1975 Conference on Information Sciences and Systems, April 1975

# MULTIPARAMETER FILTERING WITH QUANTUM MEASUREMENTS

John S. Baras and R. O. Harger
Electrical Engineering Department
University of Maryland
College Park, Maryland 20742

## Abstract

We analyze here the problem of estimating a member of a vector discrete time process utilizing past and present quantum mechanical measurements. The minimum variance linear estimator based on optimal present measurement selection combined with optimal linear processing of past measurements is studied. A necessary condition for the optimal extended measurement and the optimal coefficient matrices is given which leads to the generation of several more specialized necessary conditions. Certain operator equations related to these necessary conditions are studied and it is found that the conditions are necessary and sufficient in the commutative case. Finally when the average quantum measurement is linear in the random signal and the signal process is pair wise Gaussian, the filter separates: the optimal measurement can be taken the same as the optimal measurement with no regard to past data and the past and present data are processed classically. The results are illustrated by considering the estimator of the amplitude and phase of a laser received in a single-mode cavity along with thermal noise; when the random signal sequence satisfies a linear recursion, the estimate can be computed recursively.

## Introduction

Detection and estimation problems have recently been studied [1, 2, 3] employing measurement models correctly incorporating quantum mechanics. Such work applies directly, e.g., to establishing fundamental limitations in optical communication systems [4]. More recently the analogue of filtering a random signal sequence has been considered [5, 6, 13, 14]; here the problem of estimating $X_k$, a member of a "signal" sequence $\{X_0, X_1, \dots, X_k, \dots\}$ of vector random variables is considered; the parameter k is conveniently regarded as discrete time. To be chosen is the "best" measurement at time k and the "best" linear combination of present and past measurements at times $k' = 0, 1, \dots, k-1$. The random sequence so obtained is defined precisely below but is simply described in the optical communication setting as follows.

At time k a laser modulated in some fashion by $X_k$ is received in a cavity containing otherwise only an electromagnetic field due to thermal noise: the total field is in a state described by a density operator $\rho(X_k)$ that depends on $X_k$ (but not otherwise on k). If $X_k$ is a scalar, the measurement (whose outcome is denoted $v_k$) at time k will correspond to a self-adjoint operator $V_k$ [7]; if $X_k$ is a vector the essential quantum problem of simultaneous measurement arises and a more general concept of measurement [1, 2, 8, 9] must be resorted to.

By "best" is meant "minimum mean-square error"; the implied average is over the ("classical") distributions of $\{X_k\}$ and the distributions due to quantum mechanical measurement.

An ultimate objective would include efficient computation; e.g., suppose that $X_k$ is a "dynamical state" generated by the recursive equation

$$X_{k+1} = \varphi_k X_k + W_k$$

where $\{\varphi_k\}$ is a sequence of matrices and $\{W_k\}$ is a sequence of independent, Gaussian random vectors, with zero mean and covariance matrix $Q_k$: solutions in a form that compute recursively the best estimate and measurement at time k would be highly desirable. In a specific situation below this is achieved.

## 1. Extended quantum measurements

Following Holevo [1, 2] let $\mathcal{H}$ be a Hilbert space and $\mathcal{R}$ the algebra of all bounded operators on $\mathcal{H}$. An extended (quantum) state is the linear functional $\hat{\rho}: \mathcal{R} \to R^1$ such that, $\forall A \in \mathcal{R}$, $\hat{\rho}(A) \equiv Tr\{\rho A\}$ and is positive and normed $(\hat{\rho}(I) = 1)$; here $\rho$ is the d.o. corresponding to the state $\hat{\rho}$. An extended measurement $M: B^N \to \mathcal{R}$ ($B^N$ the Borel $\sigma$-algebra of $\mathbb{R}^N$) is such that:

(i) $\forall B \in B^N$, $M(B) \geq 0$;

(ii) $\forall$ partition $\{B_i\}$ of $\mathbb{R}^N, B_i \epsilon \mathcal{B}^N, \sum_i M(B_i) = I$;

that is, an extended measurement is a positive operator-valued measure. Given an extended state, an extended measurement induces a probability measure $\mu$ on $\mathcal{B}^N$: $\forall B \epsilon \mathcal{B}^N$, $\mu(B) = \hat{\rho}[M(B)] = \mathrm{Tr}\{\rho M(B)\}$.

