

Review Identification of Linear and Bilinear Systems: A Unified Study

Jacob Benesty¹, Constantin Paleologu^{2,*}, Laura-Maria Dogariu² and Silviu Ciochină²

¹ INRS-EMT, University of Quebec, Montreal, QC H5A 1K6, Canada; Jacob.Benesty@inrs.ca

² Department of Telecommunications, University Politehnica of Bucharest, 1-3, Iuliu Maniu Blvd.,

061071 Bucharest, Romania; ldogariu@comm.pub.ro (L.-M.D.); silviu@comm.pub.ro (S.C.)

* Correspondence: pale@comm.pub.ro

Abstract: System identification problems are always challenging to address in applications that involve long impulse responses, especially in the framework of multichannel systems. In this context, the main goal of this review paper is to promote some recent developments that exploit decomposition-based approaches to multiple-input/single-output (MISO) system identification problems, which can be efficiently solved as combinations of low-dimension solutions. The basic idea is to reformulate such a high-dimension problem in the framework of bilinear forms, and to then take advantage of the Kronecker product decomposition and low-rank approximation of the spatiotemporal impulse response of the system. The validity of this approach is addressed in terms of the celebrated Wiener filter, by developing an iterative version with improved performance features (related to the accuracy and robustness of the solution). Simulation results support the main theoretical findings and indicate the appealing performance of these developments.

Keywords: system identification; linear system; bilinear system; best approximation; singular value decomposition; optimal filtering; Wiener filter; multichannel acoustic echo cancellation

1. Introduction

Solving a system identification problem represents a key step in many important real-world applications [1,2]. In general, such a problem can be formulated in terms of estimating or modeling the parameters of an unknown system when a set of data is available, which is usually related to the input and output of the system. Depending on the specific particularities of the problem or application, we can deal with different types of systems, according to their numbers of inputs and outputs. The simplest formulation is the well-known single-input/single-output (SISO) system. Furthermore, in some applications we can deal with more elaborated structures, such as multiple-input/single-output (MISO) and multiple-input/multiple-output (MIMO) systems.

The linearity is an important feature of a system, which can significantly simplify the overall identification problem. Even if many real-world systems face nonlinear behaviors, it is always desirable to address or reformulate the framework such that it has a linear approach to some extent. In this context, a useful topic is related to bilinear forms, which have been addressed in the literature in different ways and contexts [3–22]; most often they are related to approximating nonlinear systems.

In general, a bilinear model can approximate a large class of nonlinear systems via a finite sum of the Volterra series expansion between the inputs and outputs of the system. Therefore, in this context, bilinear systems behave similarly (to some extent) to linear models. This could further simplify the analysis, as outlined before. Due to this simplicity, bilinear systems have been involved in a wide range of applications, such as digital filter synthesis [7], prediction problems [8], channel equalization [9], echo cancellation [10], chaotic communications [16], neural networks [20], and active noise control [21]. Nevertheless, in all these frameworks, the bilinear term is defined with respect to the data, i.e., in terms of an input-output relation.



Citation: Benesty, J.; Paleologu, C.; Dogariu, L.-M.; Ciochină, S. Identification of Linear and Bilinear Systems: A Unified Study. *Electronics* 2021, *10*, 1790. https://doi.org/ 10.3390/electronics10151790

Academic Editor: Juan M. Corchado

Received: 4 June 2021 Accepted: 21 July 2021 Published: 26 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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 (https:// creativecommons.org/licenses/by/ 4.0/). In this study, we focus on a different approach by defining the bilinear term with respect to the impulse responses of a spatiotemporal model, in the context of MISO systems. Several similar frameworks can be found in the literature, in the context of particular applications, such as channel equalization [13], nonlinear acoustic echo cancellation [15], and target detection [18]. However, most of these works were not associated with or analyzed in conjunction with bilinear forms. Usually, they were referred to as joint adaptation processes or cascaded systems, which are similar to the Hammerstein model [23].

