JB 77-02

Paper Entitled

Infinite Dimensional Filtering Problems in Optcal Communication Systems

From the Proceedings of

Distributed Parameter Systems: Modelling and Identification

Springer-Verlag Lecture Notes in Control and Information Sciences, Vol 1

pp 128-137

Rome, Italy
June 1976

INFINITE DIMENSIONAL FILTERING PROBLEMS IN OPTICAL COMMUNICATION SYSTEMS

John S. Baras
Electrical Engineering Department
University of Maryland
College Park, Maryland 20742/USA

ABSTRACT

Several filtering problems utilizing quantum mechanical measurements are discussed and formulated as optimization problems in infinite dimensional spaces. The solution to some of these problems and their physical interpretation is given. Two examples illustrate the implementation of the mathematical results.

I. Introduction

With the advent of lasers detection and estimation problems in quantum electronics became of primary importance [1, 4]. The optical frequencies necessitate quantum mechanical modeling of the underlying system and measurement processes. If one utilizes classical approaches in modeling such systems, the resulting estimators and detectors are suboptimal. More recently linear filtering of a random signal sequence utilizing quantum mechanical measurements has been considered [5, 10, 11, 12]. In this paper I want to describe the results obtained to date within our program, at the University of Maryland, on Quantum Filtering Theory. I believe this to be a rich and fruitful area of research which has deep roots in fundamental mathematical physics and adresses nonclassical infinite dimensional filtering problems. In addition to optical communication problems there are other areas where quantum mechanical modelling becomes necessary.

The basic problem we consider is the linear filtering of a random sequence $\{x(k)\}$ which influences a quantum field, utilizing quantum mechanical measurements. These problems arise typically in laser communication systems. Here is a concrete example. At each time k a laser modulated in some fashion by $\{x(j)\}$ is received in a cavity and a device is used to perform a measurement on the captured field. Then the cavity is cleared and reopens to repeat the process at time $k+1$. The problem is to select optimally the measuring device at each time along with the postprocessing scheme of past and current measurement outcomes in order to estimate $\{x(k)\}$. The optimality criterion here is error covariance, but others can be utilized as well (see for example the general results of Holevo [2] in Quantum Decision problems). We would like to emphasize the nonclassical character of this filtering problem,

---

This work was partially supported by the National Science Foundation under grant ENG 75-20900.

which is due to the optimization over possible measurement processes in addition to the usual optimization over signal processing schemes.

There are two distinct cases that have been anlayzed to date: 1) $\{x(k)\}$ is a scalar sequence; 2) $\{x(k)\}$ is a vector sequence. The complexity of the mathematical methods used, and of the final solution differ considerably for these two cases. We are going to present the results separately for each case, to illustrate the similarities and differences more clearly. When modelling a quantum mechanical system, a priori statistical information is represented by a density operator ρ on a Hilbert space H (ρ is a self-adjoint, positive definite operator with unit trace and represents the state of the quantum system [9, p. 94, p. 132]. In our case there is a signal process that influences the state of the quantum system. This is described by making the density operator a function of x(k), $\rho(x(k))$ in such a way that ρ does not depend explicity on k. This latter property is crucial for the results obtained to-date, because it allows us to avoid complications due to the Liouville-von Neumann equation that $\rho(x(k))$ satisfies in general [9, p. 158]. Through x(k), therefore, $\rho(x(k))$ becomes an operator-valued stochastic process.

In modelling the measurement process it is necessary to distinguish two cases:

1) Scalar processes: We need only make one measurement at a time, and therefore the measurement process at time j is represented by a self-adjoint operator V_j on H (an observable, [8], [9]), with outcome a classical scalar random variable v(j) with distribution function [8]

$$F_{v(j)}(\xi) = \int_{R} Tr[\rho(\zeta)E_{V_j}(-\infty,\xi]]\, F_{x(j)}(d\zeta) . \tag{1}$$

Here E_{V_j} is the spectral measure associated with the self-adjoint operator V_j [8], and Tr indicates the operation of trace on H.

