Some results in abstract optimal linear filtering

Vladimir Fomin Michael Ruzhansky

Abstract

The linear optimal filtering problems in infinite dimensional Hilbert spaces and their extensions are discussed. The quality functional is allowed to be a general quadratic functional defined by a possibly degenerate operator. We describe the solution of the stable and the causal filtering problems. In the case of causal filtering, we establish the relation with a relaxed causal filtering problem in the extended space. We solve the last problem in continuous and discrete cases and give the necessary and sufficient conditions for the solvability of the original causal problem as well as the conditions for the analogue of Bode–Shannon formula to define an optimal filter.

1 Introduction

We consider the linear optimal filtering problems in infinite dimensional Hilbert spaces and their extensions. Briefly, the problem is as follows. Let H', H'' be Hilbert spaces and $z = \begin{bmatrix} x \\ y \end{bmatrix}$ a random element in $H = H' \times H''$, where x and y are unobservable and observable components of z in H' and H'' respectively. The correlation operator of z is assumed to be bounded in H and we denote by $\mathbf{H}$ a subset of all linear operators $h : H'' \to H'$. The $\mathbf{H}$–optimal linear filtering problem is a problem of the estimations of the unobservable component x based on the realizations of the observable component y in the form

$$\hat{x} = hy \tag{1}$$

solving the minimization problem in $\mathbf{H}$

$$J(h) \to \inf_{h \in \mathbf{H}}, \tag{2}$$

where the quality functional J is defined by

$$J(h) = E||D(x - \hat{x})||^2 \tag{3}$$

with a suitable norm in (3) and a linear operator $D : H' \to H'$. The operator D here is an arbitrary operator in general. In the sequel we will sometimes assume that it has an adjoint D^* and that it is continuous if the other operators in the problem are continuous. The choice of operator D allows to perform the minimization of the quality functional $J(h)$ with respect to some of the variables. In this case D is not bijective and its choice is dictated by the problem at hand. If D is bijective then, as we shall see, the minimization in (2) is performed with respect to *all* variables with certain weights assigned. This leads to the linear transformation of the solution for $D = I$, I being the identity operator. If D is degenerate, the solution of (2) is not unique and we will provide formulas for it.

If $\mathbf{H}$ consists of all continuous linear operators h, the problem (1), (2), (3) is called *stable*. If H' and H'' are Hilbert resolution spaces, one has a time structure in H and in its terms defines an "independent of the future" class of the causal continuous operators $\mathbf{H}$. In this case the problem is called *causal*.

If H' and H'' are finite dimensional (the time set $\mathbf{T}$ is discrete and finite), the problem (1), (2), (3) is quite trivial: the causal operators become the lower triangular matrices in a natural basis. In the case of the nondegenerate correlation matrix R_y of the random vector y and the identity matrix D, the problem (2) is solvable in the class of the causal weight operators, the solution is unique and can be effectively expressed in the terms of the Holetsky factorization of R_y. This result finds various applications ([12], [9]).

In the case of the discrete time $\mathbf{T} = \mathbf{Z}$ and a stationary partially observable process (time series) z the problem (2) was first treated by Kolmogorov ([8]). In the case of the continuous time its solution was first obtained by Wiener ([12]), who also developed a method for the synthesis of the transfer function of the optimal filter. The Wiener–Kolmogorov optimal filtering theory of the stationary processes was universally accepted partly due to the interpretation of the optimal filter given by Bode and Shannon ([1]), where the signal y is being "prewhitened" first and the result is optimally processed. The solution representation for optimization problem (2), allowing the mentioned interpretation, bears the name of the Bode–Shannon formula.

However, in many applications one estimates only the specific components of x or their combination, which is represented by the degenerate matrix D in the quality functional (3) and the solutions of the generalized finite dimensional problems can be found in [10]. In this case the solution need not be unique and there are conditions on the degeneracy of D for which the Bode–Shannon formula still defines an optimal filter.

On the other hand, the infinite dimensional applications required the development of the filtering theory in Hilbert ([3]) and sometimes Banach spaces ([4], [7]). The unobservable component of a process can be formed by the infinite dimensional filter. For example it can be given by several differential equations with delay or by differential integral equations. In this case it is not possible to use the methods developed for the finite dimensional state space. The method discussed in the paper allows to formulate such infinite dimensional problems and provides techniques for their solutions. For the applications of this theory to the problem of the linear estimation of the parameters of a signal based on the observations of its realizations see for example [4]. In this paper the stable filtering problem will be solved for the general quadratic quality functional (3). The solution of the causal filtering problem need not exist in general. We will establish necessary and sufficient conditions of solvability by relaxing the problem, allowing a slightly general class of the weight operators in (2). The relaxed problem can be solved and the analysis of its solution can be used for the construction of minimizing sequences. The solutions will be given for continuous and discrete resolutions. The conditions for the analogue of the Bode–Shannon formula to define an optimal filter will be also given.

