axioms

MDPI

Article

Constructive Approximation of Nonlinear Operators Based on Piecewise Interpolation Technique

Anatoli Torokhti * and Peter Pudney

School of Mathematical Sciences, Adelaide University, Mawson Lakes, SA 5095, Australia
* Correspondence: anatoli.torokhti@gmail.com

Abstract

Suppose K_Y and K_X are the image and the preimage of a nonlinear operator $\mathcal{F} : K_Y \to K_X$. It is supposed that the cardinality of each K_Y and K_X is N and N is large. We provide an approximation to the map $\mathcal{F}$ that requires prior information only on a few elements p from K_Y, where $p \ll N$, but still effectively represents $\mathcal{F}(K_Y)$. It is achieved under Lipschitz continuity assumptions. The device behind the proposed method is based on a special extension of the piecewise linear interpolation technique to the case of sets of stochastic elements. The proposed technique provides a single operator that transforms any element from the arbitrarily large set K_Y. The operator is determined in terms of pseudo-inverse matrices so that it always exists.

Keywords: approximation of nonlinear mappings; error minimization; optimization; interpolation

MSC: 15A99; 65F05

1. Introduction

A purpose of the proposed methodology is to provide an effective way to transform large data sets. The methodology is motivated by the problems arising in signal processing, where the nonlinear operator $\mathcal{F} : K_Y \to K_X$ is interpreted as a nonlinear system (or a nonlinear filter) transforming a set of stochastic signals K_Y to the set of stochastic signals K_X. Therefore, below we refer to this terminology.

The device behind the proposed method is based on a special extension of the piecewise linear interpolation technique to the case of stochastic signal sets. The device is not straightforward and requires the careful substantiation presented in Sections 2.2, 3.4, 4.2 and 4.4 below.

1.1. Motivations

1.1.1. Filtering of Large Arrays of Stochastic Signals

We consider the case when the cardinality N of signal sets K_Y and K_X is large. Members of K_Y and K_X are represented by stochastic vectors. The problem of finding a filter that effectively transforms the large set K_Y to the large set K_X has been considered in a number of works such as those represented in [1–14]. The approaches considered in these references require prior information on each reference signal. This requirement poses a significant difficulty on the applicability of the methods mentioned above. For processing of large signal arrays, such restrictions are quite inconvenient. The exception is the techniques provided in [15,16], which exploit the initial information in the form of the vector derived from averaging over the reference signal set.

Check for updates

Academic Editor: Ricardo Roque Enguiça

Received: 3 December 2025
Revised: 31 December 2025
Accepted: 14 January 2026
Published: 26 January 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

In this paper, we show how to construct a filter $\mathcal{F}$ that requires prior information only on few signals, $p \ll N$, from K_X but still performs effectively compared to known filters based on prior information on every reference signal from K_X. Such a filter is denoted by $\mathcal{F}^{(p-1)}$.

1.1.2. Basic Idea of the Proposed Methodology

The specific feature of the proposed filter is an adjustment and extension of the piecewise function interpolation technique [17] to filtering of stochastic signals. It is well known that the piecewise function interpolation [17] often provides better accuracy and faster performance than those associated with the linear and polynomial approximation used in known filtering techniques (such as, for example, those in [5,9]).

The device of the filter under consideration is given in Sections 2.2, 3.1 and 4.2.

1.1.3. Pseudo-Inverse Matrices and Related Matters

The known filter models proposed in [1–3,6–8,10,13,14], use inverse matrices. In the cases of grossly corrupted signals or erroneous measurements, those inverse matrices may not exist, and thus, those filters cannot be applied. The examples in Section 5 illustrate this case.

The filter proposed here avoids this drawback since its model is based on exploiting pseudo-inverse matrices. As a result, the proposed filter always exists. That is, it processes any kind of noisy signals. Exploiting pseudo-inverse matrices is not straightforward and requires the development of a specific technique. Perhaps it was a reason to not use the pseudo-inverse matrices in the above-mentioned references. Here, an extension of the proposed filtering techniques to the case of implementation of the pseudo-inverse matrices is performed on the basis of theory represented in [5].

1.1.4. Computational Load

Let $\mathbf{x} = [x_1, \ldots, x_m]^T \in K_X$ and $\mathbf{y} = [y_1, \ldots, y_n]^T \in K_Y$. An implementation of the filters provided in [1–8,10,13,14] leads to computation of an $n \times n$ inverse or pseudo-inverse matrix for each pair of signals $\mathbf{x} \in K_X$ and $\mathbf{y} \in K_Y$, and a matrix product of $m \times n$ and $n \times n$ matrices. Thus, for the processing of arrays in K_X and K_Y, the filters in [1–8,10,13,14] require $O(2mn^2N) + O(26n^3N)$ flops [18].

Arrays K_X and K_Y can be represented by vectors, χ and γ, each with mN and nN components, respectively. Then an application of the techniques in [1–8,10,13,14] requires $O(2mn^2N^2)$ and $O(26n^3N^3)$ flops [18].

Thus, for the case when N is large, the computational work associated with the methods in [1–8,10,13,14] becomes, in practice, unacceptable.

The required computational load for the filter $\mathcal{F}^{(p-1)}$ considered below is substantially less. This is implied by an ability to use only p pseudo-inverse matrices, where p is much less than cardinality N. For this reason, $\mathcal{F}^{(p-1)}$ requires only $O(2mn^2p) + O(26n^3p)$ flops. Clearly, this is less than that required by the filters mentioned above. This issue is illustrated in Section 5.

1.2. Relevant Works

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

1.2.1. Generic Optimal Linear (GOL) Filter [5]

The generic optimal linear (GOL) filter in [5] is a generalization of the Wiener filter to the case when the covariance matrix is not invertible and the 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 [5] are used in the proof of Theorem 1 below.

1.2.2. Simplicial Canonical Piecewise Linear Filter [13]

A complex Wiener adaptive filter was developed in [13] from the two-dimensional complex-valued simplicial canonical piecewise linear filter [19]. The filter in [13] 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 considered, for example, in Section 5: the matrices used in [13] are not invertible for the signals as those in Section 5. Similarly, the filters studied in [8,10] were developed for the processing of a single signal when the covariance matrices are invertible.

For the filter proposed here, these restrictions are removed.

1.2.3. Adaptive Piecewise Linear Filter [12]

A piecewise linear filter in [12] 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 [12].

1.2.4. Averaging Polynomial Filter [15,16]

The averaging polynomial filter proposed in [15,16] 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. This effect is illustrated in Section 5 below.

1.2.5. Other Relevant Filters

The technique developed in [11] is an extension of the GOL filter to the constraint problem with respect to the filter rank. It concerns data compression.

The methods in [6,7,20,21] have been developed for deterministic signals. Motivated by the results achieved in [20,21], adaptive filters were elaborated in [22]. A theoretical basis for the device proposed in [20,21] is provided in [23].

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 [12,13,19–26]). 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 [24] have been obtained by arranging linear filters and thresholds in a tree structure. Piecewise linear filters discussed in [25] 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 [26] were based on a piecewise Lyapunov function.

1.3. 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 as follows:

(i) They require information on each reference signal (in the form of a sample, for example).
(ii) Matrices used in the known filters can be non-invertible (as in the simulations considered below in Section 5) and then the filter does not exist.
(iii) The associated computation work may require a very long time. For example, in simulations (Section 5), MATLAB was out of memory for computing the GOL filter [5]

when each of the sets K_X and K_Y was represented by a long vector (this option has been discussed in Section 1.1.4 above). The PC used for the simulations was a Dell Latitude 7400 with a CPU Intel Core i5 (8th Gen) 8265U/1.6 GHz and 8 GB RAM.

1.4. 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 [1–14,19–26] 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 1.1 and 1.3 above.

(ii) As a result, our piecewise linear filter model (Section 3), the statement of the problem (Section 3.3 below), and consequently, the device of its solution (Section 4 below) are different from those considered in [12,13,19–26]. In this regard, see also Section 1.2.5.

(iii) The above naturally leads to a new structure of the filter (presented in Sections 3.4 and 4.2 below), which is very different from the known ones.

1.5. 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 [1–17,19–26]. In particular, it allows us to achieve the following:

(i) Achieve a desired accuracy in signal estimation This means that any desired accuracy is achieved theoretically, as is shown in Section 4.4 below. In practice, of course, the accuracy is increased to a prescribed reasonable level.

(ii) Exploit prior 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.

(iii) Find a single filter to process any signal from the arbitrarily large signal set.

(iv) Determine the filter in terms of pseudo-inverse matrices so that the filter always exists.

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

2. Some Preliminaries

2.1. Notation

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

Let (Ω, Σ, μ) be a probability space (usually, $\Omega = \{\omega\}$ is the set of outcomes, Σ is a σ-field of measurable subsets in Ω and $\mu : \Sigma \to [0,1]$ is an associated probability measure on Σ; in particular, $\mu(\Omega) = 1$.) and K_X and K_Y be arbitrarily large sets of signals such that

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

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. 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. The variable $t \in T \subseteq \mathbb{R}$ represents time. 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. Then, for example, the stochastic 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 Equation (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 $j = 1, \ldots, p$,

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

Therefore, K_Y and K_X can now be represented as

$$K_X = \bigcup_{j=1}^{p-1} K_{X,j} \quad \text{and} \quad K_Y = \bigcup_{j=1}^{p-1} K_{Y,j}. \tag{3}$$

2.2. Brief Description of the Method

We wish to determine a single filter that estimates $\mathbf{x} \in K_X$ from $\mathbf{y} \in K_Y$ with a controlled, associated error. Array K_Y can be finite or infinite.

To this end, the proposed filter $\mathcal{F}^{(p-1)} : K_Y \to K_X$ is represented by a sum of sub-filters $\mathcal{F}_1, \ldots, \mathcal{F}_{p-1}$ (see Equations (4) and (5) below). The sub-filter $\mathcal{F}_j$ transforms signals of subset $K_{Y,j}$ of array K_Y to signals of subset $K_{X,j}$ of array K_X, i.e., $\mathcal{F}_j : K_{Y,j} \to K_{X,j}$. The prime idea is to determine $\mathcal{F}_j$ separately, for each $j = 1, \ldots, p-1$, from an associated minimization problem (Equation (11)) (see Sections 3.4 and 4.2 below). This procedure 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 $t = t_{j+1}$. It is natural to expect that the processing of a 'smaller' subset, $K_{Y,j}$, may lead to a smaller error than that associated with the processing of the whole array K_Y.

