Appl Math Optim 18:191-214 (1988)

Applied Mathematics and Optimization

An Optimization Problem from Linear Filtering with Quantum Measurements*

John S. Baras

Electrical Engineering Department and Systems Research Center,
University of Maryland, College Park, MD 20742, USA

Communicated by A. V. Balakrishnan

Abstract. We consider the problem of optimal (in the sense of minimum error variance) linear filtering a vector discrete-time signal process, which influences a quantum mechanical field, utilizing quantum mechanical measurements. The nonclassical characteristic of the problem is the joint optimization over the measurement process and the linear signal processing scheme. The problem is formulated as an optimization problem of a functional over a set of operator-valued measures and matrices. We prove existence of optimal linear filters and provide necessary and sufficient conditions for optimality.

1. Introduction

The motivation for the development of detection and estimation theory with quantum statistics is well known [14], [15], [9], [10]. It stems primarily from the need to improve the design of optical communication systems and evaluate the performance of existing systems in this area. We have previously analyzed [3] the linear filtering problem for a scalar signal process. Here we consider the linear filtering problem for a vector signal process $\{x(i)\}$, the discrete parameter i conveniently regarded as time. Due to the well-known quantum mechanical limitation on simultaneous measurements [23, p. 260], [19, p. 101] the vector process problem is more complicated than the scalar one, and necessitates the use of *generalized quantum measurements* in the sense of Holevo [15].

* This work was partially supported by the National Science Foundation under Grant ENG 75-20900 and by U.S. ARO Contract DAAG-39-83-C-0028.

The customary formulation of quantum mechanics [23, p. 258] associates a self-adjoint operator $\mathbf{V}$ on a Hilbert space H with a measurement and incorporates *a priori* statistical information with a density operator $\boldsymbol{\rho}$ on H ($\boldsymbol{\rho}$ is a self-adjoint, positive definite operator with unit trace and represents *the state* of the quantum system [19, pp. 94 and 132]). The measurement represented by $\mathbf{V}$ produces a real number v (the outcome) whose expectation is $E\{v\} = \mathrm{Tr}[\boldsymbol{\rho}\mathbf{V}]$ (where Tr denotes the operation of taking the trace of an operator on H [23, p. 374]). This formulation is adequate for restricted estimation problems only, particularly here the estimation of a scalar. When a vector is to be estimated, the essentially quantum mechanical problem of simultaneous measurements arises and a more general concept of measurement must be resorted to [15, p. 341].

To assist in motivating the concept of a generalized quantum measurement we first elaborate slightly on the customary formulation. The spectral theorem [23, p. 249] associates with each self-adjoint operator $\mathbf{V}$ on H a unique spectral measure $\mathbf{E}_{\mathbf{V}}(\cdot)$, a mapping of the Borel sets of the real line into projection operators on H. The distribution function of the outcome v is then $F_v(\xi) = \mathrm{Tr}[\boldsymbol{\rho}\mathbf{E}_{\mathbf{V}}(-\infty, \xi]]$. The spectral theorem yields the moments $E\{v^m\} = \mathrm{Tr}[\boldsymbol{\rho}\mathbf{V}^m]$, $m = 1, 2, \ldots$. The spectral measure $\mathbf{E}_{\mathbf{V}}(\cdot)$ is fundamental and is termed a *simple measurement* [15]. Following Holevo [15, p. 341] a *generalized* measurement is a map $\mathbf{M}$ from the σ-algebra of Borel sets $\mathscr{B}^n$ of the n-dimensional space $\mathbb{R}^n$, to the algebra $\mathfrak{B}(H)$ of all bounded linear operators on H, such that:

$$\begin{aligned} &\text{(i) } \mathbf{M}(B) \geq 0 \text{ for every } B \in \mathscr{B}^n, \\ &\text{(ii) if } \{B_i\} \subseteq \mathscr{B}^n \text{ is a partition of } \mathbb{R}^n \text{ then } \textstyle\sum_i \mathbf{M}(B_i) = \mathbf{I}, \end{aligned} \tag{1.1}$$

where the series converges weakly in $\mathfrak{B}(H)$ [13, p. 53], and $\mathbf{I}$ is the identity operator on H. That is, a measurement is a *positive operator-valued measure* (p.o.m.) [7, p. 6], or a *generalized resolution of the identity* [7, p. 121]. It is worth noting that if $\mathbf{M}$ is an *orthogonal* resolution of the identity, i.e., if in addition to (1.1) we have that $B \cap C = \varnothing$ for $B, C \in \mathscr{B}^n$, implies $\mathbf{M}(B)M(C) = 0$, then $\mathbf{M}$ is necessarily a spectral measure [7, p. 12] and thus we have a simple measurement. A p.o.m. $\mathbf{M}$ induces a probability measure $\mu_{\mathbf{M}}$ on $\mathscr{B}^n$ via

$$\mu_{\mathbf{M}}(B) = \mathrm{Tr}[\boldsymbol{\rho}\mathbf{M}(B)] \qquad \text{for} \quad B \in \mathscr{B}^n, \tag{1.2}$$

as is readily verified; thus $\mathbf{M}$ is also sometimes termed a *probability* operator measure. The interpretation of this mathematical construct is that a generalized measurement $\mathbf{M}$ represents a physical measurement process with outcomes $u \in \mathbb{R}^n$, with distribution function

$$F_u(\xi) = F_u(\xi_1, \ldots, \xi_n) = \mathrm{Tr}[\boldsymbol{\rho}\mathbf{M}(-\infty, \xi]], \tag{1.3}$$

where $(-\infty, \xi] \equiv (-\infty, \xi_1] \times (-\infty, \xi_2] \times \cdots \times (-\infty, \xi_n]$.

Consider now the moment $E\{u_i\}$, the expectation of the ith component of the outcome u, of the measurement represented by the p.o.m. $\mathbf{M}$:

$$E\{u_i\} = \int_{\mathbb{R}^n} u_i F_u(du_1, \ldots, du_n) = \int_{\mathbb{R}^n} u_i \,\mathrm{Tr}[\boldsymbol{\rho}\mathbf{M}(du)], \qquad i = 1, \ldots, n.$$

Assuming the interchange is permitted—this is discussed carefully in Holevo [15, Section 6]—we have

$$E\{u_i\} = \mathrm{Tr}\left[\boldsymbol{\rho} \int_{\mathbb{R}^n} u_i \mathbf{M}(du)\right].$$

The integral is a well-defined, self-adjoint operator on H [15], [16], which we denote by $\mathbf{U}_i$. Then

$$E\{u_i\} = \mathrm{Tr}[\boldsymbol{\rho}\mathbf{U}_i], \qquad i = 1, \ldots, n. \tag{1.4}$$

Consider next the second-order moment

$$E\{u_i u_j\} = \int_{\mathbb{R}^n} u_i u_j F_u(du_1, du_2, \ldots, du_n) = \mathrm{Tr}\left[\boldsymbol{\rho} \int_{\mathbb{R}^n} u_i u_j \mathbf{M}(du)\right], \tag{1.5}$$

where again the operator integral is a well-defined, self-adjoint operator on H [15], [16], which we denote by $\mathbf{U}_{ij}$. Clearly, $\mathbf{U}_{ij} = \mathbf{U}_{ji}$ but, unlike the (special) case when $\mathbf{M}$ is a spectral measure, $\mathbf{U}_{ij} \neq \mathbf{U}_i\mathbf{U}_j$. Holevo termed the operators

$$\mathbf{U}_{i_1,\ldots,i_k} = \int_{\mathbb{R}^n} u_{i_1} \cdots u_{i_k} \mathbf{M}(du) \tag{1.6}$$

the operator moments of the p.o.m. $\mathbf{M}$. In particular, U_i are the *first operator moments* and $\mathbf{U}_{ij}$ the *second operator moments.* Observe that, in the case of a simple measurement, the p.o.m. is uniquely defined by its first operator moment.

As pointed out by Holevo [15, p. 343] this generalization of the concept of a quantum measurement is well justified in view of Naimark's theorem [1, p. 124] which asserts that, given a generalized measurement $\mathbf{M}$ in H, there exist an auxiliary Hilbert space H_e, a (pure) density operator $\boldsymbol{\rho}_e$ on $\mathfrak{B}(H_e)$, and a simple measurement $\mathbf{E_M}$ in $H \otimes H_e$ (the tensor product of Hilbert spaces H, H_e) [23, p. 144] such that

$$\mathrm{Tr}[\boldsymbol{\rho}\mathbf{M}(B)] = \mathrm{Tr}[(\boldsymbol{\rho} \otimes \boldsymbol{\rho}_e)\mathbf{E_M}(B)] \tag{1.7}$$

for every $B \in \mathscr{B}^n$ and every density operator $\boldsymbol{\rho}$ on H. That is, the distribution functions of the measurement outcomes induced by the generalized measurement $\mathbf{M}$ and the simple measurement $\mathbf{E_M}$ are the same. The physical interpretation is [15], [11] that a generalized quantum measurement is realized by the measurement of compatible observables (i.e., a simple measurement) on a composite quantum system produced by adjoining to the original system, characterized by $(\boldsymbol{\rho}, \mathbf{H})$, an auxiliary system, characterized by $(\boldsymbol{\rho}_e, \mathbf{H}_e)$. Thus justifiably the triple $(H_e, \boldsymbol{\rho}_e, \mathbf{E_M})$ is called a *realization* of the measurement represented by the p.o.m. $\mathbf{M}$. For simplicity we shall refer to generalized quantum measurements as quantum measurements for the rest of this paper, unless explicitly stated otherwise.

Let x be a vector random variable, with a distribution function F_x, on which the density operator $\boldsymbol{\rho}$ depends; i.e., $\boldsymbol{\rho} = \boldsymbol{\rho}(x)$. Then the distribution function $F_u(\xi)$ (1.3) of the vector outcome u of the generalized quantum measurement $\mathbf{M}$ becomes a conditional distribution function $F_{u|x}(\xi, \zeta)$. The first moments of u are

$$E\{u_i\} = E\{E\{u_i|x\}\} = \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} u_i \,\mathrm{Tr}[\boldsymbol{\rho}(\xi)\mathbf{M}(du)]\right] F_x(d\xi).$$

Using the results of Holevo [15] we can justify interchanging the order of $\mathbf{M}(du)$ integration and trace to obtain

$$E\{u_i\} = \int_{\mathbb{R}^n} \operatorname{Tr}[\boldsymbol{\rho}(\xi)\mathbf{U}_i] F_x(d\xi).$$

Similarly,

$$E\{u_i u_j\} = \int_{\mathbb{R}^n} \operatorname{Tr}[\boldsymbol{\rho}(\xi)\mathbf{U}_{ij}] F_x(d\xi).$$

The problem of interest to us is described briefly below; for further details we refer to [3] and [4]. We consider a quantum system characterized by a density operator $\boldsymbol{\rho}(x(k))$ on a Hilbert space H, which is influenced in some fashion by a vector stochastic process $x(k)$, $k=0,1,\ldots$, with values in $\mathbb{R}^n$. At each time i, $i=0,1,\ldots$, a measurement represented by the p.o.m. $\mathbf{M}_i$ is made with outcome $v(i)\in\mathbb{R}^n$. At time k we have available the outcomes $v(i)$, $i=0,\ldots,k-1$, of the measurements represented by the p.o.m.s $\mathbf{M}_i$, performed at times $t=0,\ldots,k-1$. A new measurement, represented by $\mathbf{M}_k$, is to be performed at time $t=k$ and the present and past measurement outcomes are to be processed linearly to give the estimator

$$\hat{x}(k) = \sum_{i=0}^{k} C_i(k) v(i), \tag{1.8}$$

where $C_i(k)$, $i=0,1,\ldots,k$, are $n\times n$ matrices. The problem is to find a p.o.m. $\hat{\mathbf{M}}_k$ and matrices $\hat{C}_i(k)$, $i=0,\ldots,k$, to minimize the mean square error

$$\mathcal{J}(\mathbf{C}(k),\mathbf{M}_k) = E\{(x(k)-\hat{x}(k))'(x(k)-\hat{x}(k))\}, \tag{1.9}$$

where

$$\mathbf{C}(k) = [C_0(k), C_1(k), \ldots, C_k(k)]. \tag{1.10}$$

The expectation in (1.9) is taken with respect to the distributions of $x(i)$ and the measurement outcomes' distributions which are described below. We assume that the measurement outcomes $\{v(i)\}$ (classical vector random variables) are independent, conditioned upon the sequence $\{x(i)\}$. This assumption, together with (1.3), yields the following expression for the joint distribution of the measurement outcomes $v(0), v(1), \ldots, v(k)$:

$$\begin{aligned} &F_{v(0),\ldots,v(k)}(\nu(0),\ldots,\nu(k)) \\ &\quad = \int_{\mathbb{R}^n}\cdots\int_{\mathbb{R}^n} F_{v(0)|x(0)}(\nu(0),\xi(0))\cdots F_{v(k)|x(k)}(\nu(k),\xi(k)) \\ &\qquad \times F_{x(0),\ldots,x(k)}(d\xi(0),\ldots,d\xi(k)), \end{aligned} \tag{1.11}$$

where

$$F_{v(i)|x(i)}(\nu(i),\xi(i)) = \operatorname{Tr}[\boldsymbol{\rho}(\xi(i))\mathbf{M}_i(-\infty,\nu(i)]]. \tag{1.12}$$

A convenient example for physical motivation is provided by the following optical communication setting [9], [10], [3]: at time k a laser field 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 $\boldsymbol{\rho}(x(k))$ that depends on $x(k)$ (but not otherwise on k). The filtering problem consists of estimating $x(k)$ based on quantum mechanical measurements via the procedure described above. In this example the conditional independence assumption corresponds to "clearing" the receiver cavity prior to each reception [3].