For details and proofs we refer to [11].

2 Linear filtering

Let $H = H' \times H''$, where H' and H'' are Hilbert spaces with inner products $\langle \cdot, \cdot \rangle_{H'}$ and $\langle \cdot, \cdot \rangle_{H''}$ respectively. Let $F' \subset H'$ and $F'' \subset H''$ be linear dense subsets. The elements $\phi \in H$ can be interpreted as $\phi = \begin{bmatrix} \phi' \\ \phi'' \end{bmatrix}$ with $\phi' \in H'$, $\phi'' \in H''$. Let $F = F' \times F''$. We will consider random H_F elements $z = \begin{bmatrix} x \\ y \end{bmatrix}$, with x and y random $H'_{F'}-$ and $H''_{F''}-$ elements respectively. The correlation operator R_z will be assumed continuous on H (which is natural) and have the following block form:

$$R_z = \begin{bmatrix} R_x & R_{xy} \\ R_{yx} & R_y \end{bmatrix},$$

where we write $R_x = Exx^*$, $R_y = Eyy^*$, $R_{xy} = R_{yx}^* = Exy^*$.

Let $h : H''_{F''} \to H'_{F'}$ be linear. We assume now that there exist an operator $h^* : H'_{F'} \to H''_{F''}$ defined on the whole of $H'_{F'}$, such that for every $\phi' \in F'$, $\phi'' \in F''$ one has

$$(h\phi'')^* \phi' = (\phi'')^* (h^* \phi'). \tag{4}$$

Relation (4) defines h^* uniquely and h^* is the adjoint to h operator.

Let x and y be the unobservable and observable components of z respectively. We define the random $H'_{F'}-$ element $\hat{x}$ by

$$\hat{x} = hy. \tag{5}$$

One readily checks that $\hat{x}$ is a random element in view of our assumptions on h. Then $R_{\hat{x}} = hR_y h^* : H' \to H'$ is involutive in H' as the correlation operator of a random element $\hat{x}$. The element $\hat{x}$ is interpreted as a linear estimate of the nonobservable component x of a random H_F element z, based on the realizations of its observable component y. The relation (5) is called *a linear filter with weight operator h.*

Let a linear operator $D : H'_{F'} \to H'_{F'}$ have an adjoint D^*. We define *the quality functional* as

$$J_{\phi'}(h) = E|\langle \phi', D(x - \hat{x})\rangle|^2, \ \phi' \in F'. \tag{6}$$

Let $\mathbf{H}$ be a given subset of linear operators $h : H''_{F''} \to H'_{F'}$. Then the $\mathbf{H}$–*optimal filtering problem* is defined as a problem of the minimization of the functionals

$$J_{\phi'}(h) \to \inf_{h \in \mathbf{H}}, \tag{7}$$

defined by (6), (5) for every $\phi' \in F'$.

We will need a notion of the pseudo inversion of an operator. Let $A : H \to H$ be a Hermitian linear operator in a Hilbert space H. Let Q_A be an orthogonal projection on the image of A, $Q_A : H \to \mathrm{Im}A$. The space $Q_A H$ is invariant for A and we write $A^{-1}Q_A$ for the inverse of A in $Q_A H$. The operator

$$A^+ = Q_A \circ A^{-1} Q_A \circ Q_A$$

is called the *pseudo inverse of A*. It follows that

$$A^+ A = AA^+ = Q_A. \tag{8}$$

One readily checks that (8) determines A^+ uniquely. Assume that $\mathbf{H}$ is a space of all linear operators $h : H''_{F''} \to H'_{F'}$. Then the solution of the $\mathbf{H}$–optimal filtering problem is given by