As pointed out by Holevo [2], this extension is justified in view of Naimark's theorem: $\exists$ a Hilbert space $\mathcal{K}_e$, a state $\hat{\rho}_e$ on $\mathcal{R}(H_e)$, and a simple measurement $M_V$ in $\mathcal{K}\otimes\mathcal{K}_e$ such that, $\forall B \epsilon \mathcal{B}^N$ and $\forall \hat{\rho}$ on $\mathcal{R}(\mathcal{K})$,

$$\hat{\rho}[M(B)] = (\hat{\rho}\otimes\hat{\rho}_e)[M_V(B)].$$

Now let X be a vector r.v. on $(\Omega, \mathcal{F}, P)$ on which the extended state $\hat{\rho}$ depends: then $\mu$ becomes a conditional probability measure and, $\forall B \epsilon \mathcal{B}^N$,

$$\mu(B|X) = \hat{\rho}[M(B)|X] = \mathrm{Tr}\{\rho(X)M(B)\}. \quad (1)$$

Then the unconditional measure would be

$$\mu(B) = \int \mathrm{Tr}\{\rho(X)M(B)\}\, F_X(dX).$$

We shall call the triple $\{\mathcal{K}_e, \rho_e, M_V\}$ a realization of the measurement represented by the p.o.m. M. The physical motivation for this is well known; see [2] [9].

We consider now a sequence of measurements represented by the p.o.m. $M_i$. The outcome of each measurement which is made at time i, will be an N-vector v(i). Considering the d.o. $\rho(X(0))\otimes\ldots\otimes\rho(X(k))$ on the Hilbert space $\mathcal{K}_0\otimes\mathcal{K}_1\otimes\ldots\otimes\mathcal{K}_k$ we have that the probability measure characterizing the outcome of the first $k+1$ measuring experiments is given by

$$\mu[B(0)\times B(1)\times\ldots\times B(k)] = \int\ldots\int\ldots\mu_{k'}[B(k')|X(k')]\cdot\ldots F_{X(0),\ldots X(k)}(dX(0),\ldots dX(k))$$

where $\mu_{k'}[B(k')|X(k')]$ is given by (1) and $B(i)\epsilon\mathcal{B}^N$, $i=0,\ldots k$. This is due to the conditional (on the signal) independence of the measurement outcomes [6,13,14]. The (unconditional) density function is

$$F(a(0),\ldots a(k)) = \int\cdots\int\cdots \mathrm{Tr}\{\rho(X(k'))M_{k'}(-\infty, a(k')]\cdot F_{X(0)\cdots X(k)}(dX(0),\ldots dX(k)) \quad (2)$$

where by $(-\infty, a(k')] \triangleq (-\infty, a_1(k')]\times\ldots\times(-\infty, a_N(k')]$.

## 2. Formulation of the filtering problem and existence of the optimal linear filter.

In this paper our objective is to find a p.o.m. $M_k$ and N×N matrices $C_{k'}(k)$ $k'=0,1,\ldots k$ so that if we let

$$\hat{X}(k) = \sum_{k'=0}^{k} C_{k'}(k)v(k') \quad (3)$$

the mean square error

$$MSE = E\{\|X(k)-\hat{X}(k)\|^2_{\mathbb{R}^N}\} \quad (4)$$

is minimized. The interpretation of (3) (generalizing [13],[14]) is that at each time the optimal estimator is computed as a linear function of current and past measurement outcomes $v(k'), k'=0,\ldots k$. The average in (4) is over the distribution (2). Due to space limitations we will describe the main results here with only a sketch of the proofs. A detailed description of the results will be given in [16]. In the sequel we will use the notion of the integral and trace-integral with respect to a positive operator valued measure developed by Holevo [2, pp. 354-361]. Following [2] we let $\mathcal{T}_h$ denote the set of all trace-class (finite trace) selfadjoint (hermitean) operators on a Hilbert space $\mathcal{K}$. Then we shall denote by

$$\langle F, X\rangle_A \quad (5)$$

the trace integral over $A\subseteq\mathbb{R}^N$ of the $\mathcal{T}_h$-valued function F with respect to the p.o.m. X. In all cases described in this paper F will be of the form

$$F(u) = \sum_{i=1}^{\ell} K_i g_i(u) \quad (6)$$

where $K_i \epsilon \mathcal{T}_h$ and $g_i$ are continuous real valued functions. Then [2, p. 356], for such F the trace integral always exists and

$$\langle F, X\rangle_A = \sum_i \int_A g_i(u)\mathrm{Tr}\, K_i X(du) \quad (7)$$

