World Academy of Science, Engineering and Technology
International Journal of Mathematical and Computational Sciences
Vol:6, No:7, 2012

Piecewise Interpolation Filter for Effective Processing of Large Signal Sets

Anatoli Torokhti and Stanley Miklavcic

Open Science Index, Mathematical and Computational Sciences Vol:6, No:7, 2012 publications.waset.org/12739/pdf

Abstract—Suppose K_Y and K_X are large sets of observed and reference signals, respectively, each containing N signals. Is it possible to construct a filter $\mathcal{F} : K_Y \to K_X$ that requires a priori information only on *few signals*, $p \ll N$, from K_X but *performs better* than the known filters based on a priori information on *every* reference signal from K_X? It is shown that the positive answer is achievable under quite unrestrictive assumptions. The device behind the proposed method is based on a special extension of the piecewise linear interpolation technique to the case of random signal sets. The proposed technique provides a single filter to process any signal from the arbitrarily large signal set. The filter is determined in terms of pseudo-inverse matrices so that it always exists.

Keywords—Wiener filter, filtering of stochastic signals.

I. Introduction

A. Motivations

A PURPOSE of the proposed new filtering methodology is to provide an effective way to process large signal sets. The device behind the proposed method is quite simple and is based on a special extension of the piecewise linear interpolation technique to the case of random signal sets. At the same time, such a device is not straightforward and requires the careful substantiation presented in Sections II-C, III-D, IV-B and IV-D below.

B. Motivations

The problem under consideration is motivated by the following observations.

1) FILTERING OF LARGE SETS OF SIGNALS; LESS INITIAL INFORMATION FOR BETTER FILTERING: Suppose we need to transform a set of signals K_Y to another set of signals K_X. The signals are represented by *finite* random vectors[1]. A major difficulty and inconvenience common to many known filtering methodologies (see, for example, [2]–[10], [12], [14], [23], [24], [26]) is that they require a priori information on *each reference signal* to be estimated[2]. In particular, the filters in [23], [24], [26] are based on the use of either the reference signal $\mathbf{x} \in K_X$ itself, as in [23], [24], or its estimate, as in [26]. The Wiener filtering approach (see, for example, [2]–[14], [24], [26]) assumes that a covariance matrix formed from a reference signal, $\mathbf{x} \in K_X$, and an observed signal, $\mathbf{y} \in K_Y$, is known or can be estimated. The latter can be done, for instance, from samples of $\mathbf{x}$ and $\mathbf{y}$. In particular, this means that the reference signal $\mathbf{x}$ can be measured.

In the case of processing large signal sets, such restrictions become much more inconvenient.

The major motivating question for this work is as follows. Let $\mathcal{F} : K_Y \to K_X$ denote a filter that estimates a large set of reference signals, K_X, from a large set of observed signals, K_Y. Each set contains N signals. Is it possible to construct a filter $\mathcal{F}$ that requires a priori information only on *few signals*, $p \ll N$, from K_X but *performs better* than the known filters based on a prior information on *every* reference signal from K_X? We denote such a filter by $\mathcal{F}^{(p-1)}$.

It is shown in Sections II-C and IV-D that the positive answer is achievable under quite unrestrictive assumptions. The required features of filter $\mathcal{F}^{(p-1)}$ are satisfied by its special structure described in Sections II-C, III-A and III-D. The related conditions are also considered in those Sections.

2) FILTERING BASED ON IDEA OF PIECEWISE FUNCTION INTERPOLATION: The specific structure of the proposed filter follows from the extension of piecewise function interpolation [15]. This is because the technique of piecewise function interpolation [15] has significant advantages over the methods of linear and polynomial approximation used in known filtering techniques (such as, for example, those in [6], [10]).

The structure of the proposed filter is presented in Sections II-C, III-A and IV-B below.

3) EXPLOITING PSEUDO-INVERSE MATRICES IN THE FILTER MODEL: Most of the known filtering techniques, for example, those ones in [2]–[4], [7]–[9], [12], [24], [26], are based on exploiting inverse matrices in their mathematical models. In the cases of grossly corrupted signals or erroneous measurements those inverse matrices may not exist and, thus, those filters cannot be applied.

The filter proposed here avoids this drawback since its model is based on exploiting pseudo-inverse matrices. As a result, the proposed filter always exist. That is, it processes any kind of noisy signals. An extension of the filtering techniques to the case of implementation of the pseudo-inverse matrices is done on the basis of theory presented in [6].

4) COMPUTATIONAL WORK: Let m and n be the number of components of $\mathbf{x} \in K_X$ and of $\mathbf{y} \in K_Y$, respectively, where K_X and K_Y each contains N signals. The known filtering techniques (e.g. see [2]–[9], [12], [24], [26]), applied to $\mathbf{x}$ and $\mathbf{y}$, require the computation of a product of an $m \times n$

A. Torokhti is with the CIAM, School of Mathematics & Statistics, University of South Australia, SA, Australia, e-mail: anatoli.torokhti@unisa.edu.au.
S. Miklavcic is with the Phenomics and Bioinformatics Research Centre, School of Mathematics and Statistics, University of South Australia, SA, Australia, e-mail: stan.miclavcic@unisa.edu.au.
Manuscript received February 29, 2012.

[1]We say a random vector $\mathbf{x}$ is finite if its realization has a finite number of scalar components.

[2]To the best of our knowledge, the exception is the methodology in [11], [13] where the filtering techniques exploit information on reference signals in the form of the vector obtained from averaging over reference signal sets.

Open Science Index, Mathematical and Computational Sciences Vol:6, No:7, 2012 publications.waset.org/12739/pdf

matrix and an $n \times n$ matrix, as well as the computation of an $n \times n$ inverse or pseudo-inverse matrix for *each pair* of signals $\mathbf{x} \in K_X$ and $\mathbf{y} \in K_Y$. This requires $O(2mn^2)$ and $O(26n^3)$ flops, respectively [27]. Thus, for the processing of all signals in K_X and K_Y, the filters in [2]–[9], [12], [24], [26] require $O(2mn^2N) + O(26n^3N)$ operations.

Alternatively, K_X and K_Y can be represented by vectors, $\boldsymbol{\chi}$ and $\boldsymbol{\gamma}$, each with mN and nN components, respectively. In such a case, the techniques in [2]–[9], [12], [24], [26] can be applied to $\boldsymbol{\chi}$ and $\boldsymbol{\gamma}$ as opposed to each signals in K_X and K_Y. The computational requirement is then $O(2mn^2N^2)$ and $O(26n^3N^3)$ operations, respectively [27].

In both cases, but especially when N is large, the computational work associated with the approaches [2]–[9], [12], [24], [26] becomes unreasonable hard.

For the filter $\mathcal{F}^{(p-1)}$ to be introduced below, the associated computational work is substantially less. This is because $\mathcal{F}^{(p-1)}$ requires the computation of only p pseudo-inverse matrices associated with p selected signals in K_X, where p is much less than the number of signals in K_X. Therefore, for processing of the signal sets, K_X and K_Y, $\mathcal{F}^{(p-1)}$ requires only $O(2mn^2p) + O(26n^3p)$ flops where $p \ll N$.

C. Relevant works

Some particular filtering techniques relevant to the method proposed below are as follows.

1) GENERIC OPTIMAL LINEAR (GOL) FILTER [6]: The generic optimal linear (GOL) filter in [6] is a generalization of the Wiener filter to the case when covariance matrix is not invertible and observable signal is arbitrarily noisy (i.e. when, in particular, noise is not necessarily additive and Gaussian). The GOL filter has been developed for processing an *individual* stochastic signal. Some ideas from [6] are used in the proof of Theorem 1 below.

2) SIMPLICIAL CANONICAL PIECEWISE LINEAR FILTER [24]: A complex Wiener adaptive filter was developed in [24] from the two-dimensional complex-valued simplicial canonical piecewise linear filter [25]. The filter in [24] was developed for the processing of an *individual* stochastic signal and can be exploited when the reference signal is known and a 'covariance-like' matrix is invertible. The latter precludes an application to the signal types when the matrices used in [24] are not invertible for the signals. Similarly, the filters studied in [9], [12] were developed for the processing of a single signal when the covariance matrices are invertible.

For the filter proposed here, these restrictions are removed.

3) ADAPTIVE PIECEWISE LINEAR FILTER [23]: A piecewise linear filter in [23] was proposed for a fixed image denoising (given by a matrix), corrupted by an *additive* Gaussian noise. That is, the method involved a non stochastic reference signal and required its knowledge. No theoretical justification for the filter was given in [23].