Therefore, the advantages of $\mathcal{F}^{(p-1)}[\mathbf{y}(t,\cdot)]$ are similar to those that follow from an application of the piecewise interpolation procedure, such as, for example, the high accuracy of estimation.

Section 4.4 confirms this observation. Importantly, it is shown in Sections 4.5 and 5 that the proposed technique avoids the bottlenecks of the known methods discussed in Section 1.3 above.

3. Description of the Problem

3.1. 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

$$\mathcal{F}_j[\mathbf{y}(t,\cdot)] = \alpha_j + \mathcal{B}_j[\mathbf{y}(t,\cdot)] \quad \text{and} \quad \delta_j = \begin{cases} 1, & \text{if } t_j \le t \le t_{j+1}, \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

Here, $\mathcal{F}_j$ is a sub-filter defined for $t_j \le t \le t_{j+1}$. In Equation (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

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

Thus, $\mathcal{F}_j$ is defined by an operator $F_j : \mathbb{R}^n \to \mathbb{R}^m$ such that

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

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

3.2. Assumptions

In the known approaches related to filtering of stochastic signals (e.g., see [1–11,13–16]), 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 $j = 1, \ldots, p$ and p to be a small number. 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 $t \in [a,\ b]$), $p \ll N$, where N is the number of signals in K_X or K_Y.

3.3. The Problem

In Equations (4)–(6), parameters of the filter $\mathcal{F}^{(p-1)}$, i.e., vector α_j and matrix B_j, for $j = 1, \ldots, p-1$, are unknown. Therefore, under the assumptions described in Section 3.2, the problem is to determine α_j and B_j, for $j = 1, \ldots, p-1$. The related problem is to estimate an error associated with the filter $\mathcal{F}^{(p-1)}$.

Solutions to both problems are given in Sections 4.2 and 4.4, respectively. In particular, in the following Section 3.4, interpolation Equations (8) and (11) are introduced that lead to a determination of α_j and B_j.

3.4. Interpolation Conditions

Let us denote

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

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 [1–11,13–16]. This is the initial condition of the proposed technique.

For $j = 1, \ldots, p-1$, each sub-filter F_j in Equations (5) and (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

$$\widehat{\mathbf{x}}(t_1, \cdot) = \alpha_1 + \mathcal{B}_1[\mathbf{y}(t_1, \cdot)] \quad \text{and} \quad \min_{\mathcal{B}_1} \|[\mathbf{x}(t_2, \cdot) - \alpha_1] - \mathcal{B}_1[\mathbf{y}(t_2, \cdot)]\|_\Omega^2, \tag{8}$$

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

$$\widehat{\mathbf{x}}(t, \cdot) = \mathcal{F}_1[\mathbf{y}(t, \cdot)] = \alpha_1 + \mathcal{B}_1[\mathbf{y}(t, \cdot)] = \widehat{\mathbf{x}}(t_1, \cdot) + \mathcal{B}_1[\mathbf{y}(t, \cdot) - \mathbf{y}(t_1, \cdot)] \tag{9}$$

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

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

Extending this procedure up to $j = k-1$, where $k = 3, \ldots, p$, 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

$$\widehat{\mathbf{x}}(t_{k-1}, \cdot) = \mathcal{F}_{k-2}[\mathbf{y}(t_{k-1}, \cdot)]. \tag{10}$$

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

$$\widehat{\mathbf{x}}(t_{k-1},\cdot) = \alpha_{k-1} + \mathcal{B}_{k-1}[\mathbf{y}(t_{k-1},\cdot)] \quad \text{and} \quad \min_{\mathcal{B}_{k-1}} \|[\mathbf{x}(t_k,\cdot) - \alpha_{k-1}] - \mathcal{B}_{k-1}[\mathbf{y}(t_k,\cdot)]\|_{\Omega}^2, \quad (11)$$

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

$$\widehat{\mathbf{x}}(t,\cdot) = \mathcal{F}_{k-1}[\mathbf{y}(t,\cdot)] = \alpha_{k-1} + \mathcal{B}_{k-1}[\mathbf{y}(t,\cdot)] = \widehat{\mathbf{x}}(t_{k-1},\cdot) + \mathcal{B}_1[\mathbf{y}(t,\cdot) - \mathbf{y}(t_{k-1},\cdot)]. \quad (12)$$

Equations (8) and (11) are motivated by the device of piecewise function interpolation and associated advantages [17].

Filter $\mathcal{F}^{(p-1)}$ of the form of Equations (4) and (5) with α_j and $\mathcal{B}_j$ satisfying Equations (8) and (11) 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 Equation (1) is called the *interpolation pair*.

Remark 1. *In general, the above procedure might be sensitive to the choice of the initial estimate* $\mathbf{x}(t_1,\cdot)$ *determined by known methods [1–11,13–16]. If the initial estimate is poor, then the associated error increases.*

4. Main Results

4.1. General Device

In accordance with the scheme presented in Sections 3.1 and 3.4 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

$$\widehat{\mathbf{x}}(t,\cdot) = \mathcal{F}^{(p-1)}[\mathbf{y}(t,\cdot)] = \sum_{j=1}^{p-1} \delta_j \mathcal{F}_j[\mathbf{y}(t,\cdot)], \quad (13)$$

where, for each $j = 1,\ldots,p-1$, the sub-filter $\mathcal{F}_j$ is given by Equation (5), and is defined from interpolation Equations (8) and (11).

Below, we show how to determine $\mathcal{F}_j$ to satisfy the conditions of Equations (8) and (11).

4.2. Filter Model

We write $\mathbf{z}(t_j,t_{j+1},\cdot) = [\mathbf{z}^{(1)}(t_j,t_{j+1},\cdot),\ldots,\mathbf{z}^{(m)}(t_j,t_{j+1},\cdot)]^T$ and $\mathbf{w}(t_j,t_{j+1},\cdot) = [\mathbf{w}^{(1)}(t_j,t_{j+1},\cdot),\ldots,\mathbf{w}^{(n)}(t_j,t_{j+1},\cdot)]^T$, where

$$\mathbf{z}(t_j,t_{j+1},\cdot) = \mathbf{x}(t_{j+1},\cdot) - \widehat{\mathbf{x}}(t_j,\cdot) \quad \text{and} \quad \mathbf{w}(t_j,t_{j+1},\cdot) = \mathbf{y}(t_{j+1},\cdot) - \mathbf{y}(t_j,\cdot). \quad (14)$$

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

The associated covariance matrix is defined by

$$E_{z_j w_j} = \left\{ \left\langle \mathbf{z}^{(i)}(t_j,t_{j+1},\cdot), \mathbf{w}^{(k)}(t_j,t_{j+1},\cdot) \right\rangle \right\}_{i,k=1}^{m,n}, \quad (15)$$

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

The Moore–Penrose generalized inverse of a matrix M is denoted by M^+.

The main results are represented below.

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* $\mathbf{x}(t_1,\cdot)$. *As has been mentioned in Section 3.4,* $\widehat{\mathbf{x}}(t_1,\cdot)$ *can be determined by the known methods. Then, for any* $t \in [a,\ b]$, *the proposed piecewise linear interpolation filter* $\mathcal{F}^{(p-1)} : L^2(\Omega,\mathbb{R}^n) \to L^2(\Omega,\mathbb{R}^m)$ *transforms 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*

$$\widehat{\mathbf{x}}(t,\cdot) = \mathcal{F}^{(p-1)}[\mathbf{y}(t,\cdot)] = \sum_{j=1}^{p-1} \delta_j \mathcal{F}_j[\mathbf{y}(t,\cdot)] \tag{16}$$

where

$$\mathcal{F}_j[\mathbf{y}(t,\cdot)] = \widehat{\mathbf{x}}(t_j,\cdot) + \mathcal{B}_j[\mathbf{y}(t,\cdot) - \mathbf{y}(t_j,\cdot)], \tag{17}$$

$$\widehat{\mathbf{x}}(t_j,\cdot) = \mathcal{F}_{j-1}[\mathbf{y}(t_j,\cdot)] \quad (\textit{for } j = 2,\ldots,p-1), \tag{18}$$

$$B_j = E_{z_j w_j} E^{\dagger}_{w_j w_j} + M_{B_j}[I_n - E_{w_j w_j} E^{\dagger}_{w_j w_j}], \tag{19}$$

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 Appendix A. □

It is worthwhile to observe that, due to an arbitrary matrix M_{B_j} in Equation (19), 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 [5] (which is also not unique for the same reason).

4.3. Numerical Realization of Filter $\mathcal{F}^{(p-1)}$ *and Associated Algorithm*

4.3.1. Numerical Realization

In practice, the set $T = [a,\ b]$ (see Section 2.1) 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

$$\widehat{X}^{(k)} = [\widehat{\mathbf{x}}(\tau_k,\omega_1),\ldots,\widehat{\mathbf{x}}(\tau_k,\omega_q)] \quad \text{and} \quad Y^{(k)} = [\mathbf{y}(\tau_k,\omega_1),\ldots,\mathbf{y}(\tau_k,\omega_q)]. \tag{20}$$

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

$$\tau_1 = t_1 < \ldots < t_p = \tau_N, \tag{21}$$

where $t_1 = \tau_{n_0}$, $t_2 = \tau_{n_0+n_1}$, $\ldots$, $t_p = \tau_{n_0+n_1\ldots+n_{p-1}}$, 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 $j = 1,\ldots,p$, vectors $\widehat{\mathbf{x}}(tj,\cdot)$ and $\mathbf{y}(tj,\cdot)$ associated with t_j in Equation (21) are represented, respectively, by

$$\widehat{X}_j^{(k)} = [\widehat{\mathbf{x}}(t_j, \omega_1), \ldots, \widehat{\mathbf{x}}(t_j, \omega_N)] \quad \text{and} \quad Y_j = [\mathbf{y}(t_j, \omega_1), \ldots, \mathbf{y}(t_j, \omega_N)].$$

4.3.2. Algorithm

As it has been mentioned in Section 3.4, 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 3.4 and 4.2, the performance algorithm of the proposed filter consists of the following steps (Algorithm 1). For $j = 1 \ldots, p$, we write $N_j = n_0 + n_1 + \ldots + n_{j-1}$.