After some calculations (which we omit) the M S E can be written as

$$\begin{aligned} MSE &= E\{(X(k)-\sum_{k'=0}^{k} C_{k'}(k)v(k'))^t(X(k) - \sum_{k'=0}^{k} C_{k'}(k)v(k'))\} \\ &= E_X\int\{[X(k)-C_k(k)v(k)]^t[X(k)-C_k(k)v(k)] - 2[X(k)-C_k(k)v(k)]^t\sum_{i=0}^{k-1}C_i(k)E[v(i)|X(i)]\}\cdot \\ &\cdot \mathrm{Tr}\,\rho(X(k))M_k(dv(k)) + \\ &+ \mathrm{tr}[C_0(k)\ldots C_{k-1}(k)]E\begin{bmatrix}v(0)\\ \vdots\\ v(k-1)\end{bmatrix}\begin{bmatrix}v(0)\\ \vdots\\ v(k-1)\end{bmatrix}^t\begin{bmatrix}C_0^t(k)\\ \vdots\\ C_{k-1}^t(k)\end{bmatrix} \end{aligned} \quad (8)$$

where tr denotes the trace of matrices (while Tr denotes the trace of operators). Let us define the following operators:

$$\Lambda(k) = \int X(k)^t X(k)\,\rho(X(k))\,F_{X(k)}(dX(k)) \tag{9}$$

$$\eta(k) = \int \rho(X(k))\,F_{X(k)}(dX(k)) \tag{10}$$

$$\delta_i(k) = \int X_i(k)\rho(X(k))F_{X(k)}(dX(k));\, i=1,\dots N \tag{11}$$

By $\delta(k)$ we understand the N-vector of operators with components $\delta_i(k)$. Consider also (in obvious notation) the N-vector of operators

$$\zeta(k,i) = \int E[\,v(i)\,|\,X(i)\,]\,\rho(X(k))F_{X(k)}(dX(k)) \tag{12}$$

and the NxN matrix of operators

$$\Pi(k,i) = \int X(k)E[\,v(i)\,|\,X(i)\,]^t\,\rho(X(k))F_{X(k)}(dX(k)). \tag{13}$$

Then as shown in [16] letting $C(k)=[\,C_0(k), \dots C_{k-1}(k)\,]$ and $v(0,k)=[\,v(0)^t, \dots v(k-1)^t\,]^t$ we have:

$$\text{MSE} \triangleq J(C(k), M_k) = \langle \mathcal{F}_{C(k)}, M_k \rangle_{\mathbb{R}^N} + \operatorname{tr} C(k) E v(0,k) v(0,k)^t C(k)^t \tag{14}$$

where

$$\mathcal{F}_{C(k)}(v(k)) = \Lambda(k) - 2v(k)^t C_k(k)\delta(k) + v(k)^t C_k(k)^t C_k(k) v(k)\eta(k) + 2\sum_{i=0}^{k-1} v(k)^t C_k(k)^t C_i(k)\zeta(k,i) - 2\sum_{i=0}^{k-1} \operatorname{tr} C_i(k)^t \Pi(k,i) \tag{15}$$

In (15) we use the notation that vectors and matrices over $\mathbb{R}^N$ and vectors and matrices of operators are multiplied in a formal fashion. It is easily seen [16] that $\mathcal{F}_{C(k)}(v(k))$ is of the type described in (6).

Now let $\mathcal{M}(k)$ be the set of p.o.m. on $\mathcal{H}_k$, which is convex [2]. Let $\mathbb{R}^{N\times N}$ denote the set of NxN matrices. Then our problem can be formulated as follows:

Find $\hat{C}_i(k) \in \mathbb{R}^{N\times N}$, $i=0,\dots k$, and $\hat{M}_k \in \mathcal{M}(k)$ to minimize (14). Generalizing the work of Holevo we have then:

Theorem 1: There exist p.o.m. $\hat{M}_k$ and matrices $\hat{C}_i(k)$, $i=0,\dots k$, which minimize $J(C(k), M_k)$ over $\mathbb{R}^{N\times N} \times \dots \times \mathbb{R}^{N\times N} \times \mathcal{M}(k)$.

The proof is somewhat lengthy and is given in [16].

## 3. Necessary Conditions

We can think of $\mathcal{M}(k)$ as embedded in the vector space of operator measures (not positive). Then the minimization problem of the previous section can be solved by a standard application of Gateaux differentials, see [11, p. 178]:

Theorem 2: Let $\hat{C}_0(k), \hat{C}_1(k), \dots \hat{C}_k(k)$ and $\hat{M}_k$ be the optimal p.o.m. at time k and the optimal processing coefficient matrices. Then

i) $\langle \mathcal{F}_{\hat{C}(k)}, X \rangle_{\mathbb{R}^N} \geq \langle \mathcal{F}_{\hat{C}(k)}, \hat{M}(k) \rangle_{\mathbb{R}^N}$, $\forall X \in \mathcal{M}(k)$

ii)

$$\begin{bmatrix} E(v(0)v(0)^t) & \dots & E(v(0)\hat{v}(k)^t) \\ \vdots & & \vdots \\ E(\hat{v}(k)v(0)^t) & \dots & E(\hat{v}(k)\hat{v}(k)^t) \end{bmatrix} \begin{bmatrix} \hat{C}_0(k)^t \\ \vdots \\ \hat{C}_k(k)^t \end{bmatrix} = \begin{bmatrix} E(v(0)X(k)^t) \\ \vdots \\ E(\hat{v}(k)X(k)^t) \end{bmatrix}$$