Theorem 1 *Let the correlation operator R_z of a random H element z be continuous in H and let R_y^+ denote the pseudo inverse operator for the correlation operator R_y of y in H''. Then the minimization problem (7) in the class $\mathbf{H}$ of all weight operators $h : H''_{F''} \to H'_{F'}$ is solvable and any solution is of the form*

$$h_{\mathrm{opt}} = R_{xy} R_y^+ + Q, \tag{9}$$

where $Q : H''_{F''} \to H'_{F'}$ is any linear operator satisfying $DQQ_{R_y} = 0$ and Q_{R_y} is the kernel of R_y. Moreover, one has

$$\inf_{h \in \mathbf{H}} J_{\phi'}(h) = J_{\phi'}(h_{\mathrm{opt}}) = \langle \phi', D[R_x - R_{xy} R_y^+ R_{xy}^*] D^* \phi' \rangle.$$

3 Linear stable filtering

If $\mathbf{H}$ is a space of all *continuous* linear operators from H'' to H', then the linear filters of the form (5) with weight operator in $\mathbf{H}$ are called *stable* and $\mathbf{H}$–optimal filtering problem is called the *stable filtering problem*. In this case one allows $\phi' \in H'$ in (6) and the minimization problem can be reformulated for scalar functionals

$$J(h) = \sup_{\phi' \in H'} \frac{E|\langle \phi', D(x-\hat{x})\rangle|^2}{|\phi'|^2_{H'}}. \qquad (10)$$

Now we are ready to describe the solution of the linear stable filtering problem. Let us assume that R_y is continuously invertible in its image $R_y H''$, which means that there exist a neighborhood U of zero such that $\sigma(R_y) \cap U = \{0\}$, $\sigma(R_y)$ being the spectrum of R_y. We will also assume that the operator D in the quality functional (6) is continuous in H' and has an adjoint D^*.

Theorem 2 *Let the correlation operator R_z of a random H element z be continuous in H and assume that the correlation operator R_y of y has the continuous pseudo inverse operator R_y^+ in H''. Then the minimization problem (7) in the class $\mathbf{H}$ of all continuous weight operators $h : H'' \to H'$ is solvable and any solution is of the form*

$$h_{\text{opt}} = R_{xy} R_y^+ + Q, \qquad (11)$$

where $Q : H'' \to H'$ is any linear continuous operator satisfying $DQQ_{R_y} = 0$. Moreover, one has

$$\inf_{h \in \mathbf{H}} J_{\phi'}(h) = J_{\phi'}(h_{\text{opt}}) = \langle \phi', D[R_x - R_{xy} R_y^+ R_{xy}^*] D^* \phi' \rangle_{H'}.$$

The operators (11) are also optimal in the problem with quality functional (10) and

$$\inf_{h \in \mathbf{H}} J(h) = J(h_{\text{opt}}) = |D[R_x - R_{xy} R_y^+ R_{xy}^*] D^*|_{H'}.$$

These methods can be applied for the problems of linear estimation of the parameters of a signal based on the observations of its realizations.

4 Linear causal filtering

In this section we will give the solution of the generalized linear causal filtering problem. However, we need some preliminary notions and results first.

4.1 Hilbert resolution spaces and causal operators

Let H be a Hilbert space, $\mathbf{T} = (t_s, t_f)$, $-\infty \le t_s < t_f \le +\infty$, and let $\mathbf{P}_T = \{P_t, t \in \mathbf{T}\}$ be a family of commutative projectors $P_t : H \to H$, $P_t^2 = P_t$, $P_t P_s = P_s P_t$, $t, s \in \mathbf{T}$. Let $\mathbf{P}_T$ satisfy the following two properties

(i) *monotonicity*: $P_t P_s = P_s$ for $t \ge s$, $t, s \in \mathbf{T}$.

(ii) *completeness*: $\lim_{t \to t_s} P_t = 0_H$, $\lim_{t \to t_f} P_t = I_H$, where the limits are taken in the strong operator topology.

Note, that condition (i) is equivalent to the fact that $P_s H \subset P_t H$, $t \ge s$. We assume the family $\mathbf{P}_T$ to be bounded uniformly in t: $\sup_{t \in \mathbf{T}} |P_t| < \infty$ and strongly continuous from the left: $\lim_{\epsilon \to 0+} P_{t-\epsilon}\phi = P_t \phi$ for every $\phi \in H$. Such family $\mathbf{P}_T$ is called a *resolution of the identity of* H and $(H, \mathbf{P}_T)$ is called a *Hilbert resolution space*. If $\mathbf{P}_T$ consists of the orthogonal projectors: $P_t = P_t^*$, then it is called a *Hermitian resolution of the identity*. In this case the condition of the uniform boundedness in t is automatically satisfied since $|P_t| \le 1$.