2) Vector processes: We need now more than one measurement simultaneously and the essentially quantum mechanical problem of compatible simultaneous measurements arises [8], [9]. It has been shown by Holevo [2], that the correct model for a measurement with outcomes in R^n, is provided by a positive operator valued measure p.o.m., M which is a map from the Borel σ-algebra $\mathcal{B}^n$ of R^n to the algebra $\mathfrak{B}(H)$ of all bounded linear operators on H such that

i) $M(B) \geq 0 \quad , \quad \forall B \in \mathcal{B}^n$

ii) if $\{B_i\} \subseteq \mathcal{B}^n$ is a partition of R^n then $\sum_i M(Bi) = I$. (2)

So at each time j the measurement process is represented by a p.o.m. M_j, and gives as outcome a classical vector random variable v(j) with distribution function

$$F_{v(j)}(\xi_1,\ldots,\xi_n) = \int_{R^n} Tr[\rho(\zeta)M_j(-\infty,\xi]]\, F_{x(j)}(d\zeta) \tag{3}$$

where $(-\infty, \xi] = (-\infty, \xi_1] \times \dots \times (-\infty, \xi_n]$. This generalization of the concept of a quantum measurement corresponds to approximate measurement of incompatible observables on the original system and, as pointed out by Holevo [2, p. 393], is well justified in view of Naimark's theorem [14, p. 124]. The latter theorem asserts that, given a p.o.m. M on H there exist an auxiliary Hilbert space H_e, a (pure) density operator ρ_e in $\mathfrak{B}(H_e)$ and a spectral measure E_M on $H \otimes H_e$ (the tensor product of Hilbert spaces H, H_e [8, p. 144] such that

$$\mathrm{Tr}[\rho\, M(B)] = \mathrm{Tr}[(\rho \otimes \rho_e)\, E_M(B)]$$

for every $B \in \mathcal{B}^n$ and every density operator ρ on H. The physical interpretation of this result is that a measurement represented by a p.o.m. corresponds to the simultaneous measurement of compatible observables on an appropriately augumented quantum system. The triple (H_e, ρ_e, E_M) is called a realization of the measurement represented by the p.o.m. M.

The final assumption we make is that the measurement outcomes v(j) (scalar or vectors) conditioned on x(j) are independent. This assumption facilitates the analysis, and one possible physical interpretation is the clearing of the receiver cavity after each measurement, described in the optical communication example cited in the beginning of this section.

The linear filtering problem is then the following. At time k, $k = 0, 1, \dots$ the previous measurement outcomes v(j), $j = 0, \dots, k-1$ are available, a current measurement is to be chosen, with outcome v(k), as well as processing coefficients $C_i(k)$, $i = 0, \dots, k$ so that the estimator

$$\hat{x}(k) = \sum_{i=0}^{k} C_i(k)\, v(i) \tag{4}$$

becomes the minimum variance estimator of x(k).

II. Scalar Signal Processes

This case has been completely resolved in joint work with R. O. Harger and Y.H. Park and we refer to [5] for details. Due to the conditional independence assumption the joint distribution function of the outcomes $v(0), \dots, v(k)$ is given by

$$F_{v(0), \dots, v(k)}(\nu(0), \dots, \nu(k)) = \int_R \dots \int_R F_{v(0)|x(0)}(\nu(0), \xi(0)) \dots F_{v(k)|x(k)}(\nu(k), \xi(k)) \cdot F_{x(0), \dots, x(k)}(d\xi(0), \dots, d\xi(k)) \tag{5}$$

where

$$F_{v(i)|x(i)}(\nu(i), \xi(i)) = \mathrm{Tr}[\rho(\xi(i))\, E_{V_i}(-\infty, \nu(i)]] .$$

Following [5] the problem reduces to minimization of

$$J(\underset{\sim}{C}(k), V_k) = E\{Tr[\rho(x(k))(x(k)I - V_k - I\sum_{j=0}^{k-1} C_j(k)v(j))^2]\} \qquad (6)$$

over all self-adjoint operators V_k on H and all k vectors

$$\underset{\sim}{C}(k) = \begin{bmatrix} C_0(k) \\ C_1(k) \\ \vdots \\ C_{k-1}(k) \end{bmatrix} \qquad (7)$$

By application of the projection theorem on an appropriate space of operator valued functions the optimizing solution is characterized by, [5]:

Theorem 1: There exist optimum observable $\hat{V}_k$ and optimal processing coefficients $\hat{C}_i(k)$, i = 0, ... , k-1 if and only if there exists a solution to the following equations

$$\eta(k)\hat{V}_k + \hat{V}_k\eta(k) = 2\delta(k) - 2\sum_{j=0}^{k-1} \hat{C}_j(k)\,\gamma(k,j) \qquad (8)$$

$$\sum_{j=0}^{k} \hat{C}_j(k)\, E\{v(i)v(j)\} = E\{v(i)x(k)\}, \quad i = 0, \ldots, k \qquad (9)$$

where

$$\eta(k) = E\{\rho(x(k))\} \qquad (10)$$

$$\delta(k) = E\{x(k)\rho(x(k))\} \qquad (11)$$

$$\gamma(k,i) = E\{v(i)\rho(x(k))\} \qquad (12)$$

are self-adjoint operators on H.

This result settles the existence of optimal linear filters in the scalar case and provides necessary and sufficient conditions for optimality. It should be noted that the operators $\eta(k)$, $\delta(k)$, $\gamma(k,i)$ appearing in (8) can be computed knowing the functional expression of $\rho(x(k))$ and the a priori statistics of $\{x(i)\}$. Finally (9) are just the usual normal equations [15] of linear mean square error estimation of x(k) based on the classical random variables $v(0), \ldots, v(k-1)$, $\hat{v}(k)$.

The resulting filter is clearly very complex, and equations (8) and (9) indicate that a new measuring device may be needed at each time. So it is crucial to discover assumptions that simplify the filter structure. One such case is described in [5] and utilizes Gaussian statistics for signal and measurement outcomes. To clearly describe the separation theorem in the case of Gaussian statistics, we introduce the following measurements that we shall call intrinsic. Let T_k be the observable whose outcome $\tau(k)$ at time k provides the minimum error variance estimator of x(k) without regard to past data. As a result of Theorem 1 (put $\hat{C}_j(k) = 0$, j=0, ..., k-1)

T_k if the solution of

$$\eta(k)\, T_k + T_k \eta(k) = 2\, \delta(k) \ . \tag{13}$$

These operators are intrinsic to the quantum system and can be computed a priori. Then in [5] the following separation theorem is proven:

Theorem 2: Suppose i) $\{x(i)\}$ is a Gaussian process
ii) the intrinsic measurement outcome $\tau(j)$ and $x(j)$ are jointly Gaussian for each j.

We form the linear minimum variance quantum estimator $\hat{x}(k)$ of $x(k)$ utilizing observables $\hat{V}_j$ and coefficients $\hat{C}_j(k)$ that satisfy the optimality conditions of Theorem 1. Then the intrinsic measurement outcomes $\tau(j)$, $j = 0, 1, \dots, k$ are a sufficient statistic for $\hat{x}(k)$.

The physical interpretation of this theorem is that the optimal quantum measurements are chosen separately from the optimal classical postprocessing of the measurement outcomes. This is why we call this the separation theorem. This is illustrated in Fig. 1.