2. Formulation of the Optimization Problem

In this section we formulate the optimal linear filtering problem described above as an optimization problem of a functional over a set of p.o.m.s and $n \times n$ matrices. Before proceeding in this direction we present a brief summary of the mathematical concepts and techniques necessary for the development. Much of the mathematical machinery used here has been developed by Holevo [15, Section 6], where we refer the reader for details. Our formulation utilizes the integration theory with respect to operator-valued measures developed by Holevo in [15]. This integration theory is more akin to Riemann integration. Since the original submission of the present paper, Mitter and Young [21], have developed a different, more satisfactory integration theory which is analogous to Lebesgue integration.

The latter permits formulation of optimization problems arising in quantum detection and estimation theory as convex analysis problems in certain Banach spaces of operator-valued functions. In addition, it facilitates the development of a duality theory for such problems. Due to the special nature of our problem we do not need Holevo's techniques in their full generality. The problem treated here is more general than the estimation problems treated by Holevo [15] or Mitter and Young [21], in the sense that we are studying recursive estimation. Furthermore, the techniques of Mitter and Young [21] must be extended in order to apply to the problem treated here, because our cost function does not satisfy their assumptions. Nevertheless, their work provides an alternate route, and perhaps more significantly reinforces the validity of the results obtained here.

Following [15] let $\mathfrak{T}_h$ denote the set of all trace-class [23, p. 374], self-adjoint operators on a Hilbert space H. Let $\mathbf{M}$ be a p.o.m., representing a measurement with values in $\mathbb{R}^n$ and $\mathbf{G}(u)$ a $\mathfrak{T}_h$-valued function, $u \in \mathbb{R}^n$. Let B be a bounded set in $\mathscr{B}^n$ and $\{p_n\}$ a sequence of finite partitions $p_n = \{B_{ni}\}$, $B_{ni} \in \mathscr{B}^n$ of B, and d_n the mesh of the partition p_n such that $\lim_{n\to\infty} d_n = 0$. We then introduce the integral sums

$$\boldsymbol{\sigma}_n = \sum_i \mathbf{G}(u_i)\mathbf{M}(B_{ni}),$$

where $u_i \in B_{ni}$. Clearly, $\boldsymbol{\sigma}_n \in \mathfrak{T}$, the space of trace class operators on H which is equipped with the norm

$$\|\mathbf{A}\|_{\mathfrak{T}} \triangleq \mathrm{Tr}[(\mathbf{A}^*\mathbf{A})^{1/2}]$$

(where the asterisk denotes adjoint). If the sequence $\boldsymbol{\sigma}_n$ converges in the norm of $\mathfrak{T}$ as $n \to \infty$ and the limit does not depend on the choice of $u_i \in B_{ni}$, then $\mathbf{G}$ is called *left integrable with respect to $\mathbf{M}$ over B*, and the limit is called the *left integral* and is denoted by

$$\int_B \mathbf{G}(u)\mathbf{M}(du). \tag{2.1}$$

We define similarly right integrals, and it is useful to note that right integrability is equivalent to left integrability. $\mathbf{G}$ is called *trace integrable with respect to $\mathbf{M}$* if the sequence $\operatorname{Tr} \boldsymbol{\sigma}_n$ converges. The limit is called the *trace integral of $\mathbf{G}$ with respect to $\mathbf{M}$ over B* and is denoted by

$$\lim_{n \to \infty} \operatorname{Tr} \boldsymbol{\sigma}_n = \langle \mathbf{G}, \mathbf{M} \rangle_B. \tag{2.2}$$

In all cases described in this paper $\mathbf{G}$ will be of the form

$$\mathbf{G}(u) = \sum_{i=1}^{l} \boldsymbol{\varkappa}_i g_i(u), \tag{2.3}$$

where $\boldsymbol{\varkappa}_i \in \mathfrak{T}_h$ and g_i are continuous real-valued functions on $\mathbb{R}^n$. Then for such $\mathbf{G}$ the trace integral with respect to any p.o.m. $\mathbf{M}$ and over any bounded set $B \in \mathscr{B}^n$ exists [15, p. 356] and equals

$$\langle \mathbf{G}, \mathbf{M} \rangle_B = \sum_{i=1}^{l} \int_B g_i(u) \operatorname{Tr}[\boldsymbol{\varkappa}_i \mathbf{M}(du)]. \tag{2.4}$$

The operator-valued function $\mathbf{G}$ is *locally trace integrable* if it is trace integrable over any bounded $B \in \mathscr{B}^n$. Let B_i be a nondecreasing sequence of subsets in $\mathscr{B}^n$ and such that $\bigcup_i B_i = \mathbb{R}^n$. Then one writes $B_i \uparrow \mathbb{R}^n$. If the limit

$$\lim_{B_i \uparrow \mathbb{R}^n} \int_{B_i} \mathbf{G}(u)\mathbf{M}(du)$$

exists in $\mathfrak{T}$, and does not depend on the choice of the sequence $\{B_i\}$, $\mathbf{G}$ is called *integrable over* $\mathbb{R}^n$ and the limit is denoted by

$$\int_{\mathbb{R}^n} \mathbf{G}(u)\mathbf{M}(du). \tag{2.5}$$

Let $\mathbf{G}$ be locally trace integrable with respect to $\mathbf{M}$. Then $\mathbf{G}$ is *trace integrable* if the limit

$$\lim_{B_i \uparrow \mathbb{R}^n} \langle \mathbf{G}, \mathbf{M} \rangle_{B_i}$$

exists, finite or infinite. The limit is denoted by

$$\langle \mathbf{G}, \mathbf{M} \rangle_{\mathbb{R}^n}$$

and is called the *trace integral of $\mathbf{G}$ with respect to $\mathbf{M}$ over* $\mathbb{R}^n$. Notice that if $\mathbf{G}$ is left or right integrable with respect to $\mathbf{M}$ over $\mathbb{R}^n$ then

$$\langle \mathbf{G}, \mathbf{M} \rangle_{\mathbb{R}^n} = \operatorname{Tr} \int_{\mathbb{R}^n} \mathbf{G}(u)\mathbf{M}(du) = \operatorname{Tr} \int \mathbf{M}(du)\mathbf{G}(u). \tag{2.6}$$

When $\mu(du)$ is a measure on $\mathbb{R}^l$ and $\mathbf{G}$ a $\mathfrak{T}$-valued function we say that $\mathbf{G}$ is *Bochner integrable* (in the $\mathfrak{T}$-norm sense) [13, p. 78] with respect to μ if the scalar function $\|\mathbf{G}\|_{\mathfrak{T}}$ is integrable with respect to μ, i.e., when

$$\int_{\mathbb{R}^l} \|\mathbf{G}(u)\|_{\mathfrak{T}}\mu(du) < \infty.$$

Here $\mathfrak{T}$ and $\mathfrak{T}_h$ refer to spaces of operators on H.

Before proceeding with the actual computation of the functional $\mathcal{J}(\mathbf{C}(k), \mathbf{M}_k)$ in (1.9) we want to make two remarks. First let us observe that we can set $C_k(k) = I_n$ (i.e., the identity matrix on $\mathbb{R}^n$) without loss of generality. Indeed, consider any pair of p.o.m. $\mathbf{X}$ and $n \times n$ matrix C, and let $v \in \mathbb{R}^n$ be the outcome of the measurement represented by $\mathbf{X}$. Let $c(x) = Cx$ be a linear map from $\mathbb{R}^n$ into $\mathbb{R}^n$ and define for every $A \in \mathcal{B}^n$

$$\mathbf{X}'(A) = \mathbf{X}(c^{-1}(A)). \tag{2.7}$$

It is easy to verify that $\mathbf{X}'$ is a p.o.m. which represents the measurement with outcome Cv. So for the rest of the paper we shall take $C_k(k) = I_n$. Second, a direct consequence here of $\mathbf{U}_{ij} \neq \mathbf{U}_i\mathbf{U}_j$ (cf. (1.4)–(1.6)) is that the mean square error will not be expressible directly in terms of self-adjoint operators in a quadratic form (as in the scalar filtering problem [3]) but rather will remain expressed in terms of p.o.m.s. Since the set of self-adjoint operators on H is a linear space while the set of p.o.m.s is only a convex set, the nature of the optimization problem will be different.

Since we are dealing with a second-order problem it is customary to assume that each $x(i)$ and the past measurement outcomes $v(0), v(1), \ldots, v(k-1)$ have finite second moments. Then since

$$\int_{\mathbb{R}^n} \|\xi'\xi\boldsymbol{\rho}(\xi)\|_{\mathfrak{T}} F_{x(k)}(d\xi) \leq \int_{\mathbb{R}^n} |\xi'\xi| F_{x(k)}(d\xi) < \infty$$

the Bochner integral

$$\boldsymbol{\lambda}(k) = \int_{\mathbb{R}^n} \xi'\xi\boldsymbol{\rho}(\xi) F_{x(k)}(d\xi) \tag{2.8}$$

exists and is a nonnegative operator in $\mathfrak{T}_h$. Similarly, since

$$\int_{\mathbb{R}^n} \|\xi_i\boldsymbol{\rho}(\xi)\|_{\mathfrak{T}} F_{x(k)}(d\xi) \leq \int_{\mathbb{R}^n} |\xi_i| F_{x(k)}(d\xi) < \infty,$$

the Bochner integral

$$\boldsymbol{\delta}_i(k) = \int_{\mathbb{R}^n} \xi_i\boldsymbol{\rho}(\xi) F_{x(k)}(d\xi), \qquad i = 1, 2, \ldots, n, \tag{2.9}$$

exists and is an operator in $\mathfrak{T}_h$. We introduce the following n-vector of operators:

$$\boldsymbol{\delta}(k) = \begin{bmatrix} \boldsymbol{\delta}(k)_1 \\ \boldsymbol{\delta}(k)_2 \\ \vdots \\ \boldsymbol{\delta}(k)_n \end{bmatrix}. \tag{2.10}$$

