

Noncommuting Observables in  
Quantum Detection and Estimation Theory

Carl W. Helstrom\*

Department of Applied Physics and Information Science  
University of California, San Diego, La Jolla, Calif. 92037

and

Robert S. Kennedy

Department of Electrical Engineering  
Massachusetts Institute of Technology, Cambridge, Mass. 02139

Abstract

**CASE FILE  
COPY**

Basing decisions and estimates on simultaneous approximate measurements of noncommuting observables in a quantum receiver is shown to be equivalent to measuring commuting projection operators on a larger Hilbert space than that of the receiver itself. The quantum-mechanical Crámer-Rao inequalities derived from right logarithmic derivatives and symmetrized logarithmic derivatives of the density operator are compared, and it is shown that the latter give superior lower bounds on the error variances of individual unbiased estimates of arrival time and carrier frequency of a coherent signal. For a suitably weighted sum of the error variances of simultaneous estimates of these, the former yield the superior lower bound under some conditions.

\* The research of C.W. Helstrom was supported by Grant NGL 05-009-079, that of R.S. Kennedy by Grant NGL 22-009-013, both from the National Aeronautics and Space Administration.

## 1. Quantum Measurement

Quantum detection theory has been developed within the conventional framework of quantum mechanics, one of the principal tenets of which is that only observables associated with commuting operators can be simultaneously measured on the same system.<sup>[1-3]</sup> It has been suggested that this formulation is too restrictive, that noncommuting operators can be at least approximately measured on the same system, and that to include this possibility may permit more effective detection, appraised at a lower average Bayes cost.<sup>[4,5]</sup> An explication of this proposal requires prior discussion of quantum measurement in general. A more thorough treatment can be found in many a textbook on quantum mechanics.<sup>[6-8]</sup>

In quantum mechanics an observable is associated with an Hermitian operator, say  $F$ , which possesses an array of orthonormal eigenstates  $|f_n\rangle$  and eigenvalues  $f_n$ , defined by

$$F|f_n\rangle = f_n|f_n\rangle. \quad (1)$$

The operator  $|f_n\rangle\langle f_n|$  projects an arbitrary state vector onto the eigenvector  $|f_n\rangle$ , and these projection operators  $|f_n\rangle\langle f_n|$  commute for different indices  $n$  and sum to the identity operator,

$$\sum_n |f_n\rangle\langle f_n| = 1; \quad (2)$$

we say that they form a complete, commuting resolution of the identity.

The primary measurement concept of quantum mechanics envisions the measurement of an observable  $F$  as performed by an ideal apparatus that applies to the system such a resolution of the identity. The state of the system, originally--let us say--described by a density operator<sup>[9]</sup>  $\rho$ , is projected onto one and only one eigenstate  $|f_n\rangle$ , and the apparatus registers the associated eigen-

value  $f_n$ , which is termed the outcome of the measurement of  $F$ ; the probability of this event is

$$\text{Pr}(f_n) = \text{Tr}[\rho |f_n\rangle\langle f_n|] = \langle f_n | \rho | f_n \rangle, \quad (3)$$

where 'Tr' stands for the trace of an operator. Because  $\text{Tr} \rho = 1$ , (2) shows that these probabilities sum to 1. (An operator may be "degenerate", possessing a set of several eigenstates for each eigenvalue  $f_n$ . A practical measurement of  $F$  may then project the initial state onto the subspace spanned by that set.)

More generally, any finite or infinite resolution of the identity into commuting projection operators  $E_n$ ,

$$\sum_n E_n = 1, \quad E_n E_m = E_n \delta_{nm}, \quad (4)$$

is in principle measurable by some apparatus. A projection operator has eigenvalues 0 and 1, and the apparatus measuring the set  $\{E_n\}$  will register the value 1 for one of them, say  $E_k$ , and the value 0 for the rest; the probability that this happens is  $\text{Tr}(\rho E_k)$ . The fundamental role of projection operators in quantum measurement was emphasized by von Neumann.<sup>[10]</sup>

Commuting operators share a common resolution of the identity into commuting projection operators, and to say that commuting operators are being measured simultaneously on a system is only to assert that the apparatus is applying to it their common resolution of the identity. The eigenstates composing the resolution bear multiple parameters, each of which is associated with one of the commuting operators.<sup>[11]</sup> The simultaneous eigenstates  $|xyz\rangle$  of the operators  $X$ ,  $Y$ , and  $Z$  corresponding to the three rectilinear coordinates provide one example; another is the set of simultaneous eigenstates of the energy  $H$ , the total angular momentum  $L$ , and the component  $L_z$  of angular momentum along an arbitrary axis, for a particle in a spherically symmetrical potential. The observables  $X$ ,  $Y$ , and  $Z$ --or  $H$ ,  $L$ , and  $L_z$ --are said to be simultaneously

measurable or compatible. The simultaneous measurement of noncommuting observables has no meaning in the sense of this primary concept of measurement.

Gordon and Louisell have advanced a secondary concept of measurement that may lend meaning to the at least approximate simultaneous measurement of noncommuting observables.<sup>[12]</sup> They remind us that there exist overcomplete resolutions of the identity in terms of noncommuting projection operators. The most familiar example involves the coherent states  $|\alpha\rangle$ , which are the right eigenstates of the annihilation operator  $a$  that plays an important role in the quantum theories of harmonic oscillators and boson fields [13], [8, §3.7],

$$\begin{aligned} a|\alpha\rangle &= \alpha|\alpha\rangle, \\ a &= a_x + ia_y. \end{aligned} \quad (5)$$

The real part  $a_x$  of  $a$  is proportional to the coordinate operator, the imaginary part  $a_y$  to the momentum operator, and these do not commute, nor does the non-Hermitian operator  $a$  commute with its conjugate  $a^+$ , the creation operator,

$$aa^+ - a^+a = [a, a^+] = 1. \quad (6)$$

The eigenvalues  $\alpha = \alpha_x + i\alpha_y$  are complex, and the states  $|\alpha\rangle$  are overcomplete in the sense that

$$\pi^{-1} \int |\alpha\rangle\langle\alpha| d^2\alpha = \underline{1}, \quad d^2\alpha = d\alpha_x d\alpha_y. \quad (7)$$

As these states are not orthogonal,

$$\langle\alpha|\beta\rangle = \exp(\alpha^*\beta - \frac{1}{2}|\alpha|^2 - \frac{1}{2}|\beta|^2), \quad (8)$$

the projection operators  $|\alpha\rangle\langle\alpha|$  do not commute.

Because of (7) it is tempting to assert that the quantity  $P(\alpha_x, \alpha_y) = \pi^{-1} \langle\alpha|\rho|\alpha\rangle$  represents the joint probability density function (p.d.f.) of the outcomes  $\alpha_x$  and  $\alpha_y$  of simultaneous approximate measurements of the noncommuting operators  $a_x$  and  $a_y$  when the system has the density operator  $\rho$ , for

$$\iint P(\alpha_x, \alpha_y) d\alpha_x d\alpha_y = 1 . \quad (9)$$

Generalizing to arbitrary multiparameter noncommuting resolutions of the identity, Gordon and Louisell defined an 'ideal' measurement as one yielding outcomes having a joint p.d.f. calculated in this way, and the nonorthogonal states corresponding to the coherent states  $|\alpha\rangle$  were termed 'measurement states'. They showed how such an ideal measurement can be carried out on a system S by causing it to interact for a time with a second system A, termed the apparatus. After the interaction, commuting observables are measured on the apparatus A, or more generally on both S and A; and from the outcomes of these measurements, values of  $\alpha_x$  and  $\alpha_y$  are deduced that have a joint p.d.f. given by  $P(\alpha_x, \alpha_y)$  as required.