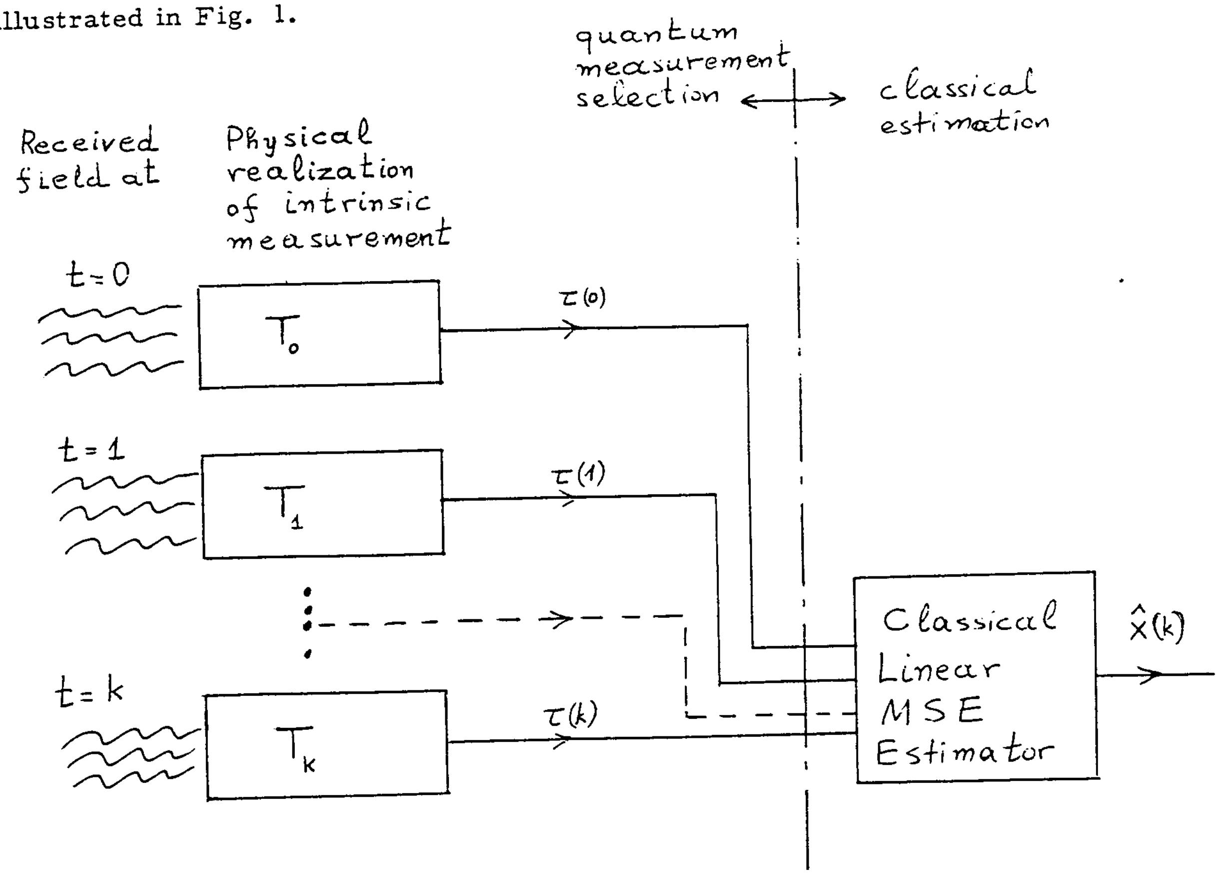


Fig. 1: Illustrating the separation of the optimal quantum linear filter with Gaussian statistics in the signal and intrinsic measurement processes.

It is also important to note that if the intrinsic measurement outcomes $\tau(j)$, $j=0,1,\ldots$ allow a (classical recursive estimate of $\hat{x}(k)$, the quantum filtering problem will also have a recursive solution.

III. Vector Signal Processes

As it was mentioned in the introduction this is a considerably more difficult and delicate problem than the scalar case. Again due to the conditional independence assumption the joint distribution function of the vector outcomes $v(0),\ldots,v(k)$ is given by the first of (5) while the second becomes

$$F_{v(i)|x(i)}(\nu(i),\xi(i)) = \mathrm{Tr}\,[\,\rho(\xi(i))M_i(-\infty,\nu(i)]\,]. \tag{14}$$

The linear filtering problem here cannot be formulated as a minimum norm problem and projection theorem methods do not work. Utilizing some mathematical techniques developed by Holevo [2] the problem has been formulated in [12] as the minimization of the functional

$$J(\underset{\sim}{C}(k), M_k) = \mathrm{Tr}\int_{R^n} \mathfrak{F}(u,\underset{\sim}{C}(k))M_k(du) \tag{15}$$

where $C(k) = [C_0(1),\ldots,C_{k-1}(k), I_n]$ is an $n\times n(k+1)$ matrix and

$$\mathfrak{F}(u,\underset{\sim}{C}(k)) = E\{E\{\,||x(k)-u-\sum_{i=0}^{k-1} C_i(k)v(i)||^2\,\rho(x(k))\,|\,x(0),\ldots,x(k-1)\}\} \tag{16}$$

is a nonnegative, self-adjoint, trace class operator valued function for each $\underset{\sim}{C}(k)$ [12]. The set over which the minimization is considered in [12] is the set

$$\mathcal{M}\mathcal{G} \times (R^{n\times n})^k$$

where $\mathcal{M}\mathcal{G}$ is the set of all p.o.m.'s on H. The existence of optimal linear filters and necessary and sufficient conditions for optimality have been rigorously established in [12]. The proofs are very technical, and rather long and essentially solve a nonlinear programming problem on a space of p.o.m.'s. We have the following analogue of Theorem 1 from [12] in the vector case:

<u>Theorem 3:</u> Suppose the signal sequence $\{x(i)\}$ and the measurement outcomes $v(i)$, $i=0,\ldots,k-1$ have finite second moments. Then there exist optimal p.o.m. $\hat{M}_k$ and $n\times n$ matrices $\hat{C}_i(k)$, $i=0,\ldots,k-1$. Moreover the optimal measurement outcome has also finite second moments. Necessary and sufficient conditions for optimality are

i) $\sum_{j=0}^{k} E\{v(i)v(j)^t\}\,\hat{C}_j^T(k) = E\{v(i)x(k)^T\}$ $i=0,1,\ldots,k$

ii) $\mathfrak{F}(\cdot,\hat{\underset{\sim}{C}}(k))$ is integrable with respect to $\hat{M}_k$

iii) $\mathfrak{F}(u, \hat{\underset{\sim}{C}}(k)) \geq \hat{\tau}$ for all $u \in R^n$ where $\hat{\tau} = \int_{R^n} \mathfrak{F}(u, \hat{\underset{\sim}{C}}(k)) \hat{M}_k(du)$.

We immediately observe that the optimal filter equations are much more complex and indirect than in the scalar case. The reasons for that are explained in detail in joint work with R. O. Harger in [13]. It is also important to recall that in the multiparameter case the implementation is more difficult, because we have to find the auxiliary system (ρ_e, H_e) and the set of compatible measurements on the augmented system that realize the optimal measurement (see Introduction). As in the scalar case we obtain a separation theorem under Gaussian statistics. This is more delicate here and it has been completely worked out in joint work with R. O. Harger [13]. We denote by $\hat{Z}_j$ the p.o.m. that represent the intrinsic measurements, i.e. those that provide the minimum error variance estimator of x(k) without regard to past data. We give here a special case of the separation theorem proven in [13] due to the complexity of the general case.

Theorem 4: Suppose i) $\{x(i)\}$ is a Gaussian process, and the component processes $\{x_j(i)\}$ $j = 1, \ldots, n$, are uncorrelated and identically distributed.

ii) The intrinsic measurement outcome $\hat{z}(i)$ and x(i) are jointly Gaussian for each i, and that the component processes $\{\hat{z}_j(i)\}$ $j = 1, \ldots, n$ are uncorrelated with the same covariance function.

Then the intrinsic measurement outcomes $\hat{z}(i)$, $i = 0, 1, \ldots, k$, are a sufficient statistic for the linear minimum error variance quantum estimator $\hat{x}(k)$ of x(k).

We have then in this case also the analoque of Fig. 1. The additional complication of the separation theorem is due to the quantum mechanical constraints in the multiparameter case.

IV. Examples

We give here two examples that illustrate the optimal filters obtained from the theory described above.

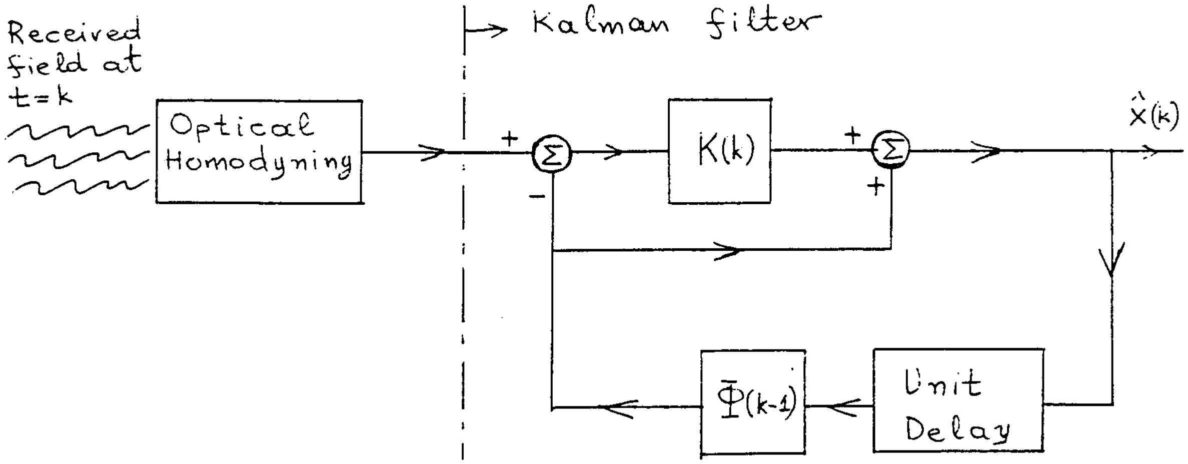
Example 1: This is worked out in detail in [5]. Suppose x(k) is a Gaussian process transmitted as the real amplitude of a laser signal (assumed monochromatic) and received, along with thermal noise, in a single-mode cavity upon which an optimal measurement is to be made. The density operator in the coherent state or P-representation is then [4]