Clearly, since

$$\int_{\mathbb{R}^n} \|\boldsymbol{\rho}(\xi)\|_{\mathfrak{T}} F_{x(k)}(d\xi) = 1,$$

the Bochner integral

$$\boldsymbol{\eta}(k) = \int_{\mathbb{R}^n} \boldsymbol{\rho}(\xi) F_{x(k)}(d\xi) \tag{2.11}$$

exists and is a nonnegative operator in $\mathfrak{T}_h$. Since

$$\int_{\mathbb{R}^n} \cdots \int_{\mathbb{R}^n} \|E\{v_l(i)|x(i) = \zeta\}\boldsymbol{\rho}(\xi)\|_{\mathfrak{T}} F_{x(i),x(k)}(d\zeta, d\xi)$$
$$\leq \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} E\{v_l(i)|x(i) = \zeta\} F_{x(i)}(d\zeta) \leq E\{v_l(i)\} < \infty,$$

the Bochner integral

$$\boldsymbol{\gamma}_l(k, i) = \int_{\mathbb{R}^n} \cdots \int_{\mathbb{R}^n} E\{v_l(i)|x(i) = \zeta\}\boldsymbol{\rho}(\xi) F_{x(i),x(k)}(d\zeta, d\xi),$$
$$l = 1, 2, \ldots, n, \tag{2.12}$$

exists and is an operator in $\mathfrak{T}_h$. We introduce the n-vector of operators

$$\boldsymbol{\gamma}(k, i) = \begin{bmatrix} \boldsymbol{\gamma}_1(k, i) \\ \boldsymbol{\gamma}_2(k, i) \\ \vdots \\ \boldsymbol{\gamma}_n(k, i) \end{bmatrix}. \tag{2.12a}$$

Finally, since

$$\int_{\mathbb{R}^n} \cdots \int_{\mathbb{R}^n} \|\xi_l E\{v_j(i)|x(i) = \zeta\}\boldsymbol{\rho}(\xi)\|_{\mathfrak{T}} F_{x(i),x(k)}(d\zeta, d\xi)$$
$$\leq \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |\xi_l E\{v_j(i)|x(i) = \zeta\}| F_{x(i),x(k)}(d\zeta, d\xi) \leq E\{|x_l(k) v_j(i)|\}$$
$$\leq (E\{|x_l(k)|^2\})^{1/2} (E\{|v_j(i)|^2\})^{1/2} < \infty,$$

the Bochner integral

$$\boldsymbol{\pi}_{l,j}(k, i) = \int_{\mathbb{R}^n} \cdots \int_{\mathbb{R}^n} \xi_l E\{v_j(i)|x(i) = \zeta\}\boldsymbol{\rho}(\xi) F_{x(i),x(k)}(d\zeta, d\xi),$$
$$l = 1, \ldots, n, \quad j = 1, \ldots, n, \tag{2.13}$$

exists and is an operator in $\mathfrak{T}_h$. We introduce the $n \times n$ matrix of operators:

$$\boldsymbol{\pi}(k, i) = \begin{bmatrix} \boldsymbol{\pi}_{1,1}(k, i) & \cdots & \boldsymbol{\pi}_{1,n}(k, i) \\ \vdots & \ddots & \vdots \\ \boldsymbol{\pi}_{n,1}(k, i) & \cdots & \boldsymbol{\pi}_{n,n}(k, i) \end{bmatrix}. \tag{2.14}$$

From (1.9) the mean-square error can be rewritten as

$$\begin{aligned}\mathcal{J}(\mathbf{C}(k), \mathbf{M}_k) &= E\{(x(k)-v(k))'(x(k)-v(k))\} \\ &\quad -2E\left\{(x(k)-v(k))'\left(\sum_{j=0}^{k-1} C_j(k)v(j)\right)\right\} \\ &\quad +E\left\{\left(\sum_{i=0}^{k-1} C_i(k)v(i)\right)'\left(\sum_{j=0}^{k-1} C_j(k)v(j)\right)\right\}. \end{aligned} \tag{2.15}$$

The first two terms now become

$$\begin{aligned}\int_{\mathbb{R}^n}\cdots\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\Bigg[&\xi(k)'\xi(k)-2u'\xi(k)+u'u \\ &-2\xi(k)'\sum_{i=1}^{k-1} C_i(k)E\{v(i)|x(i)=\xi(i)\} \\ &+2\sum_{i=0}^{k-1} u'C_i(k)E\{v(i)|x(i)=\xi(i)\}\Bigg] \\ \times \operatorname{Tr}[\boldsymbol{\rho}(\xi(k))\mathbf{M}_k(du)]&F_{x(0),\ldots,x(k)}(d\xi(0),\ldots,d\xi(k)). \end{aligned} \tag{2.16}$$

Proceeding in a fashion similar to that of Holevo [15] we let $W(u, \xi(0), \ldots, \xi(k), \mathbf{C}(k))$ denote the integrand in (2.16), where $\mathbf{C}(k)$ is (as usual) the matrix

$$\mathbf{C}(k) = [C_0(k), C_1(k), \ldots, I_n].$$

In view of (2.8), (2.9), (2.11), (2.12), and (2.13) the Bochner integral

$$\begin{aligned}\mathbf{G}(u, \mathbf{C}(k)) = \int_{\mathbb{R}^n}\cdots\int_{\mathbb{R}^n} & W(u, \xi(0), \ldots, \xi(k), \mathbf{C}(k)) \\ &\times \boldsymbol{\rho}(\xi(k))F_{x(0),\ldots,x(k)}(d\xi(0),\ldots,d\xi(k)) \end{aligned} \tag{2.17}$$

exists and is an operator-valued function with values in $\mathfrak{T}_h$. Utilizing the notation introduced in (2.8)–(2.14), we have

$$\begin{aligned}\mathbf{G}(u, \mathbf{C}(k)) &= \boldsymbol{\lambda}(k) - 2\sum_{i=1}^{n} u_i \boldsymbol{\delta}(k)_i + u'u\boldsymbol{\eta}(k) \\ &\quad -2\sum_{i=0}^{k-1}\sum_{l=1}^{n}\sum_{j=1}^{n}[C_i(k)]_{l,j}\boldsymbol{\pi}_{l,j}(k,i) \\ &\quad +2\sum_{i=0}^{k-1}\sum_{l=1}^{n}[C_i(k)'u]_l\boldsymbol{\gamma}_l(k,i), \end{aligned} \tag{2.18}$$

where the notation $[\ ]_i$, $[\ ]_{l,j}$ indicate the ith element of a vector and the l, j element of a matrix, respectively.

To simplify this expression for $\mathbf{G}(u, \mathbf{C}(k))$ we introduce the following notation:

(a) For $a \in \mathbb{R}^n$ and $\boldsymbol{\beta}$ an n-vector of operators,

$$a'\boldsymbol{\beta} = \boldsymbol{\beta}'a = \sum_{i=1}^{n} a_i \boldsymbol{\beta}_i. \tag{2.19}$$

(b) For A an $n \times n$ matrix and $\boldsymbol{\sigma}$ an $n \times n$ matrix of operators,

$$\operatorname{tr} A\boldsymbol{\sigma} = \operatorname{tr} \boldsymbol{\sigma} A = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} \boldsymbol{\sigma}_{ji}. \tag{2.20}$$

As a result we can rewrite (2.18) as

$$\mathbf{G}(u, \mathbf{C}(k)) = \boldsymbol{\lambda}(k) - 2u'\boldsymbol{\delta}(k) + u'u\boldsymbol{\eta}(k) + 2 \sum_{i=0}^{k-1} u'C_i(k)\boldsymbol{\gamma}(k, i) - 2 \sum_{i=0}^{k-1} \operatorname{tr} C_i(k)'\boldsymbol{\pi}(k, i). \tag{2.21}$$

It is immediately seen from (2.21) that for any $\mathbf{C}(k)$, $\mathbf{G}(u, \mathbf{C}(k))$ is a $\mathfrak{T}_h$-valued function of the form described in (2.3). Therefore from (2.4) $\mathbf{G}(u, \mathbf{C}(k))$ is locally trace integrable with respect to any p.o.m. $\mathbf{X}$ on H. Let now B be a bounded set in $\mathscr{B}^n$. Then since $\mathbf{G}(\cdot, \mathbf{C}(k))$ is in the class (2.3), its trace integral with respect to any p.o.m. $\mathbf{X}$ on H over B is given, according to (2.4) and (2.21), by

$$\langle \mathbf{G}(\cdot, \mathbf{C}(k)), \mathbf{X} \rangle_B = \int_{\mathbb{R}^n} \cdots \int_{\mathbb{R}^n} \int_B W(u, \xi(0), \ldots, \xi(k), \mathbf{C}(k)) \times \operatorname{Tr}[\boldsymbol{\rho}(\xi(k))\mathbf{X}(du)] F_{x(0),\ldots,x(k)}(d\xi(0), \ldots, d\xi(k)). \tag{2.22}$$

That is so because we can interchange integrals in the right-hand side of (2.22) by Fubini's theorem [24, p. 140] in view of the finite second moments assumptions made above, and because

$$\operatorname{Tr}[\boldsymbol{\lambda}(k)\mathbf{X}(du)] = \int_{\mathbb{R}^n} \xi'\xi \operatorname{Tr}[\boldsymbol{\rho}(\xi)\mathbf{X}(du)] F_{x(k)}(d\xi),$$

$$\operatorname{Tr}[\boldsymbol{\delta}(k)_i\mathbf{X}(du)] = \int_{\mathbb{R}^n} \xi_i \operatorname{Tr}[\boldsymbol{\rho}(\xi)\mathbf{X}(du)] F_{x(k)}(d\xi),$$

$$\operatorname{Tr}[\boldsymbol{\eta}(k)\mathbf{X}(du)] = \int_{\mathbb{R}^n} \operatorname{Tr}[\boldsymbol{\rho}(\xi)\mathbf{X}(du)] F_{x(k)}(d\xi),$$

$$\operatorname{Tr}[\boldsymbol{\gamma}_l(k, i)\mathbf{X}(du)] = \int_{\mathbb{R}^n} \cdots \int_{\mathbb{R}^n} E\{v_l(i) | x(i) = \zeta\} \operatorname{Tr}[\boldsymbol{\rho}(\xi)\mathbf{X}(du)] \cdot F_{x(i),x(k)}(d\zeta, d\xi),$$

$$\operatorname{Tr}[\boldsymbol{\pi}_{l,j}(k, i)\mathbf{X}(du)] = \int_{\mathbb{R}^n} \cdots \int_{\mathbb{R}^n} \xi_l E\{v_j(i) | x(i) = \zeta\} \operatorname{Tr}[\boldsymbol{\rho}(\xi)\mathbf{X}(du)] \cdot F_{x(i),x(k)}(d\zeta, d\xi).$$

Since the outcomes $v(0), \ldots, v(k-1)$ have finite second moments

$$\int_{\mathbb{R}^n} \cdots \int_{\mathbb{R}^n} \left\| E\left\{ \left(\sum_{i=0}^{k-1} C_i(k) v(i) \right)' \left(\sum_{j=0}^{k-1} C_j(k) v(j) \right) \middle| x(0) = \xi(0), \ldots, \right.\right.$$
$$\left.\left. x(k-1) = \xi(k-1) \right\} \boldsymbol{\rho}(\xi(k)) \right\| F_{x(0),\ldots,x(k)}(d\xi(0), \ldots, d\xi(k))$$
$$\leq E\left\{ \left(\sum_{i=0}^{k-1} C_i(k) v(i) \right)' \left(\sum_{j=0}^{k-1} C_j(k) v(j) \right) \right\} < \infty$$

and therefore the Bochner integral

$$\boldsymbol{\zeta}(\mathbf{C}(k)) = \int_{\mathbb{R}^n} \cdots \int_{\mathbb{R}^n} E\left\{ \left(\sum_{i=0}^{k-1} C_i(k) v(i) \right)' \left(\sum_{i=0}^{k-1} C_j(k) v(j) \right) \middle| x(0) = \xi(0), \ldots, \right.$$
$$\left. x(k-1) = \xi(k-1) \right\}$$
$$\times \boldsymbol{\rho}(\xi(k)) F_{x(0),\ldots,x(k)}(d\xi(0), \ldots, d\xi(k)) \tag{2.23}$$

exists and defines a nonnegative operator in $\mathfrak{T}_h$. For any p.o.m. $\mathbf{X}(du)$ on H we obviously have that

$$\langle \boldsymbol{\zeta}(\mathbf{C}(k)), \mathbf{X} \rangle_{\mathbb{R}^n} = \int_{\mathbb{R}^n} \mathrm{Tr}[\boldsymbol{\zeta}(\mathbf{C}(k)) \mathbf{X}(du)]$$
$$= E\left\{ \left(\sum_{i=0}^{k-1} C_i(k) v(i) \right)' \left(\sum_{j=0}^{k-1} C_j(k) v(j) \right) \right\} < \infty. \tag{2.24}$$