More recently, an iterative Wiener filter for such bilinear forms was developed in the framework of a MISO system identification problem [24]. As compared to the conventional Wiener filter, the iterative version can obtain good accuracy even when a only small amount of information is available for the estimation of the statistics. Following the Wiener benchmark, another category of solutions relies on adaptive filtering [25,26]. Several adaptive filters tailored for the identification of bilinear forms have also been developed, following the main categories of algorithms. For example, the least-mean-square (LMS) and normalized LMS (NLMS) versions can be found in [27,28]. In addition, several recursive least-squares (RLS) algorithms for bilinear forms were developed in [29]. Moreover, a Kalman filter tailored for the identification of bilinear forms was proposed in [30].

In the previously mentioned approaches, the spatiotemporal impulse response of the MISO system is considered perfectly separable, and its components are combined using the Kronecker product. The identification of such linearly separable systems can be efficiently exploited in the frameworks of different applications, such as source separation [31,32], array beamforming [33,34], and object recognition [35,36]. In these contexts, the basic solution relies on the decomposition and modeling techniques of rank-1 tensors [37–42]. Nevertheless, it is highly useful to exploit the decomposition-based approach for the identification of more general forms of impulse responses.

Several recent works have followed this idea by exploiting the nearest Kronecker product decomposition and low-rank approximations [43–47]. In this context, the basic concept is to reformulate a high-dimension system identification problem as a combination of low-dimension solutions, thereby gaining in terms of both performance and complexity. Due to its features, this approach can be used in different practical applications—e.g., [48–55], among which we can mention acoustic feedback cancellation, adaptive beamforming, speech dereverberation, multichannel linear prediction, and nonlinear system identification.

A unified study on the efficient identification of linear and bilinear systems exploiting the decomposition-based approach is provided in this review paper. First, in Section 2, we present different system models, in the context of linear and bilinear forms. Then, in Section 3, we show how these models are related, thereby outlining the equivalence among the systems. The ideas behind the decomposition-based approach together with the optimal low-rank approximation are presented in Section 4. Since the Wiener filter represents a benchmark tool for the system identification problems, we illustrate its behavior in Section 5, wherein we also introduce an iterative version with improved performance features. Simulation results are provided in Section 6, in order to support the main theoretical findings. Finally, several conclusions and perspectives of this study are outlined in Section 7.

2. Different Input Output Linear/Bilinear System Models

In this study, we assume that the input and output, and the noise signals, take real values and have zero means. The most popular input output system is the so-called SISO system given by

$$d(k) = \mathbf{h}_{\mathbf{t}}^{T} \mathbf{x}(k) + w(k)$$
(1)
= y(k) + w(k),

where d(k) denotes the desired (or reference) signal at discrete-time index k, \mathbf{h}_t is the system's temporal impulse response of length L, and the superscript ^{*T*} denotes the transpose operator. The vector

$$\mathbf{x}(k) = \begin{bmatrix} x(k) & x(k-1) & \cdots & x(k-L+1) \end{bmatrix}^T$$
(2)

contains the most recent *L* samples of the input signal, x(k); w(k) is the additive noise; and $y(k) = \mathbf{h}_t^T \mathbf{x}(k)$ is the linear form in \mathbf{h}_t . A typical assumption that can be made is that x(k) and w(k) are uncorrelated (or even independent, which is not really required if we only handle second-order statistics). We refer to (1) as the linear SISO (LSISO) system. Its general block diagram is provided in Figure 1.



Figure 1. A general block diagram of the LSISO system from (1).