## 2. Detection

Let us apply to quantum detection the scheme of allowing the system S, which is now our ideal receiver, to interact with a suitable apparatus A. The decision about the received signal will be based on the outcome of measuring commuting projection operators on the combined system S+A. We suppose that an observer is to decide among M hypotheses  $H_1, H_2, \dots, H_M$  about the state of the receiver. It interacts with the apparatus A for a while, and at a later time t the density operator for the combined system under hypothesis  $H_j$  is  $\rho_j^{S+A}(t)$ . Let  $\{\Pi_j\}$  be a set of commuting projection operators forming an M-fold resolution of the identity,

$$\sum_{i=1}^M \Pi_i = 1_{S+A} \quad (10)$$

On the combination S + A we are to measure these M projection operators at time t, and if the k-th yields the value 1, hypothesis  $H_k$  is selected as true. The

average cost is then [1]

$$\bar{C} = \sum_{i=1}^M \sum_{j=1}^M \zeta_j C_{ij} \text{Tr}[\rho_j^{S+A}(t) \Pi_i] , \quad (11)$$

where  $\zeta_j$  is the prior probability of hypothesis  $H_j$  and  $C_{ij}$  is the cost of choosing  $H_i$  when  $H_j$  is true. Let  $\{\Pi_j(t)\}$  be the projection operators that minimize  $\bar{C}$  when the system  $S + A$  is observed at time  $t$ ; we call these optimum.

In the Schrödinger picture the density operator  $\rho_j^{S+A}(t)$  is related to the density operator  $\rho_j^{S+A}(t_0)$  at an earlier time  $t_0$  by<sup>[14]</sup>

$$\rho_j^{S+A}(t) = U(t, t_0) \rho_j^{S+A}(t_0) U^\dagger(t, t_0), \quad (12)$$

with  $U(t, t_0)$  a unitary operator obeying the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} U(t, t_0) = H U(t, t_0),$$

$$U(t_0, t_0) = \underline{1}, \quad U U^\dagger = \underline{1}, \quad (13)$$

where  $H$  is the Hamiltonian (energy) operator for the combined system  $S + A$  and  $\hbar$  is Planck's constant  $h/2\pi$ .

As the set  $\{\Pi_i(t)\}$  of projection operators minimizes the Bayes cost  $\bar{C}$  when measured at time  $t$ , by (12) the operators

$$\Pi_j(t_0) = U^\dagger(t, t_0) \Pi_j(t) U(t, t_0) \quad (14)$$

minimize  $\bar{C}$  when  $S + A$  is observed at time  $t_0$ . Because of the unitarity of  $U(t, t_0)$ , the set  $\{\Pi_j(t_0)\}$  also forms an  $M$ -fold resolution of the identity into commuting projection operators, and the  $\Pi_j(t_0)$  are optimum at time  $t_0$ . Since the minimization is carried out over all possible  $M$ -fold resolutions of the identity, the minimum Bayes cost  $\bar{C}_{\min}$  must be independent of the observation time  $t$ . To this independence Helstrom<sup>[15]</sup> and Liu<sup>[16]</sup> have previously adverted.

Now let us roll time back to an epoch  $t_0$  before the system  $S$  has come

into contact with the apparatus A. In the Schrödinger picture this amounts to applying the inverse unitary transformation  $U^+(t, t_0)$  to the state vectors of the combined system  $S + A$ . Because  $S$  and  $A$  are independent at this time  $t_0$ , the density operators  $\rho_j^{S+A}$  must now have the factored form

$$\rho_j^{S+A}(t_0) = \rho_j^S(t_0) \otimes \rho^A(t_0), \quad j = 1, 2, \dots, M, \quad (15)$$

on the product of the Hilbert spaces for system  $S$  and apparatus  $A$ . Furthermore, as the apparatus  $A$  before the interaction has no information about which hypothesis is true,  $\rho^A(t_0)$  in (15) must be independent of  $j$ . The Bayes cost is now

$$\bar{C} = \text{Tr} \left\{ \sum_{i=1}^M \sum_{j=1}^M \zeta_j C_{ij} \left[ \rho^A(t_0) \otimes \rho_j^S(t_0) \right] \Pi_i \right\} \quad (16)$$

Let us define positive definite Hermitian operators  $Q_i$  acting on the Hilbert space  $\mathcal{H}_S$  of the system  $S$  alone by

$$Q_i = \text{Tr}_A [\rho^A(t_0) \Pi_i(t_0)], \quad (17)$$

where  $\text{Tr}_A$  indicates a trace with respect to the Hilbert space  $\mathcal{H}_A$  of the apparatus  $A$ . Then the Bayes cost can be written

$$\bar{C} = \text{Tr}_S \sum_{i=1}^M \sum_{j=1}^M \zeta_j C_{ij} \rho_j^S Q_i ; \quad (18)$$

we consider henceforth only the epoch  $t_0$  and dispense with labeling our operators therewith. The operators  $Q_i$  form a resolution of the identity  $\mathbb{1}_S$  for the system  $S$ ,

$$\sum_{i=1}^M Q_i = \text{Tr}_A \left[ \rho^A \sum_{i=1}^M \Pi_i \right] = \text{Tr}_A [\rho^A \mathbb{1}_{S+A}]$$

or

$$\sum_{i=1}^M Q_i = \mathbb{1}_S, \quad (19)$$

but they do not necessarily commute and thus may not be measurable on system  $S$  alone.

Although it is not immediately obvious, any set of suitably well-behaved positive-definite Hermitian operators that satisfy (19) can be expressed in the form of (17) [17,18]. Specifically, it is possible to find a Hilbert space  $\mathcal{H}_A$  and a density operator  $\rho^A$  such that in the product space  $\mathcal{H}_S \otimes \mathcal{H}_A$  there exists a set  $\{\Pi_i\}$  of commuting projectors that when substituted into (17) yield the given  $Q_i$ . There is considerable arbitrariness in the choice of  $\rho^A$ ; the space  $\mathcal{H}_A$  must in general have infinite dimension. The density operator  $\rho^A$  may be taken as representing a pure state of the apparatus,

$$\rho^A = |\psi_A\rangle\langle\psi_A|. \quad (20)$$

Let  $|w_i\rangle$  be a complete orthonormal set spanning the Hilbert space  $\mathcal{H}_S$ . The vectors  $|w_i\rangle|\psi_A\rangle$  then span a linear manifold in  $\mathcal{H}_S \otimes \mathcal{H}_A$ , within which the operators  $Q_i \otimes \rho^A$  form a noncommuting resolution of the identity. By virtue of a theorem of M.A. Najmark's [18], this resolution can be extended to a resolution of the identity into commuting projection operators satisfying (10) and (17).

The equivalence of the constraints provided by (17) and (19) allows us to view the cost minimization problem as one of minimizing (18) over all sets of positive-definite Hermitian operators that satisfy (19). A special form of this problem was first posed and solved by Yuen [20,5], but its equivalence to the original minimization problem was not then established.