$$\rho(x(k)) = \frac{1}{\pi N} \int \exp\left(-\frac{|\alpha - x(k)|^2}{N}\right) |\alpha\rangle\langle\alpha| \, d^2\alpha \qquad (17)$$

where N defines the thermal noise level. Suppose that the signal sequence is generated by the recursive equation

$$x(k+1) = \Phi(k)x(k) + w(k) \quad (18)$$

where w(k) is a white noise process. Then the optimal quantum linear filter (see [5] for details) is described in Fig. 2.



$$P(k) = \Phi^2(k-1)\{P(k-1)[1 - K(k-1)]\} + Q(k-1)$$

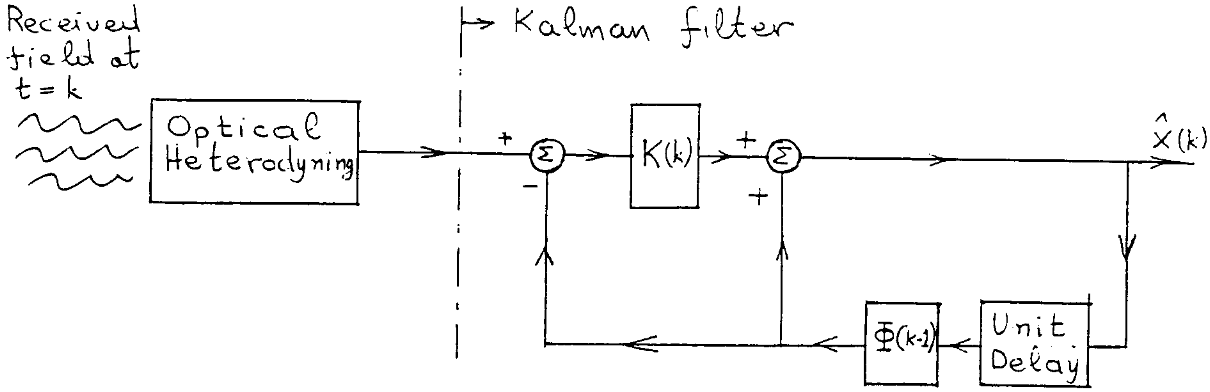
$$K(k) = P(k)[P(k) + (\frac{N}{2} + \frac{1}{4})]^{-1}$$

Fig. 2: Optimal linear filter for Gaussian signal sequence received as amplitude of a laser signal in single-mode cavity along with thermal noise.

Example 2: This is worked out in detail in [13]. Suppose that x(k) is an R^2 Gaussian process with identically distributed, independent, component sequences, which is transmitted as the in phase $(x_1(k))$ and quadrature $(x_2(k))$ amplitudes of a monochromatic laser, that is received along with thermal noise, in a single mode cavity upon which measurements can be made. The density operator in the P-representation is now

$$\rho(x(k)) = \frac{1}{\pi\eta_0}\int e^{-|\alpha - x_1(k) - ix_2(k)|^2/\eta_0} |\alpha\rangle\langle\alpha| d^2\alpha \quad (19)$$

The signal process is generated by the vector analogue of (18). Then the optimal quantum linear filter (see [13] for details) is described in Fig. 3.



$$P(k) = \Phi(k-1)[P(k-1) - K(k-1)P(k-1)]\,\Phi^T(k-1) + Q(k-1)$$

$$K(k) = P(k)[P(k) + \frac{n_0+1}{2} I_2]^{-1}$$

Fig. 3: Optimal linear filter for 2 dim. Gaussian signal sequence received as in phase and quadrature amplitudes of a laser in single-mode cavity along with thermal noise.