4) AVERAGING POLYNOMIAL FILTER [11], [13]: The averaging polynomial filter proposed in [11], [13] was developed for the purpose of processing infinite signal sets. The filter was based on an argument involving the'averaging' over sets of signals under consideration. This device allows one to determine a single filter for the processing of infinite signal sets. At the same time, it leads to an increase in the associated error when signals differ considerably from each other.

5) OTHER RELEVANT FILTERS: The technique developed in [14] is an extension of the GOL filter to the constraint problem with respect to the filter rank. It concerns data compression.

The methods in [7], [8], [16], [17] have been developed for deterministic signals. Motivated by the results achieved in [16], [17], adaptive filters were elaborated in [18]. A theoretical basis for the device proposed in [16], [17] is provided in [19].

We note that the idea of piecewise linear filtering has been used in the literature in several *very different conceptual frameworks*, despite exploiting some very similar terms (as in [16]–[25]). At the same time, a common feature of those techniques is that they were developed for the processing of a single signal, not of large signal sets as in this paper. In particular, piecewise linear filters in [20] have been obtained by arranging linear filters and thresholds in a tree structure. Piecewise linear filters discussed in [21] were developed using so-called threshold decomposition, which is a segmentation operator exploited to split a signal into a set of multilevel components. Filter design methods for piecewise linear systems proposed in [22] were based on a piecewise Lyapunov function.

D. Difficulties associated with the known filtering techniques

Basic difficulties associated with applying the known filtering techniques to the case under consideration (i.e. to processing of large signal sets, K_X and K_Y) are that:

(i) they require an information on *each* reference signal (in the form of a sample, for example),

(ii) matrices used in the known filters can be not invertible and then the filter does not exist, and

(ii) the associated computation work may require a very long time. MATLAB can be out of memory for computing the GOL filter [6] when each of sets K_X and K_Y is represented by a long vector (this option has been discussed in Section I-B4 above).

E. Differences from the known filtering techniques

The differences from the known filtering techniques discussed above are as follows.

(i) We consider a single filter that processes *arbitrarily large* input-output sets of stochastic signal-vectors. The known filters [2]–[10], [12], [14], [16]–[26] have been developed for the processing of an individual signal-vector only. In the case of their application to arbitrarily large signal sets, they imply difficulties described in Sections I-B and I-D above.

(ii) As a result, our piecewise linear filter model (Section III), the statement of the problem (Section III-C below) and consequently, the device of its solution (Section IV below) are different from those considered in [16]–[25]. In this regard, see also Section I-C5.

(iii) The above naturally leads to a new structure of the filter (presented in Sections III-D and IV-B below) which is very different from the known ones.

Open Science Index, Mathematical and Computational Sciences Vol:6, No:7, 2012 publications.waset.org/12739/pdf

F. Contribution

In general, for the processing of large data sets, the proposed filter allows us to achieve better results in comparison with the known techniques in [2]–[26]. In particular, it allows us to

(i) achieve a desired accuracy in signal estimation[3],

(ii) exploit a priori information only on *few reference signals*, p, from the set K_X that contains $N \gg p$ signals or even infinite number of signals,

(iii) find *a single* filter to process any signal from the *arbitrarily large* signal set,

(vi) determine the filter in terms of pseudo-inverse matrices so that the filter always exists, and

(v) decrease the computational load compared to the related known techniques.

II. Some preliminaries

A. Notation

The signal sets we consider are, in fact, special representations of time series.

Let (Ω, Σ, μ) be a probability space[4], and K_X and K_Y be arbitrarily large sets of signals such that

$$K_X = \{\mathbf{x}(t,\cdot) \in L^2(\Omega, \mathbb{R}^m) \mid t \in T\} \quad \text{and}$$

$$K_Y = \{\mathbf{y}(t,\cdot) \in L^2(\Omega, \mathbb{R}^n) \mid t \in T\}$$

where $T := [a,\ b] \subseteq \mathbb{R}$. We interpret $\mathbf{x}(t,\cdot)$ as a reference signal and $\mathbf{y}(t,\cdot)$ as an observable signal, an input to the filter $\mathcal{F}$ studied below[5]. The variable $t \in T \subseteq \mathbb{R}$ represents time.[6] Then, for example, the random signal $\mathbf{x}(t,\cdot)$ can be interpreted as an arbitrary stationary time series.

Let $\{t_k\}_1^p \subset T$ be a sequence of fixed time-points such that

$$a = t_1 < \ldots < t_p = b. \tag{1}$$

Because of the partition (1), the sets K_Y and K_X are divided in 'smaller' subsets $K_{X,1}, \ldots, K_{X,p-1}$ and $K_{Y,1}, \ldots, K_{Y,p-1}$, respectively, so that, for each $j = 1, \ldots, p$,

$$K_{X,j} = \{\mathbf{x}(t,\cdot) \mid t_j \le t \le t_{j+1}\} \quad \text{and} \tag{2}$$

$$K_{Y,j} = \{\mathbf{y}(t,\cdot) \mid t_j \le t \le t_{j+1}\}. \tag{3}$$

Therefore, K_Y and K_X can now be represented as $K_X = \bigcup_{j=1}^{p-1} K_{X,j}$ and $K_Y = \bigcup_{j=1}^{p-1} K_{Y,j}$.

[3]This means that any desired accuracy is achieved theoretically, as is shown in Section IV-D below. In practice, of course, the accuracy is increased to a prescribed reasonable level.

[4]As usually, $\Omega = \{\omega\}$ is the set of outcomes, Σ a σ–field of measurable subsets in Ω and $\mu : \Sigma \to [0,1]$ an associated probability measure on Σ. In particular, $\mu(\Omega) = 1$.

[5]In an intuitive way $\mathbf{y}$ can be regarded as a noise-corrupted version of $\mathbf{x}$. For example, $\mathbf{y}$ can be interpreted as $\mathbf{y} = \mathbf{x} + \mathbf{n}$ where $\mathbf{n}$ is white noise. In this paper, we do not restrict ourselves to this simplest version of $\mathbf{y}$ and assume that the dependence of $\mathbf{y}$ on $\mathbf{x}$ and $\mathbf{n}$ is arbitrary.

[6]More generally, T can be considered as a set of parameter vectors $\alpha = (\alpha^{(1)}, \ldots, \alpha^{(q)})^T \in C^q \subseteq \mathbb{R}^q$, where C^q is a q-dimensional cube, i.e., $\mathbf{y} = \mathbf{y}(\alpha,\cdot)$ and $\mathbf{x} = \mathbf{x}(\alpha,\cdot)$. One coordinate, say $\alpha^{(1)}$ of α, could be interpreted as time.

B. Brief description of the problem

Given two *arbitrarily large* sets of random signals, K_Y and K_X, find a single filter $\mathcal{F} : K_Y \to K_X$ that estimates the signal $\mathbf{x} \in K_X$ with a controlled, associated error. Note that in our formulation the set K_Y can be finite or infinite.

C. Brief description of the method

The solution of the above problem is based on the representation of the proposed filter in the form of a sum with $p-1$ terms $\mathcal{F}_1, \ldots, \mathcal{F}_{p-1}$ where each term, $\mathcal{F}_j$, is interpreted as a particular sub-filter (see (4) and (5) below). Such a filter is denoted by $\mathcal{F}^{(p-1)} : K_Y \to K_X$.

The sub-filter $\mathcal{F}_j$ transforms signals that belong to 'piece' $K_{Y,j}$ of set K_Y to signals in 'piece' $K_{X,j}$ of K_X, i.e. $\mathcal{F}_j : K_{Y,j} \to K_{X,j}$. Each sub-filter $\mathcal{F}_j$ depends on two parameters, α_j and $\mathcal{B}_j$.

The prime idea is to determine $\mathcal{F}_j$ (i.e. α_j and $\mathcal{B}_j$) separately, for each $j = 1, \ldots, p-1$. The required α_j and $\mathcal{B}_j$ follow from the solutions of the equation (12) and an associated minimization problem (12) (see Sections III-D and IV-B below). This procedure adjusts $\mathcal{F}_j$ so that the error associated with the estimation of $\mathbf{x}(t,\cdot) \in K_{X,j}$ is minimal.