Algorithm 1 Computation of $\widehat{X}^{(2)}$, $\widehat{X}^{(3)}$,..., $\widehat{X}^{(N)}$ given by Theorem 1

Initial parameters: $Y^{(1)}, \ldots, Y^{(N)}$, $\{t_j\}_{j=1}^p$ (see Equation (21)), $\{E_{z_j w_j}\}_{j=1}^p$, $\{E_{w_j w_j}\}_{j=1}^p$ (see Equations (14) and (15)), $\widehat{X}_1$, $n_0 = 1$, $N_0 = 0$ and $M_{B_j} = \mathbb{O}$, for $j = 1, \ldots, p-1$.)
The condition $M_{B_j} = \mathbb{O}$, for $j = 1, \ldots, p-1$, is motivated by a desire to consider the minimal norm representation of B_j, for $j = 1, \ldots, p-1$, in Equation (19). Possible ways to obtain estimates of $E_{z_j w_j}$ and $E_{w_j w_j}$ are discussed below in Section 4.5.
Final parameters: $\widehat{X}^{(2)}$, $\widehat{X}^{(3)}$,..., $\widehat{X}^{(N)}$.
for $j \leftarrow 1, \ldots, p$ **do**
  $B_j \leftarrow E_{z_j w_j} E^{\dagger}_{w_j w_j}$
  for $k \leftarrow N_{j-1} + 1, \ldots, N_j$ **do**
    $\widehat{X}^{(k)} \leftarrow \widehat{X}_j + B_j(Y^{(k)} - Y_j)$
    $\widehat{X}_{j+1} \leftarrow \widehat{X}^{(N_j)}$
  end for
end for

4.4. Error Analysis

It is natural to expect that the error associated with the piecewise interpolating filter $\mathcal{F}^{(p-1)}$ decreases when $\max\limits_{j=1,\ldots,p-1} \Delta t_j$ 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

$$\|\mathbf{x}(t, \cdot)\|^2_{T,\Omega} = \frac{1}{b-a} \int_T \|\mathbf{x}(t, \cdot)\|^2_{\Omega} dt. \tag{22}$$

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 $t \in [t_j,\ t_{j+1}]$,

$$\|\mathbf{x}(t, \omega) - \mathbf{x}(t_j, \omega)\|^2_{T,\,\Omega} \leq \lambda_j \Delta t_j \quad \text{and} \quad \|\mathbf{y}(t, \omega) - \mathbf{y}(t_{j+1}, \omega)\|^2_{T,\,\Omega} \leq \gamma_j \Delta t_j \tag{23}$$

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

Theorem 2. *Under the conditions (Equation (23)) 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:*

$$\begin{aligned} &\|\mathbf{x}(t, \omega) - F^{(p-1)}[\mathbf{y}(t, \omega)]\|^2_{T,\,\Omega} \\ &\quad \leq \max_{j=1,\ldots,p-1} \Big[(\lambda_j + \gamma_j \|B_j\|^2) \Delta t_j + \|E^{1/2}_{z_j z_j}\|^2 - \|E_{z_j w_j} (E^{1/2}_{w_j w_j})^{\dagger}\|^2 \Big]. \end{aligned} \tag{24}$$

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

Further, to show that the error of the reference signal estimate tends to 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 Δt_1, i.e. that, for some constant $c_1 \geq 0$,

$$\|\mathbf{x}(t,\omega) - \widehat{\mathbf{x}}(t_1,\omega)\|^2_{\Omega} \leq c_1 \Delta t_1, \quad \text{for } t \in [t_1,\ t_2]. \tag{25}$$

Theorem 3. *Let the conditions (23) and (25) 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:*

$$\|\mathbf{x}(t,\omega) - F^{(p-1)}[\mathbf{y}(t,\omega)]\|^2_{T,\,\Omega} \to 0 \quad as \quad \max_{j=1,\ldots,p-1} \Delta t_j \to 0 \quad and \quad p \to \infty. \tag{26}$$

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

Remark 2. *We would like to emphasize that the statement of Theorem 3 is fulfilled only under Equations (23) and (25). At the same time, Equations (23) and (25) are not restrictive from a practical point of view. Equation (23) is true for Lipschitz continuous signals* $\mathbf{x}$ *and* $\mathbf{y}$*, i.e., for a very wide class of signals. Equation (25) is achieved by choosing an appropriate known method (e.g., see [1–11,13–16]) to find the estimate $\widehat{\mathbf{x}}(t_1,\omega)$ used in the proposed filter $\mathcal{F}^{(p-1)}$ (see Equation (8) and Theorem 1).*

4.5. Estimation of Covariance Matrices in Equation (19)

Matrix $E_{z_j w_j}$ used in Equation (19), for $j = 1,\ldots,p$, can be estimated as follows:

1. A popular method of estimating $E_{z_j w_j}$ is provided, for example, in [27]—it is based on the use of samples of z_j and w_j, for $j = 1,\ldots,p$.
2. In the case of incomplete observations, the method proposed in [28,29] can be used.
3. Let $E_{\hat{z}_j w_j}$ be a matrix obtained from matrix $E_{z_j w_j}$ 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 $E_{\hat{z}_j w_j}$ can be considered as an estimate of $E_{z_j w_j}$.
4. In the important case of an *additive* noise, $E_{z_j w_j}$ 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 $E_{z_j w_j}$ can be represented as follows:
$$E_{z_j w_j} = E_{(y_{j+1}-\xi_{j+1})(y_{j+1}-y_j)} - E_{\widehat{x}_j(y_{j+1}-y_j)} \tag{27}$$

We note that the RHS of Equation (27) 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 Equation (27), 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 Equation (27), $E_{\widehat{x}_j(y_{j+1}-y_j)}$, 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 Equation (27), $E_{(y_{j+1}-\xi_{j+1})(y_{j+1}-y_j)}$, is similar to the related covariance matrix in the Wiener filtering approach [5].

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

In general, an estimation of covariance matrices is a special research topic that is not a subject of this paper. The relevant references can be found, for example, in [5,29].

5. Simulations

5.1. General Consideration

The simulations represented below have been performed by MATLAB R14 (Version 7.0). In accordance with Section 4.3.1, signal sets K_X and K_Y (see Section 2.1) are given by

$$K_X = \{\mathbf{x}(\tau_1,\cdot),\mathbf{x}(\tau_2,\cdot),\ldots,\mathbf{x}(\tau_N,\cdot)\} \quad \text{and} \quad K_Y = \{\mathbf{y}(\tau_1,\cdot),\mathbf{y}(\tau_2,\cdot,\ldots,\mathbf{y}(\tau_N,\cdot)\},$$

where, for $k = 1,\ldots,N$, $\mathbf{x}(\tau_k,\cdot) \in L^2(\Omega,\mathbb{R}^m)$ and $\mathbf{y}(\tau_k,\cdot) \in L^2(\Omega,\mathbb{R}^n)$. In many practical problems (arising, for example, in a DNA analysis the number N is quite large; for instance, $N = \mathcal{O}(10^4)$.

We set $N = 141$ and $m = n = 116$. Thus, in these simulations, the interval $T = [a,\ b]$ (see Sections 2.1 and 4.3.1) is modeled as 141 points τ_k with $k = 1,\ldots,141$ so that $[a,\ b] = [\tau_1,\tau_2,\ldots,\tau_{141}]$.

The sequence of fixed time-points $\{t_k\}_1^p \subset T$ in Equation (1) is now such that

$$\tau_1 = t_1 < \ldots < t_p = \tau_{141}. \tag{28}$$

Below, in Examples 1–12, four particular choices of the specific interpolation signal pairs $\{\mathbf{x}(t_j,\cdot),\mathbf{y}(t_j,\cdot)\}_1^p$ (introduced in Section 3.4) are considered, for $p = 5,8,15$ and 28. Points $t_1,\ldots,t_p$ are as follows.

For $p = 5,8,15$, if $j = 1,\ldots,p$, then $t_j = t_j^{(p)} = \tau_1 + (j-1)\Delta_p$, respectively, where $\Delta_5 = 35$, $\Delta_8 = 20$ and $\Delta_{15} = 10$.

For $p = 28$, if $j = 1,\ldots,p-1$, then $t_j = t_j^{(p)} = \tau_1 + (j-1)\Delta_{28}$, and if $j = p$, then $t_{28} = t_{28}^{(28)} = t_{27} + 6 = 141$, where $\Delta_{28} = 5$.

Signals $\mathbf{x}(\tau_k,\cdot)$ and $\mathbf{y}(\tau_k,\cdot)$ have been simulated as digital images represented by 116×256 matrices

$$X^{(k)} = [\mathbf{x}(\tau_k,\omega_1),\ldots,\mathbf{x}(\tau_k,\omega_{256})] \quad \text{and} \quad Y^{(k)} = [\mathbf{y}(\tau_k,\omega_1),\ldots,\mathbf{y}(\tau_k,\omega_{256})], \tag{29}$$

respectively, for $k = 1,\ldots,141$, so that $X^{(k)}$ represents an image that should be estimated from an observed image $Y^{(k)}$. A column of matrices $X^{(k)}$ and $Y^{(k)}$, $\mathbf{x}(\tau_k,\omega_i) \in \mathbb{R}^{116}$ and $\mathbf{y}(\tau_k,\omega_i) \in \mathbb{R}^{116}$, for $i = 1,\ldots,256$, represents a realization of signals $\mathbf{x}(\tau_k,\cdot)$ and $\mathbf{y}(\tau_k,\cdot)$, respectively.

Note that $X^{(1)},\ldots,X^{(141)}$ did not used in the piecewise linear filter $F^{(p-1)}$ below since they are not supposed to be known. They are represented here for illustration purposes only. In particular, $X^{(1)},\ldots,X^{(141)}$ are used to compare their estimates by different filters.

Observed noisy signals $Y^{(1)},\ldots,Y^{(141)}$ have been simulated in different forms presented by Equations (40) and (49)–(51) in the Examples 1–12 below. We note that the considered observed signals are grossly corrupted.

To estimate the signals $X^{(1)}$, ..., $X^{(141)}$ from the observed signals $Y^{(1)}$, ..., $Y^{(141)}$, the proposed piecewise linear filter $F^{(p-1)}$, the generic optimal linear (GOL) filters [5] and the averaging polynomial filter [16] have been used.

The filters proposed in [11–13,16] have not been applied here by the reasons discussed in Section 1. In particular, the filter in [13] cannot be applied to signals represented by $Y^{(1)}$, ..., $Y^{(141)}$ in the form Equations (40), (49)–(51) below because the associated inverse matrices used in [13] do not exist.

