_{\substack{\text{steady}\\ \text{state}}} = \frac{2}{1+\left(1+2c^2/\lambda_o k\right)^{1/2}} = \sigma_{ss}^2.$$

It is interesting that in both cases the error is decreased as B is increased. For large values of B

$$\text{Homodyne} \qquad \sigma_{ss}^2 = \frac{2}{1+\left(1+\dfrac{8\beta A^2}{k}\right)^{1/2}}$$

$$\text{Heterodyne} \qquad \sigma_{ss}^2 = \frac{2}{1+\left(1+\dfrac{4\beta A^2}{k}\right)^{1/2}}.$$

In this limit both performance and structure of the estimator are precisely those for the classical observation model with additive Gaussian noise.[15,16]

References

1. J. R. Clark, "Estimation for Poisson Processes with Applications in Optical Communications," Ph.D. Thesis, Department of Electrical Engineering, M.I.T., September 1971.
2. D. L. Snyder, "Filtering and Detection for Doubly Stochastic Poisson Processes," IEEE Trans., Vol. IT-18, No. 1, pp. 91-102, January 1972.
3. H. L. Van Trees, Detection, Estimation, and Modulation Theory: Part 1. Detection, Estimation, and Linear Modulation Theory (John Wiley and Sons, Inc., New York, 1968).
4. A. Gelb, Applied Optimal Estimation (The M. I. T. Press, Cambridge, Mass., 1974).
5. R. E. Kalman and R. Bucy, "New Results in Linear Filtering and Prediction Theory," Trans. ASME (J. Basic Engrg.), Vol. 83D, pp. 95-108, March 1961.
6. R. W. Brockett, Finite Dimensional Dynamical Systems (John Wiley and Sons, Inc., New York, 1970).
7. D. L. Synder and I. B. Rhodes, "Filtering and Control Performance Bounds with Implications on Asymptotic Separation," Automatica, Vol. 8, November 1972.
8. M. Athans and P. L. Falb, Optimal Control (McGraw-Hill Book Company, New York, 1966).

9. D. L. Kleinman, "On the Linear Regulator Problem and the Matrix Riccati Equation," Report ESL-R-271, Electronic Systems Laboratory, M. I. T., June 1966.

10. T. P. McGarty, Stochastic Systems and State Estimation (John Wiley and Sons, Inc., New York, 1974).

11. W. K. Pratt, Laser Communication Systems (John Wiley and Sons, Inc., New York, 1969).

12. E. V. Hoversten, "Optical Communication Theory," in Laser Handbook (North-Holland Publishing Co., Amsterdam, 1972).

13. D. L. Snyder and I. B. Rhodes, "Phase and Frequency Tracking Accuracy in Direct-Detection Optical-Communication Systems," IEEE Trans., Vol. COM-20, No. 6, pp. 1139-1142, December 1972.

14. R. H. Forrester and D. L. Snyder, "Phase-Tracking Performance of Direct-Detection Optical Receivers," IEEE Trans., Vol. COM-21, No. 9, pp. 1037-1039, September 1973.

15. H. S. Van Trees, Detection, Estimation, and Modulation Theory: Part 2. Nonlinear Modulation Theory (John Wiley and Sons, Inc., New York, 1971).

16. D. L. Snyder, The State Variable Approach to Continuous Estimation with Applications to Analog Communication Theory (The M.I.T. Press, Cambridge, Mass., 1969).