Let $H = H' \times H''$, where $(H', \mathbf{P}'_T)$, $(H'', \mathbf{P}''_T)$ are Hilbert resolution spaces. Then H may be equipped with the canonical resolution of the identity

$$P_t = \begin{bmatrix} P'_t & 0_{12} \\ 0_{21} & P''_t \end{bmatrix}, \; t \in \mathbf{T}, \qquad (12)$$

where $0_{12} : H'' \to H'$, $0_{21} : H' \to H''$ are zero operators.

Definition 1 *Let $A : D(A) \to H$ be a linear densely defined operator. A is called* finite from above *if there exists a measurable, essentially bounded function $\tau : \mathbf{T} \to \mathbf{T}$, such that for almost all $t \in \mathbf{T}$ the operator $P_t A$ is bounded in H and if $t - \tau(t) \in \mathbf{T}$, then*

$$P_t A = P_t A P_{t-\tau(t)} \qquad (13)$$

on $D(A) \cap P_{t-\tau(t)} D(A)$. The function $\tau = \tau_+(\cdot)$ is called the upper characteristic of A. *A finite from above operator A with characteristic $\tau_+(\cdot)$ is called τ–*causal *or τ_+–*finite.

The space of all τ_+–finite operators will be denoted by $\mathbf{A}^\tau$ and $\mathbf{A}^0 = \cup_\tau \mathbf{A}^\tau$. 0–causal operators are called *causal*. For $\phi \in H$ one can consider a trajectory $\{P_t \phi, \; t \in \mathbf{T}\}$ connecting ϕ and zero in H. Then (13) means that a τ–causal operator A considered as a shift operator along these trajectories does not depend on a future with respect to the resolution, namely it follows from the completeness of $\mathbf{P}_T$ that $P_t A \phi$ is independent of $P_s \phi$ for $s > t - \tau(t)$.

4.2 Linear causal filtering problem

Let $(H', \mathbf{P}'_T)$, $(H'', \mathbf{P}''_T)$ be Hermitian resolution spaces. Let $H = H' \times H''$ be equipped with the resolution defined by (12). We denote by $\mathbf{H}^\tau$ the space of all linear continuous τ–causal operators $h : H'' \to H'$. Let $D : H' \to H'$ be continuous with the adjoint $D^* : H' \to H'$. Then *the optimal linear causal filtering problem* is the minimization problem

$$J_{\phi'}(h) \to \inf_{h \in \mathbf{H}^\tau} \tag{14}$$

for every $\phi' \in H'$, where $J_{\phi'}(h)$ is defined by

$$J_{\phi'}(h) = E|\langle \phi', D(x - hy)\rangle|^2, \; h \in \mathbf{H}^\tau. \tag{15}$$

It turns out that the condition of the continuity of weight operators is very restrictive for the solution of the problem (14). In general, the optimal filtering problem in the class of continuous weight operators can be unsolvable or can be very complicated. At the same time if we drop the continuity condition the problem can be solved. It is relatively simple to check whether the weight operator of the determined optimal filter is continuous. In this way one can obtain solutions for the original continuous problem. We will apply the methods presented in [4], namely first we relax the problem (14) allowing h to be unbounded. Analyzing the solution of the relaxed problem we derive the conditions for the solvability of (14).

4.3 Generalized linear causal filtering problem

Let $\bar{H}'$, $\bar{H}''$ be the t–completions of H' and H'' respectively. Let $\bar{\mathbf{H}}^\tau$ be the space of all linear τ–causal operators $\bar{h} : \bar{H}'' \to \bar{H}'$, such that for every $t \in \mathbf{T}$ the operators $\bar{P}_t' \bar{h} \bar{h}^* \bar{P}_t' : P_t' H' \to P_t' H'$ are continuous. Assume z to be a random $\bar{H}$ element, and, therefore, $R_z = Ezz^*$, $z = \begin{bmatrix} x \\ y \end{bmatrix}$, is bounded on the space F of finite elements in H and can be then continuously extended to the whole of H. The problem is to find linear estimates of a random H' element x based on the realizations of a random H'' element y of the form

$$\hat{x} = \bar{h} y, \tag{16}$$

minimizing for every $t \in \mathbf{T}$ the functional

$$J^{(t)}(\bar{h}) = E|D\bar{P}_t(x - \hat{x})|^2_{H'}. \tag{17}$$

Note that $J^{(t)}(\bar{h})$ is finite for $\bar{h} \in \bar{\mathbf{H}}^\tau$, $t \in \mathbf{T}$, therefore the problem of the minimization