We define now a new $\mathfrak{T}_h$-valued function

$$\mathfrak{F}(u, \mathbf{C}(k)) = \mathbf{G}(u, \mathbf{C}(k)) + \boldsymbol{\zeta}(\mathbf{C}(k)). \tag{2.25}$$

It then follows directly from the local trace integrability of $\mathbf{G}(\cdot, \mathbf{C}(k))$ and (2.24) that $\mathfrak{F}(\cdot, \mathbf{C}(k))$ is locally trace integrable with respect to any p.o.m. on H (cf. Proposition 6.1(1) in [15]). Utilizing (2.18), (2.8)-(2.14), and (2.23) we have

$$\mathfrak{F}(u, \mathbf{C}(k))$$
$$= \int_{\mathbb{R}^n} \cdots \int_{\mathbb{R}^n} \Bigg[\xi(k)'\xi(k) - 2u'\xi(k) + u'u$$
$$+ 2 \sum_{i=0}^{k-1} u' C_i(k) E\{v(i) | x(i) = \xi(i)\}$$
$$- 2 \sum_{i=0}^{k-1} \xi(k)' C_i(k) E\{v(i) | x(i) = \xi(i)\}$$
$$+ E\left\{ \left(\sum_{i=0}^{k-1} C_i(k) v(i) \right)' \left(\sum_{j=0}^{k} C_j(k) v(j) \right) \middle| x(0) = \xi(0), \ldots, \right.$$
$$\left. x(k-1) = \xi(k-1) \right\} \Bigg]$$
$$\times \boldsymbol{\rho}(\xi(k)) F_{x(0),\ldots,x(k)}(d\xi(0), \ldots, d\xi(k))$$

$$= \int_{\mathbb{R}^n} \cdots \int_{\mathbb{R}^n} E\left\{\left(\xi(k) - u - \sum_{i=0}^{k-1} C_i(k)v(i)\right)' \times \left(\xi(k) - u - \sum_{j=0}^{k-1} C_j(k)v(j)\right) \middle| x(0) = \xi(0), \ldots, x(k-1) = \xi(k-1)\right\}$$
$$\times \boldsymbol{\rho}(\xi(k)) F_{x(0),\ldots,x(k)}(d\xi(0), \ldots, d\xi(k)). \tag{2.26}$$

Therefore $\mathfrak{F}(\cdot, \mathbf{C}(k))$ is a nonnegative $\mathfrak{T}_h$-valued function, and since it is locally trace integrable it follows from Proposition 6.1(2) of [15] that its trace integral with respect to any p.o.m. $\mathbf{X}$ on H over $\mathbb{R}^n$ is well defined. Furthermore, we see immediately from (2.15), (2.22), and (2.24) that

$$\mathcal{J}(\mathbf{C}(k), \mathbf{M}_k) = \langle \mathfrak{F}(\cdot, \mathbf{C}(k)), \mathbf{M}_k \rangle_{\mathbb{R}^n}. \tag{2.27}$$

We summarize the above in the following.

Lemma 2.1. *If the signal sequence $\{x(i)\}$ and the past measurement outcomes at times $0, 1, \ldots, k-1$ have finite second moments, then the mean-square error may be expressed as in (2.27) above. For each $n \times (k+1)n$ matrix $\mathbf{C}(k)$, and $u \in \mathbb{R}^n$, $\mathfrak{F}(u, \mathbf{C}(k))$ is a nonnegative, self-adjoint operator, with finite trace on H.*

If we now let $\mathfrak{M}$ be the convex set of p.o.m.s on H, we see that the linear filtering problem becomes: find a p.o.m. $\hat{\mathbf{M}}_k$ and $n \times n$ matrices $\hat{C}_i(k)$, $i = 0, \ldots, k-1$, which minimize (2.27) over the set $\mathfrak{M} \times (\mathbb{R}^{n \times n})^k$. This is the optimization problem we analyze in this paper. We would also like to point out that all operators appearing in (2.18) or (2.25) (i.e., in $\mathbf{G}$ or $\mathfrak{F}$) are known at the time instant k and are computable from the signal statistics and the known p.o.m. $\mathbf{M}_0, \ldots, \mathbf{M}_{k-1}$.

The optimization problem (2.27) is more general than the one considered by Holevo [15] and Mitter and Young [21], in that we must jointly optimize over a set of parameters and p.o.m.s.

3. Existence of Optimal Linear Filters

In this section we establish the existence of solutions to the optimization problem formulated in the previous section. We need some preparations first. Let $\mathfrak{M}$ be the set of p.o.m.s on H. This is a convex set. Following Holevo [15, p. 363] we have a convergence notion on $\mathfrak{M}$. A sequence of p.o.m.s $\mathbf{X}_n$ on H *converges*, if for any $\varphi \in H$, the sequence of scalar probability measures $\mu_\varphi^n(du) = \langle \varphi, \mathbf{X}_n(du)\varphi \rangle_H$ converges weakly [26, p. 19], i.e., the limit

$$\lim_{n\to\infty} \int_{\mathbb{R}^n} g(u)\mu_\varphi^n(du)$$

exists for any bounded continuous function g. If this limit equals $\int_{\mathbb{R}^n} g(u)\mu_\varphi(du)$, where $\mu_\varphi(du) = \langle \varphi, \mathbf{X}(du)\varphi \rangle_H$ for some p.o.m. $\mathbf{X}$ on H, we shall say that $\mathbf{X}_n$

converges to $\mathbf{X}$. Let $\mathfrak{M}^2$ be the set of p.o.m. on H which represent measurements with outcomes $v(k)$ having finite second moments. The set $\mathfrak{M}^2$ is convex. This follows from the fact that $M \in \mathfrak{M}^2$ if and only if

$$\int_{\mathbb{R}^n} \cdots \int_{\mathbb{R}^n} u'u \operatorname{Tr}[\boldsymbol{\rho}(\xi)\mathbf{M}(du)] F_{x(k)}(d\xi) < \infty \tag{3.1}$$

which is closed under convex combinations. Clearly, $\mathcal{J}(\mathbf{C}(k), \mathbf{M}_k)$ is a nonnegative functional on $\mathfrak{M} \times (\mathbb{R}^{n\times n})^k$, where $\mathbb{R}^{n\times n}$ is the set of all $n \times n$ matrices over the reals. We endow $\mathbb{R}^{n\times n}$ with the standard Euclidean topology induced by the norm

$$\|C\|_{\mathbb{R}^{n\times n}} = (\operatorname{tr}[C'C])^{1/2}$$

and $\mathfrak{M}$ with the topology induced by the convergence notion described above. From the discussion in [15, p. 363] we have that $\mathfrak{M}$ with this topology is *sequentially complete*; that is, every convergent sequence of p.o.m. converges to a p.o.m. in $\mathfrak{M}$. We give $\mathfrak{M} \times (\mathbb{R}^{n\times n})^k$ the product topology and it also becomes sequentially complete.

Recall [20, p. 40] that a real-valued function f on a topological space N is *lower semicontinuous* at $y_0 \in N$ if $\lim_{n\to\infty} \inf f(y_n) \geq f(y_0)$ for any sequence $\{y_n\}$ converging to y_0. It is lower semicontinuous on a subset S if it is lower semicontinuous at every point of S.

Lemma 3.1.

(a) *For each* $\mathbf{M}_k \in \mathfrak{M}$, $\mathcal{J}(\mathbf{C}(k), \mathbf{M}_k)$ *is continuous on* $(\mathscr{B}^{n\times n})^k$.
(b) *For each* $\mathbf{C}(k) \in (\mathbb{R}^{n\times n})^k$, $\mathcal{J}(\mathbf{C}(k), \mathbf{M}_k)$ *is lower semicontinuous on* $\mathfrak{M}$.

Proof. We are certainly interested only in those p.o.m. $\mathbf{M}_k$ which give finite mean-square error. Part (a) follows immediately from (2.15). As for part (b) it follows directly from the nonnegativity of $\mathfrak{F}(\cdot, \mathbf{C}(k))$ for each $\mathbf{C}(k)$, (2.27), and Lemma 7.5 in [15]. □

Lemma 3.2. $\mathcal{J}(\mathbf{C}(k), \mathbf{M}_k)$ *is a lower semicontinuous functional on* $\mathfrak{M} \times (\mathbb{R}^{n\times n})^k$.

Proof. Immediate from Lemma 3.1. □

Let c be a positive number such that

$$\inf_{\mathbf{M}_k \in \mathfrak{M}} \mathcal{J}(\mathbf{C}(k), \mathbf{M}_k) \leq c, \qquad \mathbf{C}(k) \in (\mathbb{R}^{n\times n})^k.$$

We are obviously only interested in the case where $c < \infty$. Let

$$\mathscr{A}_c = \{\mathbf{M}_k \in \mathfrak{M}, \mathbf{C}(k) \in (\mathbb{R}^{n\times n})^k \mid \mathcal{J}(\mathbf{C}(k), \mathbf{M}_k) \leq c\}. \tag{3.2}$$

Recall that a subset S of a topological space N is *conditionally compact* if every sequence in S has a convergent subsequence [22], [13, p. 5].

Lemma 3.3. *The set $\mathscr{A}_c \subseteq \mathfrak{M} \times (\mathbb{R}^{n\times n})^k$ is conditionally compact.*

Proof. Since $\mathscr{J}(\mathbf{C}(k), \mathbf{M}_k)$ is nonnegative and continuous in $\mathbf{C}(k)$, for each $\mathbf{M}_k$, the set

$$\mathscr{C}^c_{\mathbf{M}_k} = \{\mathbf{C}(k) \in (\mathbb{R}^{n\times n})^k | \mathscr{J}(\mathbf{C}(k), \mathbf{M}_k) \leq c\} \tag{3.3}$$

is compact. Since for any real α, β

$$\begin{aligned}
&\left\| x(k) - \sum_{i=0}^{k} C_i(k)v(i)\right\|^2_{\mathbb{R}^n} \\
&\quad \geq (1-\alpha^2)\|x(k)-v(k)\|^2_{\mathbb{R}^n} + (1-\alpha^{-2})\left\|\sum_{i=0}^{k-1} C_i(k)v(i)\right\|^2_{\mathbb{R}^n} \\
&\quad \geq (1-\alpha^2)[(1-\beta^2)\|x(k)\|^2_{\mathbb{R}^n} + (1-\beta^{-2})\|v(k)\|^2_{\mathbb{R}^n}] \\
&\qquad + (1-\alpha^{-2})\left\|\sum_{i=0}^{k-1} C_i(k)v(i)\right\|^2_{\mathbb{R}^n} \\
&\quad = (1-\alpha^2)(1-\beta^{-2})\|v(k)\|^2_{\mathbb{R}^n} + (1-\alpha^2)(1-\beta^2)\|x(k)\|^2_{\mathbb{R}^n} \\
&\qquad + (1-\alpha^{-2})\left\|\sum_{i=0}^{k-1} C_i(k)v(i)\right\|^2_{\mathbb{R}^n}
\end{aligned} \tag{3.4}$$

we find from (2.26) that

$$\begin{aligned}
\mathfrak{F}(u, \mathbf{C}(k)) \geq{}& (1-\alpha^2)(1-\beta^{-2})\|u\|^2_{\mathbb{R}^n}\boldsymbol{\eta}(k) + (1-\alpha^2)(1-\beta^2)\boldsymbol{\lambda}(k) \\
&+ (1-\alpha^{-2})\boldsymbol{\zeta}(\mathbf{C}(k)).
\end{aligned} \tag{3.5}$$