Without loss of generality, let us assume that $L = L_1L_2$, with $L_1 \ge L_2$. A shorter version of the input signal vector, $\mathbf{x}(k)$, may be written as

$$\mathbf{x}'(k) = \begin{bmatrix} x(k) & x(k-1) & \cdots & x(k-L_1+1) \end{bmatrix}^{T}$$
. (3)

As a result, we can express (2) as

$$\mathbf{x}(k) = \begin{bmatrix} \mathbf{x}'^{T}(k) & \mathbf{x}'^{T}(k-L_{1}) & \cdots & \mathbf{x}'^{T}[k-(L_{2}-1)L_{1}] \end{bmatrix}^{T},$$
(4)

from which we deduce the matrix of size $L_1 \times L_2$:

$$\mathbf{X}(k) = \begin{bmatrix} \mathbf{x}'(k) & \mathbf{x}'(k-L_1) & \cdots & \mathbf{x}'[k-(L_2-1)L_1] \end{bmatrix}.$$
 (5)

In other terms, we have

$$\mathbf{x}(k) = \operatorname{vec}[\mathbf{X}(k)],\tag{6}$$

where vec[·] denotes vectorization, i.e., the operation of converting a matrix into a vector. It may also be convenient to use the inverse of the vectorization operator [40], i.e., $\mathbf{X}(k) = \text{ivec}[\mathbf{x}(k)]$, which is equivalent to (6). Therefore, the most straightforward bilinear system that follows from the previous development results as

$$d(k) = \mathbf{h}_{t1}^T \mathbf{X}(k) \mathbf{h}_{t2} + w(k)$$

$$= y(k) + w(k),$$
(7)

where \mathbf{h}_{t1} and \mathbf{h}_{t2} are the first and second temporal impulse responses, of lengths L_1 and L_2 , respectively; and $y(k) = \mathbf{h}_{t1}^T \mathbf{X}(k) \mathbf{h}_{t2}$ is now bilinear in \mathbf{h}_{t1} and \mathbf{h}_{t2} . We call (7) the bilinear SISO (BSISO1) system. The equivalency between the LSISO and BSISO1 systems is explained and detailed in Section 3, together with the connections among different models that are discussed in the current section.

An obvious generalization of (7) is

$$d(k) = \sum_{l=1}^{L_2} \mathbf{h}_{t1,l}^T \mathbf{X}(k) \mathbf{h}_{t2,l} + w(k),$$
(8)

where $\mathbf{h}_{t1,l}$, $l = 1, 2, ..., L_2$ and $\mathbf{h}_{t2,l}$, $l = 1, 2, ..., L_2$ are the first and second sets of the system temporal impulse responses of lengths L_1 and L_2 , respectively. We refer to (8) as the BSISO2 system. Expression (8) can be rewritten as

$$d(k) = \boldsymbol{h}_{t1}^T \boldsymbol{X}(k) \boldsymbol{h}_{t2} + w(k)$$

$$= y(k) + w(k),$$
(9)

where

$$\boldsymbol{h}_{t1} = \begin{bmatrix} \mathbf{h}_{t1,1}^T & \mathbf{h}_{t1,2}^T & \cdots & \mathbf{h}_{t1,L_2}^T \end{bmatrix}^T,$$
(10)

$$\boldsymbol{h}_{t2} = \begin{bmatrix} \mathbf{h}_{t2,1}^T & \mathbf{h}_{t2,2}^T & \cdots & \mathbf{h}_{t2,L_2}^T \end{bmatrix}^T,$$
(11)

and

$$\mathbf{X}(k) = \mathrm{bdiag}[\mathbf{X}(k), \mathbf{X}(k), \dots, \mathbf{X}(k)]$$
(12)

is a block-diagonal matrix with L_2 diagonal blocks. We can see that $y(k) = h_{t1}^T X(k) h_{t2}$ is bilinear in h_{t1} and h_{t2} .