Proof: We only sketch the proof. The details are in [16]. We compute the Gateaux differential of $J(\cdot,\cdot)$ at $\hat{C}(k), \hat{M}_k$ and let us denote it by $\delta J(\hat{C}(k), \hat{M}_k; A, X)$. Then from [11, p. 178] we must necessarily have $\delta J(\hat{C}(k), \hat{M}_k; A-\hat{C}(k), X-\hat{M}_k) \geq 0$ for all k+1 tuples of NxN matrices A and all p.o.m. X. Writting down this condition and taking first $A=\hat{C}(k)$ and then $X=\hat{M}_k$ result in i), ii).

Two important observations should be emphasized: a) notice that ii) is just the normal equations for the minimum variance linear estimate of X(k) based on the random variables $v(0), \dots v(k-1), \hat{v}(k)$. The optimum measurement $\hat{M}_k$ enters through the statistics of $\hat{v}(k)$. b) Notice that i) states that $\hat{M}_k$ minimizes the linear functional $\langle \mathcal{F}_{\hat{C}(k)}, X \rangle_{\mathbb{R}^N}$ (where $\hat{C}(k)$ are the optimal matrices) over $\mathcal{M}(k)$. This problem has been investigated by Holevo and thus we can utilize his results [2, p. 368-371].

Notice that without loss of generality we can assume that $\hat{C}_k(k) = I_N$ (identity on $\mathbb{R}^N$).

We derive here new necessary conditions for the problem stated in i) of Theorem 2, which include those of Holevo [2, p.368]. Recall that since $\mathcal{F}_{\hat{C}(k)}$ is of the type described in (6) above its trace integral exists with respect to any p.o.m. X. Suppose in addition that $\mathcal{F}_{\hat{C}(k)}$ is integrable with respect to $\hat{M}_k$ and let

$$\Delta_{\hat{C}(k)} = \int \mathcal{F}_{\hat{C}(k)} \hat{M}_k(dv) \quad . \tag{16}$$

Then $\langle \mathcal{F}_{\hat{C}(k)}, X \rangle_{\mathbb{R}^N} \geq \langle \mathcal{F}_{\hat{C}(k)}, \hat{M}_k \rangle_{\mathbb{R}^N}$ for any X in $\mathcal{M}(k)$ implies

$$\int_A \operatorname{Tr}(\mathcal{F}^{(v)}_{\hat{C}(k)} - \Delta_{\hat{C}(k)})X(dv) \quad \geq 0;\ \forall A \in B^N. \tag{17}$$

Now

$$\int (\mathcal{F}^{(v)}_{\hat{C}(k)} - \Delta_{\hat{C}(k)})\hat{M}_k(dv) \quad = 0 \tag{18}$$

and

$$\int_A (\mathcal{F}^{(v)}_{\hat{C}(k)} - \Delta_{\hat{C}(k)})\hat{M}_k(dv) \quad \geq 0,\ \forall A \in B^N: \tag{19}$$

for if $\exists\ A \in B^N$ such that $\int_A (\mathcal{F}^{(v)}_{\hat{C}(k)} - \Delta_{\hat{C}(k)})\hat{M}_k(dv) < 0$ Then $\int_A \operatorname{Tr}(\mathcal{F}^{(v)}_{\hat{C}(k)} - \Delta_{\hat{C}(k)})\hat{M}_k(dv) < 0$ which is a contradiction to (17). But then (18) and (19) imply that

$$\int_A (\mathcal{F}^{(v)}_{\hat{C}(k)} - \Delta_{\hat{C}(k)}) \hat{M}_k(dv) = 0; \forall A \in B^N. \quad (20)$$

We have thus:

Theorem 3: Let $\hat{M}_k$ and $\hat{C}_0(k), \hat{C}_1(k), \ldots \hat{C}_{k-1}(k)$ be the optimal p.o.m. and the optimal processing coefficient matrices at time k. Then if $\mathcal{F}_{\hat{C}(k)}$ is $\hat{M}_k$ integrable and $\Delta_{\hat{C}(k)} = \int \mathcal{F}^{(v)}_{\hat{C}(k)} \hat{M}_k(dv)$

i) $\int_A (\mathcal{F}^{(v)}_{\hat{C}(k)} - \Delta_{\hat{C}(k)}) \hat{M}_k d(v) = 0; \forall A \in B^N$

ii) the normal equations of Theorem 2 are satisfied.