If the  $Q_i$  that minimize the Bayes cost  $\bar{C}$  in (18) do happen to commute, they must possess at least one common set of eigenstates  $|v_j\rangle$  and can be written



$$Q_i = \sum_j q_{ij} |v_j\rangle\langle v_j|, \quad (21)$$

$$0 \leq q_{ij} \leq 1, \quad \sum_{i=1}^M q_{ij} = 1.$$

Then the optimum strategy is to measure the projection operators  $|v_j\rangle\langle v_j|$  in  $\mathcal{H}_S$  and to choose hypothesis  $H_i$  with probability  $q_{ik}$  when the measurement of  $|v_k\rangle\langle v_k|$  yields the outcome 1. This would in general be a randomized strategy. However, once the set  $\{|v_k\rangle\langle v_k|\}$  has been specified, the decision process can be treated by classical decision theory in terms of the probabilities  $\Pr\{v_k|H_j\} = \langle v_k|\rho_j^S|v_k\rangle$  and the likelihood ratios formed from them. As we know that the minimum Bayes cost can be attained by a pure strategy, the  $q_{ik}$ 's must be either 0 or 1, and the  $Q_i$ 's must be projection operators. If we then take  $\Pi_i = Q_i \otimes \mathbb{1}_A$  as the decision operators in the product space of  $S + A$ , we can attain the same minimum Bayes cost  $\bar{C}_{\min}$ . Furthermore,  $\bar{C}_{\min}$  will be independent of the density operator  $\rho^A$  of the apparatus, and measurements on both system and apparatus cannot lead to a lower Bayes cost than measurements made on the system  $S$  alone.

There are three cases in which the optimum Bayes strategy is known to require measuring commuting projection operators on the space  $\mathcal{H}_S$  of the receiver alone: (a) binary decisions ( $M = 2$ )<sup>[1]</sup>, (b) decisions among  $M$  commuting density operators<sup>[1]</sup>, and (c) decisions among  $M$  linearly independent pure states<sup>[19]</sup>. In general, however, there is no guarantee that the positive-definite Hermitian operators  $Q_i$  that minimize the Bayes cost  $\bar{C}$  of (18) subject to (19) will commute. To the contrary, A.S. Holevo<sup>[17]</sup> has shown, by example, that minimum Bayes cost may sometimes be attained by a set of positive-definite noncommuting operators  $Q_i$  in the Hilbert space  $\mathcal{H}_S$ . He also showed that, in general, the minimizing operators obey the conditions

$$(K_i - \Lambda) Q_i = 0, \quad K_i - \Lambda \geq 0, \quad i = 1, 2, \dots, M, \quad (22)$$

where

$$K_i = \sum_{j=1}^M \zeta_j C_{ij} \rho_j^S \quad (23)$$

and

$$\Lambda = \sum_{i=1}^M K_i Q_i. \quad (24)$$

For  $\zeta_j \equiv M^{-1}$ ,  $C_{ij} = 1$ ,  $i \neq j$ ,  $C_{ii} = 0$ , these reduce to the conditions derived by H. Yuen [20]. If the minimizing operators  $Q_i$  turn out to commute, they can be measured on the receiver itself. Otherwise an extension to a set of commuting projection operators  $\Pi_i$  in a product space  $\mathcal{H}_S \otimes \mathcal{H}_A$  must be determined, and if the minimum Bayes cost is to be actually attained, some way of measuring them must be invented.

### 3. Estimation

Determining the values of signal parameters such as amplitude, carrier frequency, and time of arrival corresponds in quantum theory to estimating the parameters  $\theta_1, \theta_2, \dots, \theta_m$  of the density operator  $\rho^S(\theta_1, \theta_2, \dots, \theta_m)$  of a receiver of the signal. Parameter estimation is a continuous version of multiple hypothesis testing; costs of errors as functions of the true values and the estimates lead to an average cost similar to  $\bar{C}$  in (11); and by an argument like that in § 2 the minimum cost can be shown to be independent of the epoch at which the receiver is measured. One can again expect that attaining the minimum cost may require measurements of commuting operators in a product Hilbert space  $\mathcal{H}_S \otimes \mathcal{H}_A$  of the receiver  $S$  and an auxiliary apparatus  $A$ . Holevo<sup>[17]</sup> has treated both estimation and detection from this point of view. In general, the operators that must be measured in order to minimize an arbitrary cost function will be difficult to calculate. The minimization of a quadratic cost function in estimating a single parameter  $\theta$  has been achieved by Personick<sup>[21]</sup>, and Yuen and Lax have similarly derived the optimum estimates of a pair of parameters combined into a complex number, the cost being the sum of the mean-square errors in each.<sup>[22]</sup>

Bounds can be set below the mean-square errors of unbiased estimates of parameters of an ordinary probability density function by means of the Cra mer-Rao inequality of statistics. A quantum-mechanical counterpart to it has been formulated in terms of the symmetrized logarithmic derivatives of the density operator  $\rho^S$ .<sup>[3]</sup> Yuen and Lax have produced a similar inequality, but have restricted it to the simultaneous estimates of pairs of parameters combined into complex numbers, one pair for each mode of a multimode quantum receiver.<sup>[22]</sup> We wish to show how their method can be applied to setting bounds below the

variances of unbiased estimates of arbitrary real parameters of a density operator. The role of noncommuting observables measured on the receiver itself will become apparent, and we shall see that sometimes the one, sometimes the other type of bound is superior.

A quantum-mechanical system  $S$  is in a state described by a density operator  $\rho^S(\underline{\theta})$  that is a function of a number  $m$  of unknown real parameters  $\underline{\theta} = (\theta_1, \theta_2, \dots, \theta_m)$ , whose values are to be estimated. The procedure for doing so may require creating an auxiliary system  $A$ , which will be called the apparatus and which is in a state described by the density operator  $\rho^A$  independent of  $\underline{\theta}$ ; the density operator for the combination of system and apparatus is the tensor product

$$\rho(\underline{\theta}) = \rho^S(\underline{\theta}) \otimes \rho^A. \quad (25)$$

On this combination a set of commuting Hermitian operators  $\Theta_1, \Theta_2, \dots, \Theta_m$  are measured, and the outcomes  $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_m$  of the measurement are taken as the estimates of the parameters  $\underline{\theta}$ . These estimates are simultaneously eigenvalues of the  $m$  operators  $\Theta_j$  with eigenstates  $|\hat{\underline{\theta}}\rangle$  defined in the product space of  $S + A$ ,

$$\Theta_j |\hat{\underline{\theta}}\rangle = \hat{\theta}_j |\hat{\underline{\theta}}\rangle, \quad \hat{\underline{\theta}} = (\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_m), \quad j=1, 2, \dots, m. \quad (26)$$

We postulate that the estimates are unbiased,

$$\text{Tr}(\rho \Theta_j) = \theta_j, \quad j = 1, 2, \dots, m. \quad (27)$$

An inequality of the Cramér-Rao type is now derived by a method used by Yuen and Lax.<sup>[22]</sup> The right-logarithmic-derivative (r.l.d.) operators  $L_j$  are defined by

$$\frac{\partial \rho}{\partial \theta_j} = \rho L_j = L_j^+ \rho. \quad (28)$$

Equation (27) and the fact that  $\text{Tr} \rho = 1$  give as usual

$$\text{Tr} \frac{\partial \rho}{\partial \theta_k} (\theta_j - \theta_j) = \text{Tr} \rho (\theta_j - \theta_j) L_k^+ = \delta_{kj} = \begin{cases} 1, & k=j, \\ 0, & k \neq j. \end{cases} \quad (29)$$