An important extension of the LSISO system in (1) is the so-called linear MISO (LMISO) system:

$$d(k) = \sum_{m=1}^{M} \mathbf{h}_m^T \mathbf{x}_m(k) + w(k), \qquad (13)$$

where *M* denotes the number of system inputs (or channels), \mathbf{h}_m , m = 1, 2, ..., M are the *M* channel impulse responses of length *L*, and the vector

$$\mathbf{x}_m(k) = \begin{bmatrix} x_m(k) & x_m(k-1) & \cdots & x_m(k-L+1) \end{bmatrix}^T$$
(14)

contains the most recent *L* samples of the *m*th (m = 1, 2, ..., M) input signal, $x_m(k)$. The general block diagram of the LMISO system is provided in Figure 2. Equation (13) can be rewritten as

$$d(k) = \overline{\mathbf{h}}^{1} \overline{\mathbf{x}}(k) + w(k)$$

$$= y(k) + w(k),$$
(15)

where

$$\overline{\mathbf{h}} = \begin{bmatrix} \mathbf{h}_1^T & \mathbf{h}_2^T & \cdots & \mathbf{h}_M^T \end{bmatrix}^T,$$
(16)

$$\overline{\mathbf{x}}(k) = \begin{bmatrix} \mathbf{x}_1^T(k) & \mathbf{x}_2^T(k) & \cdots & \mathbf{x}_M^T(k) \end{bmatrix}^T.$$
(17)

Clearly, $y(k) = \overline{\mathbf{h}}^T \overline{\mathbf{x}}(k)$ is linear in $\overline{\mathbf{h}}$. Of course, the particular case of M = 1 corresponds to the LSISO system.



Figure 2. A general block diagram of the LMISO system from (13).

As in the single-channel case, let $L = L_1L_2$ but with $ML_1 \ge L_2$. We can decompose $\mathbf{x}_m(k)$, similarly to (4), as

$$\mathbf{x}_m(k) = \begin{bmatrix} \mathbf{x}_m^{\prime T}(k) & \mathbf{x}_m^{\prime T}(k-L_1) & \cdots & \mathbf{x}_m^{\prime T}[k-(L_2-1)L_1] \end{bmatrix}^T,$$
(18)

where

$$\mathbf{x}'_m(k) = \begin{bmatrix} x_m(k) & x_m(k-1) & \cdots & x_m(k-L_1+1) \end{bmatrix}^T.$$
(19)

Then, we concatenate the *M* input signals as

$$\underline{\mathbf{x}}(k) = \begin{bmatrix} \underline{\mathbf{x}}^{T}(k) & \underline{\mathbf{x}}^{T}(k-L_{1}) & \cdots & \underline{\mathbf{x}}^{T}[k-(L_{2}-1)L_{1}] \end{bmatrix}^{T},$$
(20)

where

$$\underline{\mathbf{x}}'(k) = \begin{bmatrix} \mathbf{x}_1'^T(k) & \mathbf{x}_2'^T(k) & \cdots & \mathbf{x}_M'^T(k) \end{bmatrix}^T$$
(21)

is a vector of length ML_1 . Consequently, the LMISO system in (13) or (15) can be expressed in an equivalent manner as

$$d(k) = \underline{\mathbf{h}}^T \underline{\mathbf{x}}(k) + w(k), \tag{22}$$

where $\underline{\mathbf{h}}$ (of length *ML*) represents the spatiotemporal impulse response of the system, with the same coefficients as $\overline{\mathbf{h}}$, resulting through simple permutations, according to the inputs.

The first bilinear MISO (BMISO1) system can be derived from the LSISO system in (1), according to (15) [24]:

$$d(k) = \mathbf{h}_{\mathbf{t}}^{T} \overline{\mathbf{X}}(k) \mathbf{h}_{\mathbf{s}} + w(k)$$

$$= y(k) + w(k),$$
(23)

where \mathbf{h}_t (of length *L*) represents the temporal impulse response of the system, $\overline{\mathbf{X}}(k) = \text{ivec}[\overline{\mathbf{x}}(k)]$, \mathbf{h}_s (of length *M*) represents the spatial impulse response of the system, and $y(k) = \mathbf{h}_t^T \overline{\mathbf{X}}(k) \mathbf{h}_s$ is the bilinear form in \mathbf{h}_t and \mathbf{h}_s . For M = 1, (23) is equivalent to the LSISO system in (1); this also means that the bilinear structure is lost in the single-channel particular case.