## 4. Average observation linear in the signal

To obtain more detailed results about the structure of the optimal filter we assume that the signal process is pair wise Gaussian. Thus we obtain a generalization in the multiparameter case of some of the results of [14]. Following Holevo [2, p. 356] we shall consider only p.o.m. that have a base. This class is very important for applications (cf. P-representations). A p.o.m. X is a p.o.m. measure with base $\mu$ if there exist a positive operator valued function P and a measure $\mu$ on $B^N$ such that for every $B \in B^N$ we have

$$X(B) = \int_B P(u)\mu(du) \quad (21)$$

where the integral is the Bochner $\mathcal{k}$ -integral. For such p.o.m. our necessary condition becomes

$$(\mathcal{F}_{\hat{C}(k)}(v) - \Delta_{\hat{C}(k)})P(v) = 0 \quad \mu.\text{a.e.} \quad (22)$$

Clearly from the form of $\mathcal{F}_{\hat{C}(k)}$, see (Eq. 15), $\mathcal{F}_{\hat{C}(k)}$ is differentiable as an operator valued function. Let us assume that P is continuous, and then from (22) we have

$$P(u)(\mathcal{F}_{\hat{C}(k)}(u) - \mathcal{F}_{\hat{C}(k)}(v))P(v) = 0 \quad \mu.\text{a.e.} \quad (23)$$

Dividing by u-v and taking the limit as $u-v \to 0$ we recapture the necessary condition originally due to Holevo [2, Th. 9.2]:

$$P(u) \frac{\partial \mathcal{F}_{\hat{C}(k)}(u)}{\partial u_\ell} P(v) = 0; \ \ell = 1, \ldots N, \ \mu.\text{a.e.} \quad (24)$$

We want now to compare the optimal filter that utilizes past measurements with the one that utilizes only present measurements. Let $\hat{Z}_i$ be the optimal measurements for the latter. All p.o.m. considered are p.o.m. with a base and with continuous densities (i.e. P). Then from (24) (15) we have that $\hat{M}_i$ satisfies

$$\hat{P}_i(u)(\delta(i) - u\eta(i) - \sum_{j=0}^{i-1} \hat{C}_j(i)\zeta(i,j))\hat{P}_i(v) = 0; \mu_i.\text{a.e.} \quad (25)$$

Notice that the optimal measurements for the filter without post processing satisfy (set $\hat{C}_j(i)=0$ in (25)) the necessary condition:

$$\hat{\Xi}_i(z)(\delta(i) - z\eta(i))\hat{\Xi}_i(z') = 0; \nu_i \text{ a.e.} \quad (26)$$

where $\hat{Z}_i = \int \hat{\Xi}_i(z)\nu_i(dz)$. (27)

Let us denote by z(i) the outcomes of the measurement represented by $\hat{Z}(i)$. Suppose that there exists matrices $\Gamma(i)$ such that $E[z(i)|X(i)] = \Gamma(i)X(i)$. Certainly this is so whenever $(z(i), X(i))$ are jointly Gaussian. On the other hand since $X(k'), X(k)$ are jointly Gaussian, there exists matrices $A(k',k)$ such that $E[X(k')|X(k)] = A(k',k)X(k)$.

Clearly $\hat{Z}_0$ satisfies also the necessary condition (25) for the optimal measurement of the filter with post processing. Now define a new measure $M_1$ via the equation

$$M_1(A) = \hat{Z}_1(\beta_1^{-1}(A)), \ \forall A \in B^N \quad (28)$$

where $\beta_1(X) = B(1)X$ and $B(1) = I_N - \hat{C}_0(1)\Gamma(0)A(0,1)$.

Without loss of generality (as will be explained later) we can assume that B(1) is nonsingular. Then from (26) we have

$$\hat{\Xi}_1(z)(B(1)\delta(1) - B(1)z\eta(1))\hat{\Xi}_1(z') = 0; \nu_1 \text{ a.e.} \quad (29)$$

But from (12) $\zeta(1,0) = E_X[E[v(0)|X(0)]\rho(X(1))] =$

$$= \Gamma(0)E_X[E[X(0)|X(1)]\rho(X(1))] = \Gamma(0)A(0,1)\delta(1) \quad (30)$$

So (29) reads

$$\hat{\Xi}_1(z)(\delta(1) - \hat{C}_0(1)\zeta(1,0) - B(1)z\eta(1))\hat{\Xi}_1(z') = 0; \nu_1 \text{ a.e.} \quad (31)$$