A motivation for such a structure of the filter $\mathcal{F}^{(p-1)}$ is as follows. The method of determining α_j and $\mathcal{B}_j$ provides an estimate $\mathcal{F}_j[\mathbf{y}(t,\cdot)]$ that interpolates $\mathbf{x}(t,\cdot) \in K_{X,j}$ at $t = t_j$ and $t = t_{j+1}$. In other words, the filter is flexible to variations in the sets of observed and reference signals K_Y and K_X, respectively. Due to this way of determining $\mathcal{F}_j$, it is natural to expect that the processing of a 'smaller' signal set, $K_{Y,j}$, may lead to a *smaller associated error* than that for the processing of the whole set K_Y by a filter which is not specifically adjusted to each particular piece $K_{Y,j}$.

As a result, $\mathcal{F}^{(p-1)}[\mathbf{y}(t,\cdot)]$ represents a special *piecewise interpolation procedure* and, thus, should be attributed with the associated advantages such as, for example, the high accuracy of estimation.

In Section IV-D, this observation is confirmed. In Section IV-E, it is also shown that the proposed technique allows us to avoid the difficulties discussed in Section I-D above.

III. Description of the problem.

A. Piecewise linear filter model

Let $\mathcal{F}^{(p-1)} : K_Y \to K_X$ be a filter such that, for each $t \in T$,

$$\mathcal{F}^{(p-1)}[\mathbf{y}(t,\cdot)] = \sum_{j=1}^{p-1} \delta_j \mathcal{F}_j[\mathbf{y}(t,\cdot)], \tag{4}$$

where

$$\mathcal{F}_j[\mathbf{y}(t,\cdot)] = \alpha_j + \mathcal{B}_j[\mathbf{y}(t,\cdot)] \tag{5}$$

and

$$\delta_j = \begin{cases} 1, & \text{if } t_j \le t \le t_{j+1}, \\ 0, & \text{otherwise.} \end{cases}$$

Here, $\mathcal{F}_j$ is a sub-filter defined for $t_j \le t \le t_{j+1}$. In (5), $\alpha_j = [\alpha_j^{(1)}, \ldots, \alpha_j^{(m)}]^T \in \mathbb{R}^m$ and $\mathcal{B}_j : L^2(\Omega, \mathbb{R}^n) \to L^2(\Omega, \mathbb{R}^m)$ is a linear operator given by a matrix $B_j \in \mathbb{R}^{m \times n}$, so that

$$[\mathcal{B}_j(\mathbf{y})](t,\omega) = B_j[\mathbf{y}(t,\omega)].$$

Open Science Index, Mathematical and Computational Sciences Vol:6, No:7, 2012 publications.waset.org/12739/pdf

Thus, $\mathcal{F}_j$ is defined by a matrix $F_j \in \mathbb{R}^{m\times n}$ such that

$$F_j[\mathbf{y}(t,\omega)] = \alpha_j + B_j[\mathbf{y}(t,\omega)]. \quad (6)$$

Filter $\mathcal{F}^{(p-1)}$ defined by (4)–(6) is called the *piecewise* filter[7].

B. Assumptions

In the known approaches related to filtering of stochastic signals (e.g. see [2]–[14], [24], [26]), it is assumed that covariance matrices formed from the reference signal and observed signal are known or can be estimated.

The assumption used here is similar. The covariance matrices that are assumed to be known or can be estimated, are formed from *selected* signal pairs $\{\mathbf{x}(t_j,\cdot),\mathbf{y}(t_j,\cdot)\}$ with $j = 1,\dots,p$ and p to be a small number[8], $p \ll N$, where N is the number of signals in K_X or K_Y.

C. The problem

In (4)–(6), parameters of the filter $\mathcal{F}^{(p-1)}$, i.e. vector α_j and matrix B_j, for $j = 1,\dots,p-1$, are unknown. Therefore, under the assumptions described in Section III-B, the problem is to determine α_j and B_j, for $j = 1,\dots,p-1$. The related problem is to estimate an error associated with the filter $\mathcal{F}^{(p-1)}$.

Solutions to the both problems are given in Sections IV-B and IV-D, respectively. In particular, in the following Section III-D, interpolation conditions (8) and (12) are introduced that lead to a determination of α_j and B_j.

D. Interpolation conditions

Let us denote

$$\|\mathbf{x}(t_j,\cdot)\|_\Omega^2 = \int_\Omega \|\mathbf{x}(t_j,\omega)\|_2^2 d\mu(\omega) \quad (7)$$

where $\|\mathbf{x}(t_j,\omega)\|_2$ is the Euclidean norm of $\mathbf{x}(t_j,\omega) \in \mathbb{R}^m$.

For $t = t_1$, let $\widehat{\mathbf{x}}(t_1,\cdot)$ be an estimate of $\mathbf{x}(t_1,\cdot)$ determined by known methods [2]–[14], [24], [26]. This is *the initial condition* of the proposed technique.

For $j = 1,\dots,p-1$, each sub-filter F_j in (5)–(6) is defined so that α_j and $\mathcal{B}_j$ satisfy the conditions as follows.

Sub-filter $\mathcal{F}_1$: For $j = 1$, α_1 and $\mathcal{B}_1$ solve

$$\widehat{\mathbf{x}}(t_1,\cdot) = \alpha_1 + \mathcal{B}_1[\mathbf{y}(t_1,\cdot)] \quad \text{and} \quad (8)$$

$$\min_{\mathcal{B}_1} \|[\mathbf{x}(t_2,\cdot) - \alpha_1] - \mathcal{B}_1[\mathbf{y}(t_2,\cdot)]\|_\Omega^2, \quad (9)$$

respectively. Then an estimate of $\mathbf{x}(t,\cdot)$, $\widehat{\mathbf{x}}(t,\cdot)$, for $t \in [t_1,\ t_2]$, is determined as

$$\widehat{\mathbf{x}}(t,\cdot) = \mathcal{F}_1[\mathbf{y}(t,\cdot)] = \widehat{\mathbf{x}}(t_1,\cdot) + \mathcal{B}_1[\mathbf{y}(t,\cdot) - \mathbf{y}(t_1,\cdot)] \quad (10)$$

where α_1 and $\mathcal{B}_1$ satisfy (8). In particular, $\alpha_1 = \widehat{\mathbf{x}}(t_1,\cdot) - \mathcal{B}_1[\mathbf{y}(t_1,\cdot)]$ and

$$\widehat{\mathbf{x}}(t_2,\cdot) = \mathcal{F}_1[\mathbf{y}(t_2,\cdot)].$$

[7]Hereinafter, we will use a non-curly symbol to denote an operator and associated matrix (e.g., the operator $\mathcal{F}_j : L^2(\Omega,\mathbb{R}^n) \to L^2(\Omega,\mathbb{R}^m)$ and the associated matrix $F_j \in \mathbb{R}^{m\times n}$ are denoted by F_j).

[8]It is worthwhile to note that it is *not* assumed that the covariance matrices are known for *each* signal pair from $K_X \times K_Y$, $\{\mathbf{x}(t,\cdot),\mathbf{y}(t,\cdot)\}$ with $t \in [a,\ b]$.

Extending this procedure up to $j = k-1$, where $k = 3,\dots,p$, we set the following. Let $\widehat{\mathbf{x}}(t_{k-1},\cdot)$ be an estimate of $\mathbf{x}(t_{k-1},\cdot)$ defined by the preceding steps as

$$\widehat{\mathbf{x}}(t_{k-1},\cdot) = \mathcal{F}_{k-2}[\mathbf{y}(t_{k-1},\cdot)]. \quad (11)$$

Then sub-filter $\mathcal{F}_{k-1}$ is defined as follows.

Sub-filter $\mathcal{F}_{k-1}$: For $j = k-1$, α_{k-1} and $\mathcal{B}_{k-1}$ solve

$$\widehat{\mathbf{x}}(t_{k-1},\cdot) = \alpha_{k-1} + \mathcal{B}_{k-1}[\mathbf{y}(t_{k-1},\cdot)] \quad \text{and} \quad (12)$$

$$\min_{\mathcal{B}_{k-1}} \|[\mathbf{x}(t_k,\cdot) - \alpha_{k-1}] - \mathcal{B}_{k-1}[\mathbf{y}(t_k,\cdot)]\|_\Omega^2, \quad (13)$$

respectively. Then an estimate of $\mathbf{x}(t,\cdot)$, $\widehat{\mathbf{x}}(t,\cdot)$, for $t \in [t_{k-1},\ t_k]$, is determined as

$$\begin{aligned}\widehat{\mathbf{x}}(t,\cdot) &= \mathcal{F}_{k-1}[\mathbf{y}(t,\cdot)] \\ &= \widehat{\mathbf{x}}(t_{k-1},\cdot) + \mathcal{B}_1[\mathbf{y}(t,\cdot) - \mathbf{y}(t_{k-1},\cdot)]. \quad (14)\end{aligned}$$

The conditions (8) and (12) are motivated by the device of piecewise function interpolation and associated advantages [15].