Multiplying by the complex numbers  $y_j^*$  and  $z_k$  and summing we get

$$\text{Tr} \sum_{j=1}^m \sum_{k=1}^m y_j^* \rho (\theta_j - \theta_j) L_k^+ z_k = \sum_{k=1}^m y_k^* z_k, \quad (30)$$

which can be written

$$\underline{Y}^+ \underline{Z} = \text{Tr} P Q^+,$$

where  $\underline{Y}$  and  $\underline{Z}$  are column vectors of the  $y_j$ 's and  $z_j$ 's respectively, and  $P$  and

$Q$  are the operators

$$P = \rho^{1/2} \sum_{j=1}^m y_j^* (\theta_j - \theta_j), \quad (31)$$

$$Q = \rho^{1/2} \sum_{j=1}^m z_j^* L_j. \quad (32)$$

The Schwarz inequality for traces,

$$|\text{Tr} P Q^+|^2 \leq \text{Tr} P P^+ \text{Tr} Q Q^+, \quad (33)$$

yields

$$|\underline{Y}^+ \underline{Z}|^2 \leq (\underline{Z}^+ \underline{A} \underline{Z}) (\underline{Y}^+ \underline{B} \underline{Y}), \quad (34)$$

where  $\underline{A}$  and  $\underline{B}$  are Hermitian matrices whose elements are

$$A_{ij} = \text{Tr} \rho L_i L_j^+, \quad (35)$$

$$\begin{aligned} B_{ij} &= \text{Tr} \rho (\theta_i - \theta_i) (\theta_j - \theta_j) \\ &= \text{cov} (\hat{\theta}_i, \hat{\theta}_j). \end{aligned} \quad (36)$$

The diagonal element  $B_{ii}$  is the variance  $\text{Var} \hat{\theta}_i$  of the estimate  $\hat{\theta}_i$ ; "cov" stands for covariance. Equality holds in (33) if  $P$  and  $Q$  are proportional, that is if

$$\rho \sum_{j=1}^m y_j^* (\theta_j - \theta_j) = k(\theta) \rho \sum_{j=1}^m z_j^* L_j, \quad (37)$$

where we have multiplied both sides by  $\rho^{1/2}$  for later use.

Here  $k(\theta)$  is a c-number function, not an operator.

If we put  $\underline{Y} = \underline{B}^{-1} \underline{Z}$ , we obtain from (34)

$$\underline{Z}^+ \underline{B}^{-1} \underline{Z} \leq \underline{Z}^+ \underline{A} \underline{Z} \quad (38)$$

the left-hand side of which, when the  $z_j$ 's are real variables, is a quadratic form that when set equal to a constant  $(m + 2)$  specifies the concentration ellipsoid of the estimates<sup>[23]</sup>; and (38) states that that concentration ellipsoid lies outside the ellipsoid specified by the equation  $\underline{Z}^+ \underline{A} \underline{Z} = m + 2$ . Alternatively we can put  $\underline{Z} = \underline{A}^{-1} \underline{Y}$  to obtain the multivariate Cramér-Rao inequality

$$\underline{Y}^+ \underline{B} \underline{Y} \geq \underline{Y}^+ \underline{A}^{-1} \underline{Y}, \quad (39)$$

which for various choices of the  $y_j$ 's yields inequalities among linear combinations of the variances and covariances of the estimates. Furthermore, a quadratic risk function can be set up in terms of a positive-definite matrix  $\underline{G}$  as

$$R = \text{Tr } \underline{G} \underline{B}, \quad (40)$$

and by writing  $\underline{G}$  in terms of its eigenvalues and eigenvectors one can show that the risk  $R$  must exceed  $\text{Tr } \underline{G} \underline{A}^{-1}$ .

Because the density operator  $\rho$  has the product form in (25) with  $\rho^A$  independent of  $\theta$ , the r.l.d. operators  $L_j$  must have the form

$$L_j = L_j^S \otimes \underline{1}_A, \quad (41)$$

where the r.l.d. operators  $L_j^S$  defined by

$$\partial \rho^S / \partial \theta_j = \rho^S L_j^S \quad (42)$$

act only on the Hilbert space of the system  $S$ , and  $\underline{1}_A$  is the identity operator for the Hilbert space of the apparatus  $A$ . Thus the matrix elements  $A_{ij}$  defined by (35) become

$$A_{ij} = \text{Tr}_S (\rho^S L_i^S L_j^{S+}) \quad (43)$$

and are independent of the density operator  $\rho^A$  for the apparatus. These bounds, therefore, have a universal validity in the sense that they hold no matter what auxiliary apparatus A is brought up and no matter what state  $\rho^A$  it is in. They apply also to measurements of commuting estimators on the system S alone.

If we put into (37)  $\underline{Z} = \underline{A}^{-1}\underline{Y}$ , we find that equality in (39) is attained if there is a set of commuting estimating operators  $\theta_j$  such that

$$\rho \sum_{j=1}^m y_j^* (\theta_j - \theta_j) = k(\underline{\theta}) \rho \sum_{j=1}^m \sum_{k=1}^m y_j^* (\underline{A}^{-1})_{jk} L_k$$

or

$$\begin{aligned} & \rho^S \otimes \rho^A \sum_{j=1}^m y_j^* (\theta_j - \theta_j) \\ &= k(\underline{\theta}) \sum_{j=1}^m \sum_{k=1}^m y_j^* (\underline{A}^{-1})_{jk} \left( \rho^S L_k^S \right) \otimes \rho^A. \end{aligned} \quad (44)$$

Now, following Holevo<sup>[17]</sup>, we suppose that there exists a mapping of the parameter space  $\underline{\Theta}$  into a set of non-negative Hermitian measurement operators  $X(F)$  in the Hilbert space  $\mathcal{H}_S$  of the system S such that  $F \rightarrow X(F)$ , where F is any region of the space  $\underline{\Theta}$ ,  $\underline{\Theta} \rightarrow \underline{1}_S$ , and

$$\sum_i X(F_i) = \underline{1}_S, \quad (45)$$

when  $\{F_i\}$  is a countable measurable decomposition of the space  $\underline{\Theta}$  into disjoint regions  $F_i$ . Let  $X(\underline{\theta}; d^m \underline{\theta})$  be the operator corresponding to the differential region of volume  $d^m \underline{\theta}$  in the neighborhood of the point  $\underline{\theta}$  of the parameter space, and define the operators

$$\theta_j^S = \int_{\underline{\Theta}} \theta_j X(\underline{\theta}; d^m \underline{\theta}), \quad \underline{\theta} = (\theta_1, \dots, \theta_m). \quad (46)$$

The operators  $\theta_j^S$  may not commute, but suppose they satisfy the equation

$$\sum_{j=1}^m y_j^* \left( \theta_j^S - \theta_j \right) = k(\theta) \sum_{j=1}^m \sum_{k=1}^m y_j^* \binom{A^{-1}}{jk} L_k^S. \quad (47)$$