Letting $u = B(1)z$, $P_1(u) = \hat{\Xi}_1(B(1)^{-1}u)$ and $\mu_1(du) = \nu_1(B(1)^{-1}du)$ we see that (31) reduces to

$$P_1(u)(\delta(1) - \hat{C}_0(1)\zeta(1,0) - u\eta(1))P_1(v) = 0; \mu_1 \text{ a.e.} \quad (32)$$

and therefore the p.o.m. $M_1$ defined in (28) has a base and satisfies the necessary condition for the optimal measurements of the filter with postprocessing. The physical meaning of the relation between $\hat{Z}_1$ and $M_1$ is illustrated below:

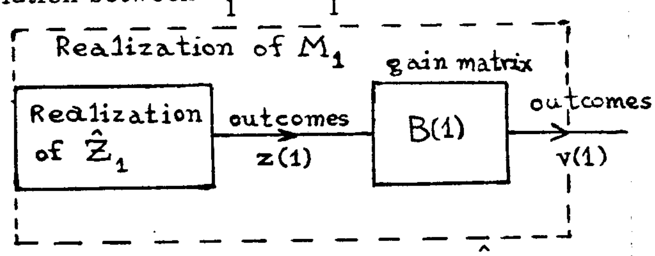


Fig. 1: Relation between the p.o.m $\hat{Z}_1$ and $M_1$.

Repeating the same process we have:

Theorem 4: Consider a sequence of measurements $\hat{Z}_i$ that satisfy the necessary conditions (26), for the optimal filter without postprocessing. Suppose the signal process is pairwise Gaussian and that the average observation from the measurement $\hat{Z}_i$ is linear in the signal (i.e. $E[z(i)|X(i)]=\Gamma(i)X(i)$ for some matrices $\Gamma(i)$). Let the matrices B(k) be defined recursively via

$$B(k)=I_N-\sum_{k'=0}^{k-1}\hat{C}_{k'}(k)B(k')\Gamma(k')A(k',k);\ B(0)=I_N, \quad (33)$$

and define the sequence of p.o.m.

$$M_i(A)=\hat{Z}_i(\beta_i^{-1}(A)),\ \forall A\in B^N, \quad (34)$$

where $\beta_i(X)=B(i)X$. Then the p.o.m. $M_i$ have base and satisfy the necessary conditions (25) for the optimal measurements of the filter with postprocessing.

Proof: The only thing we need to complete the proof is the inductive step. Assume that the theorem is true at time k-1. Then for j=0,... k-1 we have

$$\zeta(k,j)=E_X[E[v(j)|X(j)]\rho(X(k))]=B(j)E_X[[z(j)|X(j)]\rho(X(k))]=$$

$$=B(j)\Gamma(j)A(j,k)\delta(k) \quad (35)$$

From the necessary condition for $\hat{Z}_k$ we get

$$\hat{\mathfrak{X}}_k(z)(B(k)\delta(k)-B(k)z\eta(k))\hat{\mathfrak{X}}_k(z')=0;\ \nu_k \text{ a.e.} \quad (36)$$

Therefore letting $P_k(u)=\hat{\mathfrak{X}}_k(B(k)^{-1}u)$ and $\mu_k(du)=\nu_k(B(k)^{-1}du)$ we have

$$P_k(u)(\delta(k)-\sum_{k'=0}^{k-1}\hat{C}_{k'}(k)\zeta(k,k')-u\eta(k))P_k(v)=0;\ \mu_k \text{ a.e.} \quad (37)$$

and therefore the theorem holds at time k.

In view of Theorem 4 and the discussion that preceded it, the optimal estimator takes the form

$$\hat{X}(k)=v(k)+\sum_{k'=0}^{k-1}\hat{C}_{k'}(k)v(k')=B(k)z(k)+\sum_{k'=0}^{k-1}\hat{C}_{k'}(k)B(k')z(k') \quad (38)$$

where the vector random variables $\{z(k')\}$ result from the measurements represented by the p.o.m.'s $\{\hat{Z}_k\}$. The normal equations of Theorem (2) then become

$$\begin{bmatrix} B(0)E[z(0)z(0)^t]B(0)^t & \cdots & B(0)E[z(0)z(k)^t]B(k)^t \\ \vdots & & \\ B(k)E[z(k)z(0)^t]B(0)^t & \cdots & B(k)E[z(k)z(k)^t]B(k)^t \end{bmatrix}\begin{bmatrix}\hat{C}_0(k)^t\\ \vdots \\ \hat{C}_{k-1}(k)^t \\ I_N\end{bmatrix} = \begin{bmatrix} B_0E[z(0)X(k)^t] \\ \vdots \\ B(k)E[z(k)X(k)^t]\end{bmatrix} \quad (39)$$