For signals under consideration (given by matrices $X^{(k)}$ and $Y^{(k)}$ with $k = 1,\ldots,141$), the filter $F^{(p-1)}$, the generic optimal linear (GOL) filters [5] and the averaging polynomial filter [15,16] are represented as follows:

(i) Piecewise linear filter $F^{(p-1)}$. For $j = 1, \ldots, p$, $\{X_j, Y_j\}$ designates an interpolation pair defined similarly to that in Section 3.4. Each X_j and Y_j is associated with t_j in Equation (28) so that

$$X_j = [\mathbf{x}(t_j, \omega_1), \ldots, \mathbf{x}(t_j, \omega_{256})] \quad \text{and} \quad Y_j = [\mathbf{y}(t_j, \omega_1), \ldots, \mathbf{y}(t_j, \omega_{256})].$$

The estimate $\widehat{X}^{(k)}$ of $X^{(k)}$ by the filter $F^{(p-1)}$ is given by

$$\widehat{X}^{(k)} = F^{(p-1)}[Y^{(k)}], \tag{30}$$

where Equations (16)–(19) are presented in Section 4.2,

$$F^{(p-1)}[Y^{(k)}] = \sum_{j=1}^{p-1} \delta_j F_j[Y^{(k)}], \quad \delta_j = \begin{cases} 1, & \text{if } t_j \le \tau_k \le t_{j+1}, \\ 0, & \text{otherwise,} \end{cases} \tag{31}$$

$$F_j^{(p-1)}[Y^{(k)}] = \widehat{X}_j + B_j[Y^{(k)} - Y_j], \tag{32}$$

$$\widehat{X}_j = F_{j-1}[Y_j], \quad \widehat{X}_1 \text{ is given}, \tag{33}$$

$$B_j = E_{Z_j W_j}(E_{W_j W_j})^{\dagger}, \tag{34}$$

and where $E_{Z_j W_j}$ and $E_{W_j W_j}$ are estimates of matrices $E_{z_j w_j}$ and $E_{w_j w_j}$ in Equation (19), respectively. In particular, $E_{W_j W_j}$ can be represented in the form

$$E_{W_j W_j} = W_j W_j^T, \quad \text{where } W_j = Y_{j+1} - Y_j. \tag{35}$$

Further, matrix $E_{Z_j W_j}$ depends on $Z_j = X_{j+1} - \widehat{X}_j$ where X_{j+1} is unknown. Therefore, a determination of $E_{Z_j W_j}$ is reduced, in fact, to finding an estimate of X_{j+1}. Since it is customary to find $E_{Z_j W_j}$ in terms of signal samples [5], $E_{Z_j W_j}$ has been presented as

$$E_{Z_j W_j} = \widetilde{Z}_j W_j^T, \quad \text{where } \widetilde{Z}_j = \widetilde{X}_{j+1} - \widehat{X}_j \tag{36}$$

and $\widetilde{X}_{j+1}$ has been constructed from a sample of X_{j+1} as follows. The sample of X_{j+1} is a 116×128 matrix presented by odd columns of X_{j+1}. Then an estimate of X_{j+1} is chosen as a 116×256 matrix $\widetilde{X}_{j+1}$ where each odd column is a related odd column of X_{j+1}, and each even column is an average of two adjacent columns. The last column in $\widetilde{X}_{j+1}$ is the same as its preceding column.
This way of estimating $E_{Z_j W_j}$ was chosen for illustration purposes only. Other related methods have been considered in Section 4.5.
The errors associated with the filter $F^{(p-1)}$ are given by

$$\varepsilon_{k,F}^{(p-1)} = \left\| X^{(k)} - F^{(p-1)}[Y^{(k)}] \right\|_F^2, \quad \text{for } k = 1, \ldots, 141. \tag{37}$$

(ii) Generic optimal linear (GOL) filters [5]. To each signal $Y^{(k)}$, an individual GOL filter W_k has also been applied, so that W_k estimates $X^{(k)}$ from $Y^{(k)}$ in the form

$$W_k Y^{(k)} = E_{X^{(k)} Y^{(k)}} E_{Y^{(k)} Y^{(k)}}^{\dagger} Y^{(k)},$$

for each $k = 1, \ldots, 141$. Thus, the GOL filter W_k requires an estimate of 141 matrices $E_{X^{(k)} Y^{(k)}}$, for each $k = 1, \ldots, 141$.
Similarly to matrix $E_{Z_j W_j}$ in the filter $F^{(p-1)}$ above, the matrix $E_{X^{(k)} Y^{(k)}}$ has been estimated from samples of each $X^{(k)}$, $\widetilde{X}^{(k)}$, for each $k = 1, \ldots, 141$.

One of the advantages of the proposed filter $F^{(p-1)}$ is that $F^{(p-1)}$ requires a smaller number, p, of samples of X_j, $\widetilde{X}_j$, to be known (where $j = 1, \ldots, p$).
The errors associated with filters W_k are given by

$$\epsilon_{k,w} = \|X^{(k)} - W_k Y^{(k)}\|_F^2. \tag{38}$$

(iii) Averaging polynomial filters [15,16]. By the methodology in [15], the averaging polynomial filter W is based on the use of the estimates of the covariance matrices, E_{XY} and E_{YY}, in the form

$$E_{XY} = \frac{1}{141}\sum_{k=1}^{141} \widetilde{X}^{(k)}(Y^{(k)})^T \quad \text{and} \quad E_{YY} = \frac{1}{141}\sum_{k=1}^{141} Y^{(k)}(Y^{(k)})^T.$$

Then, for each, $k = 1, \ldots, 141$, the estimate of $X^{(k)}$ is given by

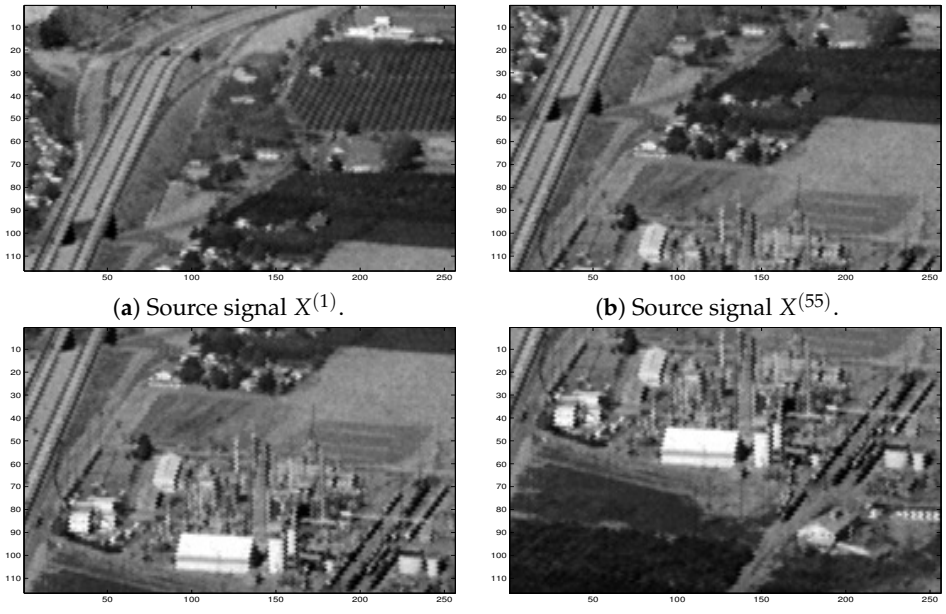
$$WY^{(k)} = E_{XY}E_{YY}^{\dagger}Y^{(k)}.$$

The errors associated with the filter W are given by

$$\varepsilon_{k_W} = \|X^{(k)} - WY^{(k)}\|_F^2, \quad \text{for } k = 1, \ldots, 141. \tag{39}$$

5.2. Simulations with Signals Modelled from Images 'Plant': Application of Piecewise Interpolation Filter and GOL Filters

Here, results of simulations for reference signals represented by matrices $X^{(1)}, \ldots, X^{(141)}$ (see Equation (29) above) formed from images 'plant' from http://sipi.usc.edu/services/database.html (accessed on 21 February 2025). Typical selected images $X^{(k)}$ are shown in Figure 1.



(**a**) Source signal $X^{(1)}$. (**b**) Source signal $X^{(55)}$.

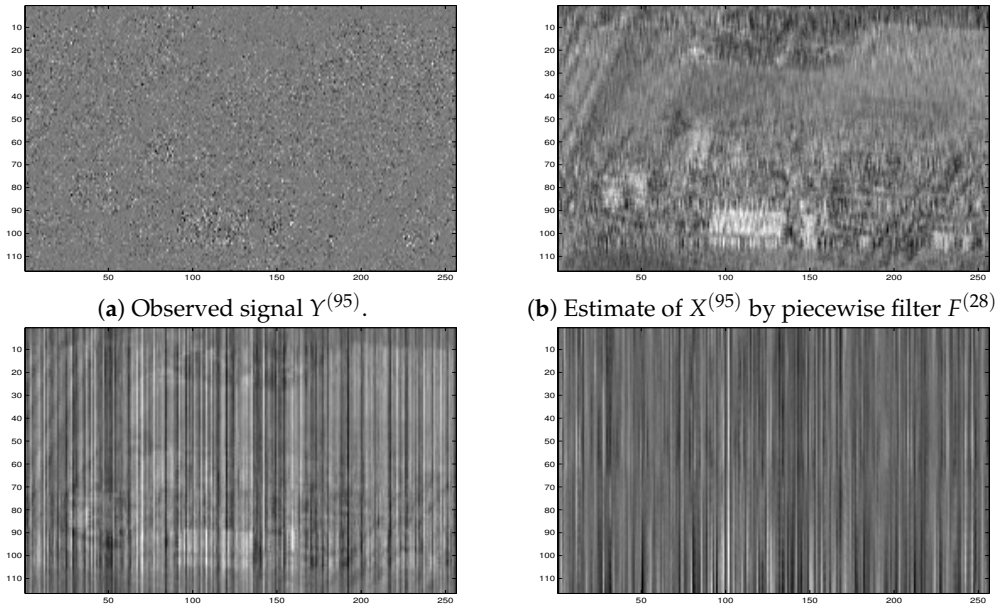
(**c**) Source signal $X^{(95)}$. (**d**) Source signal $X^{(141)}$.

Figure 1. Examples of selected signals to be estimated from observed data.