Then one can interpret the results of Holevo's work as asserting the existence

of an apparatus A, a product space  $\mathcal{H}_S \otimes \mathcal{H}_A$ , a pure state  $\rho^A = |\Psi_A\rangle\langle\Psi_A|$  of the apparatus, and a set of commuting estimators  $\{\theta_j\}$  in  $\mathcal{H}_S \otimes \mathcal{H}_A$  such that

(44) holds and the equality in (39) is attained. This set of estimators is

constructed by starting with the set of operators  $X(F) \otimes \rho^A$  acting in the

subspace  $\mathcal{H}_S \otimes |\Psi_A\rangle$  of  $\mathcal{H}_S \otimes \mathcal{H}_A$  that is spanned by the tensor products

$|w_i\rangle|\Psi_A\rangle$  of  $|\Psi_A\rangle$  and each of a complete orthonormal set of vectors  $|w_i\rangle$  spanning

$\mathcal{H}_S$ . By using Naimark's procedure<sup>[18]</sup> this noncommuting set of operators can

be extended to a commuting resolution of the identity in  $\mathcal{H}_S \otimes \mathcal{H}_A$ ,

$$X(F) \otimes \rho^A \rightarrow X'(F),$$

where

$$\sum_i X'(F_i) = \mathbb{1}_{S+A} \quad (48)$$

for any countable measurable decomposition of the parameter space  $\Theta$  into dis-

joint regions  $F_i$ . Projection of the operators  $X'(F)$  into the subspace

$\mathcal{H}_S \otimes |\Psi_A\rangle$  yields the operators  $X(F) \otimes \rho^A$ . Then the estimators defined by

$$\theta_j = \int_{\Theta} \theta_j X'(\theta; d^m \theta) \quad (49)$$

will commute and satisfy (44), for

$$\begin{aligned} \left( \rho^S \otimes \rho^A \right) \cdot \theta_j &= \rho^S \int_{\Theta} \theta_j X(\theta; d^m \theta) \otimes \rho^A \\ &= \rho^S \theta_j^S \otimes \rho^A, \end{aligned} \quad (50)$$



which when applied to (47) yields (44). In order to carry out this extension a product space  $\mathcal{H}_S \times \mathcal{H}_A$  of infinite dimension is required. As Holevo pointed out, the state  $\rho^A$  and the apparatus are not necessarily unique; many such states in many such product spaces  $\mathcal{H}_S \otimes \mathcal{H}_A$  may exist that permit achieving equality in (39).

#### 4. Estimation of Complex Mode Amplitudes

As an example we use the well-worn problem of estimating the real and imaginary parts of the complex amplitude  $\mu = \mu_x + i\mu_y$  of a coherent signal in a harmonic oscillator representing a single mode of a quantum receiver with thermal noise. The density operator is, in Glauber's P-representation [13]

$$\rho^S(\mu) = (\pi N)^{-1} \int \exp\left(-|\alpha - \mu|^2/N\right) |\alpha\rangle \langle \alpha| d^2\alpha, \quad (51)$$

where  $|\alpha\rangle$  is a coherent state,  $\alpha = \alpha_x + i\alpha_y$ ,  $d^2\alpha = d\alpha_x d\alpha_y$ , integration is taken over the entire complex  $\alpha$ -plane, and  $N$  is the mean number of thermal photons in the mode. The normalization is such that  $|\mu|^2$  is the mean number of photons contributed by the signal.

By (28) and eq. (3.20) of [1] the r.l.d. operators corresponding to  $\partial\rho^S/\partial\mu_x$  and  $\partial\rho^S/\partial\mu_y$  are ..

$$\begin{aligned} L_x^S &= \frac{a'^+}{N} + \frac{a'}{N+1}, \\ L_y^S &= i\left(\frac{a'^+}{N} - \frac{a'}{N+1}\right), \end{aligned} \quad (52)$$

where  $a' = a_S - \mu$  and  $a_S$  is the annihilation operator for the mode, obeying the usual commutation rule  $a_S a_S^+ - a_S^+ a_S = 1_S$ . Using this and  $\text{Tr } \rho^S a'^+ a' = N$ , we can easily derive the matrix  $\underline{A}$  of (35) and its inverse,

$$\underline{A} = \frac{1}{N(N+1)} \begin{bmatrix} 2N+1 & -i \\ i & 2N+1 \end{bmatrix}, \quad \underline{A}^{-1} = \frac{1}{4} \begin{bmatrix} 2N+1 & i \\ -i & 2N+1 \end{bmatrix} \quad (53)$$

and

$$\underline{A}^{-1} \begin{bmatrix} L_x^S \\ L_y^S \end{bmatrix} = \begin{bmatrix} Q_S - \mu_x \\ P_S - \mu_y \end{bmatrix}, \quad (54)$$

where

$$Q_S = \frac{1}{2} (a_S + a_S^\dagger), \quad P_S = \frac{1}{2} i(a_S^\dagger - a_S) \quad (55)$$

are proportional to the coordinate and momentum operators for the mode. If we now take  $\underline{y}^+ = [1 \ 0]$  and  $\underline{y}^+ = [0 \ 1]$  we obtain from the inequality (39) the bounds<sup>[3]</sup>

$$\text{Var } \hat{\mu}_x \text{ and } \text{Var } \hat{\mu}_y \geq \frac{1}{2} (N + \frac{1}{2}). \quad (56)$$

By (47) the operator achieving the minimum value of the variance  $\text{Var } \hat{\mu}_x$  is  $Q_S$ , and the operator minimizing  $\text{Var } \hat{\mu}_y$  is  $P_S$ . These do not commute and cannot be measured simultaneously on the same system S in such a way that both attain the minimum variance  $\frac{1}{2}(N + \frac{1}{2})$ .

If on the other hand we take  $\underline{y}^+ = [1 \ -i]$ , we find from (39) for simultaneous commuting measurement of the estimators  $M_x$  of  $\mu_x$  and  $M_y$  of  $\mu_y$ .

$$\begin{aligned} \text{Var } \hat{\mu}_x + \text{Var } \hat{\mu}_y = \\ \text{Tr } \rho \left[ \left( M_x - \mu_x \right)^2 + \left( M_y - \mu_y \right)^2 \right] \geq N + 1, \end{aligned} \quad (57)$$

an inequality discovered by Yuen and Lax.<sup>[22]</sup> If we use (54) with  $\underline{y}^+ = [1 \ -i]$  in (47) we find that inequality is achieved in (57) by the noncommuting operators  $M_x' = Q_S$ ,  $M_y' = P_S$ . The noncommuting resolution of the identity corresponding to (45) is provided by the operators

$$X(\mu; d^2 \mu) = |\mu\rangle \langle \mu| d^2 \mu / \pi, \quad (58)$$

in terms of the coherent states  $|\mu\rangle$ , and the real and imaginary parts of

$$Q_S + i P_S = a_S = \int \mu |\mu\rangle \langle \mu| d^2 \mu / \pi \quad (59)$$

correspond to (46).