$$J^{(t)}(\bar{h}) \to \inf_{\bar{h} \in \bar{\mathbf{H}}^\tau} \tag{18}$$

for every $t \in \mathbf{T}$ is correctly posed. Let us reformulate the problem (18) now. For $\phi' \in H'$ we define

$$\begin{aligned} J^{(t)}_{\phi'}(\bar{h}) &= E|\langle \phi', DP_t'(x - \hat{x})\rangle_{H'}|^2 \\ &= E|\langle \phi', DP_t'(x - \bar{h}x)\rangle_{H'}|^2 \\ &= \langle \phi', DP_t'[R_x - R_{xy}\bar{h}^* - \bar{h}R_{yx} \\ &\quad + \bar{h}R_y\bar{h}^*]P_t'D^*\phi'\rangle_{H'}. \end{aligned} \tag{19}$$

Now, the problem (18) is equivalent to the problem

$$J^{(t)}_{\phi'}(\bar{h}) \to \inf_{\bar{h} \in \bar{\mathbf{H}}^\tau} \tag{20}$$

for every $\phi' \in H'$.

Theorem 3 *Let $R_z = Ezz^*$ satisfy*

(i) *The operators $R_z^{(t,t)} = \bar{P}_t R_z \bar{P}_t : P_t H \to P_t H$ are continuous for every $t \in \mathbf{T}$.*

(ii) *The operators $P_t'' R_y P_t'' : H'' \to H''$ are positive in the invariant subspace $P_t'' H''$ for every $t \in \mathbf{T}$.*

Then for every $t \in \mathbf{T}$ there exist $\hat{x}_t \in P_t' H'$ such that for every $\phi \in H'$ one has

$$E|\langle \phi', D(x - \hat{x}_t)\rangle_{H'}|^2 = \inf_{h \in \bar{\mathbf{H}}^\tau} E|\langle \phi', D(x - P_t' hy)\rangle_{H'}|^2.$$

The estimates $\hat{x}_t$ are given by

$$\hat{x}_t = R_{xy}^{(t,t-\tau(t))} (R_y^{(t-\tau(t),t-\tau(t))})^{-1} P''_{t-\tau(t)} y + Q_t P''_{t-\tau(t)} y, \tag{21}$$

where $R_{xy}^{(t,t-\tau(t))} = P_t' R_{xy} P''_{t-\tau(t)}$, $R_y^{(t,t)} = P_t'' R_y P_t''$, $(R_y^{(t,t)})^{-1}$ means the inverse of $R_y^{(t,t)}$ in the invariant subspace $P''_{t-\tau(t)} H''$ and any $Q_t : H'' \to H'$ such that $DQ_t Q_{R_y} = 0$. Moreover,

$$E|\langle \phi', D(x - \hat{x}_t)\rangle_{H'}|^2 \langle \phi', D[P_t' R_x P_t' - R_{xy}^{(t,t-\tau(t))} (R_y^{(t-\tau(t),t-\tau(t))})^{-1} R_{yx}^{(t-\tau(t),t)}] D^* \phi'\rangle_{H'}.$$

In view of the discussion above it is easy to see that problem (14) is solvable if and only if the solution of (18) is a continuous τ-causal operator. In this case the restriction of this operator to H'' gives a solution to (14). The detailed discussion and the solutions of these problems for $D = I_{H'}$ can be found in [4], [5], [7]. We will treat further the spaces with the discrete resolution of the identity.

4.4 Discrete resolutions of the identity

We assume now that $\mathbf{P}_T$ is a piecewise constant operator valued functional on $\mathbf{T}$ with at most countable number of discontinuity points without accumulations in $\mathbf{T}$. Let $\mathbf{t} = \{t_k, k \in \mathbf{K}\}$ be a finite or a countable ordered subset of $\mathbf{T}$ without accumulation points, $\mathbf{K} = \mathbf{Z} \cap (0, K)$, $t_0 = t_s$, $t_K = t_f$, K finite or $K = +\infty$. The *discrete resolution of the identity in H corresponding to* $\mathbf{t} \subset \mathbf{T}$ is the set $\mathbf{P_t} = \{P_t, t \in \mathbf{t}\}$. The family of the orthogonal projectors $Q_k = P_{t_k} - P_{t_{k-1}}$, $k \in \mathbf{K}$, determines the resolution $\mathbf{P_t}$ uniquely due to the relation $P_t = \sum_{k:t_k \le t} Q_k$. These projectors are mutually orthogonal: $Q_k Q_l = Q_l Q_k = 0_H$ for $k \neq l$.