Without loss of generality we can assume B(k') nonsingular, for if for any k', B(k') is singular, B(k') may be restricted to the complement of its null space without effecting either X(k) or the normal equations (39). Hence the normal equations may be written

$$\begin{bmatrix} E[z(0)z(0)^t] & \cdots & E[z(0)z(k)^t] \\ \vdots & & \\ E[z(k)z(0)^t] & \cdots & E[z(k)z(k)^t]\end{bmatrix}\begin{bmatrix}[\hat{C}_0(k)B(0)]^t \\ \vdots \\ [\hat{C}_{k-1}(k)B(k-1)]^t\end{bmatrix} = \begin{bmatrix}E[z(0)X(k)^t]\\ \vdots \\ E[z(k)X(k)^t]\end{bmatrix} \quad (40)$$

Comparing equations (38) and (40) we have:

Theorem 5: Let the assumptions of Theorem 4 hold. Then the measurements $z(k'), k'=0,1,\dots k$, represented from the p.o.m.'s $\{\hat{Z}_{k'}, k'=0,1,\dots k\}$ are a sufficient statistic for the L.M.V.E. of X(k).

Proof: The only thing needed to complete the proof is certain sufficiency results that will be given in [16], and which we omit from here due to space limitations.

Notice that Theorem 5 establishes the "separation" of the filter (c.f. [14]).

## 5. Recursive filtering and examples

The advantage gained with the result described in Theorem 5 of the previous section is that whenever the signal process X(k) allows a recursive solution for (40), we will obtain a truly recursive filter. The measurements $\hat{Z}_k$ should be thought of as intrinsic to the quantum field at hand and can be found apriori. Another advantage is that this "separation" produces a considerable reduction on the number of measuring devices needed. Whenever X(k) satisfies a recursion of the type

$$X(k+1)=\phi(k)X(k)+W(k) \quad (41)$$

where $\phi(k)$ is a sequence of NxN matrices and W(k) is a sequence of independent, zero mean Gaussian random vectors with covariance matrices Q(k) the processing coefficient matrices can be computed recursively. Kalman Bucy filtering [11] may be directly used on the observations to obtain a recursive filtering: such an example follows.

Example: Suppose that the two dimensional dynamical state X(k) is transmitted as the in-phase $X_1(k)$ and quadrature $X_2(k)$ amplitudes of a laser (assumed monochromatic) and received, along with thermal noise, in a single mode cavity upon which an optimal extended measurement is to be made. The d.o. in the coherent state, or P-, representation is then [1]

$$\rho(X(k))=\frac{1}{\pi n_0}\int e^{-|\alpha-[X_1(k)+iX_2(k)]|^2/n_0}|\alpha\rangle\langle\alpha|\, d^2\alpha$$

Assume for simplicity that $X_1(k)$ and $X_2(k)$ are independent, zero mean Gaussian random variables with identical variances $\lambda(k)$. The solution to (26) are known to be [8, 9, 12] the p.o.m. $\hat{Z}_k$, where

$$T_1(k)=\int \mathrm{Re}\,\alpha\,\hat{Z}_k(d\alpha)=D(k)\frac{a_k+a_k^+}{2},\ T_2(k)=\int \mathrm{Im}\,\alpha\,\hat{Z}_k(d\alpha)= = D(k)\frac{a_k-a_k^+}{2i},\quad D(k)=\frac{2\lambda(k)}{n_0+2\lambda(k)+1}, \tag{42}$$

which has a base with continuous density. The outcome of this measurement, assuming fixed $X(k)$, are independent Gaussian random variables with means $X_1(k)$ and $X_2(k)$, resp. and variances $(n_0/2+1/2)$ and is realized by heterodyning [12]. Also [1, 8, 9] the extended measurement is realized by simultaneously measuring the commuting operators $[(a_k+a_k^+)/2-(a_{ek}+a_{ek}^+)/2]$ and $[(a_k-a_k^+)/2i + (a_{ek}-a_{ek}^+)/2i]$ on the Hilbert space $\mathcal{H}_k\otimes\mathcal{H}_{ek}$ for the receiver cavity adjoined by an harmonic oscillator in the ground state $|0_e\rangle\langle 0_e|$. Then Theorem 5 applies here with the 2x2 matrix $\Gamma(k)=D(k)I_2$ and the outcome of the measurement are proportional to the outcome $Y(k)$ of the extended measurement represented by $\left((a_k+a_k^+)/2,\ (a_k-a_k^+)/2i\right)$ that is structurally independent of k: only one type of device is required - an important practical simplification. But then the estimator $\hat{X}(k)$ becomes