Choosing $\alpha < 1$ and $\beta > 1$ we have that $(1-\alpha^2)(1-\beta^{-2})\boldsymbol{\eta}(k)$ is a nonnegative operator in $\mathfrak{T}_h$ and $(1-\alpha^2)(1-\beta^2)\boldsymbol{\lambda}(k) + (1-\alpha^{-2})\boldsymbol{\zeta}(\mathbf{C}(k))$ is an operator in $\mathfrak{T}_h$. Clearly then for each $\mathbf{C}(k) \in (\mathbb{R}^{n\times n})^k$, $\mathfrak{F}(\cdot, \mathbf{C}(k))$ satisfies the conditions of Theorem 7.1 in [15] and therefore by Lemma 7.4 of [15], for each $\mathbf{C}(k)$ the set

$$\mathfrak{M}^c_{\mathbf{C}(k)} = \{\mathbf{M}_k \in \mathfrak{M} | \mathscr{J}(\mathbf{C}(k), \mathbf{M}_k) \leq c\} \tag{3.6}$$

is conditionally compact. So now let $\{(\mathbf{C}^n(k), \mathbf{M}^n_k)\} \subseteq \mathscr{A}_c$. Since $\{(\mathbf{C}^n(k), \mathbf{M}^1_k)\} \subseteq \mathscr{C}^c_{\mathbf{M}^1_k}$ there exists a subsequence $\{\mathbf{C}^{1,n}(k)\}$ of $\{\mathbf{C}^n(k)\}$ such that $\{(\mathbf{C}^{1,n}(k), \mathbf{M}^1_k)\}$ converges. Since $\{(\mathbf{C}^{1,n}(k), \mathbf{M}^2_k)\} \subseteq \mathscr{C}^c_{\mathbf{M}^2_k}$ there exists a subsequence $\{\mathbf{C}^{2,n}(k)\}$ of $\{\mathbf{C}^{1,n}(k)\}$ such that $\{(\mathbf{C}^{2,n}(k), \mathbf{M}^2_k)\}$ converges. Thus we construct a countable family of sequences $\{\mathbf{C}^{j,n}(k)\}$, each of which is a subsequence of the previous. Then for each $\mathbf{M}^j_k$ of the original sequence $\{(\mathbf{C}^n(k), \mathbf{M}^n_k)\}$ the sequence $\{(\mathbf{C}^{n,n}(k), \mathbf{M}^j_k)\}$ converges. Now since $\{(\mathbf{C}^{1,1}(k), \mathbf{M}^j_k)\} \subseteq \mathfrak{M}^c_{\mathbf{C}^{1,1}(k)}$ there exists a subsequence $\{\mathbf{M}^{1,j}_k\}$ of $\{\mathbf{M}^j_k\}$ such that $\{(\mathbf{C}^{1,1}(k), \mathbf{M}^{1,j}_k)\}$ converges. Since $\{(\mathbf{C}^{2,2}(k), \mathbf{M}^{1,j}_k)\} \subseteq \mathfrak{M}^c_{\mathbf{C}^{2,2}(k)}$ there exists a subsequence $\{\mathbf{M}^{2,j}_k\}$ of $\{\mathbf{M}^{1,j}_k\}$ such that $\{(\mathbf{C}^{2,2}(k), \mathbf{M}^{2,j}_k)\}$ converges. Thus we construct again a countable family of sequences $\{\mathbf{M}^{j,n}_k\}$, each of which is a subsequence of the previous. Then for each $\mathbf{C}^{j,j}(k)$ the sequence $\{(\mathbf{C}^{j,j}(k), \mathbf{M}^{n,n}_k)\}$ converges. Consider now the diagonal subsequence $\{(\mathbf{C}^{n,n}(k), \mathbf{M}^{n,n}_k)\}$ of the original sequence. It clearly converges, establishing that the set $\mathscr{A}_c$ is conditionally compact. □

We now give our existence theorem.

Theorem 3.4. *Suppose that the signal sequence $\{x(i)\}$ and the measurement outcomes at time $0, 1, \ldots, k-1$ have finite second moments. Then there exist p.o.m.*

$\hat{\mathbf{M}}_k$ *and* $n\times n$ *matrices* $\hat{C}_i(k)$, $i=0,\ldots,k-1$, *which minimize* $\mathcal{J}(\mathbf{C}(k),\mathbf{M}_k)$. *Moreover, the optimal measurement outcome also has finite second moments.*

Proof. Let c be chosen as above. Then the set $\mathcal{A}_c$ is conditionally compact by Lemma 3.3. Now $\mathcal{J}(\mathbf{C}(k),\mathbf{M}_k)$ is lower semicontinuous on $\mathcal{A}_c$ by Lemma 3.2. Suppose $\{\mathbf{C}^n(k),\mathbf{M}_k^n\}\subseteq \mathbf{A}_c$ and converges say to $(\mathbf{C}^0(k),\mathbf{M}_k^0)$. Then the lower semicontinuity of $\mathcal{J}(\mathbf{C}(k),\mathbf{M}_k)$ implies that $(\mathbf{C}^0(k),\mathbf{M}_k^0)\in\mathcal{A}_c$, i.e., $\mathcal{A}_c$ is also closed. Therefore $\mathcal{A}_c$ is *sequentially compact* [13, p. 8]. But then it is well known [20, p. 40] that a lower semicontinuous functional attains its minimum on a sequentially compact subset of a topological space. It follows clearly from the expression (2.27) for $\mathcal{J}(\mathbf{C}(k),\mathbf{M}_k)$, the inequality (3.5), and the finite second moments assumption for $\{x(k)\}$ and $v(0),\ldots,v(k-1)$ that a measurement $\mathbf{M}_k$ results in finite $\mathcal{J}(\mathbf{C}(k),\mathbf{M}_k)$ if and only if $\mathbf{M}_k\in\mathfrak{M}^2$. Therefore the optimal measurement $\hat{\mathbf{M}}_k\in\mathfrak{M}^2$ (provided $\inf\mathcal{J}(\mathbf{C}(k),\mathbf{M}_k)$ is finite, which is the only interesting case). □

4. Necessary and Sufficient Conditions for Optimality

In this section we derive necessary and sufficient conditions for the optimal measurement $\hat{\mathbf{M}}_k$ and optimal processing coefficient matrices $\hat{C}_i(k)$, $i=0,1,\ldots,k-1$. Our first result is given by

Theorem 4.1. *Necessary and sufficient conditions for* $\hat{C}_0(k),\hat{C}_1(k),\ldots,\hat{C}_{k-1}(k)$ *and* $\hat{\mathbf{M}}_k$ *to be optimal processing coefficients and optimal measurement at time k are:*

(i) $\langle\mathfrak{F}(\cdot,\hat{\mathbf{C}}(k)),\mathbf{X}\rangle_{\mathbb{R}^n}\geq\langle\mathfrak{F}(\cdot,\hat{\mathbf{C}}(k)),\hat{\mathbf{M}}_k\rangle_{\mathbb{R}^n}$ for every $\mathbf{X}\in\mathfrak{M}$.

(ii) $$\begin{bmatrix} E\{v(0)v(0)'\} & \cdots & E\{v(0)v(k)'\}\\ \vdots & \ddots & \vdots\\ E\{v(k)v(0)'\} & \cdots & E\{v(k)v(k)'\}\end{bmatrix}\begin{bmatrix}\hat{C}_0'(k)\\ \vdots\\ \hat{C}_{k-1}'(k)\\ I_n\end{bmatrix}=\begin{bmatrix}E\{v(0)x(k)'\}\\ \vdots\\ E\{v(k)x(k)'\}\end{bmatrix},$$

where the distributions for $v(k)$ *are induced by* $\hat{\mathbf{M}}_k$.

In the proof of the sufficiency of this theorem we will need the following lemma.

Lemma 4.2. *Suppose* $n\times n$ *matrices* $\hat{C}_0(k),\ldots,\hat{C}_{k-1}(k)$ *and p.o.m.* $\hat{\mathbf{M}}_k$ *satisfy conditions* (i) *and* (ii) *of Theorem* 4.1. *Let* $\mathbf{X}$ *be any p.o.m. and define* $n\times n$ *matrices* $D_0(k),\ldots,D_{k-1}(k)$ *so that*

$$\begin{bmatrix} E\{v(0)v(0)'\} & \cdots & E\{v(0)v(k)'\}\\ \vdots & \ddots & \vdots\\ E\{v(k)v(0)'\} & \cdots & E\{v(k)v(k)'\}\end{bmatrix}\begin{bmatrix}D_0'(k)\\ \vdots\\ D_{k-1}'(k)\\ I_n\end{bmatrix}=\begin{bmatrix}E\{v(0)x(k)'\}\\ \vdots\\ E\{v(k)x(k)'\}\end{bmatrix}, \quad (4.1)$$

where the distributions for $v(k)$ *are induced by* $\mathbf{X}$. *Then*

$$\mathcal{J}(\mathbf{D}(k),\mathbf{X})\geq\mathcal{J}(\hat{\mathbf{C}}(k),\hat{\mathbf{M}}_k). \quad (4.2)$$

Proof. Condition (i) of Theorem 4.1 implies

$$\mathscr{J}(\hat{\mathbf{C}}(k), \mathbf{X}) \geq \mathscr{J}(\hat{\mathbf{C}}(k), \hat{\mathbf{M}}_k) \qquad \text{for any} \quad \mathbf{X} \in \mathfrak{M}. \tag{4.3}$$

Consider the partitioned covariance matrix

$$R_{\mathbf{X}} \triangleq \left[\begin{array}{c|cccc|c} \overbrace{E\{x(k)x(k)'\}}^{n} & E\{x(k)v(0)'\} & \cdots & E\{x(k)v(k-1)'\} & \overbrace{E\{x(k)v(k)'\}}^{n} \\ \hline E\{v(0)x(k)'\} & E\{v(0)v(0)'\} & \cdots & E\{v(0)v(k-1)'\} & E\{v(0)v(k)'\} \\ \vdots & \vdots & & \vdots & \vdots \\ E\{v(k-1)x(k)'\} & E\{v(k-1)v(0)'\} & \cdots & E\{v(k-1)v(k-1)'\} & E\{v(k-1)v(k)'\} \\ \hline \underbrace{E\{v(k)x(k)'\}}_{n} & E\{v(k)v(0)'\} & \cdots & E\{v(k)v(k-1)'\} & \underbrace{E\{v(k)v(k)'\}}_{n} \end{array}\right] \tag{4.4}$$

and define the various blocks via

$$R_{\mathbf{X}} \triangleq \left[\begin{array}{c|c} A & B_{\mathbf{X}} \\ \hline B'_{\mathbf{X}} & \underbrace{\Gamma_{\mathbf{X}}}_{n} \end{array}\right]\}n \tag{4.5}$$

and

$$R_{\mathbf{X}} \triangleq \begin{array}{r} n\{ \\ \\ n\{ \end{array}\left[\begin{array}{c|c|c} \Sigma & \Phi & B^0_{\mathbf{X}} \\ \hline \Phi' & A_0 & B^1_{\mathbf{X}} \\ \hline \underbrace{B^{0\prime}_{\mathbf{X}}}_{n} & B^{1\prime}_{\mathbf{X}} & \underbrace{\Gamma_{\mathbf{X}}}_{n} \end{array}\right]\begin{array}{l} \}n \\ \\ \}n \end{array} \tag{4.6}$$

where the subindex $_{\mathbf{X}}$ refers to the p.o.m. which induces the distributions for the random variable $v(k)$. Then we have

$$\mathscr{J}(\hat{\mathbf{C}}(k), \mathbf{X}) = \operatorname{tr}[I_n, -\hat{C}_0(k), \ldots, -\hat{C}_{k-1}(k), -I_n]\left[\begin{array}{c|c} A & B_{\mathbf{X}} \\ \hline B'_{\mathbf{X}} & \Gamma_{\mathbf{X}} \end{array}\right]\left[\begin{array}{c} I_n \\ \hline -\hat{C}'_0(k) \\ \vdots \\ -\hat{C}'_{k-1}(k) \\ \hline -I_n \end{array}\right]$$

$$= \operatorname{tr}\left\{[I_n, -\hat{C}_0(k), \ldots, -\hat{C}_{k-1}(k)]A\begin{bmatrix} I_n \\ -\hat{C}'_0(k) \\ \vdots \\ -\hat{C}'_{k-1}(k) \end{bmatrix}\right.$$

$$\left. -[I_n, -\hat{C}_0(k), \ldots, -\hat{C}_{k-1}(k)]B_{\mathbf{X}} - B'_{\mathbf{X}}\begin{bmatrix} I_n \\ -\hat{C}'_0(k) \\ \vdots \\ -\hat{C}'_{k-1}(k) \end{bmatrix} + \Gamma_{\mathbf{X}}\right\}. \tag{4.7}$$