Filter $\mathcal{F}^{(p-1)}$ of the form (4)–(5) with α_j and $\mathcal{B}_j$ satisfying (8) and (12) is called the *piecewise linear interpolation* filter. The pair of signals $\{\mathbf{x}(t_k,\cdot),\mathbf{y}(t_k,\cdot)\}$ associated with time t_k defined by (1) is called the *interpolation pair*.

IV. Main results

A. General device

In accordance withe the scheme presented in Sections III-A and III-D above, an estimate of the reference signal $\mathbf{x}(t,\cdot)$, for any $t \in T = [a,\ b]$, by the *piecewise linear interpolation* filter $\mathcal{F}^{(p-1)}$, is given by

$$\widehat{\mathbf{x}}(t,\cdot) = \mathcal{F}^{(p-1)}[\mathbf{y}(t,\cdot)] = \sum_{j=1}^{p-1} \delta_j \mathcal{F}_j[\mathbf{y}(t,\cdot)], \quad (15)$$

where, for each $j = 1,\dots,p-1$, the sub-filter $\mathcal{F}_j$ is given by (5), and is defined from the interpolation conditions (8) and (12).

Below, we show how to determine $\mathcal{F}_j$ to satisfy the conditions (8) and (12).

B. Determination of piecewise linear interpolation filter

Let us denote

$$\mathbf{z}(t_j,t_{j+1},\cdot) = \mathbf{x}(t_{j+1},\cdot) - \widehat{\mathbf{x}}(t_j,\cdot), \quad (16)$$

$$\mathbf{w}(t_j,t_{j+1},\cdot) = \mathbf{y}(t_{j+1},\cdot) - \mathbf{y}(t_j,\cdot). \quad (17)$$

We need to represent $\mathbf{z}(t_j,t_{j+1},\cdot)$ and $\mathbf{w}(t_j,t_{j+1},\cdot)$ in terms of their components as follows:

$$\mathbf{z}(t_j,t_{j+1},\cdot) = [\mathbf{z}^{(1)}(t_j,t_{j+1},\cdot),\dots,\mathbf{z}^{(m)}(t_j,t_{j+1},\cdot)]^T$$

and

$$\mathbf{w}(t_j,t_{j+1},\cdot) = [\mathbf{w}^{(1)}(t_j,t_{j+1},\cdot),\dots,\mathbf{w}^{(n)}(t_j,t_{j+1},\cdot)]^T,$$

where $\mathbf{z}^{(j)}(t_j,t_{j+1},\cdot) \in L^2(\Omega,\mathbb{R})$ and $\mathbf{w}^{(i)}(t_j,t_{j+1},\cdot) \in L^2(\Omega,\mathbb{R})$ are random variables, for all $j = 1,\dots,m$.

Then we can introduce the covariance matrix

$$E_{z_j w_j} = \left\{\left\langle \mathbf{z}^{(i)}(t_j,t_{j+1},\cdot), \mathbf{w}^{(k)}(t_j,t_{j+1},\cdot)\right\rangle\right\}_{i,k=1}^{m,n}, \quad (18)$$

Open Science Index, Mathematical and Computational Sciences Vol:6, No:7, 2012 publications.waset.org/12739/pdf

where

$$\left\langle \mathbf{z}^{(i)}(t_j, t_{j+1}, \cdot), \mathbf{w}^{(k)}(t_j, t_{j+1}, \cdot) \right\rangle = \int_\Omega \mathbf{z}^{(i)}(t_j, t_{j+1}, \omega) \mathbf{w}^{(k)}(t_j, t_{j+1}, \omega)\ d\mu(\omega).$$

Below, $M^\dagger$ is the Moor-Penrose generalized inverse of a matrix M.

Now, we are in a position to establish the main results.

Theorem 1: Let

$$K_{_X} = \{\mathbf{x}(t, \cdot) \in L^2(\Omega, \mathbb{R}^m) \mid t \in T = [a,\ b]\}$$

and

$$K_{_Y} = \{\mathbf{y}(t, \cdot) \in L^2(\Omega, \mathbb{R}^n) \mid t \in T = [a,\ b]\}$$

be sets of reference signals and observed signals, respectively. Let $t_j \in [a,\ b]$, for $j = 1, \ldots, p$, be such that

$$a = t_1 < \ldots < t_p = b.$$

For $t = t_1$, let $\widehat{\mathbf{x}}(t_1, \cdot)$ be a known estimate of $\mathbf{x}(t_1, \cdot)$[9]. Then, for any $t \in [a,\ b]$, the proposed piecewise linear interpolation filter $\mathcal{F}^{(p-1)} : L^2(\Omega, \mathbb{R}^n) \to L^2(\Omega, \mathbb{R}^m)$ transforming any signal $\mathbf{y}(t, \cdot) \in L^2(\Omega, \mathbb{R}^m)$ to an estimate of $\mathbf{x}(t, \cdot)$, $\widehat{\mathbf{x}}(t, \cdot)$, is given by

$$\widehat{\mathbf{x}}(t, \cdot) = \mathcal{F}^{(p-1)}[\mathbf{y}(t, \cdot)] = \sum_{j=1}^{p-1} \delta_j \mathcal{F}_j[\mathbf{y}(t, \cdot)] \qquad (19)$$

where

$$\mathcal{F}_j[\mathbf{y}(t, \cdot)] = \widehat{\mathbf{x}}(t_j, \cdot) + \mathcal{B}_j[\mathbf{y}(t, \cdot) - \mathbf{y}(t_j, \cdot)], \qquad (20)$$

$$\widehat{\mathbf{x}}(t_j, \cdot) = \mathcal{F}_{j-1}[\mathbf{y}(t_j, \cdot)] \quad (\text{for } j = 2, \ldots, p-1), \qquad (21)$$

$$B_j = E_{z_j w_j} E^\dagger_{w_j w_j} + M_{B_j}[I_n - E_{w_j w_j} E^\dagger_{w_j w_j}], \qquad (22)$$

and where I_n is the $n \times n$ identity matrix and M_{B_j} is an $m \times n$ arbitrary matrix.

Proof: The proof of Theorem 1 is given in the Appendix. □

It is worthwhile to observe that, due to an arbitrary matrix M_{B_j} in (22), the filter $\mathcal{F}^{(p-1)}$ is not unique. In particular, M_{B_j} can be chosen as the zero matrix $\mathbb{O}$ similarly to the generic optimal linear [6] (which is also not unique by the same reason).

C. Numerical realization of filter $\mathcal{F}^{(p-1)}$ and associated algorithm

1) Numerical realization : In practice, the set $T = [a,\ b]$ (see Section II-A) is represented by a finite set $\{\tau_k\}_{k=1}^N$, i.e. $[a,\ b] = [\tau_1, \tau_2, \ldots, \tau_N]$ where $a \le \tau_1 < \tau_2 < \ldots < \tau_N \le b$.

For $k = 1, \ldots, N$, the estimate of $\mathbf{x}(\tau_k, \cdot)$, $\widehat{\mathbf{x}}(\tau_k, \cdot)$, and observed signal $\mathbf{y}(\tau_k, \cdot)$ are represented by $m \times q$ and $n \times q$ matrices

$$\widehat{X}^{(k)} = [\widehat{\mathbf{x}}(\tau_k, \omega_1), \ldots, \widehat{\mathbf{x}}(\tau_k, \omega_q)]$$

and

$$Y^{(k)} = [\mathbf{y}(\tau_k, \omega_1), \ldots, \mathbf{y}(\tau_k, \omega_q)].$$

[9]As it has been mentioned in Section III-D, $\widehat{\mathbf{x}}(t_1, \cdot)$ can be determined by the known methods.

The sequence of fixed time-points $\{t_k\}_1^p \subset [a,\ b]$ introduced in (1) is such that

$$\tau_1 = t_1 < \ldots < t_p = \tau_N, \qquad (23)$$

where

$$t_1 = \tau_{n_0}, \quad t_2 = \tau_{n_0+n_1}, \quad \ldots, \quad t_p = \tau_{n_0+n_1 \ldots + n_{p-1}},$$

and where $n_0 = 1$ and $n_1, \ldots, n_{p-1}$ are positive integers such that $N = n_0 + n_1 + \ldots + n_{p-1}$.