V. Some Open Problems

What I described in this paper, is just an initial attempt to understand some filtering problems with nonclassical character, motivated by important problems in quantum physics as well as practical problems in optical (laser) communication systems. The area is very rich and I believe that important contributions can be made in this class of distributed parameter problems. I would like to mention just a few of the immediate open problems:

1) What is the solution, and what modifications are necessary when the conditional independence assumption is dropped?
2) Continuous time problems, that bring into the picture the dynamics of the evolution of ρ via the Liouville-von Neumann equation [9, p. 158] which makes the problem clearly a distributed parameter filtering problem.
3) More examples are needed where the computations can be carried through.
4) The nonlinear filtering problem and its relations to Fock space representations and conditional expectations on von Neumann Algebras [16].

Acknowledgement:

As it is obvious from the references and the discussion in the paper, many of the results described here are joint work with R.O. Harger and are part of our joint continuing efforts in the area of Quantum Filtering Theory.

References

1. Holevo, A. S., "Statistical Problems in Quantum Physics", Proceedings of 1972 Soviet-Japanese Symposium on Probability and Statistics, Vol. 1, 1972, pp. 22-40.

2. Holevo, A. S., "Statistical Decision Theory for Quantum Systems", Journal of Multivariate Analysis, Vol. 3, pp. 337-394, 1973.

3. Helstrom, C. W., "Detection Theory and Quantum Mechanics", Information and Control, Vol. 10, pp. 254-291, March 1967.

4. Helstrom, C. W., Liu, J. W., and Gordon, J. P., "Quantum Mechanical Communication Theory", Proceedings of the IEEE, Vol. 58, pp. 1578-1598, October 1970.

5. Baras, J. S., Harger, R. O., and Park, Y. H., "Quantum Mechanical Linear Filtering of Random Signal Sequences", IEEE Trans. on Info. Theory, January 1976, pp. 59-64.

6. Yuen, H. P. and Lax, M., "Multiple-Parameter Quantum Estimation and Measurement of Non-selfadjoint Observables", IEEE Trans. on Info. Theory, Vol. IT-19, pp. 740-749, November 1973.

7. Helstrom, C. W. and Kennedy, R. S., "Non-commuting Observables in Quantum Detection and Estimation Theory", IEEE Trans. on Info. Theory, Vol. IT-20, pp. 16-24, January 1974.

8. Prugovecki, E., Quantum Mechanics in Hilbert Space, New York: Academic Press, 1971.

9. Jauch, J. M., Foundations of Quantum Mechanics, Reading: Addison-Wesley, 1968.

10. Baras, J. S. and Harger, R. O., "Multiparameter Filtering with Quantum Measurements", Proceedings of 1975 Conference on Information Sciences and Systems, Johns Hopkins University, April 1975, pp. 178-184.

11. Baras, J. S. and Harger, R. O., "Nonclassical Filtering Problems in Quantum Electronics", Proc. of 1975 IEEE Conference on Decision and Control, Houston, Texas, December 1975, pp. 804-806.

12. Baras, J. S., "An Optimization Problem from Linear Filtering with Quantum Measurements", to appear.

13. Baras, J. S. and Harger, R. O., "Quantum Mechanical Linear Filtering of Vector Signal Processes", to appear.

14. Akhiezer, N. I. and Glazman, I. M., Theory of Operators in Hilbert Space, Vol. II, New York: Frederick Ungar Pub. Co., 1963.

15. Luenberger, D. G., Optimization by Vector Space Methods, New York: John Wiley & Sons, 1969.

16. Baras, J. S. and Harger, R. O., "Conditional Expectations and Fock Space Representations in Quantum Filtering", Proc. of 4th International Symposium on Information Theory, Leningrad 1976, Part I, pp. 180-185.