Definition 2 *A family $\mathbf{Q}_K$ of the mutually orthogonal projectors Q_k is called* the orthogonal resolution of the identity *if $\mathbf{Q}_K$ is complete in a sense that $Q_k \to O_H$ for $k \to k_s$ and $\sum_{l \le k} Q_l \to I_H$ for $k \to k_f$. The pair $(H, \mathbf{Q}_K)$ is called the* discrete resolution space.

Every linear operator $R : H \to H$ can be decomposed with respect to $\mathbf{Q}_K$ into blocks $R_{kl} = Q_k R Q_l$ and $R = \sum_{k,l \in \mathbf{K}} R_{kl}$. The definitions of finiteness, causality and anticausality can be reformulated in terms of the discrete structure $\mathbf{Q}_K$. The function τ in Definition 1 is replaced by $\tau : \mathbf{t} \to \mathbf{t}$ with a property that $\tau(t_k) = t_l,\ k, l \in \mathbf{K}$ and the latter corresponds to a function $\kappa : \mathbf{K} \to \mathbf{K}$ such that $\tau(t_k) = t_{\kappa(k)}$. In analogy to the continuous case one has

Definition 3 *A linear operator $R : H \to H$ is called* κ–causal (strictly κ–causal, κ–anticausal) *if $R_{kl} = 0_H$ for $l > k - \kappa(k)$, ($l \ge k - \kappa(k)$, $l < k - \kappa(k)$) respectively. It is called* neutral *if its causal and anticausal.*

For a linear operator $R : H \to H$ we denote its κ–causal, anticausal and neutral components by $R_{[\kappa]} = \sum_{l \le k - \kappa(k)} R_{kl},\ R^{[\kappa]} = \sum_{l \ge k - \kappa(k)} R_{kl},\ R_{[[\kappa]]} = \sum_{l = k - \kappa(k)} R_{kl}$ respectively.

Now we are ready to formulate the optimal causal filtering problem for the discrete resolution space $H = H' \times H''$, H', H'' equipped with the orthogonal resolutions of the identity $\mathbf{Q}'_K$ and $\mathbf{Q}''_K$ respectively. Let $\mathbf{H}^\kappa$ denote the space of all κ–causal continuous operators $h : H'' \to H'$ and $\hat{x}_k = Q'_k \hat{x},\ y_k = Q''_k y,\ h_{kl} = Q'_k h Q''_l$. Then the problem is the linear estimation

$$\hat{x}_k = \sum_{l \le k - \kappa(k)} h_{kl} y_l \tag{22}$$

minimizing the functionals

$$J_{\phi'}(h) = E|\langle \phi', D(x - hy) \rangle|^2 \to \inf_{h \in \mathbf{H}^\kappa} \tag{23}$$

for every $\phi \in H'$. Note, that this is the same as the minimization of

$$J_k(h) = E|D(x_k - \hat{x}_k)|^2 \to \inf_{h \in \mathbf{H}^\kappa} \tag{24}$$

for every $k \in \mathbf{K}$, where $x_k = Q'_k x$.

In analogy with the continuous case we will treat the relaxed problem first, replacing the condition of the continuity of h by the continuity of $h_k = Q'_k h = \sum_{l \in \mathbf{K}} h_{kl} : H'' \to H'$ for every $k \in \mathbf{K}$. The space of all linear κ–causal operators for which all the correspondent operators h_k are continuous will be denoted by $\bar{\mathbf{H}}^\kappa$. Note that because J_k are finite when R_z is bounded, the problem

$$J_k(h) = E|D(x_k - \hat{x}_k)|^2 \to \inf_{h \in \bar{\mathbf{H}}^\kappa},\ k \in \mathbf{K} \tag{25}$$

is correctly posed. Note that if $\bar{H}', \bar{H}''$ are the completions of H', H'' in t–topology, then the space $\bar{\mathbf{H}}^\kappa$ is isomorphic to the space of all κ–causal operators from $\bar{H}''$ to $\bar{H}'$. The problem now becomes

$$J_k(h) = E|DQ'_k(x - \bar{h}y)|^2 \to \inf_{\bar{h} \in \bar{\mathbf{H}}^\kappa},\ k \in \mathbf{K}. \tag{26}$$