For $j = 1, \ldots, p$, signal $\mathbf{y}(tj, \cdot)$ associated with t_j in (23) is represented by

$$Y_j = [\mathbf{y}(t_j, \omega_1), \ldots, \mathbf{y}(t_j, \omega_N)].$$

2) Algorithm : As it has been mentioned in Section III-D, it is supposed that, for $t = t_1$, an estimate of X_1, $\widehat{X}_1$, is known and can be determined by the known methods. This is the initial condition of the proposed technique.

On the basis of the results obtained in Sections III-D and IV-B, the performance algorithm of the proposed filter consists of the following steps. For $j = 1 \ldots, p$, we write $N_j = n_0 + n_1 + \ldots + n_{j-1}$.

Initial parameters: $Y^{(1)}, \ldots, Y^{(N)}$, $\{t_j\}_{j=1}^p$ (see (23)), $\{E_{z_j w_j}\}_{j=1}^p$, $\{E_{w_j w_j}\}_{j=1}^p$ (see (16) and (18)), $\widehat{X}_1$, $n_0 = 1$ and $M_{B_j} = \mathbb{O}$, for $j = 1, \ldots, p-1$.

(Possible ways to get estimates of $E_{z_j w_j}$ and $E_{w_j w_j}$ are discussed below in Section IV-E.)

Final parameters: $\widehat{X}^{(2)}$, $\widehat{X}^{(3)}$,..., $\widehat{X}^{(N)}$.

Algorithm:

- for $j = 1$ to p do
  begin
  $$B_j = E_{z_j w_j} E^\dagger_{w_j w_j};$$
  - for $k = N_{j-1} + 1$ to N_j do
    begin
    $$\widehat{X}^{(k)} = \widehat{X}_j + B_j(Y^{(k)} - Y_j);$$
    end
  end

D. Error analysis

It is natural to expect that the error associated with the piecewise interpolating filter $\mathcal{F}^{(p-1)}$ decreases when $\max\limits_{j=1,\ldots,p-1} {}_\triangle t_j$ decreases. Below, in Theorem 3, we justify that this observation is true. To this end, first, in the following Theorem 2, we establish an estimate of the error associated with the filter F.

Let us introduce the norm by

$$\|\mathbf{x}(t, \cdot)\|^2_{T,\Omega} = \frac{1}{b-a} \int_T \|\mathbf{x}(t, \cdot)\|^2_\Omega dt. \qquad (24)$$

We also denote $\|\mathbf{x}(t, \omega)\|^2_{T,\Omega} = \|\mathbf{x}(t, \cdot)\|^2_{T,\Omega}$.

Let us suppose that $\mathbf{x}(\cdot, \omega)$ and $\mathbf{y}(\cdot, \omega)$ are Lipschitz continuous signals, i.e. that there exist real non-negative constants λ_j and γ_j, with $j = 1, \ldots, p$, such that, for $t \in [t_j,\ t_{j+1}]$,

$$\|\mathbf{x}(t, \omega) - \mathbf{x}(t_j, \omega)\|^2_{T,\ \Omega} \le \lambda_j {}_\triangle t_j \qquad (25)$$

and

$$\|\mathbf{y}(t, \omega) - \mathbf{y}(t_{j+1}, \omega)\|^2_{T,\ \Omega} \le \gamma_j {}_\triangle t_j \qquad (26)$$

Open Science Index, Mathematical and Computational Sciences Vol:6, No:7, 2012 publications.waset.org/12739/pdf

where $_{\vartriangle}t_j = |t_{j+1} - t_j|$.

Theorem 2: Under the conditions (25) the error associated with the piecewise interpolation filter, $\|\mathbf{x}(t,\omega) - F^{(p-1)}[\mathbf{y}(t,\omega)]\|^2_{T,\,\Omega}$, is estimated as follows:

$$\|\mathbf{x}(t,\omega) - F^{(p-1)}[\mathbf{y}(t,\omega)]\|^2_{T,\,\Omega} \tag{27}$$

$$\leq \max_{j=1,\ldots,p-1}\Big[(\lambda_j + \gamma_j\|B_j\|^2)_{\vartriangle}t_j + \|E^{1/2}_{z_jz_j}\|^2 \tag{28}$$

$$-\|E_{z_jw_j}(E^{1/2}_{w_jw_j})^{\dagger}\|^2\Big]. \tag{29}$$

Proof: The proof of Theorem 2 is given in the Appendix. □

Further, to show that the error of the reference signal estimate tends to the zero, we need to assume that, for $t \in [t_1,\ t_2]$, the known estimate $\widehat{\mathbf{x}}(t_1,\omega)$ differs from $\mathbf{x}(t,\omega)$ for the value of the order $_{\vartriangle}t_1$, i.e. that, for some constant $c_1 \geq 0$,

$$\|\mathbf{x}(t,\omega) - \widehat{\mathbf{x}}(t_1,\omega)\|^2_{\Omega} \leq c_1{}_{\vartriangle}t_1, \quad \text{for } t \in [t_1,\ t_2]. \tag{30}$$

Theorem 3: Let the conditions (25) and (30) be true. Then the error associated with the piecewise interpolating filter F, $\|\mathbf{x}(t,\omega) - F^{(p-1)}[\mathbf{y}(t,\omega)]\|^2_{T,\,\Omega}$, decreases in the following sense:

$$\|\mathbf{x}(t,\omega) - F^{(p-1)}[\mathbf{y}(t,\omega)]\|^2_{T,\,\Omega} \to 0 \quad \text{as} \tag{31}$$

$$\max_{j=1,\ldots,p-1}{}_{\vartriangle}t_j \to 0 \quad \text{and} \quad p \to \infty. \tag{32}$$

Proof: The proof of Theorem 3 is given in the Appendix. □

Remark 1: We would like to emphasize that the statement of Theorem 3 is fulfilled only under assumptions (25) and (30). At the same time, the assumptions (25) and (30) are not restrictive from a practical point of view. The condition (25) is true for Lipschitz continuous signals $\mathbf{x}$ and $\mathbf{y}$, i.e. for very wide class of signals. The condition (30) is achieved by a choosing an appropriate known method (e.g. see [2]–[14], [24], [26]) to find the estimate $\widehat{\mathbf{x}}(t_1,\omega)$ used in the proposed filter $\mathcal{F}^{(p-1)}$ (see (8) and Theorem 1).

E. Some remarks related to the assumptions of the method

As it has been mentioned in Section III-B, for $j = 1,\ldots,p$, matrices $E_{z_jw_j}$ and $E_{w_jw_j}$ in (22) are assumed to be known or can be estimated. Here, p is a chosen number of selected interpolation signal pairs (see Section III-D). We note that normally p is much smaller than the number of input-output signals $\mathbf{x}(t,\cdot)$ and $\mathbf{y}(t,\cdot)$. Therefore, to estimate any signal $\mathbf{x}(t,\cdot)$ from an arbitrarily large set $K_{_X}$, only a small number, p, of matrices $E_{z_jw_j}$ and $E_{w_jw_j}$ should be estimated (or be known). This issue has also been discussed in Sections I-B1 and I-B4.

By the proposed method, $\mathbf{x}(t,\cdot)$ is estimated for $t \in [t_j,\ t_{j+1}]$. While $E_{w_jw_j}$ in (22) can be directly estimated from observed signals $\mathbf{y}(t_{j+1},\cdot)$ and $\mathbf{y}(t_j,\cdot)$, an estimate of matrix $E_{z_jw_j}$ depends on the reference signal $\mathbf{x}(t_{j+1},\cdot)$ (see (16) and (18)) which is unknown (because the estimate is considered for $t \in [t_j,\ t_{j+1}]$).

Some possible approaches to an estimation of matrix $E_{z_jw_j}$ could be as follows.

1. In the general case, when $\mathbf{x}(t,\cdot)$ and $\mathbf{y}(t,\cdot)$ are arbitrary signals as discussed in Section II-A above, matrix $E_{z_jw_j}$ can be estimated as proposed, for example, in [28], from samples of z_j and w_j.

2. In the case of incomplete observations, the method proposed in [29], [30] can be used.

3. Let $E_{\hat{z}_jw_j}$ be a matrix obtained from matrix $E_{z_jw_j}$ where the term $\mathbf{x}(t_{j+1},\cdot)$ is replaced by $\widehat{\mathbf{x}}(t,\cdot)$ with $t \in [t_{j-1},\ t_j]$. Since $\widehat{\mathbf{x}}(t,\cdot)$ with $t \in [t_{j-1},\ t_j]$ is known, matrix $E_{\hat{z}_jw_j}$ can be considered as an estimate of $E_{z_jw_j}$.