On the other hand, $\mathscr{J}(\hat{\mathbf{C}}(k), \hat{\mathbf{M}}_k)$ has a similar expression where we change the subindex $_{\mathbf{X}}$ to $_{\hat{\mathbf{M}}_k}$. Therefore (4.3) implies

$$\begin{aligned}&\operatorname{tr}\{-[I_n, -\hat{C}_0(k), \ldots, -\hat{C}_{k-1}(k)]B_{\mathbf{X}} - B'_{\mathbf{X}}[I_n, -\hat{C}_0(k), \ldots, -\hat{C}_{k-1}(k)]' \\ &\quad + \Gamma_{\mathbf{X}} + [I_n, -\hat{C}_0(k), \ldots, -\hat{C}_{k-1}(k)]B_{\hat{\mathbf{M}}_k} + B'_{\hat{\mathbf{M}}_k}[I_n, -\hat{C}_0(k), \ldots, -\hat{C}_{k-1}(k)]' \\ &\quad - \Gamma_{\hat{\mathbf{M}}_k}\} \geq 0 \qquad \text{for any} \quad \mathbf{X} \in \mathfrak{M}. \end{aligned} \tag{4.8}$$

But $\hat{C}_0(k), \ldots, \hat{C}_{k-1}(k)$ satisfy the normal equations (ii) with the p.o.m. $\hat{\mathbf{M}}_k$. So

$$\left[\begin{array}{c|c} A_0 & B^1_{\hat{\mathbf{M}}_k} \\ \hline B^{1\prime}_{\hat{\mathbf{M}}_k} & \Gamma_{\hat{\mathbf{M}}_k}\end{array}\right] \begin{bmatrix} \hat{C}'_0(k) \\ \vdots \\ \hat{C}'_{k-1}(k) \\ I_n \end{bmatrix} = \begin{bmatrix} \Phi' \\ B^{0\prime}_{\hat{\mathbf{M}}_k}\end{bmatrix}. \tag{4.9}$$

So

$$[\hat{C}_0(k), \ldots, \hat{C}_{k-1}(k)]A_0 + B^{1\prime}_{\hat{\mathbf{M}}_k} = \Phi \tag{4.10}$$

and

$$[\hat{C}_0(k), \ldots, \hat{C}_{k-1}(k)]B^1_{\hat{\mathbf{M}}_k} + \Gamma_{\hat{\mathbf{M}}_k} = B^0_{\hat{\mathbf{M}}_k}. \tag{4.11}$$

Then (4.11) gives

$$[I_n, -\hat{C}_0(k), \ldots, -\hat{C}_{k-1}(k)]B_{\hat{\mathbf{M}}_k} = \Gamma_{\hat{\mathbf{M}}_k} \tag{4.12}$$

and then (4.8) becomes

$$\begin{aligned}&\operatorname{tr}\{-[I_n, -\hat{C}_0(k), \ldots, -\hat{C}_{k-1}(k)]B_{\mathbf{X}} - B'_{\mathbf{X}}[I_n, -\hat{C}_0(k), \ldots, -\hat{C}_{k-1}(k)]' \\ &\quad + \Gamma_{\mathbf{X}} + \Gamma_{\hat{\mathbf{M}}_k}\} \geq 0 \qquad \text{for any} \quad \mathbf{X} \in \mathfrak{M}. \end{aligned} \tag{4.13}$$

So assumptions (i) and (ii) imply (4.13). Now given any p.o.m. $\mathbf{X} \in \mathfrak{M}$ we define the matrices $D_0(k), \ldots, D_{k-1}(k)$ via (4.1). Then (4.1) implies (similarly as (ii) implies (4.10)-(4.12))

$$[D_0(k), \ldots, D_{k-1}(k)]A_0 + B^{1\prime}_{\mathbf{X}} = \Phi, \tag{4.14}$$

$$[D_0(k), \ldots, D_{k-1}(k)]B^1_{\mathbf{X}} + \Gamma_{\mathbf{X}} = B^0_{\mathbf{X}}, \tag{4.15}$$

and

$$[I_n, -D_0(k), \ldots, -D_{k-1}(k)]B_{\mathbf{X}} = \Gamma_{\mathbf{X}}. \tag{4.16}$$

Then

$$\begin{aligned}\mathscr{J}(\mathbf{D}(k), \mathbf{X}) &= \operatorname{tr}\left\{[I_n, -D_0(k), \ldots, -D_{k-1}(k)]\left[\begin{array}{c|c}\Sigma & \Phi \\ \hline \Phi' & A_0\end{array}\right]\right. \\ &\qquad \left.\times [I_n, -D_0(k), \ldots, -D_{k-1}(k)]' - \Gamma_{\mathbf{X}}\right\} \\ &= \operatorname{tr}\{\Sigma + \Phi[-D_0(k), \ldots, -D_{k-1}(k)]' + [-D_0(k), \ldots, -D_{k-1}(k)]\Phi' \\ &\qquad + [-D_0(k), \ldots, -D_{k-1}(k)]A_0[-D_0(k), \ldots, -D_{k-1}(k)]' - \Gamma_{\mathbf{X}}\}. \end{aligned} \tag{4.17}$$

Trivial cases apart, we can assume without loss of generality that A_0 is invertible. Then from (4.10)

$$[-\hat{C}_0(k), \ldots, -\hat{C}_{k-1}(k)] = B^{1\prime}_{\hat{\mathbf{M}}_k} A_0^{-1} - \Phi A_0^{-1}. \tag{4.18}$$

Therefore (4.13) becomes

$$\begin{aligned} &\operatorname{tr}\{-B^0_{\mathbf{X}} - B^{1\prime}_{\hat{\mathbf{M}}_k} A_0^{-1} B^1_{\mathbf{X}} + \Phi A_0^{-1} B^1_{\mathbf{X}} - B^{0\prime}_{\mathbf{X}} - B^{1\prime}_{\mathbf{X}} A_0^{-1} B^1_{\hat{\mathbf{M}}_k} \\ &\quad + B^{1\prime}_{\mathbf{X}} A_0^{-1} \Phi' + \Gamma_{\mathbf{X}} + \Gamma_{\hat{\mathbf{M}}_k}\} \geq 0 \qquad \text{for any} \quad \mathbf{X} \in \mathfrak{M}. \end{aligned} \tag{4.19}$$

From (4.14) and (4.15) it now follows that

$$-B^{1\prime}_{\mathbf{X}} A_0^{-1} B^1_{\mathbf{X}} + \Phi A_0^{-1} B^1_{\mathbf{X}} + \Gamma_{\mathbf{X}} = B^0_{\mathbf{X}}. \tag{4.20}$$

Using (4.20), (4.19) becomes

$$\begin{aligned} &\operatorname{tr}\{B^{1\prime}_{\mathbf{X}} A_0^{-1} B^1_{\mathbf{X}} - \Gamma_{\mathbf{X}} - B^{1\prime}_{\hat{\mathbf{M}}_k} A_0^{-1} B^1_{\mathbf{X}} - B^{1\prime}_{\mathbf{X}} A_0^{-1} B^1_{\hat{\mathbf{M}}_k} + B^{1\prime}_{\mathbf{X}} A_0^{-1} B^1_{\mathbf{X}} - \Gamma_{\mathbf{X}} + \Gamma_{\mathbf{X}} + \Gamma_{\hat{\mathbf{M}}_k}\} \\ &\quad = \operatorname{tr}\{(B^1_{\mathbf{X}} - B^1_{\hat{\mathbf{M}}_k})' A_0^{-1} (B^1_{\mathbf{X}} - B^1_{\hat{\mathbf{M}}_k}) + B^{1\prime}_{\mathbf{X}} A_0^{-1} B^1_{\mathbf{X}} \\ &\qquad - \Gamma_{\mathbf{X}} - B^{1\prime}_{\hat{\mathbf{M}}_k} A_0^{-1} B^1_{\hat{\mathbf{M}}_k} + \Gamma_{\hat{\mathbf{M}}_k}\} \geq 0 \qquad \text{for any} \quad \mathbf{X} \in \mathfrak{M}. \end{aligned} \tag{4.21}$$

So the assumptions of the lemma are equivalent to (4.21).

Now we transform the inequality we want to establish (i.e., (4.2)) using similar methods. Utilizing (4.14) = (4.15), (4.17) results in

$$\begin{aligned} \mathscr{J}(\mathbf{D}(k), \mathbf{X}) &= \operatorname{tr}\{\Sigma + \Phi A_0^{-1} B^1_{\mathbf{X}} - \Phi A_0^{-1} \Phi' + B^{1\prime}_{\mathbf{X}} A_0^{-1} \Phi' - \Phi A_0^{-1} \Phi' \\ &\qquad - \Gamma_{\mathbf{X}} + B^{1\prime}_{\mathbf{X}} A_0^{-1} B^1_{\mathbf{X}} - \Phi A_0^{-1} B^1_{\mathbf{X}} - B^{1\prime}_{\mathbf{X}} A_0^{-1} \Phi' + \Phi A_0^{-1} \Phi'\} \\ &= \operatorname{tr}\{\Sigma - \Phi A_0^{-1} \Phi' + B^{1\prime}_{\mathbf{X}} A_0^{-1} B^1_{\mathbf{X}} - \Gamma_{\mathbf{X}}\}. \end{aligned} \tag{4.22}$$

Similarly $\mathscr{J}(\hat{\mathbf{C}}(k), \hat{\mathbf{M}}_k)$ has an identical expression to (4.22), with subindices ${}_{\hat{\mathbf{M}}_k}$ instead of ${}_{\mathbf{X}}$. Therefore the result of the lemma (i.e., (4.2)) holds if and only if

$$\operatorname{tr}\{B^{1\prime}_{\mathbf{X}} A_0^{-1} B^1_{\mathbf{X}} - \Gamma_{\mathbf{X}} - B^{1\prime}_{\hat{\mathbf{M}}_k} A_0^{-1} B^1_{\hat{\mathbf{M}}_k} + \Gamma_{\hat{\mathbf{M}}_k}\} \geq 0 \qquad \text{for any} \quad \mathbf{X} \in \mathfrak{M}. \tag{4.23}$$

Therefore the proof of the lemma will be complete if we show that (4.21) implies (4.23). We prove it by contradiction. So assume there exists a p.o.m. $\boldsymbol{\Psi}$ such that

$$\operatorname{tr}\{B^{1\prime}_{\boldsymbol{\Psi}} A_0^{-1} B^1_{\boldsymbol{\Psi}} - \Gamma_{\boldsymbol{\Psi}} - B^{1\prime}_{\hat{\mathbf{M}}_k} A_0^{-1} B^1_{\hat{\mathbf{M}}_k} + \Gamma_{\hat{\mathbf{M}}_k}\} < 0. \tag{4.24}$$

Let now

$$\mathbf{Z}(B) = \alpha \boldsymbol{\Psi}(B) + (1-\alpha)\hat{\mathbf{M}}_k(B) \qquad \text{for any} \quad B \in \mathscr{B}^n, \tag{4.25}$$

where α is a real number $0 < \alpha < 1$. Clearly, $\mathbf{Z} \in \mathfrak{M}$ for all α in that interval. Moreover, observe that

$$B^1_{\mathbf{Z}} = \alpha B^1_{\boldsymbol{\Psi}} + (1-\alpha) B^1_{\hat{\mathbf{M}}_k} \tag{4.26}$$

and

$$\Gamma_{\mathbf{Z}} = \alpha \Gamma_{\boldsymbol{\Psi}} + (1-\alpha) \Gamma_{\hat{\mathbf{M}}_k}. \tag{4.27}$$