It is known that equality in (57) is achieved by measuring the commuting

operators [24]

$$M_x = Q_S + Q_A, M_y = P_S - P_A \quad (60)$$

on a combination of the system S and another harmonic oscillator A in the

ground state  $|0_A\rangle$ , the density operator being now

$$\rho = \rho^S \otimes |0_A\rangle \langle 0_A|. \quad (61)$$

Here  $Q_A$  and  $P_A$  are coordinate and momentum operators for A. Equation (44) requires,

with  $k(\mu_x, \mu_y) = 1$ ,  $\tilde{y}^+ = [1 \quad -i]$ ,

$$\left( \rho^S \otimes \rho^A \right) \left( M_x + iM_y \right) = \rho^S a_S \otimes \rho^A$$

which is satisfied by (60) because

$$\rho^A \left( Q_A - i P_A \right) = \rho^A a_A^+ = |0_A\rangle \langle 0_A| a_A^+ = 0.$$

## 5. Estimation of Essential Signal Parameters

By essential signal parameters we mean those intrinsic to the form of the signal, as distinct from its complex amplitude; the arrival time  $\tau$  and carrier frequency  $\Omega$  are the essential parameters we shall deal with here. We shall see that when these are measured separately, the Cramér-Rao inequality based on the symmetrized logarithmic derivatives<sup>[3]</sup> (s.l.d.)  $\mathcal{L}_j$  yields greater lower bounds on the error variances than does the inequality (34) based on the right logarithmic derivatives (r.l.d.)  $L_j$ .

The s.l.d.  $\mathcal{L}_j$  is an Hermitian operator defined by

$$\frac{\partial \rho}{\partial \theta_j} = \frac{1}{2} (\rho \mathcal{L}_j + \mathcal{L}_j \rho). \quad (62)$$

The corresponding Cramér-Rao inequality has the form<sup>[3]</sup>

$$\tilde{Y}' B' Y \geq \tilde{Y}' A'^{-1} Y, \quad (63)$$

where the column vector  $Y$  has real elements and  $\tilde{Y} = (y_1 \ y_2 \ \dots \ y_m)$ . The elements of the matrices  $A'$  and  $B'$  are

$$A'_{ij} = \frac{1}{2} \text{Tr } \rho (\mathcal{L}_i \mathcal{L}_j + \mathcal{L}_j \mathcal{L}_i), \quad (64)$$

$$B'_{ij} = \frac{1}{2} \text{Tr } \rho (\theta'_i \theta'_j + \theta'_j \theta'_i); \quad \theta'_i = \theta_i - \theta_i. \quad (65)$$

If one goes through the derivation of (63) as in reference [3], one sees that it does not require that the estimating operators  $\theta_i$  commute. Furthermore, the density operator  $\rho$  may have the product form  $\rho^S(\theta) \otimes \rho^A$  corresponding to a combination of the receiver  $S$  and an ignorant apparatus  $A$ . The s.l.d. operators will then take the form

$$\mathcal{L}_j = \mathcal{L}_j^S \otimes \mathbb{1}_A \quad (66)$$

where the s.l.d. operators  $\mathcal{L}_j^S$  are given by (62) with  $\rho$  replaced by  $\rho^S(\theta)$ . The matrix  $A'$  then depends only on  $\rho^S$ , and the bound in (63) has a universal validity,

applying no matter what apparatus  $A$  is used nor what state  $\rho^A$  it has.

We consider estimates of parameters  $\theta$  of a coherent signal with complex envelope

$$g(\theta, t) = \beta f(t-\tau) e^{i\Omega(t-\tau)},$$

$$\beta = \beta_x + i\beta_y, \theta = (\beta_x, \beta_y, \tau, \Omega), \quad (67)$$

$\tau$  being the arrival time and  $\Omega$  the deviation from a reference carrier frequency  $\Omega_0$ . The signal field is spatially coherent over the aperture of the observing instrument, so that (67) represents the amplitude of a single spatial mode.<sup>[25]</sup>

The signal occupies such a narrow band of frequencies that the thermal noise can be considered to have a spectral density independent of frequency, and the observation interval  $(-\frac{1}{2}T, \frac{1}{2}T)$  is assumed so long that the possibility that the signal overlaps either end of it can be disregarded. We then expand the signal into temporal modes

$$\gamma_k(t) = T^{-1/2} \exp i\omega_k t, \quad \omega_k = 2\pi k/T, \quad (68)$$

whose amplitudes are

$$\mu_k(\theta) = \int_{-T/2}^{T/2} g(\theta, t) \gamma_k^*(t) dt =$$

$$\beta T^{-1/2} \exp(-i\omega_k \tau) F(\omega_k - \Omega), \quad (69)$$

where

$$F(\omega) = \int_{-\infty}^{\infty} f(s) e^{-i\omega s} ds \quad (70)$$

is the spectrum of the signal. The density operator  $\rho^S(\theta)$  now has a Gaussian  $P$ -representation, and as in (38) of reference [3],

$$\partial \rho / \partial \theta_j = \sum_k N^{-1} \left[ \frac{\partial \mu_k}{\partial \theta_j}^* (a_k - \mu_k) \rho + \rho \frac{\partial \mu_k}{\partial \theta_j} (a_k^\dagger - \mu_k^*) \right], \quad (71)$$

where  $N$  is the mean number of background photons per mode and  $a_k$  and  $a_k^+$  are the annihilation and creation operators for the  $k$ -th mode. The right logarithmic derivative is now [26]

$$L_j = \sum_k \left[ N^{-1} \frac{\partial \mu_k}{\partial \theta_j} (a_k^+ - \mu_k^*) + (N+1)^{-1} \frac{\partial \mu_k^*}{\partial \theta_j} (a_k - \mu_k) \right] \quad (72)$$

from which, with  $\text{Tr} \rho a_k'^+ a_k' = N$ ,  $a_k' = a_k - \mu_k$ , we find for the matrix  $\underline{A}$  of (35) the elements

$$\begin{aligned} A_{ij} &= \sum_k \left[ N^{-1} \frac{\partial \mu_k}{\partial \theta_i} \frac{\partial \mu_k^*}{\partial \theta_j} + (N+1)^{-1} \frac{\partial \mu_k^*}{\partial \theta_i} \frac{\partial \mu_k}{\partial \theta_j} \right] \\ &= \left[ N^{-1} \frac{\partial^2 G}{\partial \theta_{1i} \partial \theta_{2j}} + (N+1)^{-1} \frac{\partial^2 G}{\partial \theta_{2i} \partial \theta_{1j}} \right] \quad \theta_1 = \theta_2 = \theta \end{aligned} \quad (73)$$

where

$$\begin{aligned} G(\theta_1, \theta_2) &= \sum_k \mu_k(\theta_1) \mu_k^*(\theta_2) = \\ &= \beta_1 \beta_2^* \int_{-\infty}^{\infty} F(\omega - \Omega_1) F^*(\omega - \Omega_2) \exp \left[ -i\omega (\tau_1 - \tau_2) \right] d\omega / 2\pi \end{aligned} \quad (74)$$

in the limit  $T \rightarrow \infty$ . This is conveniently written in terms of the ambiguity function as [27]

$$G(\theta_1, \theta_2) = \beta_1 \beta_2^* \exp \left[ -\frac{1}{2} i(\Omega_1 + \Omega_2)(\tau_1 - \tau_2) \right] \lambda(\tau_1 - \tau_2, \Omega_2 - \Omega_1). \quad (75)$$