Observed noisy images $Y^{(1)}, \ldots, Y^{(141)}$ have been simulated in the form

$$Y^{(k)} = X^{(k)} \bullet \texttt{randn}_{(k)} \bullet \texttt{rand}_{(k)}, \tag{40}$$

for each $k = 1, \ldots, 141$. Here, $\bullet$ means the Hadamard product, and $\texttt{randn}_{(k)}$ and $\texttt{rand}_{(k)}$ are 116×256 matrices with random entries. The entries of $\texttt{randn}_{(k)}$ are normally distributed with mean zero, variance one and standard deviation one. The entries of $\texttt{rand}_{(k)}$ are uniformly distributed in the interval $(0, 1)$. A typical example of such images is given in Figure 2a. Examples of signal $X^{(95)}$ estimates by different filters are represented in Figure 2b–d.



(**a**) Observed signal $Y^{(95)}$. (**b**) Estimate of $X^{(95)}$ by piecewise filter $F^{(28)}$.

(**c**) Estimate of $X^{(95)}$ by the GOL filter. (**d**) Estimate of $X^{(95)}$ by averaging polynomial filter.

Figure 2. Examples of the observed signal and the estimates obtained by different filters.

To demonstrate the effectiveness of the proposed filter $F^{(p-1)}$, sub-filters $F_j^{(p-1)}$ and associated interpolation signal pairs $\{X_j, Y_j\}_{j=1}^{p}$ have been chosen in four different ways as follows.

Example 1. *First, for $p = 5$, the interpolation signal pairs are*

$$\{X_1, Y_1\} = \{X^{(1)}, Y^{(1)}\}, \quad \{X_2, Y_2\} = \{X^{(35)}, Y^{(35)}\}, \quad \{X_3, Y_3\} = \{X^{(70)}, Y^{(70)}\}, \tag{41}$$

$$\{X_4, Y_4\} = \{X^{(105)}, Y^{(105)}\}, \quad \{X_5, Y_5\} = \{X^{(141)}, Y^{(141)}\}. \tag{42}$$

The error values $\{\varepsilon_{k,F}^{(4)}\}_1^{141}$ associated with filter $F^{(4)}$ are evaluated by Equation (37). The graph of $\{\varepsilon_{k,F}^{(4)}\}_1^{141}$ is presented in Figure 3a.

Example 2. *For $p = 8$, the interpolation signal pairs are*

$$\{X_1, Y_1\} = \{X^{(1)}, Y^{(1)}\}, \quad \{X_j, Y_j\} = \{X^{(20(j-1))}, Y^{(20(j-1))}\}, \quad \textit{for } j = 2, \ldots, 7; \tag{43}$$

$$\textit{and} \quad \{X_8, Y_8\} = \{X^{(141)}, Y^{(141)}\}. \tag{44}$$

The error magnitudes $\{\varepsilon_{k,F}^{(7)}\}_1^{141}$ associated with the piecewise interpolation filter $F^{(7)}$ constructed by Equations (31)–(36) with the interpolation signal pairs given by Equations (43) and (44) are diagrammatically shown in Figure 3b.

It follows from Figure 3b that the errors associated with filter $F^{(7)}$ is less than those of filter $F^{(4)}$. This is a confirmation of Theorem 3.

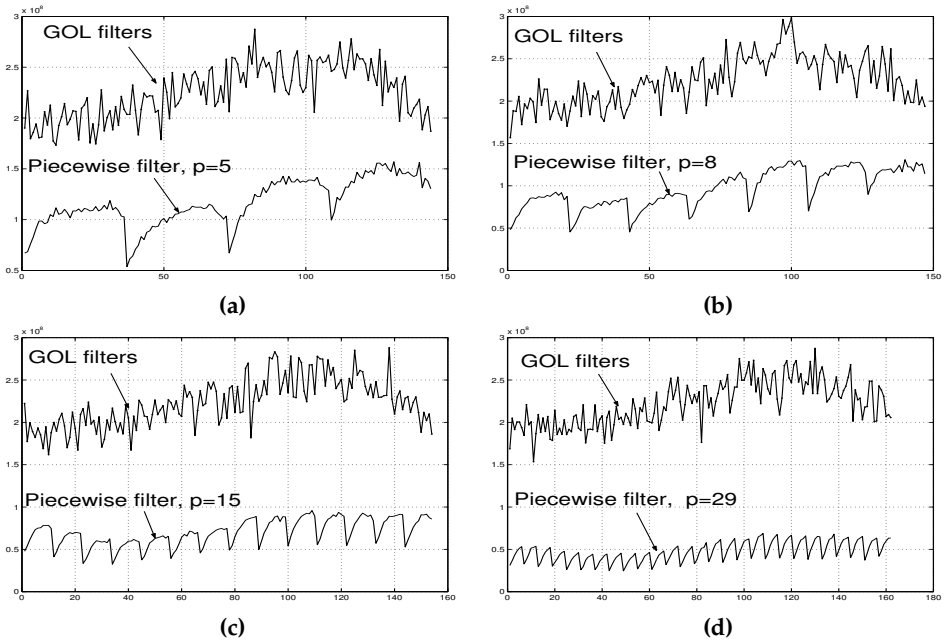


Figure 3. Diagrams of errors associated with the piecewise interpolation filters $F^{(p-1)}$ and the GOL filters [5] applied to signals described in Examples 1–4: (**a**) for $p = 5$; (**b**) for $p = 8$; (**c**) for $p = 15$; (**d**) for $p = 29$.

Example 3. *Further, for* $p = 15$*, the interpolation pairs are*

$$\{X_1, Y_1\} = \{X^{(1)}, Y^{(1)}\}, \quad \{X_j, Y_j\} = \{X^{(10(j-1))}, Y^{(10(j-1))}\} \quad \text{for } j = 2, \ldots, 14; \tag{45}$$

$$\text{and} \quad \{X_{15}, Y_{15}\} = \{X^{(141)}, Y^{(141)}\}. \tag{46}$$

In Figure 3c, the errors $\{\varepsilon_{k,F}^{(15)}\}_1^{141}$ *associated with the piecewise interpolation filter* $F^{(15)}$ *are presented. Figure 3c demonstrates a further confirmation of Theorem 3: the errors associated with the piecewise interpolation filter diminish as p increases.*

Example 4. *Finally, the number of interpolation signal pairs* $\{X_j, Y_j\}_{j=1}^{p}$ *is* $p = 29$ *so that*

$$\{X_1, Y_1\} = \{X^{(1)}, Y^{(1)}\}, \quad \{X_j, Y_j\} = \{X^{(5(j-1))}, \quad Y^{(5(j-1))}\} \quad \text{for } j = 2, \ldots, 28; \tag{47}$$

$$\text{and} \quad \{X_{29}, Y_{29}\} = \{X^{(141)}, Y^{(141)}\}. \tag{48}$$

In this case, when p is greater than in the previous Examples 1–3, the errors $\{\varepsilon_{k,F}^{(29)}\}_1^{141}$ *associated with the piecewise interpolation filter* $F^{(29)}$ *are smaller than those associated with filters* $F^{(4)}$, $F^{(8)}$ *and* $F^{(15)}$*—see Figure 3d.*

The diagrams of errors associated with the GOL filters [5] are also presented in Figure 3, which shows that the proposed filters $F^{(4)}$, $F^{(8)}$, $F^{(15)}$ and $F^{(29)}$ provide better accuracy then that of the GOL filters.

At the same time, the filter $F^{(p-1)}$ is easier to implement since it requires less initial information compared to GOL filters, as it has been discussed in Sections 1.1.1 and 1.1.4.

Results of the application of the averaging polynomial filter [15,16] are discussed in Section 5.4 below.

5.3. Simulations with Signals Modelled from Images 'Boat': Application of Piecewise Interpolation Filter and GOL Filters

In this section, results of the simulations for a different type of signal than those considered in Section 5.2 above are presented. Here, the reference signals $X^{(1)}, \ldots, X^{(141)}$ are

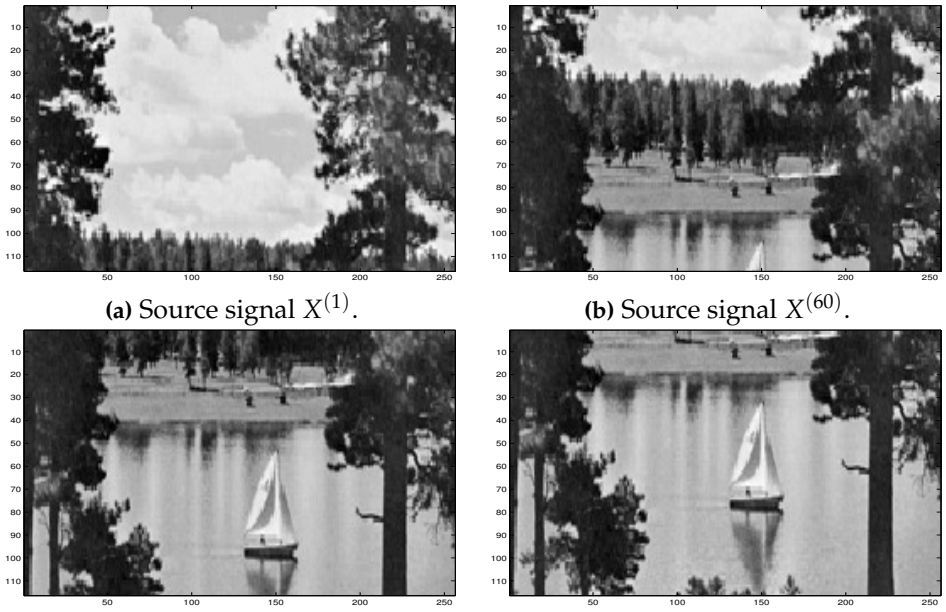
formed from images 'boat' (http://sipi.usc.edu/services/database.html, accessed 21 February 2025).

Observed noisy signals $Y^{(1)}, \ldots, Y^{(141)}$ have been simulated in the form

$$Y^{(k)} = X^{(k)} \bullet \mathtt{randn}_{(k)}, \tag{49}$$

for each $k = 1, \ldots, 141$. The noise term is different from that in Equation (40).

Typical selected images $X^{(k)}$ and $Y^{(k)}$ are shown in Figures 4 and 5, respectively.