Now, from (20), we can build the matrix of size $ML_1 \times L_2$:

$$\underline{\mathbf{X}}(k) = \operatorname{ivec}[\underline{\mathbf{x}}(k)]. \tag{24}$$

Then, our second bilinear MISO (BMISO2) system is derived according to (22). We get

$$d(k) = \mathbf{h}_{st1}^T \underline{\mathbf{X}}(k) \mathbf{h}_{t2} + w(k)$$
(25)
= y(k) + w(k),

where \mathbf{h}_{st1} (of length ML_1) is the spatiotemporal impulse response of the system, \mathbf{h}_{t2} (of length L_2) is the system temporal impulse response, and $y(k) = \mathbf{h}_{st1}^T \underline{\mathbf{X}}(k) \mathbf{h}_{t2}$ is the bilinear form in \mathbf{h}_{st1} and \mathbf{h}_{t2} . For M = 1, we obtain exactly the BSISO1 system in (7).

Our third and last bilinear MISO (BMISO3) system is just an obvious generalization of (25), i.e.,

$$d(k) = \sum_{l=1}^{L_2} \mathbf{h}_{\text{stl},l}^T \underline{\mathbf{X}}(k) \mathbf{h}_{\text{t2},l} + w(k),$$
(26)

where $\mathbf{h}_{\text{st1},l}$, $l = 1, 2, ..., L_2$ (of length ML_1) is the set of spatiotemporal impulse responses of the system, and $\mathbf{h}_{\text{t2},l}$, $l = 1, 2, ..., L_2$ (of length L_2) is the set of temporal impulse responses of the system. For M = 1, we get the BSISO2 in (8). Relation (26) can be rewritten as

$$d(k) = \boldsymbol{h}_{st1}^T \underline{\boldsymbol{X}}(k) \boldsymbol{h}_{t2} + w(k)$$

$$= y(k) + w(k),$$
(27)

where

$$\boldsymbol{h}_{\text{st1}} = \begin{bmatrix} \mathbf{h}_{\text{st1},1}^T & \mathbf{h}_{\text{st1},2}^T & \cdots & \mathbf{h}_{\text{st1},L_2}^T \end{bmatrix}^T,$$
(28)

$$\boldsymbol{h}_{t2} = \begin{bmatrix} \mathbf{h}_{t2,1}^T & \mathbf{h}_{t2,2}^T & \cdots & \mathbf{h}_{t2,L_2}^T \end{bmatrix}^T,$$
(29)

and

$$\underline{\mathbf{X}}(k) = \mathrm{bdiag}[\underline{\mathbf{X}}(k), \underline{\mathbf{X}}(k), \dots, \underline{\mathbf{X}}(k)]$$
(30)

is a block-diagonal matrix with L_2 diagonal blocks, while $y(k) = h_{st1}^T \underline{X}(k) h_{t2}$ is bilinear in h_{st1} and h_{t2} .

3. Equivalence among Systems

In this section, we show how the different linear and bilinear systems are related. Let us start with the BSISO1 system in (7). Its bilinear term can be rewritten as

$$y(k) = \mathbf{h}_{t1}^{T} \mathbf{X}(k) \mathbf{h}_{t2}$$

= tr $\left[\left(\mathbf{h}_{t1} \mathbf{h}_{t2}^{T} \right)^{T} \mathbf{X}(k) \right]$
= vec^T $\left(\mathbf{h}_{t1} \mathbf{h}_{t2}^{T} \right)$ vec $[\mathbf{X}(k)]$
= $\left(\mathbf{h}_{t2} \otimes \mathbf{h}_{t1} \right)^{T} \mathbf{x}(k)$, (31)

where tr[·] denotes the trace of a square matrix and \otimes is the Kronecker product [56]. With (31) in mind, comparing the BSISO1 system with the LSISO system in (1), we can clearly observe that the two systems are identical if $\mathbf{h}_t = \mathbf{h}_{t2} \otimes \mathbf{h}_{t1}$. Therefore, in general, we can say that BSISO1 is a particular case of LSISO. In other words, BSISO1 is also an LSISO with some structure of its temporal impulse response.