$$\hat{X}(k)=B(k)D(k)Y(k)+\sum_{k'=0}^{k-1}\hat{C}_{k'}(k)B(k')D(k')Y(k') \tag{43}$$

where the coefficient matrices $[\hat{C}_0(k)B(0)D(0), \ldots, B(k)D(k)]$ satisfy the normal equations for the L.M.V.E of $X(k)$ based on $Y(0)\ldots Y(k)$. However [12] the outcomes $Y(k')$ are statistically equivalent to the following fictitious observation process

$$Y(k')=X(k')+U(k'),\quad k'=0,1,\ldots k \tag{44}$$

where $U(k')$ is a white, zero mean Gaussian random vector sequence, with covariance matrix $(n_0/2+1/2)I_2$. If further the sequence $X(k')$ satisfies (41) then the optimal estimator is given through the well known Kalman Bucy filtering equations for the classical problem (41), (44):

$$\hat{X}(k)=\phi(k-1)\hat{X}(k-1)+K(k)[Y(k)-\phi(k-1)\hat{X}(k-1)] \tag{45}$$

$$\left.\begin{array}{l}\text{where } K(k)=P(k)[P(k)+(\frac{n_0}{2}+\frac{1}{2})I_2]^{-1}\\ P(k)=\phi(k-1)(P(k-1)-K(k-1)P(k-1))\phi(k-1)^t+Q(k-1)\end{array}\right\} \tag{46}$$

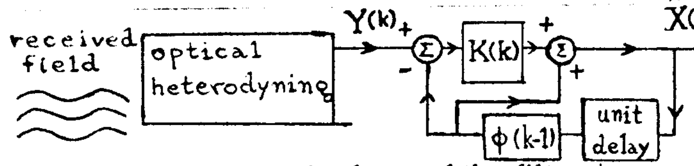


Figure 2: Illustrating the form of the filter in this example.

## References


1. A.S. Holevo, "Statistical Problems in Quantum Physics", Proceedings of 1972 Soviet-Japanese Symposium on Probability.
2. A.S. Holevo, "Statistical Decision Theory for Quantum Systems", Journal of Multivariate Analysis, Vol. 3, pp. 337-394, 1973.
3. C.W. Helstrom, "Detection Theory and Quantum Mechanics", Information & Control, Vol. 10, pp. 254-291, March 1967.
4. C.W. Helstrom, J.W. Liu and J.P. Gordon, "Quantum Mechanical Communication Theory", Proceedings of the IEEE, Vol. 58, pp. 1578-1598, October 1970.
5. Y.H. Park, "Quantum Linear Recursive Minimum Mean Square-Error Estimation", Ph.D. Thesis, Univ. of Md., August 1974.
6. J.S. Baras, R.O. Harger and A. Ephremides, "Recursive Filtering of Operator-valued Processes in Quantum Estimation", presented at 1974 Intern'l. Symposium on Information Theory, South Bend, Indiana, October 1974.
7. E. Prugovecki, Quantum Mechanics in Hilbert Space. New York: Academic Press, 1971.
8. H.P. Yuen and M. Lax, "Multiple-Parameter Quantum Estimation and Measurement of Non-selfadjoint Observables", IEEE Trans. on Info. Th., Vol. IT-19, pp. 740-749, Nov. 1973.
9. C.W. Helstrom and R.S. Kennedy, "Non-commuting Observables in Quantum Detection and Estimation Theory", IEEE Trans. on Info. Th., Vol. IT-20, pp. 16-24, Jan. 1974.
10. J.M. Jauch, Foundations of Quantum Mechanics. Reading: Addison-Wesley, 1968.
11. D.G. Luenberger, Optimization by Vector Space Methods. New York: John Wiley & Sons, 1969.
12. S.D. Personick, "Application of Quantum Estimation Theory to Analog Communication Over Quantum Channels", IEEE Trans. on Info. Th., Vol. IT-17, pp. 240-246, May 1971.
13. J.S. Baras and Y.H. Park, "Estimation of Random Signals Based on Quantum Mechanical Measurements", Proceedings of Eighth

Asilomar Conference on Circuits, Systems, and Computers, December, 1974, pp. 533-539

14. J.S. Baras, R.O. Harger and Y.H. Park, "Quantum Mechanical Linear Filtering of Random Signal Sequences", submitted for publication, December 1974.

15. S.K. Berberian, Notes on Spectral Theory. New York: D. Van Nostrand, 1966.

16. J.S. Baras and R.O. Harger, "Quantum Mechanical Linear Filtering of Vector Signal Processes", in preparation.

17. E. Hille and R.S. Phillips, Functional Analysis and Semi-groups, A.M.S. Colloquium Publication, 1957.