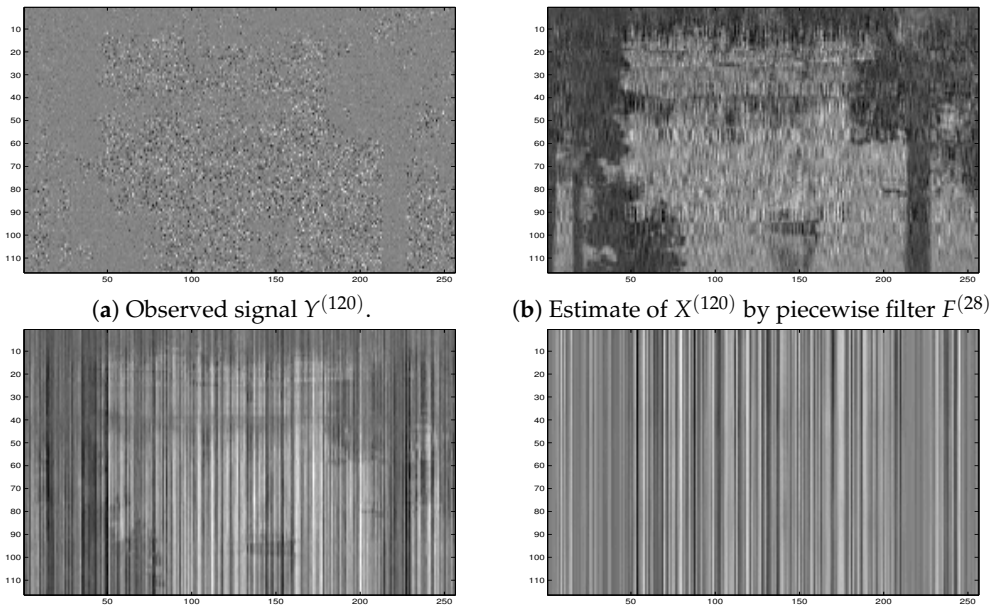


(a) Source signal $X^{(1)}$. **(b)** Source signal $X^{(60)}$.

(c) Source signal $X^{(120)}$. **(d)** Source signal $X^{(141)}$.

Figure 4. Examples of selected signals to be estimated from observed data considered in Examples 5–9.

As in Section 5.2, the piecewise interpolation filter $F^{(p-1)}$ is constructed by Equations (31)–(36). In Examples 5–8 below, the number $p-1$ of sub-filters $F_j^{(p-1)}$ and associated interpolation signal pairs $\{X_j, Y_j\}_{j=1}^{p}$ have been chosen in four different ways.



(**a**) Observed signal $Y^{(120)}$. (**b**) Estimate of $X^{(120)}$ by piecewise filter $F^{(28)}$.

(**c**) Estimate of $X^{(120)}$ by GOL filter. (**d**) Estimate of $X^{(120)}$ by averaging polynomial filter.

Figure 5. Examples of the observed signal and the estimates obtained by different filters.

Example 5. *First, similar to Example 1, the number of interpolation signal pairs* $\{X_j, Y_j\}_{j=1}^{p}$ *has been chosen as* $p = 5$*, and* X_j *and* Y_j *have been presented as in Equations (41) and (42).*

The error values $\{\varepsilon_{k,F}^{(5)}\}_1^{141}$ *associated with the piecewise interpolation filter* $F^{(5)}$ *applied to these data are presented in Figure 6a.*

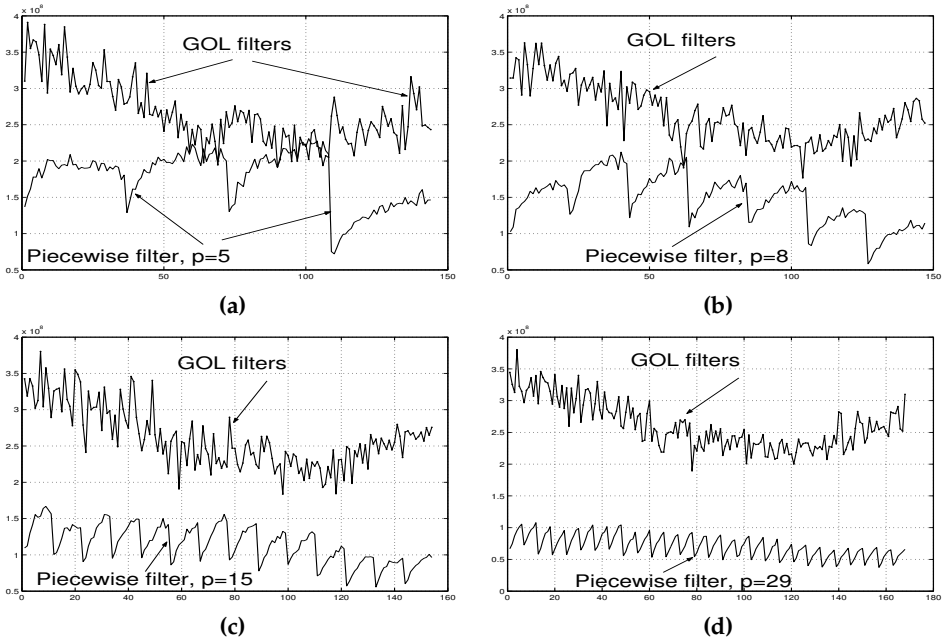


Figure 6. Diagrams of errors associated with the piecewise interpolation filter $F^{(p-1)}$ of order p and the generic optimal linear (GOL) filters [5] applied to signals described in Examples 5–8: (**a**) for $p = 5$; (**b**) for $p = 8$; (**c**) for $p = 15$; (**d**) for $p = 29$.

Example 6. *For the greater number of interpolation signal pairs than that in Example 5,* $p = 8$*, and for* X_j *and* Y_j $(j = 1, \ldots, 8)$ *chosen as in Equations (43) and (44), the error magnitudes* $\{\varepsilon_{k,F}^{(8)}\}_1^{141}$ *associated with the piecewise interpolation filter* $F^{(8)}$ *are diagrammatically shown in Figure 6b. A comparison between Figure 6a,b demonstrates that the increase in* p *implies the decrease in the errors associated with the filter* $F^{(p-1)}$*.*

Example 7. *For* $p = 15$*, and for* X_j *and* Y_j $(j = 1, \ldots, 15)$ *chosen in Equations (45) and (46), the errors* $\{\varepsilon_{k,F}^{(15)}\}_1^{141}$ *associated with the piecewise interpolation filter* $F^{(15)}$ *are further less than those for filters* $F^{(5)}$ *and* $F^{(8)}$*. See Figure 6c in this regard.*

Example 8. *The further increase in* p *to* $p = 29$ *confirms this tendency. The piecewise interpolation filter* $F^{(29)}$ *with* X_j *and* Y_j $(j = 1, \ldots, 29)$ *chosen similar to Equations (47) and (48) produces the associated errors* $\{\varepsilon_{k,F}^{(29)}\}_1^{141}$ *represented in Figure 6d. They are, clearly, less than the errors associated with filters* $F^{(5)}$*,* $F^{(8)}$ *and* $F^{(15)}$*.*

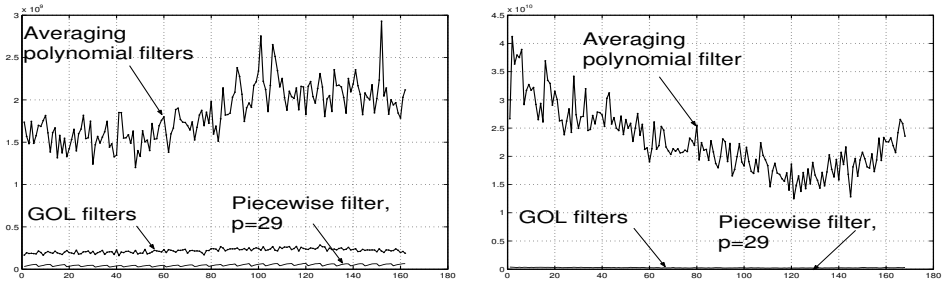
The errors associated with the GOL filters are also presented in Figure 6a–d. The figures clearly demonstrate the advantage of the piecewise interpolation filter $F^{(p-1)}$.

Results of the application of the averaging polynomial filter [16] are discussed in Section 5.4 below.

5.4. Results of Simulations for Averaging Polynomial Filter [15,16]

To further illustrate the effectiveness of the proposed piecewise interpolation filter, in this Section, results of simulations for the averaging polynomial filter [15,16] are presented. The filter has been applied to two different types of data considered in Sections 5.2 and 5.3.

Example 9. *The filter [15,16] applied to signals considered in Section 5.2 gives the associated errors* $\{\epsilon_{k_W}\}_{k=1}^{141}$ *(see Equation (39)) represented in Figure 7a. For a comparison, the errors associated with the piecewise interpolation filter* $F^{(29)}$ *and the GOL filters [5] are also given in Figure 7a.*



(a) **(b)**

Figure 7. Diagrams of errors associated with the averaging polynomial filters [15,16] in (**a**) Example 9 and (**b**) Example 10.

A typical example of the estimated signal by the averaging polynomial filter [15,16] is presented in Figure 2d above.

Example 10. *The averaging polynomial filter [16] applied to signals considered in Section 5.3 produces the associated errors* $\{\epsilon_{k_W}\}_{k=1}^{141}$ *shown in Figure 7b. The errors associated with the piecewise interpolation filter* $F^{(29)}$ *are much smaller, and they are not discerned in Figure 7b.*

The extreme failure of the averaging polynomial filter (the errors of magnitude 10^9) in Examples 9 and 10 illustrates the observation made in Section 1.2.4: The associated error increases when signals differ considerably from each other. This is the case in Examples 9 and 10.

Together, Figures 2, 3, 5, 6 and 7a,b illustrate the advantage of the piecewise interpolation filter.

5.5. Further Simulations with Different Type of Noise

In Examples 11 and 12 below, a different type of noise is considered. Unlike the multiplicative noise in Equations (40) and (49), here, the noise is additive.