By using eq. (2.11) of ch.X of reference [27], we find the derivatives

$$\begin{aligned} \partial^2 G / \partial \beta_{1x} \partial \beta_{2x} &= \partial^2 G / \partial \beta_{1y} \partial \beta_{2y} = 1, \\ \partial^2 G / \partial \beta_{1x} \partial \beta_{2y} &= -\partial^2 G / \partial \beta_{1y} \partial \beta_{2x} = -i, \\ \partial^2 G / \partial \tau_1 \partial \tau_2 &= |\beta|^2 \Delta \omega^2, \\ \partial^2 G / \partial \tau_1 \partial \omega_2 &= -|\beta|^2 \left[ \Delta(\omega\tau) + \frac{1}{2} i \right], \\ \partial^2 G / \partial \tau_2 \partial \omega_1 &= -|\beta|^2 \left[ \Delta(\omega\tau) - \frac{1}{2} i \right], \\ \partial^2 G / \partial \omega_1 \partial \omega_2 &= |\beta|^2 \Delta \tau^2, \end{aligned} \quad (76)$$

where we have set  $\theta_{\sim 1} = \theta_{\sim 2} = \theta$  and used the definitions of the mean-square duration  $\Delta t^2$ , mean-square bandwidth  $\Delta \omega^2$ , and mean duration-frequency product  $\Delta(\omega t)$  as given in reference [27], ch.I. §5, and we have taken

$$\bar{t} = 0, \bar{\omega} + \Omega = 0, \quad (77)$$

where  $\bar{t}$  is the mean epoch and  $\bar{\omega}$  the mean frequency deviation of the signal.

The matrix  $\underline{A}$  now has the block form

$$\underline{A} = \begin{bmatrix} \underline{A}_{\sim 1} & \underline{0} \\ \underline{0} & \underline{A}_{\sim 2} \end{bmatrix}, \quad (78)$$

with  $\underline{A}_{\sim 1}$  given by (53); the elements of  $\underline{A}_{\sim 2}$  are

$$A_{2\tau\tau} = (2N+1) |\beta|^2 \Delta \omega^2 / N(N+1), \quad (79)$$

$$A_{2\tau\Omega} = A_{2\Omega\tau}^* = -|\beta|^2 \left[ (2N+1) \Delta(\omega t) + \frac{1}{2} i \right] / N(N+1), \quad (80)$$

$$A_{2\Omega\Omega} = (2N+1) |\beta|^2 \Delta t^2 / N(N+1). \quad (81)$$

Thus the bounds on the error variances as given by the Cramér-Rao inequality (39) based on the r.l.d.'s  $L_j$  depend on the matrix  $\underline{A}_{\sim 2}^{-1}$ , whose elements are

$$\left( \underline{A}_{\sim 2}^{-1} \right)_{\tau\tau} = \mathcal{F} (2N+1) \Delta t^2, \quad (82)$$

$$\left( \underline{A}_{\sim 2}^{-1} \right)_{\tau\Omega} = \left( \underline{A}_{\sim 2}^{-1} \right)_{\Omega\tau}^* = \mathcal{F} \left[ (2N+1) \Delta(\omega t) + \frac{1}{2} i \right], \quad (83)$$

$$\left( \underline{A}_{\sim 2}^{-1} \right)_{\Omega\Omega} = \mathcal{F} (2N+1) \Delta \omega^2, \quad (84)$$

where

$$\mathcal{F} = |\beta|^{-2} N(N+1) \left[ (2N+1)^2 D - \frac{1}{4} \right]^{-1},$$

$$D = \Delta \omega^2 \Delta t^2 - \left[ \Delta(\omega t) \right]^2. \quad (85)$$

In particular, (82) and (84) give the lower bounds on  $\text{Var } \hat{t}$  and  $\text{Var } \hat{\Omega}$  respectively.



From the symmetrized logarithmic derivatives  $\mathcal{L}_j$  as given by (39) of reference [3] we obtain, as in (43) therein, the matrix elements of  $\underline{A}'$  in (64),

$$A'_{ij} = 4(2N+1)^{-1} \operatorname{Re} \left[ \partial^2 G / \partial \theta_{1i} \partial \theta_{2j} \right] \quad (86)$$

$\theta_{\sim 1} = \theta_{\sim 2} = \theta,$

where  $\operatorname{Re}$  indicates the real part. From (76) this yields the matrix  $\underline{A}'$  in block form

$$\underline{A}' = \begin{bmatrix} \underline{A}'_1 & \underline{0} \\ \underline{0} & \underline{A}'_2 \end{bmatrix}; \quad (87)$$

where  $\underline{A}'_1$  is diagonal with diagonal elements  $\frac{1}{4}(2N+1)$ , and the elements of  $\underline{A}'_2$  are

$$A'_{2\tau\tau} = 4(2N+1)^{-1} |\beta|^2 \Delta\omega^2, \quad (88)$$

$$A'_{2\tau\Omega} = A'_{2\Omega\tau} = -4(2N+1)^{-1} |\beta|^2 \Delta(\omega\tau), \quad (89)$$

$$A'_{2\Omega\Omega} = 4(2N+1)^{-1} |\beta|^2 \Delta\tau^2. \quad (90)$$

The elements of the matrix  $\underline{A}'_2^{-1}$  needed in (63) are now

$$\left( \underline{A}'_2^{-1} \right)_{\tau\tau} = \frac{1}{4}(2N+1) |\beta|^{-2} \Delta\tau^2 / D, \quad (91)$$

$$\left( \underline{A}'_2^{-1} \right)_{\tau\Omega} = \frac{1}{4}(2N+1) |\beta|^{-2} \Delta(\omega\tau) / D, \quad (92)$$

$$\left( \underline{A}'_2^{-1} \right)_{\Omega\Omega} = \frac{1}{4}(2N+1) |\beta|^2 \Delta\omega^2 / D, \quad (93)$$

with  $D$  given in (85). In particular, (91) and (93) give lower bounds on the variances  $\operatorname{Var} \hat{\tau}$  and  $\operatorname{Var} \hat{\Omega}$  of estimates of arrival time  $\tau$  and carrier-frequency shift  $\Omega$ . These bounds are larger than those given by (82) and (84), except when  $D = 1/4$ . The minimum value  $1/4$  is attained by  $D$  when the signal has

a Gaussian envelope [27, p.20]; the s.l.d.'s and the r.l.d.'s then yield equal lower bounds.

If, on the other hand, we invent a risk function of the form

$$R = \frac{\text{Var } \hat{\tau}}{\Delta t^2} + \frac{\text{Var } \hat{\Omega}}{\Delta \omega^2}, \quad (94)$$

we find, upon using (39), (79)-(81), and  $\underline{y}^+ = \begin{bmatrix} \Delta \omega & i\Delta t \end{bmatrix}$ , that when the operators estimating the arrival time  $\tau$  and the frequency shift  $\Omega$  commute,

$$R \geq \frac{2N(N+1)}{|\beta|^2 \Delta \omega \Delta t \left[ (2N+1)\Delta \omega \Delta t - \frac{1}{2} \right]} = R_1; \quad (95)$$

here we assumed for simplicity that  $\Delta(\omega t) = 0$ . The s.l.d. bounds in (91) and (93) yield, on the other hand,

$$R \geq \frac{1}{2} (2N+1) |\beta|^{-2} \Delta \omega^{-2} \Delta t^{-2} = R_2 \quad (96)$$

and  $R_1 \geq R_2$  for  $\frac{1}{2} \leq \Delta \omega \Delta t \leq N + \frac{1}{2}$ . Thus if the arrival time and carrier frequency of the signal are to be measured in the same receiver, the r.l.d.