4. In the important case of an *additive* noise, $E_{z_jw_j}$ can be represented in the explicit form. Indeed, if

$$\mathbf{y}(t,\cdot) = \mathbf{x}(t,\cdot) + \xi(t,\cdot)$$

where $\xi(t,\cdot) \in L^2(\Omega,\mathbb{R}^m)$ is a random noise, then $\mathbf{z}(t_j,t_{j+1},\cdot) = \mathbf{y}(t_{j+1},\cdot) - \xi(t_{j+1},\cdot) - \widehat{\mathbf{x}}(t_j,\cdot)$ and matrix $E_{z_jw_j}$ can be represented as follows:

$$E_{z_jw_j} = E_{(y_{j+1}-\xi_{j+1})(y_{j+1}-y_j)} - E_{\widehat{x}_j(y_{j+1}-y_j)} \tag{33}$$

We note that the RHS of (33) depends only on observed signals $\mathbf{y}(t_j,\cdot)$, $\mathbf{y}(t_{j+1},\cdot)$, estimated signal $\widehat{\mathbf{x}}(t_j,\cdot)$, and noise $\xi(t_{j+1},\cdot)$, not on the reference signal $\mathbf{x}(t_{j+1},\cdot)$. In particular, in (33), the term $E_{\xi_{j+1}(y_{j+1}-y_j)}$ can be estimated as $\pm(E[\xi^2_{j+1}])^{1/2}(E[(y_{j+1}-y_j)^2])^{1/2}$ where $E[\xi^2_{j+1}] = \int_{\Omega}[\xi(t_{j+1},\omega)]^2\,d\mu(\omega)$. It is motivated by the Holder's inequality for integrals. The second term in (33), $E_{\widehat{x}_j(y_{j+1}-y_j)}$, can be estimated from the samples of $\widehat{\mathbf{x}}(t_{j+1},\cdot)$ and $\mathbf{y}(t_{j+1},\cdot) - \mathbf{y}(t_j,\cdot)$.

We also note that the first term in the RHS of (33), $E_{(y_{j+1}-\xi_{j+1})(y_{j+1}-y_j)}$, is similar to the related covariance matrix in the Wiener filtering approach [6].

5. Other known ways to estimate $E_{\xi_{j+1}(y_{j+1}-y_j)}$ can be found in [6], Section 5.3.

In general, an estimation of covariance matrices is a special research topic which is not a subject of this paper. The relevant references can be found, for example, in [6], [30].

V. Conclusion

The theory for a new approach to filtering arbitrarily large sets of stochastic signals K_Y and K_X is provided. Distinctive features of the approach are as follows.

(i) The proposed filter $\mathcal{F}^{(p-1)} : K_Y \to K_X$ is nonlinear and is presented in the form of a sum with $p-1$ terms where each term, $\mathcal{F}_j : K_{Y,j} \to K_{X,j}$, is interpreted as a particular sub-filter. Here, $K_{Y,j}$ and $K_{X,j}$ are 'small' pieces of K_Y and K_X, respectively.

(ii) The prime idea is to exploit a priori information only on *few reference signals*, p, from the set $K_{_X}$ that contains $N \gg p$ signals (or even an infinite number of signals) and determine $\mathcal{F}_j$ separately, for each pieces $K_{Y,j}$ and $K_{X,j}$, so that the associated error is minimal. In other words, the filter $\mathcal{F}^{(p-1)}$ is flexible to changes in the sets of observed and reference signals K_Y and K_X, respectively.

(iii) Due to the specific way of determining $\mathcal{F}_j$, the filter $\mathcal{F}^{(p-1)}$ provides a smaller associated error than that for the processing of the whole set K_Y by a filter which is not specifically adjusted to each particular piece $K_{Y,j}$. Moreover,

Open Science Index, Mathematical and Computational Sciences Vol:6, No:7, 2012 publications.waset.org/12739/pdf

the error associated with our filter decreases when the number of its terms, $\mathcal{F}_1, \ldots, \mathcal{F}_{p-1}$, increases.

(iv) While the proposed filter $\mathcal{F}^{(p-1)}$ processes arbitrarily large (and even infinite) signal sets, the filter is nevertheless fixed for all signals in the sets.

(v) The filter $\mathcal{F}^{(p-1)}$ is determined in terms of pseudo-inverse matrices so that the filter always exists.

(vi) The computational load associated with the filter $\mathcal{F}^{(p-1)}$ is less than that associated with other known filters applied to the processing of large signal sets.

APPENDIX A

Proof of Theorem 1: It follows from (8) and (12) that α_j, for $j = 1, \ldots, p-1$, is given by

$$\alpha_j = \widehat{\mathbf{x}}(t_j, \omega) - B_j[\mathbf{y}(t_j, \omega)]. \tag{34}$$

Further, for α_j given by (34),

$$\begin{aligned}
&\|[\mathbf{x}(t_{j+1}, \cdot) - \alpha_j] - \mathcal{B}_j[\mathbf{y}(t_{j+1}, \cdot)]\|_\Omega^2 \\
&= \|\mathbf{z}(t_j, t_{j+1}, \cdot) - \mathcal{B}_j[\mathbf{w}(t_j, t_{j+1}, \cdot))]\|_\Omega^2 \\
&= \mathrm{tr}\{E_{z_j z_j} - E_{z_j w_j} B_j^T - B_j E_{w_j z_j} + B_j E_{w_j w_j} B_j^T\} \\
&= \|E_{z_j z_j}^{1/2}\|^2 - \|E_{z_j w_j}(E_{w_j w_j}^{1/2})^\dagger\|^2 \\
&\quad + \|(B_j - E_{z_j w_j} E_{w_j w_j}^\dagger) E_{w_j w_j}^{1/2}\|^2 \\
&= \|E_{z_j z_j}^{1/2}\|^2 - \|E_{z_j w_j}(E_{w_j w_j}^{1/2})^\dagger\|^2 \\
&\quad + \|E_{z_j w_j}(E_{w_j w_j}^{1/2})^\dagger - B_j E_{w_j w_j}^{1/2}\|^2,
\end{aligned} \tag{35}$$

where $\|\cdot\|$ is the Frobenius norm. The latter is true because

$$E_{w_j w_j}^\dagger E_{w_j w_j}^{1/2} = (E_{w_j w_j}^{1/2})^\dagger$$

and

$$E_{z_j w_j} E_{w_j w_j}^\dagger E_{w_j w_j} = E_{z_j w_j} \tag{36}$$

by Lemma 24 in [6]. Thus, the second expression in (12) is reduced to the problem

$$\min_{B_j} \|E_{z_j w_j}(E_{w_j w_j}^{1/2})^\dagger - B_j E_{w_j w_j}^{1/2}\|^2. \tag{37}$$

It is known (see, for example, [6], p. 304) that the solution of problem (37) is given by (22). The equation (20) follows from (6) and (34).

Theorem 1 is proven. □

Proof of Theorem 2: For $t \in [t_j,\ t_{j+1}]$ and F_j defined by (20)–(22),

$$\begin{aligned}
&\mathbf{x}(t, \omega) - F[\mathbf{y}(t, \omega)] \\
&= \mathbf{x}(t, \omega) - F_j[\mathbf{y}(t, \omega)] \\
&= \mathbf{x}(t, \omega) - \widehat{\mathbf{x}}(t_j, \omega) + B_j \mathbf{y}(t_j, \omega) - B_j \mathbf{y}(t, \omega) \\
&= [\mathbf{x}(t, \omega) - \mathbf{x}(t_{j+1}, \omega)] + \mathbf{z}(t_j, t_{j+1}, \omega) \\
&\quad - B_j \mathbf{w}(t_j, t_{j+1}, \omega) + B_j[\mathbf{y}(t_{j+1}, \omega) - \mathbf{y}(t, \omega)].
\end{aligned} \tag{38}$$