Example 11. *First, the piecewise interpolation filter* $F^{(28)}$*, the GOL filters [5] and the averaging polynomial filter [16] have been applied to the observed signals given by*

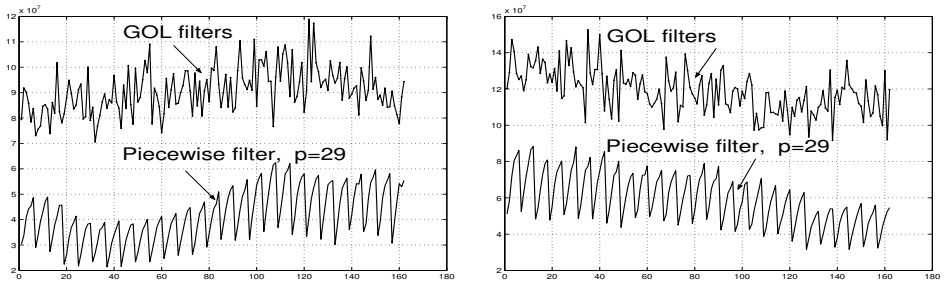
$$Y^{(k)} = X^{(k)} + 900 \times \texttt{randn}_{(k)}, \quad \text{for } k = 1, \ldots, 141. \tag{50}$$

where $X^{(k)}$ *is as in Section 5.2, i.e.,* $X^{(k)}$ *is formed from the images 'plant'. In Figure 8a, the diagrams of the errors associated with filter* $F^{(28)}$ *and the GOL filters [5] are given. The errors associated with the averaging polynomial filter [16],* $\{\epsilon_{k_W}\}_{k=1}^{141}$*, are much greater (of order* $\mathcal{O}(10^9)$*) and they are not presented in Figure 8a.*

Example 12. *In this example, the reference signals* $X^{(1)}, \ldots, X^{(141)}$ *are as those in Section 5.3, i.e., they are formed from the image 'boat'. The observed signals are given by*

$$Y^{(k)} = X^{(k)} + 1000 \times \texttt{randn}_{(k)}, \quad \text{for } k = 1, \ldots, 141. \tag{51}$$

The piecewise interpolation filter $F^{(29)}$ *and the GOL filters [5] estimate the reference signals with the associated errors represented in Figure 8b. As in Example 11 above, in this case, the errors associated with the averaging polynomial filter [15,16] are much greater (of order* $\mathcal{O}(10^{10})$*), and they are not presented in Figure 8b.*



(a) **(b)**

Figure 8. Diagrams of errors associated with the piecewise interpolation filter $F^{(p-1)}$ and the generic optimal linear (GOL) filters [5] applied to signals described in (**a**) Example 11 and (**b**) Example 12.

Examples 11 and 12 further demonstrate the advantages of the proposed piecewise interpolation filter.

5.6. Summary of Simulations

The above simulations confirm the theoretical results obtained in Theorems 1–3. In particular, Figures 3 and 6 demonstrate that the error associated with the piecewise interpolation filter $F^{(p-1)}$ decreases when the number of sub-filters $F_1, \ldots, F_p$, p, increases.

A comparison between the proposed filter $F^{(p-1)}$ and the known related filters [5,12,13,15,16] has been performed. The filter $F^{(p-1)}$ estimates the reference signals with the accuracies that are much better than those of the generic optimal linear (GOL) filters [5] and the averaging polynomial filter [15,16]. Further, the filters proposed in [12,13] fail in processing the signals under consideration. This is because the observed signals in Equations (40), (49), (50) and (51) are grossly corrupted, and, therefore, the inverse matrices used in the filter structures in [13] do not exist. The technique in [12] requires the use of the reference signal in the proposed filter, which is supposed to be unknown in the simulations above.

The filters have been applied to the different signal sets (presented in Sections 5.2 and 5.3), using different forms of noise (given in Equations (40), (49), (50) and (51)).

Computational work associated with the proposed filter $F^{(p-1)}$ is substantially less than that associated with the known filters discussed in Section 1 (in particular, with the filters in [10–13,16,17,20–26]). This is because, for the processing of a data set containing N signals, filter $F^{(p-1)}$ requires computation of p covariance matrices with $p \ll N$ while the known filters require computation of N matrices (in the above examples, $p = 5, 7, 15, 28$, respectively, and $N = 141$).

Table 1 represents a typical time in seconds used by the piecewise linear interpolation filter (PILF), GOL and the averaging polynomial filter (APF) [16] in the above examples. It follows from Table 1 that the PILF is faster than the GOL up to ten times. While the time used by the PILF and APF for $p = 29$ is almost the same, the accuracy associated with the PILF is much better than that associated with the APF.

Table 1. Time in seconds used by PILF, GOL and APF in the above examples.

Initial Parameters	**Time in Seconds**		
	PILF	**COL**	**APF**
'Plant', $p = 8, N = 256$	0.37	3.72	1.25
'Boat', $p = 8, N = 256$,	0.36	3.81	1.17
'Plant', $p = 15, N = 256$	0.70	3.72	1.25
'Boat', $p = 15, N = 256$,	0.67	3.81	1.17
'Plant', $p = 29, N = 256$	1.23	3.72	1.25
'Boat', $p = 29, N = 256$,	1.18	3.81	1.17

Table 2 summarizes the computational and structural advantages of the proposed method. The abbreviation 'PINV' means the use of the pseudo-inverse matrices (not inverse matrices), and ε_{PILF}, ε_{GOL} and ε_{APF} denote the errors associated with the PILF, GOL and APF, correspondingly. 'Stability' means numerical stability. The last column illustrates features related to filters considered in [10–13,16,17,20–26]. As mentioned above, these filters cannot be applied to the data under consideration because of the singularity of matrices used in [10–13,16,17,20–26]. Therefore, their errors and a computational complexity are marked by N/A as not applicable.

Table 2. Summary of computational and structural advantages of the proposed method.

	PILF	**COL**	**APF**	**[10–13,16,17,20–26]**
PINV	Yes	Yes	Yes	No
Error	ε_{PILF}	$\varepsilon_{GOL} \geq \varepsilon_{PILF}$	$\varepsilon_{APF} \geq \varepsilon_{PILF}$	N/A
Comp. load	p by $n \times n$ PINV, $p \ll N$	N by $n \times n$ PINV, $N \gg p$	one $n \times n$ PINV	N/A
Stability	Yes	Yes	No	No

6. Conclusions

The technique of the constructive approximation of a nonlinear operator $\mathcal{F} : K_Y \to K_X$ has been provided. Here, $K_X = \{\mathbf{x}(t,\cdot) \in L^2(\Omega, \mathbb{R}^m) \mid t \in T\}$ and $K_Y = \{\mathbf{y}(t,\cdot) \in L^2(\Omega, \mathbb{R}^n) \mid t \in T\}$ where $T := [a, b] \subseteq \mathbb{R}$ and $\Omega = \{\omega\}$ is the set of outcomes of a probability space. The device behind the proposed method is based on a special extension of the piecewise linear interpolation technique to the case of stochastic sets K_X and K_Y. The proposed methodology is motivated by the problems arising in signal processing, where the nonlinear operator $\mathcal{F}$ is interpreted as a nonlinear system (or a nonlinear filter) transforming a set of stochastic signals K_Y to the set of stochastic signals K_X. Therefore, the provided technique provides an effective way to transform large data sets.

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 prior information only on *a 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 piece $K_{Y,j}$ 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 that is not specifically adjusted to each particular piece $K_{Y,j}$. Moreover, 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) 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.

Author Contributions: Conceptualization, methodology, writing—original draft, A.T.; numerical simulations, A.T. and P.P.; algorithm, P.P.; Matlab codes, P.P.; English amelioration, P.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Proof of Theorem 1. It follows from Equations (8) and (11) that α_j, for $j = 1, \ldots, p-1$, is given by

$$\alpha_j = \widehat{\mathbf{x}}(t_j, \omega) - B_j[\mathbf{y}(t_j, \omega)]. \tag{A1}$$

Further, for α_j given by Equation (A1),

$$\begin{aligned} &\left\| [\mathbf{x}(t_{j+1}, \cdot) - \alpha_j] - \mathcal{B}_j[\mathbf{y}(t_{j+1}, \cdot)] \right\|_\Omega^2 \\ &= \left\| \mathbf{z}(t_j, t_{j+1}, \cdot) - \mathcal{B}_j[\mathbf{w}(t_j, t_{j+1}, \cdot))] \right\|_\Omega^2 \qquad \text{(A2)} \\ &= \mathrm{tr}\{E_{z_j z_j} - E_{z_j w_j} B_j^T - B_j E_{w_j z_j} + B_j E_{w_j w_j} B_j^T\} \\ &= \|E_{z_j z_j}^{1/2}\|^2 - \|E_{z_j w_j}(E_{w_j w_j}^{1/2})^\dagger\|^2 + \|(B_j - E_{z_j w_j} E_{w_j w_j}^\dagger) E_{w_j w_j}^{1/2}\|^2 \\ &= \|E_{z_j z_j}^{1/2}\|^2 - \|E_{z_j w_j}(E_{w_j w_j}^{1/2})^\dagger\|^2 + \|E_{z_j w_j}(E_{w_j w_j}^{1/2})^\dagger - B_j E_{w_j w_j}^{1/2}\|^2, \qquad \text{(A3)} \end{aligned}$$

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

$$E_{z_j w_j} E_{w_j w_j}^\dagger E_{w_j w_j} = E_{z_j w_j} \tag{A4}$$

by Lemma 24 in [5]. Thus, the second expression in Equation (11) is reduced to the problem

$$\min_{B_j} \|E_{z_j w_j}(E_{w_j w_j}^{1/2})^\dagger - B_j E_{w_j w_j}^{1/2}\|^2. \tag{A5}$$

It is known (see, for example, [5], p. 304) that the solution of Equation (A5) is given by Equation (19).

Equation (17) follows from Equations (6) and (A1).

Theorem 1 is proven. □

Proof of Theorem 2. For $t \in [t_j,\ t_{j+1}]$ and F_j defined by Equations (17)–(19),

$$\begin{aligned} &\mathbf{x}(t, \omega) - F[\mathbf{y}(t, \omega)] \\ &= \mathbf{x}(t, \omega) - F_j[\mathbf{y}(t, \omega)] \\ &= \mathbf{x}(t, \omega) - \widehat{\mathbf{x}}(t_j, \omega) + B_j \mathbf{y}(t_j, \omega) - B_j \mathbf{y}(t, \omega) \\ &= [\mathbf{x}(t, \omega) - \mathbf{x}(t_{j+1}, \omega)] + \mathbf{z}(t_j, t_{j+1}, \omega) - B_j \mathbf{w}(t_j, t_{j+1}, \omega) \\ &\quad + B_j[\mathbf{y}(t_{j+1}, \omega) - \mathbf{y}(t, \omega)]. \qquad \text{(A6)} \end{aligned}$$

Then, Equation (A6) implies

$$\begin{aligned} \|\mathbf{x}(t,\omega) - F[\mathbf{y}(t,\omega)]\|^2_{T,\Omega} &\leq \|\mathbf{x}(t,\omega) - \mathbf{x}(t_{j+1},\omega)\|^2_{T,\Omega} \\ &+ \|\mathbf{z}(t_j,t_{j+1},\omega) - B_j\mathbf{w}(t_j,t_{j+1},\omega)\|^2_{\Omega} \\ &+ \|B_j[\mathbf{y}(t_{j+1},\omega) - \mathbf{y}(t,\omega)]\|^2_{T,\Omega} \end{aligned} \tag{A7}$$