We write (4.21) for $\mathbf{Z}$

$$
\begin{aligned}
&\operatorname{tr}\{(B^1_{\mathbf{Z}} - B^1_{\hat{\mathbf{M}}_k})' A_0^{-1}(B^1_{\mathbf{Z}} - B^1_{\hat{\mathbf{M}}_k}) + B^{1\prime}_{\mathbf{Z}} A_0^{-1} B^1_{\mathbf{Z}} - \Gamma_{\mathbf{Z}} - B^{1\prime}_{\hat{\mathbf{M}}_k} A_0^{-1} B^1_{\hat{\mathbf{M}}_k} + \Gamma_{\hat{\mathbf{M}}_k}\} \\
&\quad = \operatorname{tr}\{\alpha^2 (B^1_{\boldsymbol{\Psi}} - B^1_{\hat{\mathbf{M}}_k})' A_0^{-1} (B^1_{\boldsymbol{\Psi}} - B^1_{\hat{\mathbf{M}}_k}) \\
&\qquad + (\alpha B^1_{\boldsymbol{\Psi}} + (1-\alpha) B^1_{\hat{\mathbf{M}}_k})' A_0^{-1} (\alpha B^1_{\boldsymbol{\Psi}} + (1-\alpha) B^1_{\hat{\mathbf{M}}_k}) \\
&\qquad - \alpha \Gamma_{\boldsymbol{\Psi}} - \Gamma_{\hat{\mathbf{M}}_k} + \alpha \Gamma_{\hat{\mathbf{M}}_k} + \Gamma_{\hat{\mathbf{M}}_k} - B^{1\prime}_{\hat{\mathbf{M}}_k} A_0^{-1} B^1_{\hat{\mathbf{M}}_k}\} \\
&\quad = \operatorname{tr}\{\alpha^2 B^{1\prime}_{\boldsymbol{\Psi}} A_0^{-1} B^1_{\boldsymbol{\Psi}} - \alpha^2 B^{1\prime}_{\hat{\mathbf{M}}_k} A_0^{-1} B^1_{\boldsymbol{\Psi}} - \alpha^2 B^{1\prime}_{\boldsymbol{\Psi}} A_0^{-1} B^1_{\hat{\mathbf{M}}_k} \\
&\qquad + \alpha^2 B^{1\prime}_{\hat{\mathbf{M}}_k} A_0^{-1} B^1_{\hat{\mathbf{M}}_k} + \alpha^2 B^{1\prime}_{\boldsymbol{\Psi}} A_0^{-1} B^1_{\boldsymbol{\Psi}} \\
&\qquad + \alpha(1-\alpha) B^{1\prime}_{\boldsymbol{\Psi}} A_0^{-1} B^1_{\hat{\mathbf{M}}_k} + \alpha(1-\alpha) B^{1\prime}_{\hat{\mathbf{M}}_k} A_0^{-1} B^1_{\boldsymbol{\Psi}} \\
&\qquad + (1-\alpha)^2 B^{1\prime}_{\hat{\mathbf{M}}_k} A_0^{-1} B^1_{\hat{\mathbf{M}}_k} - \alpha \Gamma_{\boldsymbol{\Psi}} + \alpha \Gamma_{\hat{\mathbf{M}}_k} - B^{1\prime}_{\hat{\mathbf{M}}_k} A_0^{-1} B^1_{\hat{\mathbf{M}}_k}\} \\
&\quad = \operatorname{tr}\{2\alpha^2 B^{1\prime}_{\boldsymbol{\Psi}} A_0^{-1} B^1_{\boldsymbol{\Psi}} + \alpha(1-2\alpha) B^{1\prime}_{\hat{\mathbf{M}}_k} A_0^{-1} B^1_{\boldsymbol{\Psi}} \\
&\qquad + \alpha(1-2\alpha) B^{1\prime}_{\boldsymbol{\Psi}} A_0^{-1} B^1_{\hat{\mathbf{M}}_k} + 2\alpha(\alpha-1) B^{1\prime}_{\hat{\mathbf{M}}_k} A_0^{-1} B^1_{\hat{\mathbf{M}}_k} - \alpha \Gamma_{\boldsymbol{\Psi}} + \alpha \Gamma_{\hat{\mathbf{M}}_k}\} \\
&\quad = \alpha(2\alpha - 1) \operatorname{tr}\{(B^1_{\boldsymbol{\Psi}} - B^1_{\hat{\mathbf{M}}_k})' A_0^{-1} (B^1_{\boldsymbol{\Psi}} - B^1_{\hat{\mathbf{M}}_k})\} \\
&\qquad + \alpha \operatorname{tr}\{B^{1\prime}_{\boldsymbol{\Psi}} A_0^{-1} B^1_{\boldsymbol{\Psi}} - \Gamma_{\boldsymbol{\Psi}} + \Gamma_{\hat{\mathbf{M}}_k} - B^{1\prime}_{\hat{\mathbf{M}}_k} A_0^{-1} B^1_{\hat{\mathbf{M}}_k}\}.
\end{aligned}
\tag{4.28}
$$

Now the second component in (4.28) is strictly negative for all $0 < \alpha < 1$ because of our assumption (4.24). On the other hand, by choosing $0 < \alpha < \frac{1}{2}$ we can obviously make the first component nonpositive. Then for such a choice of α (i.e., $0 < \alpha < \frac{1}{2}$) the corresponding $\mathbf{Z}$ defined via (4.25) will violate (4.21), as is shown by (4.28), and thus we have a contradiction. So the proof of the lemma is complete. □

We are now able to give the

Proof of Theorem 4.1. The necessity is clear. Note that (ii) are the *normal equations* for the minimum variance linear estimate of $x(k)$ based on the random variables $v(0), \ldots, v(k-1), \hat{v}(k)$, with the constraint $\hat{C}_k(k) = I_n$. The sufficiency is more complicated. It is based on the fact that $\mathcal{J}(\mathbf{C}(k), \mathbf{X})$ is a quadratic function of $\mathbf{C}(k)$ and a linear function of $\mathbf{X}$. Given any fixed p.o.m. X we define matrices $D_0(k), \ldots, D_{k-1}(k)$ which satisfy (ii) (the normal equations) when the densities are those induced by $\mathbf{X}$. Then for any set of matrices $C_0(k), \ldots, C_{k-1}(k)$ we have that

$$\mathcal{J}(\mathbf{C}(k), \mathbf{X}) \geq \mathcal{J}(\mathbf{D}(k), \mathbf{X}). \tag{4.29}$$

Now from Lemma 4.2 we have that (i) and (ii) above imply

$$\mathcal{J}(\mathbf{D}(k), \mathbf{X}) \geq \mathcal{J}(\hat{\mathbf{C}}(k), \hat{\mathbf{M}}_k) \qquad \text{for any} \quad \mathbf{X} \in \mathfrak{M}. \tag{4.30}$$

But then (4.29) and (4.30) prove the optimality of $\hat{C}_0(k), \ldots, \hat{C}_{k-1}(k)$ and $\hat{\mathbf{M}}_k$. □

We now concentrate on condition (i) of Theorem 4.1 in our effort to improve these necessary and sufficient conditions. This condition represents an optimization problem with respect to the p.o.m. $\mathbf{X}$ (while $\hat{\mathbf{C}}(k)$ is held fixed), similar to those studied extensively by Holevo in [15, pp. 368–372]. Note that in our case the operator-valued function $\mathfrak{F}(u, \hat{\mathbf{C}}(k))$ is quadratic in u.

The following lemma, an application of the Lagrange duality theorem [2, p. 94], has been announced by Holevo in a more general setting [18].

Lemma 4.3. *Let $\mathfrak{F}$ be a continuous operator-valued function on $\mathbb{R}^n$, such that for every $u \in \mathbb{R}^n$, $\mathfrak{F}(u)$ is a nonnegative, self-adjoint operator with finite trace on a Hilbert space H. Consider the set of operators on H, $S_{\mathfrak{F}} = \{\boldsymbol{\tau}$ is self-adjoint, with finite trace, $\boldsymbol{\tau} \geq 0$ and $\boldsymbol{\tau} \leq \mathfrak{F}(u)$ for all $u \in \mathbb{R}^n\}$. Then*

$$\inf_{\mathbf{X} \in \mathfrak{M}} \langle \mathfrak{F}, \mathbf{X} \rangle_{\mathbb{R}^n} = \max_{\boldsymbol{\tau} \in S_{\mathfrak{F}}} \operatorname{Tr}[\boldsymbol{\tau}]. \tag{4.31}$$

Proof. First we observe that if $\boldsymbol{\tau} \in S_{\mathfrak{F}}$ then

$$\operatorname{Tr}[\boldsymbol{\tau}] = \langle \boldsymbol{\tau}, \mathbf{X} \rangle_{\mathbb{R}^n} \leq \langle \mathfrak{F}, \mathbf{X} \rangle_{\mathbb{R}^n}. \tag{4.32}$$

Therefore

$$\sup_{S_{\mathfrak{F}}} \operatorname{Tr}[\boldsymbol{\tau}] \leq \inf_{\mathbf{X} \in \mathfrak{M}} \langle \mathfrak{F}, \mathbf{X} \rangle_{\mathbb{R}^n}. \tag{4.33}$$

Now consider the vector space $\mathfrak{X}$ of all *unconstrained* operator-valued measures on $\mathbb{R}^n$. That is, $\mathbf{X} \in \mathfrak{X}$ if:

$$\left.\begin{array}{l} \text{(i) } X(A) \in \mathfrak{B}(H) \text{ and self-adjoint, for every } A \in \mathscr{B}^n; \\ \text{(ii) for } \{B_i\} \text{ any partition of } \mathbb{R}^n \\ \qquad \displaystyle\sum_{i=1}^{\infty} \mathbf{X}(B_i) = \mathbf{X}(\mathbb{R}^n), \end{array}\right\} \tag{4.34}$$

where (ii) is interpreted in the weak sense. We let, as usual, $\mathfrak{B}(H)$ denote the space of all bounded operators on H. Then we consider the duality pair [25, p. 369], [8, p. 38]

$$\begin{aligned} &\mathfrak{T} \times \mathfrak{B}(H) \xrightarrow{\langle \cdot, \cdot \rangle} \mathbb{C}, \\ &(\boldsymbol{\tau}, \mathbf{A}) \longrightarrow \operatorname{Tr}[\boldsymbol{\tau}\mathbf{A}] = \langle \boldsymbol{\tau}, \mathbf{A} \rangle, \end{aligned} \tag{4.35}$$

where $\mathfrak{T}$ is the trace class operators on H. Then if we give $\mathfrak{B}(H)$ the *ultra weak topology* [8, p. 32] (i.e., the weakest topology that makes all these forms (4.35) induced by $\mathfrak{T}$ continuous on $\mathfrak{B}(H)$), $\mathfrak{T}$ becomes the dual of $\mathfrak{B}(H)$ [8] [25, pp. 497–498]. Now let Ω be the subset of positive unconstrained operator-valued measures in $\mathfrak{X}$. That is $\mathbf{X} \in \Omega$ if in addition to (4.34) the condition $\mathbf{X}(A) \geq 0$ for all $A \in \mathscr{B}^n$ is satisfied. Let

$$\left.\begin{array}{l} \qquad \mathscr{K}: \mathfrak{X} \to \mathfrak{B}(H) \\ \text{and} \\ \qquad \mathscr{K}(\mathbf{X}) = \mathbf{I} - \mathbf{X}(\mathbb{R}^n), \quad \mathbf{I} \text{ the identity on } H. \end{array}\right\} \tag{4.36}$$

Clearly, $\langle \mathfrak{F}, \mathbf{X} \rangle_{\mathbb{R}^n}$ is a linear functional on $\mathfrak{X}$ and $\mathscr{K}$ is affine [2, p. 93]. Then, since all hypotheses are satisfied, by a direct application of the Lagrange duality theorem [2, pp. 92–94], [20, p. 224] we have that for $\boldsymbol{\tau} \in \mathfrak{T}_h$

$$\inf_{\substack{\mathbf{X} \in \Omega \\ \mathscr{K}(\mathbf{X}) \leq \mathbf{0}}} \langle \mathfrak{F}, \mathbf{X} \rangle_{\mathbb{R}^n} = \sup_{\boldsymbol{\tau} \geq 0} \left\{ \inf_{\mathbf{X} \in \Omega} \{ \langle \mathfrak{F}, \mathbf{X} \rangle_{\mathbb{R}^n} + \operatorname{Tr}[\boldsymbol{\tau}(\mathbf{I} - \mathbf{X}(\mathbb{R}^n))] \} \right\}. \tag{4.37}$$