Then (38) and (38) imply

$$\begin{aligned}
&\|\mathbf{x}(t, \omega) - F[\mathbf{y}(t, \omega)]\|_{T,\Omega}^2 \\
\leq\ &\|\mathbf{x}(t, \omega) - \mathbf{x}(t_{j+1}, \omega)\|_{T,\Omega}^2 \\
+\ &\|\mathbf{z}(t_j, t_{j+1}, \omega) - B_j \mathbf{w}(t_j, t_{j+1}, \omega)\|_\Omega^2 \\
+\ &\|B_j[\mathbf{y}(t_{j+1}, \omega) - \mathbf{y}(t, \omega)]\|_{T,\Omega}^2
\end{aligned} \tag{39}$$

where

$$\begin{aligned}
&\|\mathbf{z}(t_j, t_{j+1}, \omega) - B_j \mathbf{w}(t_j, t_{j+1}, \omega)\|_\Omega^2 \\
&\quad = \|\mathbf{z}(t_j, t_{j+1}, \omega) - B_j \mathbf{w}(t_j, t_{j+1}, \omega)\|_{T,\Omega}^2.
\end{aligned}$$

It follows from (35) and (35) that for B_j given by (22),

$$\begin{aligned}
&\|\mathbf{z}(t_j, t_{j+1}, \omega) - B_j \mathbf{w}(t_j, t_{j+1}, \omega)\|_\Omega^2 \\
&= \|E_{z_j z_j}^{1/2}\|^2 - \|E_{z_j w_j}(E_{w_j w_j}^{1/2})^\dagger\|^2.
\end{aligned} \tag{40}$$

Then (19)–(22), (25) and (38)–(40) imply that for all $t \in [a,\ b]$ and $\omega \in \Omega$, (27) is true. □

Proof of Theorem 3: The relation (24) implies that

$$\begin{aligned}
&\|\mathbf{x}(t, \omega) - F[\mathbf{y}(t, \omega)]\|_{T,\Omega}^2 \\
&= \frac{1}{b-a} \sum_{j=1}^{p-1} \int_{t_j}^{t_{j+1}} \|\mathbf{x}(t, \omega) - F_j[\mathbf{y}(t, \omega)]\|_\Omega^2 dt,
\end{aligned} \tag{41}$$

where

$$\begin{aligned}
&\|\mathbf{x}(t, \omega) - F_j[\mathbf{y}(t, \omega)]\|_\Omega^2 \\
&\quad = \|\mathbf{x}(t, \omega) - \widehat{\mathbf{x}}(t_j, \omega) + B_j[\mathbf{y}(t_j, \omega) - B_j \mathbf{y}(t, \omega)]\|_\Omega^2 \\
&\quad \leq \|\mathbf{x}(t, \omega) - \mathbf{x}(t_j, \omega)\|_\Omega^2 + \|\mathbf{x}(t_j, \omega) - \widehat{\mathbf{x}}(t_j, \omega)\|_\Omega^2 \\
&\qquad + \|B_j[\mathbf{y}(t_j, \omega) - B_j \mathbf{y}(t, \omega)]\|_\Omega^2.
\end{aligned}$$

Then

$$\begin{aligned}
&\int_{t_j}^{t_{j+1}} \|\mathbf{x}(t, \omega) - F_j[\mathbf{y}(t, \omega)]\|_\Omega^2 dt \\
&\leq \int_{t_j}^{t_{j+1}} \|\mathbf{x}(t, \omega) - \mathbf{x}(t_j, \omega)\|_\Omega^2 dt \\
&+ \int_{t_j}^{t_{j+1}} \|\mathbf{x}(t_j, \omega) - \widehat{\mathbf{x}}(t_j, \omega)\|_\Omega^2 dt \\
&+ \|B_j\| \int_{t_j}^{t_{j+1}} \|\mathbf{y}(t_j, \omega) - \mathbf{y}(t, \omega)\|_\Omega^2 dt \\
&\leq \lambda_j (\triangle t_j)^2 + \|\mathbf{x}(t_j, \omega) - \widehat{\mathbf{x}}(t_j, \omega)\|_\Omega^2 \triangle t_j + \|B_j\| \gamma_j (\triangle t_j)^2
\end{aligned} \tag{42}$$

Let us consider an estimate of $\|\mathbf{x}(t_j, \omega) - \widehat{\mathbf{x}}(t_j, \omega)\|_\Omega^2$, for $j = 1, \ldots, p-1$. To this end, let us denote $\triangle t = \max_{j=1,\ldots,p-1} \triangle t_j$.

For $j = 1$, i.e. for $t \in [t_1,\ t_2]$,

$$\begin{aligned}
&\|\mathbf{x}(t, \omega) - F_1 \mathbf{y}(t, \omega)\|_\Omega^2 \\
&\leq \|\mathbf{x}(t, \omega) - \mathbf{x}(t_1, \omega)\|_\Omega^2 + \|\mathbf{x}(t_1, \omega) - \widehat{\mathbf{x}}(t_1, \omega)\|_\Omega^2 \\
&\qquad + \|B_1\| \|\mathbf{y}(t_1, \omega) - \mathbf{y}(t, \omega)\|_\Omega^2 \\
&\leq \lambda_1 \triangle t_1 + c_1 \triangle t_1 + \|B_1\| \gamma_1 \triangle t_1 \\
&\leq \beta_1 \triangle t,
\end{aligned}$$

where $\beta_1 = \lambda_1 + c_1 + \|B_1\| \gamma_1$. In particular, the latter implies

$$\|\mathbf{x}(t_2, \omega) - \widehat{\mathbf{x}}(t_2, \omega)\|_\Omega^2 = \|\mathbf{x}(t_2, \omega) - F_1 \mathbf{y}(t_2, \omega)\|_\Omega^2 \leq \beta_1 \triangle t$$

For $j = 2$, i.e. for $t \in [t_2,\ t_3]$,

$$\begin{aligned}
&\|\mathbf{x}(t, \omega) - F_2 \mathbf{y}(t, \omega)\|_\Omega^2 \\
&\leq \|\mathbf{x}(t, \omega) - \mathbf{x}(t_2, \omega)\|_\Omega^2 + \|\mathbf{x}(t_2, \omega) - \widehat{\mathbf{x}}(t_2, \omega)\|_\Omega^2 \\
&\qquad + \|B_2\| \|\mathbf{y}(t_2, \omega) - \mathbf{y}(t, \omega)\|_\Omega^2 \\
&\leq \lambda_2 \triangle t_2 + \beta_1 \triangle t + \|B_2\| \gamma_2 \triangle t_2 \\
&\leq \beta_2 \triangle t,
\end{aligned}$$

where $\beta_2 = \lambda_2 + \beta_1 + \|B_2\|\gamma_2$. In particular, then it follows that

$$\|\mathbf{x}(t_3,\omega) - \widehat{\mathbf{x}}(t_3,\omega)\|_\Omega^2 = \|\mathbf{x}(t_3,\omega) - F_2\mathbf{y}(t_3,\omega)\|_\Omega^2 \le \beta_2 {}_\triangle t.$$

On the basis of the above, let us assume that, for $j = k-1$ with $k = 2,\dots,p-1$, i.e. for $t \in [t_{k-1},\ t_k]$,

$$\begin{aligned}&\|\mathbf{x}(t_k,\omega) - \widehat{\mathbf{x}}(t_k,\omega)\|_\Omega^2 = \|\mathbf{x}(t_k,\omega) - F_{k-1}\mathbf{y}(t_k,\omega)\|_\Omega^2\\ &\le \beta_{k-1} {}_\triangle t\end{aligned}$$

where β_{k-1} is defined by analogy with β_2.

Then, for $j = k$ with $k = 2,\dots,p-1$, i.e. for $t \in [t_k,\ t_{k+1}]$,

$$\begin{aligned}&\|\mathbf{x}(t,\omega) - F_k\mathbf{y}(t,\omega)\|_\Omega^2\\ &\le \|\mathbf{x}(t,\omega) - \mathbf{x}(t_k,\omega)\|_\Omega^2 + \|\mathbf{x}(t_k,\omega) - \widehat{\mathbf{x}}(t_k,\omega)\|_\Omega^2\\ &\qquad + \|B_k\|\|\mathbf{y}(t_k,\omega) - \mathbf{y}(t,\omega)\|_\Omega^2\\ &\le \lambda_k {}_\triangle t_k + \beta_{k-1} {}_\triangle t + \|B_k\|\gamma_2 {}_\triangle t_k\\ &\le \beta_k {}_\triangle t,\end{aligned}$$

where $\beta_k = \lambda_k + \beta_{k-1} + \|B_k\|\gamma_k$. Thus, the following is true:

$$\begin{aligned}&\|\mathbf{x}(t_{k+1},\omega) - \widehat{\mathbf{x}}(t_{k+1},\omega)\|_\Omega^2\\ &\quad = \|\mathbf{x}(t_{k+1},\omega) - F_k\mathbf{y}(t_{k+1},\omega)\|_\Omega^2 \le \beta_k {}_\triangle t. \quad (43)\end{aligned}$$