Theorem 4 *Let $R_z = Ezz^*$ be continuous and R_y satisfy $P_{t_k} R_y P_{t_k} \ge \epsilon P_{t_k}$ for some $\epsilon > 0$ and for every $k \in \mathbf{K}$. Then all the solutions $\bar{h}_{\text{opt}} : \bar{H}'' \to \bar{H}'$ of the problem (26) are given by*

$$\bar{h}_{\text{opt}} = \sum_{k \in \mathbf{K}} Q'_k R_{xy} P''_{t_{k-\kappa(k)}} (A_k)^{-1} P''_{t_{k-\kappa(k)}} + Q, \tag{27}$$

where $A_k = P''_{t_{k-\kappa(k)}} R_y P''_{t_{k-\kappa(k)}}$, and $Q \in \bar{\mathbf{H}}^\kappa$ satisfies $DQ'_k Q = 0$. One has

$$\inf_{h \in \bar{\mathbf{H}}^\kappa} J_k(\bar{h}) = J_k(\bar{h}_{\text{opt}})$$

$$= |DP'_{t_k}[R_x - R_{xy} P''_{t_{k-\kappa(k)}} (A_k)^{-1} P''_{t_{k-\kappa(k)}} R_{yx}] P'_{t_k} D^*|.$$

Theorem 5 *Let the assumptions of Theorem 4 be satisfied. Then the problem (24) is solvable if and only if the solution $\bar{h} : \bar{H}'' \to \bar{H}'$ of (26) is κ–bounded. In this case the image of H'' under $\bar{h}$ is contained in H' and the restriction $\bar{h}|_{H''}$ is the solution of (24).*

Note that taking finite partial sums in (27) one obtains minimizing sequences for the problem. These minimizing sequences are also available in the case when $\bar{h}$ is not κ–bounded and the problem (24) is not solvable in $\mathbf{H}^\kappa$.

5 Bode–Shannon representation of optimal filters

First we will briefly review the results on the spectral factorization of the operators which we need in order to discuss the application of Bode–Shannon theory (cf.[1],[4],[9],[10]) in our setting. The detailed discussion on various types of spectral factorization and separation of the operators can be found in [4].

Let $\mathbf{P}_T$ be a Hermitian resolution of the identity in H. As in the previous section we denote by $\bar{H}$ a t–completion of H and by $\mathbf{t}$ a discrete linearly ordered subset of $\mathbf{T}$. Let $\mathbf{G}_\mathbf{t}$ be a space of all bijective operators $\bar{G} : \bar{H} \to \bar{H}$ such that $\bar{P}_t \bar{G} \bar{P}_t$ and $\bar{P}_t \bar{G}^{-1} \bar{P}_t$ are continuous as operators from $P_t H$ to $P_t H$ for every $t \in \mathbf{t}$. Note, that $\mathbf{G}_\mathbf{t}$ contains the space of all causal, causally invertible operators in H.

Definition 4 *A continuous operator $G : H \to H$ is called* strongly spectrally factorizable *if there exist a continuous causal operator $U : H \to H$ with continuous and causal inverse, such that $G = UU^*$, where U^* is the adjoint of U.*

We call operator $\bar{G} \in \mathbf{G_t}$ *positive* if the operators $\bar{P}_t \bar{G} \bar{P}_t$, $\bar{P}_t \bar{G}^{-1} \bar{P}_t : P_t H \to P_t H$ are nonnegative for every $t \in \mathbf{t}$.

It is convenient in the discrete case $(H, \mathbf{Q}_K)$ to denote by $\mathbf{G}_K$ the space of all bijective operators $\bar{G} : \bar{H} \to \bar{H}$ such that for every $k \in \mathbf{K}$ the operators

$$\sum_{l=0}^{k} \sum_{m=0}^{k} \bar{Q}_l \bar{G} \bar{Q}_m, \quad \sum_{l=0}^{k} \sum_{m=0}^{k} \bar{Q}_l \bar{G}^{-1} \bar{Q}_m$$

are continuous from $H^k = \bigoplus_{l=0}^{k} \bar{Q}_l H$ to H^k. $\bar{G} \in \mathbf{G}_K$ is called *positive* if the operators in the definition of $\mathbf{G}_K$ are nonnegative for every $k \in \mathbf{K}$.