Since the only interesting case is when the inf on the left-hand side of (4.37) is finite, we must consider only those $\boldsymbol{\tau}$ for which $\mathfrak{F}(u) \geq \boldsymbol{\tau}$ for all $u \in \mathbb{R}^n$ and therefore we conclude that

$$\inf_{\mathbf{X} \in \Omega} \{\langle \mathfrak{F} - \boldsymbol{\tau}, \mathbf{X} \rangle_{\mathbb{R}^n} + \operatorname{Tr}[\boldsymbol{\tau}]\} = \operatorname{Tr}[\boldsymbol{\tau}] \tag{4.38}$$

and

$$\inf_{\mathbf{X} \in \mathfrak{M}} \langle \mathfrak{F}, \mathbf{X} \rangle_{\mathbb{R}^n} = \sup_{\boldsymbol{\tau} \in S_{\mathfrak{F}}} \operatorname{Tr}[\boldsymbol{\tau}] = \operatorname{Tr}[\boldsymbol{\tau}_0] \tag{4.39}$$

for some $\boldsymbol{\tau}_0 \in S_{\mathfrak{F}}$, as follows from the properties of $S_{\mathfrak{F}}$. This completes the proof. □

As a consequence we have:

Corollary 4.4. *Necessary and sufficient conditions for the p.o.m. $\hat{\mathbf{M}}_k$ to solve the optimization problem described in* (i) *of Theorem* 4.1 *are*:

(i) $\mathfrak{F}(\cdot, \hat{\mathbf{C}}(k))$ *is integrable with respect to* $\hat{\mathbf{M}}_k$;
(ii) $\mathfrak{F}(u, \hat{\mathbf{C}}(k)) \geq \hat{\boldsymbol{\tau}}$ *for all* $u \in \mathbb{R}^n$,

where $\hat{\boldsymbol{\tau}} \triangleq \int_{\mathbb{R}^n} \mathfrak{F}(u, \hat{\mathbf{C}}(k)) \hat{\mathbf{M}}_k(du)$, *which is well defined in view of* (i).

Proof. Necessity. From the lemma above we have that there exists $\boldsymbol{\tau}_0 \in S_{\mathfrak{F}(\cdot, \hat{\mathbf{C}}(k))}$ such that

$$\inf_{\mathbf{X} \in \mathfrak{M}} \langle \mathfrak{F}(\cdot, \hat{\mathbf{C}}(k)), \mathbf{X} \rangle_{\mathbb{R}^n} = \langle \mathfrak{F}(\cdot, \hat{\mathbf{C}}(k)), \hat{\mathbf{M}}_k \rangle_{\mathbb{R}^n} = \max_{\boldsymbol{\tau} \in S_{\mathfrak{F}(\cdot, \hat{\mathbf{C}}(k))}} \operatorname{Tr}[\boldsymbol{\tau}] = \operatorname{Tr}[\boldsymbol{\tau}_0]. \tag{4.40}$$

We will be done if we show $\boldsymbol{\tau}_0 = \hat{\boldsymbol{\tau}}$. Since $\mathfrak{F}(\cdot, \hat{\mathbf{C}}(k))$ is a quadratic polynomial in $(u_1, \ldots, u_n)$ (and hence locally integrable with respect to $\hat{\mathbf{M}}(du)$), we have that

$$\int_A (\mathfrak{F}(u, \hat{\mathbf{C}}(k)) - \boldsymbol{\tau}_0) \hat{\mathbf{M}}_k(du) \geq 0 \qquad \text{for every bounded} \quad A \in \mathscr{B}^n. \tag{4.41}$$

Now if there exists bounded $A_0 \in \mathscr{B}^n$ such that

$$\int_{A_0} (\mathfrak{F}(u, \hat{\mathbf{C}}(k)) - \boldsymbol{\tau}_0) \hat{\mathbf{M}}_k(du) > 0 \tag{4.42}$$

we must have

$$\langle \mathfrak{F}(\cdot, \hat{\mathbf{C}}(k)), \hat{\mathbf{M}}_k \rangle_{\mathbb{R}^n} > \operatorname{Tr}[\boldsymbol{\tau}_0] \tag{4.43}$$

which is a contradiction to (4.40). So (4.41) is in fact an equality for any bounded $A \in \mathscr{B}^n$. Choosing now an increasing sequence of bounded sets $A_i \in \mathscr{B}^n$, $A_i \to \mathbb{R}^n$, we have from (4.41) that $\mathfrak{F}(\cdot, \hat{\mathbf{C}}(k))$ is integrable with respect to $\hat{\mathbf{M}}_k$ over $\mathbb{R}^n$. Therefore, from (4.41), $\boldsymbol{\tau}_0 = \int_{\mathbb{R}^n} \mathfrak{F}(u, \mathbf{C}(k)) \hat{\mathbf{M}}_k(du) \triangleq \hat{\boldsymbol{\tau}}$. This completes the proof of necessity. For the sufficiency we observe that (i) and (ii) imply that for any p.o.m. $\mathbf{X}$

$$\langle \mathfrak{F}(\cdot, \hat{\mathbf{C}}(k)), \mathbf{X} \rangle_{\mathbb{R}^n} \geq \operatorname{Tr}[\hat{\boldsymbol{\tau}}] = \langle \mathfrak{F}(\cdot, \hat{\mathbf{C}}(k)), \hat{\mathbf{M}}_k \rangle_{\mathbb{R}^n}. \tag{4.44}$$

Then clearly Lemma 4.3 implies that

$$\langle \mathfrak{F}(\cdot, \hat{\mathbf{C}}(k)), \hat{\mathbf{M}}_k \rangle_{\mathbb{R}^n} = \inf_{\mathbf{X} \in \mathfrak{M}} \langle \mathfrak{F}(\cdot, \hat{\mathbf{C}}(k)), \mathbf{X} \rangle_{\mathbb{R}^n}$$

and the proof is complete. □

As a consequence we have the following basic necessary and sufficient conditions for the optimization problem of this paper.

Theorem 4.5. *Necessary and sufficient conditions for $\hat{C}_0(k), \ldots, \hat{C}_{k-1}(k)$ and $\hat{\mathbf{M}}_k$ to be optimal processing coefficient matrices and optimal measurement at time k are*:

(i)
$$\begin{bmatrix} E\{v(0)v(0)'\} & \cdots & E\{v(0)v(k)'\} \\ \vdots & & \vdots \\ E\{v(k)v(0)'\} & \cdots & E\{v(k)v(k)'\} \end{bmatrix} \begin{bmatrix} \hat{C}_0'(k) \\ \vdots \\ \hat{C}_{k-1}'(k) \\ I_n \end{bmatrix} = \begin{bmatrix} E\{v(0)x(k)'\} \\ \vdots \\ E\{v(k)x(k)'\} \end{bmatrix},$$

where the distributions of $v(k)$ are induced by $\hat{\mathbf{M}}_k$;

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

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

Proof. Immediate from the above sequence of lemmas and theorems. □

5. Concluding Remarks

We observe that the solution to the optimal linear filtering problem in the multiparameter (or vector) case is not as explicit as the solution to the scalar case (see [3], equations (13)-(16)). This was expected since the optimization problem here cannot be formulated as a quadratic problem (cf. our previous remarks on operator moments of measurements). Observe that in the scalar case the measurement (which in that case is simple and represented by a projection-valued measure) is uniquely defined by its first operator moments. That is why the conditions of Theorem 4.5 can be transformed into the convenient form of Corollary 1 of [3]. In the vector case, however, the best that can be done generally is to derive explicit necessary and sufficient conditions which characterize the first and second operator moments of the optimal measurements. Since these moments do not determine uniquely the optimal measurement (see also p. 536 of [16]) there exists freedom in further restricting the measurements to belong to certain convenient classes. Such a route has been followed by Holevo, using canonical measurements for estimation problems concerning Gaussian states [16].

We would like to note that although the results of this paper do not generally provide an explicit closed form solution for the optimal measurement and optimal processing coefficients, they can be used to establish optimality for candidate processing and measurement schemes. This approach has been successfully employed in similar problems by Holevo in [15] and [17] and by Belavkin in [5] and [6] (including problems with non-Gaussian states). The role played by the conditions of Theorem 4.5 in linear quantum filtering theory is central. It is therefore a natural consequence to analyze these conditions in detail and discover cases that permit explicit solution. This has been done for Gaussian statistics in joint work with Harger in [4] which also includes a practical application to an optical communication problem.

Finally, there is the question of implementation. This is a hard and mostly unanswered question even for the scalar case. For very few cases we do know how to implement (with devices) the optimal measurements which result from the solution of the problem. In the vector case, we have *in addition* to find the auxiliary system and simple measurement necessary to implement a p.o.m. The only example of an explicit construction appears in [4] and further generalizations in [16].

Acknowledgment

I would like to thank Bob Harger for listening to countless arguments and providing constructive criticism which led to improvements of the original version of this paper.

References

1. Akhiezer NI, Glazman IM (1963) Theory of Operators in Hilbert Space, vol II. Frederick Ungar, New York
2. Arrow KJ, Hurwicz L, Uzawa H (1958) Studies in Linear and Non-Linear Programming. Stanford University Press, Stanford, CA
3. Baras JS, Harger RO, Park YH (1976) Quantum mechanical linear filtering of random signal sequences. IEEE Trans Inform Theory 22:59–64
4. Baras JS, Harger RO (1977) Quantum mechanical filtering of vector signal processes. IEEE Trans Inform Theory 23:683–693
5. Belavkin VP (1974) Optimal linear randomized filtration of quantum boson signals. Problems Control Inform Theory 3:47–62
6. Belavkin VP (1976) Proceedings of the 4th International Symposium on Information Theory, Leningrad, part I, pp 17–19
7. Berberian SK (1966) Notes on Spectral Theory. Van Nostrand, New York
8. Dixmier J (1969) Les algebres d'operateurs dans l'espace Hilbertien. Gauthier-Villars, Paris
9. Helstrom CW (1967) Detection theory and quantum mechanics. Inform and Control 10:254–291
10. Helstrom CW, Liu JW, Gordon JP (1970) Quantum mechanical communication theory. Proc IEEE 58:1578–1598
11. Helstrom CW, Kennedy RS (1974) Non-commuting observables in quantum detection and estimation theory. IEEE Trans Inform Theory 20:16–24
12. Helstrom CW (1974) Cramer-Rao inequalities for operator-valued measures in quantum mechanics. Internat J Theoret Phys
13. Hille E, Phillips RS (1957) Functional Analysis and Semi-groups. Colloquium Publication. American Mathematical Society, Providence, RI
14. Holevo AS (1972) Statistical problems in quantum physics. Proceedings of 1972 Soviet-Japanese Symposium on Probability and Statistics, Vol. 1, pp 22–40
15. Holevo AS (1973) Statistical decision theory for quantum systems. Multivariate Anal 3:337–394
16. Holevo AS (1975) Some statistical problems for quantum Gaussian states. IEEE Trans Inform Theory 21:533–543

17. Holevo AS (1976) Proceedings of the 4th International Symposium on Information Theory, Leningrad, part I, pp 150–152
18. Holevo AS (1974) The theory of statistical decisions on an operator algebra. Dokl Akad Nauk SSSR 218:1276–1281
19. Jauch JM (1968) Foundations of Quantum Mechanics. Addison-Wesley, Reading, MA
20. Luenberger DG (1969) Optimization by Vector Space Methods. Wiley, New York
21. Mitter SK, Young SK (1984) Integration with respect to operator-valued measures with applications to quantum estimation theory. Ann Pura Appl (III) CXXXVII:1–39
22. Parthasarathy KR (1967) Probability Measures on Metric Spaces. Academic Press, New York
23. Prugovecki E (1971) Quantum Mechanics in Hilbert Space. Academic Press, New York
24. Rudin W (1966) Real and Complex Analysis. McGraw-Hill, New York
25. Treves F (1967) Topological Vector Spaces, Distributions and Kernels. Academic Press, New York
26. Varadhan SRS (1968) Stochastic Processes, Courant Institute of Mathematical Sciences, NY University Lecture Notes

Accepted 12 January 1987