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

It follows from Equations (A2) and (A3) that for B_j given by Equation (19),

$$\|\mathbf{z}(t_j,t_{j+1},\omega) - B_j\mathbf{w}(t_j,t_{j+1},\omega)\|^2_{\Omega} = \|E^{1/2}_{z_jz_j}\|^2 - \|E_{z_jw_j}(E^{1/2}_{w_jw_j})^{\dagger}\|^2. \tag{A8}$$

Then, Equations (16)–(19), (23) and (A6)–(A8) imply that for all $t \in [a,\ b]$ and $\omega \in \Omega$, Equation (24) is true. □

Proof of Theorem 3. Equation (22) implies that

$$\|\mathbf{x}(t,\omega) - F[\mathbf{y}(t,\omega)]\|^2_{T,\Omega} = \frac{1}{b-a}\sum_{j=1}^{p-1}\int_{t_j}^{t_{j+1}} \|\mathbf{x}(t,\omega) - F_j[\mathbf{y}(t,\omega)]\|^2_{\Omega}dt, \tag{A9}$$

where

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

Then,

$$\begin{aligned} &\int_{t_j}^{t_{j+1}} \|\mathbf{x}(t,\omega) - F_j[\mathbf{y}(t,\omega)]\|^2_{\Omega}dt \qquad (A10) \\ &\quad \leq \int_{t_j}^{t_{j+1}} \|\mathbf{x}(t,\omega) - \mathbf{x}(t_j,\omega)\|^2_{\Omega}dt + \int_{t_j}^{t_{j+1}} \|\mathbf{x}(t_j,\omega) - \widehat{\mathbf{x}}(t_j,\omega)\|^2_{\Omega}dt \\ &\qquad\qquad + \|B_j\| \int_{t_j}^{t_{j+1}} \|\mathbf{y}(t_j,\omega) - \mathbf{y}(t,\omega)\|^2_{\Omega}dt \\ &\quad \leq \lambda_j(\Delta t_j)^2 + \|\mathbf{x}(t_j,\omega) - \widehat{\mathbf{x}}(t_j,\omega)\|^2_{\Omega}\Delta t_j + \|B_j\|\gamma_j(\Delta t_j)^2 \qquad (A11) \end{aligned}$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$\|\mathbf{x}(t_{k+1},\omega) - \widehat{\mathbf{x}}(t_{k+1},\omega)\|_\Omega^2 = \|\mathbf{x}(t_{k+1},\omega) - F_k\mathbf{y}(t_{k+1},\omega)\|_\Omega^2 \le \beta_k\Delta t. \tag{A12}$$

Therefore, Equations (A10), (A11) and (A12) imply

$$\begin{aligned}
&\int_{t_j}^{t_{j+1}} \|\mathbf{x}(t,\omega) - F_j[\mathbf{y}(t,\omega)]\|_\Omega^2 dt \\
&\le \lambda_j(\Delta t_j)^2 + \beta_{j-1}(\Delta t_j)^2 + \|B_j\|\gamma_j(\Delta t_j)^2 \\
&\le \eta_j(\Delta t)^2
\end{aligned} \tag{A13}$$

where $\eta_j = \lambda_j + \beta_{j-1} + \|B_j\|$, and then it follows from Equations (A9)–(A11) and (A13) that for all $t \in [a,\ b]$,

$$\|\mathbf{x}(t,\omega) - F[\mathbf{y}(t,\omega)]\|_{T,\Omega}^2 \le \frac{1}{b-a}\sum_{j=1}^{p-1}\eta_j(\Delta t)^2 = \frac{1}{b-a}\Delta t\sum_{j=1}^{p-1}\eta_j\Delta t. \tag{A14}$$

Let us now choose $c \in \mathbb{R}$ and $d \in \mathbb{R}$ so that $\Delta t = \dfrac{d-c}{p}$ and partition interval $[c,\ d] \subset \mathbb{R}$ by points $\tau_1,\ldots,\tau_p$ so that $c = \tau_1$ and $\tau_j = \tau_1 + j\Delta t$ with $j = 1,\ldots,p$. 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,

$$\lim_{\Delta t\to\infty}\sum_{j=1}^{p-1}\eta_j\Delta t = \lim_{\Delta t\to\infty}\sum_{j=1}^{p-1}\varphi(\xi_j)\Delta t = \int_c^d \varphi(\tau)d\tau < +\infty. \tag{A15}$$

Thus,

$$\frac{1}{b-a}\Delta t \sum_{j=1}^{p-1} \eta_j \Delta t \to 0 \quad \text{as} \quad \Delta t \to 0. \tag{A16}$$

As a result, Equations (A14)–(A16) imply Equation (26). □

References

1. Chen, J.; Benesty, J.; Huang, Y.; Doclo, S. New Insights Into the Noise Reduction Wiener Filter. *IEEE Trans. Audio Speech Lang. Process.* **2006**, *14*, 1218–1234. [CrossRef]
2. Spurbeck, M.; Schreier, P. Causal Wiener filter banks for periodically correlated time series. *Signal Process.* **2007**, *87*, 1179–1187. [CrossRef]
3. Goldstein, J.S.; Reed, I.; Scharf, L.L. A Multistage Representation of the Wiener Filter Based on Orthogonal Projections. *IEEE Trans. Inf. Theory* **1998**, *44*, 2943–2959. [CrossRef]
4. Hua, Y.; Nikpour, M.; Stoica, P. Optimal Reduced-Rank estimation and filtering. *IEEE Trans. Signal Process.* **2001**, *49*, 457–469.
5. Torokhti, A.; Howlett, P. *Computational Methods for Modelling of Nonlinear Systems*; Elsevier: Amsterdam, The Netherlands, 2007.
6. Sontag, E.D. *Polynomial Response Maps*; Lecture Notes in Control and Information Sciences; Springer: Cham, Switzerland, 1979; Volume 13.
7. Chen, S.; Billings, S.A. Representation of non-linear systems: NARMAX model. *Int. J. Control* **1989**, *49*, 1013–1032. [CrossRef]
8. Mathews, V.J.; Sicuranza, G.L. *Polynomial Signal Processing*; J. Wiley & Sons: Hoboken, NJ, USA, 2001.
9. Torokhti, A.; Howlett, P. Optimal Transform Formed by a Combination of Nonlinear Operators: The Case of Data Dimensionality Reduction. *IEEE Trans. Signal Process.* **2006**, *54*, 1431–1444. [CrossRef]
10. Vesma, J.; Saramaki, T. Polynomial-Based Interpolation Filters—Part I: Filter Synthesis. *Circuits Syst. Signal Process.* **2007**, *26*, 115–146. [CrossRef]
11. Torokhti, A.; Miklavcic, S. Data Compression under Constraints of Causality and Variable Finite Memory. *Signal Process.* **2010**, *90*, 2822–2834. [CrossRef]
12. Russo, F. Technique for image denoising based on adaptive piecewise linear filters and automatic parameter tuning. *IEEE Trans. Instrum. Meas.* **2006**, *55*, 1362–1367. [CrossRef]
13. Cousseau, J.E.; Figueroa, J.L.; Werner, S.; Laakso, T.I. Efficient Nonlinear Wiener Model Identification Using a Complex-Valued Simplicial Canonical Piecewise Linear Filter. *IEEE Trans. Signal Process.* **2007**, *55*, 1780–1792. [CrossRef]
14. Wigren, T. Recursive Prediction Error Identification Using the Nonlinear Wiener Model. *Automatica* **1993**, *29*, 1011–1025. [CrossRef]
15. Torokhti, A.; Howlett, P. Filtering and Compression for Infinite Sets of Stochastic Signals. *Signal Process.* **2009**, *89*, 291–304. [CrossRef]
16. Torokhti, A.; Manton, J. Generic Weighted Filtering of Stochastic Signals. *IEEE Trans. Signal Process.* **2009**, *57*, 4675–4685. [CrossRef]
17. Babuska, I.; Banerjee, U.; Osborn, J.E. Generalized finite element methods: Main ideas, results, and perspective. *Int. J. Comput. Methods* **2004**, *1*, 67–103. [CrossRef]
18. Golub, G.H.; van Loan, C.F. *Matrix Computations*; Johns Hopkins University Press: Baltimore, MD, USA, 1996.
19. Julian, P.; Desages, A.; D'Amico, B. Orthonormal high-level canonical PWL functions with applications to model reduction. *IEEE Trans. Circuits Syst. I Fundam. Theory Appl.* **2000**, *47*, 702–712. [CrossRef]
20. Kang, S.; Chua, L. A global representation of multidimensional piecewise-linear functions with linear partitions. *IEEE Trans. Circuits Syst.* **1978**, *25*, 938–940. [CrossRef]
21. Chua, L.O.; Deng, A.-C. Canonical piecewise-linear representation. *IEEE Trans. Circuits Syst.* **1988**, *35*, 101–111. [CrossRef]
22. Lin, J.-N.; Unbehauen, R. Adaptive nonlinear digital filter with canonical piecewise-linear structure. *IEEE Trans. Circuits Syst.* **1990**, *37*, 347–353. [CrossRef]
23. Lin, J.-N.; Unbehauen, R. Canonical piecewise-linear approximations. *IEEE Trans. Circuits Syst. I Fundam. Theory Appl.* **1992**, *39*, 697–699. [CrossRef]
24. Gelfand, S.B.; Ravishankar, C.S. A tree-structured piecewise linear adaptive filter. *IEEE Trans. Inf. Theory* **1993**, *39*, 1907–1922. [CrossRef]
25. Heredia, E.A.; Arce, G.R. Piecewise linear system modeling based on a continuous threshold decomposition. *IEEE Trans. Signal Process.* **1996**, *44*, 1440–1453. [CrossRef]
26. Feng, G. Robust filtering design of piecewise discrete time linear systems. *IEEE Trans. Signal Process.* **2005**, *53*, 599–605. [CrossRef]
27. Anderson, T. *An Introduction to Multivariate Statistical Analysis*; Wiley: New York, NY, USA, 1984.

28. Perlovsky, L.I.; Marzetta, T.L. Estimating a Covariance Matrix from Incomplete Realizations of a Random Vector. *IEEE Trans. Signal Process.* **1992**, *40*, 2097–2100. [CrossRef] [PubMed]
29. Ledoit, O.; Wolf, M. A well-conditioned estimator for large-dimensional covariance matrices. *J. Multivar. Anal.* **2004**, *88*, 365–411. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.