Cramér-Rao inequality (39) sometimes yields a superior lower bound on a weighted sum of the error variances to that provided by the s.l.d. inequality (63).

Efficient estimators for essential parameters such as arrival time and carrier-frequency do not exist, and the lower bounds derived here and in [3] have only an asymptotic significance. Both sets go into the classical forms in the limit  $N \gg 1$ .

For simultaneous estimation of the complex amplitude  $\beta = \beta_x + i\beta_y$  the r.l.d. bound on  $\text{Var } \hat{\beta}_x + \text{Var } \hat{\beta}_y$  is, as we have seen, greater than the s.l.d. bound, and from this we gain an instructive insight into the role of noncommuting observables. A similar result was found for the special risk function in (94).

For single measurements of other parameters than the amplitude, however, the s.l.d. bound may be superior.

### Conclusion

In order to realize minimum Bayes cost in quantum detection and estimation by measurements of commuting operators it may be necessary to couple the receiver  $S$  with an auxiliary apparatus  $A$  and perform the measurements on the combination. The resulting Bayes cost is the same as if commuting operators were somehow measured on the pair  $S + A$  before they interacted, and the optimum strategy is equivalent to a resolution of the identity into possibly noncommuting, positive-definite Hermitian operators acting in the Hilbert space  $\mathcal{H}_S$  of the receiver alone. The minimization of the Bayes cost can, therefore, be carried out over the class of such noncommuting operators in  $\mathcal{H}_S$ , which having been found can be extended to a resolution of the identity into commuting projection operators in an encompassing Hilbert space  $\mathcal{H}_S \otimes \mathcal{H}_A$  characterizing a combination of the receiver  $S$  with some auxiliary apparatus  $A$  initially ignorant of the state of the receiver. Minimum Bayes cost can be attained without violating the quantum-mechanical restriction of simultaneous measurement to the class of commuting Hermitian operators.

There are two types of Cramér-Rao inequalities setting lower bounds to mean-square errors and quadratic loss functions in unbiased estimates of parameters of the density operator of a quantum-mechanical receiver. Sometimes one, sometimes the other yields the superior bound. The difference between them lies not in how they handle the commutativity of the estimating operators, but in their definitions in terms of right-hand or symmetrized logarithmic derivatives. Both can be applied to a combination of the receiver  $S$  with an arbitrary ignorant auxiliary apparatus  $A$ , on which combination  $S + A$  commuting observables are measured. The lower bounds they assert depend only on the

parameters of the receiver S. Because commuting estimators on  $S + A$  can be reduced to possibly noncommuting estimators on the receiver S alone, both inequalities apply to these as well.

#### Acknowledgment

We are grateful to A.L. Holevo and H.P.H. Yuen for sending us preliminary copies of their papers, and to E.A. Bishop, H.P.H. Yuen, and S.D. Personick for enlightening discussions of these matters.

## References

1. C. W. Helstrom, "Detection theory and quantum mechanics", *Infor. and Control*, vol. 10, pp. 254-291; March, 1967.
2. C. W. Helstrom, "Detection theory and quantum mechanics, II", *Infor. and Control*, vol. 13, pp. 156-171; August, 1968.
3. C. W. Helstrom, "The minimum variance of estimates in quantum signal detection", *IEEE Trans. on Information Theory*, vol. IT-14, pp. 234-242; March, 1968.
4. C. W. Helstrom, J. W. S. Liu, and J. P. Gordon, "Quantum-mechanical communication theory", *Proceedings of the IEEE*, vol. 58, pp. 1578-1598; October, 1970.
5. H. P. Yuen, R. S. Kennedy, and M. Lax, "On optimal quantum receivers for digital signal detection", *Proceedings of the IEEE*, vol. 58, pp. 1770-1773; October, 1970.
6. P. A. M. Dirac, *The Principles of Quantum Mechanics*, 3rd ed., Oxford University Press; 1947. See Chs. 1-3, pp. 1-83.
7. D. Bohm, *Quantum Theory*, Prentice-Hall, Inc., Englewood Cliffs, N. J., 1951.
8. W. H. Louisell, *Radiation and Noise in Quantum Electronics*, McGraw-Hill, New York, 1964. See Ch. 1, pp. 1-68.
9. On the subject of density operators see Dirac [6], p. 132, or Louisell [8], Ch. 6, pp. 220-228.
10. J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, J. Springer Verlag, Berlin, 1932. See Ch. III, §5, pp. 130-134.
11. See Dirac [6], §13, pp. 49-52.

12. J. P. Gordon and W. H. Louisell, "Simultaneous measurement of noncommuting observables", in *Physics of Quantum Electronics*, P. L. Kelley, B. Lax, and P. E. Tannenwald, eds. New York, McGraw-Hill, 1966, pp. 833-840.
13. R. J. Glauber, "Coherent and incoherent states of the radiation field", *Phys. Rev.* vol. 131, pp. 2766-2788; September 15, 1963.
14. See Louisell [8], p. 225.
15. C. W. Helstrom, "Fundamental limitations on the detectability of electromagnetic signals", *Int. J. Theor. Phys.*, vol. 1, pp. 37-50; May, 1968.
16. J. W. S. Liu, "Reliability of Quantum-Mechanical Communication Systems", M.I.T. R.L.E. Tech. Report 468; December 31, 1968. See p. 22.
17. A.S. Holevo, "Statistical Problems in Quantum Physics", unpublished manuscript, Steklov Mathematical Institute, Moscow, U.S.S.R., to appear in *Proceedings of Soviet-Japanese Symposium on Probability and Statistics*, 1972.
18. M.A. Naimark, "Spectral Functions of a Symmetric Operator," *Izvestia Akad.Nauk, Mathematical Series*, U.S.S.R., vol. 4, no. 3, pp. 277-318; 1940.
19. C.W. Helstrom, "Some quantum-mechanical aspects of optical communications", in *Optical Space Communication*, R.S. Kennedy and S. Karp, eds., NASA Report SP-217, 1969, p. 111-120.
20. H.P.H. Yuen, "Communication Theory of Quantum Systems", Ph.D. thesis, Department of Electrical Engineering, MIT, June 1970 and MIT R.L.E. Tech. Rep't. 482, August 30, 1971. See p. 124.
21. S.D. Personick, "Application of Quantum Estimation Theory to Analog Communication Over Quantum Channels", *IEEE Transactions on Information Theory*, vol. IT-17, pp. 240-246; May, 1971.
22. H.P.H. Yuen and M. Lax, "Multiple Parameter Quantum Estimation and Measurement of Nonselfadjoint Operators," presented at IEEE International Symposium on Information Theory, Asilomar, Calif., January 31, 1972.

23. L. Schmetterer, Mathematische Statistik, New York: Springer-Verlag, 1966, p. 63.
24. S. Personick, "An image-band interpretation of optical heterodyne noise", Bell System Tech. J. vol. 50, pp. 213-216; January, 1971.
25. C.W. Helstrom, "Quantum Detection Theory", *Progress in Optics*, E. Wolf, ed., vol. 10, North-Holland Publishing Co., Amsterdam (in press).
26. Use eqs. (3.47a), (3.47b) of ref. [ 8], p. 111.
27. C.W. Helstrom, Statistical Theory of Signal Detection, 2nd ed. Oxford: Pergamon Press, 1968, p. 354.