Let $H = H' \times H''$ be equipped with the orthogonal resolution of the identity given by $Q_k = \begin{bmatrix} Q'_k & 0 \\ 0 & Q''_k \end{bmatrix}$, $Q'_k \in \mathbf{Q}'_K$, $Q''_k \in \mathbf{Q}''_K$. The κ–causal filters are given by

$$\hat{x}_k = \sum_{l=0}^{k-\kappa(k)} h_{kl} y_l,$$

where $h_{kl} : Q''_l H'' \to Q'_k H'$ are linear continuous. The corresponding filter $\bar{h} : \bar{H}'' \to \bar{H}'$ is defined by having its blocks equal to h_{kl}.

Theorem 6 *Assume that $R_z \in \mathbf{G}_K$, $R_y : H'' \to H'$ is positive and $\kappa \geq 0$. Then all optimal linear filters for the discrete generalized linear causal filtering problem (26) are of the form*

$$\bar{h}_{\text{opt}} = [R_{xy}(U^{-1})^*]_{[\kappa]} U^{-1} + Q, \qquad (28)$$

where U is a causal operator strongly factorizing R_y, $[R_{xy}(U^{-1})^*]_{[\kappa]}$ is the κ–causal component of $R_{xy}(U^{-1})^* : \bar{H}'' \to \bar{H}'$ and any $Q \in \bar{\mathbf{H}}^\kappa$ such that $DQ = 0$. One has

$$\inf_{\bar{h} \in \bar{\mathbf{H}}^\kappa} J_k(\bar{h}) = J_k(\bar{h}_{\text{opt}}) = |DQ_k[R_x - R_{xy} R_y^{-1} R_{yx}$$
$$+ [R_{xy}(U^{-1})^*]_{[\bar{\kappa}]} ([R_{xy}(U^{-1})^*]_{[\bar{\kappa}]})^*] Q_k D^*|.$$

For $D = I_{H'}$ the only operator in (28) is obtain by taking $Q = 0$. This operator is called the Bode–Shannon weight operator and the filter (22) is called the Bode–Shannon filter.

If R_z is stable, R_y^{-1} exists and is continuous in H and stable operator $R_{xy}(U^{-1})^*$ has the stable κ–causal component, then the original linear optimal causal filtering problem (24) is solvable and all optimal weight operators are the restrictions of $\bar{h}_{\text{opt}}$ in (28) to H''. One can consult [4] for the application to the finite dimensional stationary processes, where these conditions are reduced to the conditions in terms of analytic functions. Although here we give the Bode-Shannon formula only for the optimal filtering problem with the discrete time, its analogue is valid for the continuous time as well.

References

[1] H. W. Bode, C. E. Shannon, *A simplified derivation of linear least square smoothing and prediction theory*, Proc. IRE, 38 (1950), pp. 417–425.

[2] A. Feintuch, R. Saeks, *System theory. A Hilbert space approach*, Academic Press, New York, 1982.

[3] V. N. Fomin, *Operator methods of the random processes linear filtering theory*, St.Petersburg University Publisher, St.Petersburg, 1995. [Russian]

[4] V. N. Fomin, *Optimal filtering. Vol. 1: Filtering of stochastic processes*, Kluwer Acad. Publ., Dordrecht, Boston, London, 1998.

[5] O. G. Gorshkov, V. N. Fomin, *Operator approaches to the time series filtering problem*, Vestnik S.-Peterburg. Univ. Mat. Mekh. Astronom., Ser.1, 4 (22) (1993), pp. 16–21. [Russian]

[6] A. N. Kolmogorov, *Interpolation and extrapolation of the stationary random sequences*, Izv. AN SSSR, Mathematics, 5 (1941), pp. 3–14. [Russian]

[7] O. A. Petrov, V. N. Fomin, *Linear filtering of random processes*, LGU publisher, Leningrad, 1991. [Russian]

[8] M. V. Ruzhansky, V. N. Fomin, *The optimal filter construction with the quadratic quality functional of general form*, Vestnik S.-Peterburg. Univ. Mat. Mekh. Astronom., Ser. 1, 4 (22) (1995), pp. 50–56. [Russian] (English Translation in Vestnik St.Petersburg Univ. Math.)

[9] M. V. Ruzhansky, V. N. Fomin, *Abstract optimal linear filtering*, SIAM J. Control Optim 38 (2000), pp. 1334–1352.

[10] N. Wiener, *The Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications*, Academic Press, New York, 1949.