Therefore, (42), (42) and (43) imply

$$\begin{aligned}&\int_{t_j}^{t_{j+1}} \|\mathbf{x}(t,\omega) - F_j[\mathbf{y}(t,\omega)]\|_\Omega^2 dt\\ &\le \lambda_j ({}_\triangle t_j)^2 + \beta_{j-1}({}_\triangle t_j)^2 + \|B_j\|\gamma_j ({}_\triangle t_j)^2\\ &\le \eta_j ({}_\triangle t)^2 \quad (44)\end{aligned}$$

where $\eta_j = \lambda_j + \beta_{j-1} + \|B_j\|$, and then it follows from (41)–(42) and (44) that for all $t \in [a,\ b]$,

$$\begin{aligned}&\|\mathbf{x}(t,\omega) - F[\mathbf{y}(t,\omega)]\|_{T,\Omega}^2 \le \frac{1}{b-a}\sum_{j=1}^{p-1}\eta_j ({}_\triangle t)^2\\ &= \frac{1}{b-a} {}_\triangle t \sum_{j=1}^{p-1}\eta_j {}_\triangle t. \quad (45)\end{aligned}$$

Let us now choose $c \in \mathbb{R}$ and $d \in \mathbb{R}$ so that ${}_\triangle t = \dfrac{d-c}{p}$ and partition interval $[c,\ d] \subset \mathbb{R}$ by points $\tau_1,\dots,\tau_p$ so that $c = \tau_1$ and $\tau_j = \tau_1 + j {}_\triangle t$ with $j = 1,\dots,p$. There exists an integrable (bounded) function $\varphi : [c,\ d] \to \mathbb{R}$ such that, for $\xi_j \in (\tau_j,\ \tau_{j+1})$, $\varphi(\xi_j) = \eta_j$. Then

$$\lim_{{}_\triangle t\to\infty}\sum_{j=1}^{p-1}\eta_j {}_\triangle t = \lim_{{}_\triangle t\to\infty}\sum_{j=1}^{p-1}\varphi(\xi_j) {}_\triangle t = \int_c^d \varphi(\tau)d\tau < +\infty. \quad (46)$$

Thus, $\frac{1}{b-a} {}_\triangle t \sum_{j=1}^{p-1}\eta_j {}_\triangle t \to 0$ as ${}_\triangle t \to 0$.
As a result, (45)–(46) imply (31). □

REFERENCES

[1] H. Kopka and P. W. Daly, *A Guide to LaTeX*, 3rd ed. Harlow, England: Addison-Wesley, 1999.
[2] J. Chen, J. Benesty, Y. Huang, and S. Doclo, New Insights Into the Noise Reduction Wiener Filter, *IEEE Trans. on Audio, Speech, and Language Processing*, 14, No. 4, pp. 1218 - 1234, 2006.
[3] M. Spurbeck and P. Schreier, Causal Wiener filter banks for periodically correlated time series, *Signal Processing*, 87, 6, pp. 1179-1187, 2007.
[4] J. S. Goldstein, I. Reed, and L. L. Scharf, "A Multistage Representation of the Wiener Filter Based on Orthogonal Projections," *IEEE Trans. on Information Theory*, vol. 44, pp. 2943-2959, 1998.
[5] Y. Hua, M. Nikpour, and P. Stoica, "Optimal Reduced-Rank estimation and filtering," *IEEE Trans. on Signal Processing*, vol. 49, pp. 457-469, 2001.
[6] A. Torokhti and P. Howlett, *Computational Methods for Modelling of Nonlinear Systems*, Elsevier, 2007.
[7] E. D. Sontag, *Polynomial Response Maps*, Lecture Notes in Control and Information Sciences, 13, 1979.
[8] S. Chen and S. A. Billings, Representation of non-linear systems: NARMAX model, *Int. J. Control*, vol. 49, no. 3, pp. 1013–1032, 1989.
[9] V. J. Mathews and G. L. Sicuranza, *Polynomial Signal Processing*, J. Wiley & Sons, 2001.
[10] A. Torokhti and P. Howlett, Optimal Transform Formed by a Combination of Nonlinear Operators: The Case of Data Dimensionality Reduction, *IEEE Trans. on Signal Processing*, 54, No. 4, pp. 1431-1444, 2006.
[11] A. Torokhti and P. Howlett, Filtering and Compression for Infinite Sets of Stochastic Signals, *Signal Processing*, 89, pp. 291-304, 2009.
[12] J. Vesma and T. Saramaki, Polynomial-Based Interpolation Filters - Part I: Filter Synthesis, *Circuits, Systems, and Signal Processing*, Volume 26, Number 2, Pages 115-146, 2007.
[13] A. Torokhti and J. Manton, Generic Weighted Filtering of Stochastic Signals, *IEEE Trans. on Signal Processing*, 57, issue 12, pp. 4675-4685, 2009.
[14] A. Torokhti and S. Miklavcic, Data Compression under Constraints of Causality and Variable Finite Memory, *Signal Processing*, 90 , Issue 10, pp. 2822-2834, 2010.
[15] I. Babuska, U. Banerjee, J. E. Osborn, Generalized finite element methods: main ideas, results, and perspective, *International Journal of Computational Methods*, 1 (1), pp. 67-103, 2004.
[16] S. Kang and L. Chua, A global representation of multidimensional piecewise-linear functions with linear partitions, *IEEE Trans. on Circuits and Systems*, 25 Issue:11, pp. 938 - 940, 1978.
[17] L.O. Chua and A.-C. Deng, Canonical piecewise-linear representation, *IEEE Trans. on Circuits and Systems*, 35 Issue:1, pp. 101 - 111, 1988.
[18] J.-N. Lin and R. Unbehauen, Adaptive nonlinear digital filter with canonical piecewise-linear structure, *IEEE Trans. on Circuits and Systems*, 37 Issue:3, pp. 347 - 353, 1990.
[19] J.-N. Lin and R. Unbehauen, Canonical piecewise-linear approximations, *IEEE Trans. on Circuits and Systems I: Fundamental Theory and Applications*, 39 Issue:8, pp. 697 - 699, 1992.
[20] S.B. Gelfand and C.S. Ravishankar, A tree-structured piecewise linear adaptive filter, *IEEE Trans. on Inf. Theory*, 39, issue 6, pp. 1907–1922, 1993.
[21] E.A. Heredia and G.R. Arce, Piecewise linear system modeling based on a continuous threshold decomposition, *IEEE Trans. on Signal Processing*, 44 Issue:6, pp. 1440 - 1453, 1996.
[22] G. Feng, Robust filtering design of piecewise discrete time linear systems, *IEEE Trans. on Signal Processing*, 53 Issue:2, pp. 599 - 605, 2005.
[23] F. Russo, Technique for image denoising based on adaptive piecewise linear filters and automatic parameter tuning, *IEEE Trans. on Instrumentation and Measurement*, 55, Issue:4, pp. 1362 - 1367, 2006.
[24] J.E. Cousseau, J.L. Figueroa, S. Werner, T.I. Laakso, Efficient Nonlinear Wiener Model Identification Using a Complex-Valued Simplicial Canonical Piecewise Linear Filter, *IEEE Trans. on Signal Processing*, 55 Issue:5, pp. 1780 - 1792, 2007.
[25] P. Julian, A. Desages, B. D'Amico, Orthonormal high-level canonical PWL functions with applications to model reduction, *IEEE Trans. on Circuits and Systems I: Fundamental Theory and Applications*, 47 Issue:5, pp. 702 - 712, 2000.
[26] T. Wigren, Recursive Prediction Error Identification Using the Nonlinear Wiener Model, *Automatica*, 29, 4, pp. 1011–1025, 1993.
[27] G. H. Golub and C. F. van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, 1996.
[28] T. Anderson, *An Introduction to Multivariate Statistical Analysis*, New York, Wiley, 1984.
[29] L. I. Perlovsky and T. L. Marzetta, Estimating a Covariance Matrix from Incomplete Realizations of a Random Vector, *IEEE Trans. on Signal Processing*, 40, pp. 2097-2100, 1992.
[30] O. Ledoit and M. Wolf, A well-conditioned estimator for large-dimensional covariance matrices, *J. Multivariate Analysis* 88, pp. 365–411, 2004.

Open Science Index, Mathematical and Computational Sciences Vol:6, No:7, 2012 publications.waset.org/12739/pdf