Now, let us focus on the BSISO2 system in (8). Another way to express its bilinear term is

$$y(k) = \sum_{l=1}^{L_2} \mathbf{h}_{t1,l}^T \mathbf{X}(k) \mathbf{h}_{t2,l}$$

= $\sum_{l=1}^{L_2} \operatorname{vec}^T \left(\mathbf{h}_{t1,l} \mathbf{h}_{t2,l}^T \right) \operatorname{vec}[\mathbf{X}(k)]$
= $\operatorname{vec}^T \left(\sum_{l=1}^{L_2} \mathbf{h}_{t1,l} \mathbf{h}_{t2,l}^T \right) \mathbf{x}(k)$
= $\operatorname{vec}^T(\mathbf{H}_t) \mathbf{x}(k),$ (32)

where

$$\mathbf{H}_{t} = \sum_{l=1}^{L_{2}} \mathbf{h}_{t1,l} \mathbf{h}_{t2,l}^{T}$$
(33)

is a matrix of size $L_1 \times L_2$ of rank equal to L_2 in general. At the same time, the temporal impulse response of the LSISO system can be decomposed as

$$\mathbf{h}_{t} = \begin{bmatrix} \mathbf{h}_{t,1}^{T} & \mathbf{h}_{t,2}^{T} & \cdots & \mathbf{h}_{t,L_{2}}^{T} \end{bmatrix}^{T},$$
(34)

where $\mathbf{h}_{t,l}$, $l = 1, 2, ..., L_2$ are impulse responses of length L_1 each. Next, we can rewrite the linear term of the LSISO system as

$$y(k) = \mathbf{h}_{\mathbf{t}}^{T} \mathbf{x}(k)$$

= $\operatorname{vec}^{T}(\mathbf{H}_{\mathbf{t}}) \mathbf{x}(k)$, (35)

where $\mathbf{H}_t = \text{ivec}(\mathbf{h}_t)$. It can be easily seen by comparing (32) and (35) that the LSISO and BSISO2 systems are equivalent.

In the same way, we can write the bilinear term of the BMISO1 system in (23) as

$$y(k) = \mathbf{h}_{t}^{T} \overline{\mathbf{X}}(k) \mathbf{h}_{s}$$
$$= (\mathbf{h}_{s} \otimes \mathbf{h}_{t})^{T} \overline{\mathbf{x}}(k).$$
(36)

Then, by comparing the previous expression with the bilinear form of the LMISO system in (15), we can see that the two are the same if $\overline{\mathbf{h}} = \mathbf{h}_s \otimes \mathbf{h}_t$. In general, BMISO1 is a particular case of LMISO.

The bilinear term of the second bilinear MISO system, i.e., BMISO2 in (25), can also be expressed as

$$y(k) = \mathbf{h}_{st1}^{T} \underline{\mathbf{X}}(k) \mathbf{h}_{t2}$$

= $(\mathbf{h}_{t2} \otimes \mathbf{h}_{st1})^{T} \underline{\mathbf{x}}(k).$ (37)

Again, we can conclude that the BMISO2 system is a particular case of the LMISO system in (22), where $\underline{h} = h_{t2} \otimes h_{st1}$.

Finally, the bilinear form of the BMISO3 system in (26) may be written as

$$y(k) = \sum_{l=1}^{L_2} \mathbf{h}_{\text{st1},l}^T \underline{\mathbf{X}}(k) \mathbf{h}_{\text{t2},l}$$
$$= \text{vec}^T(\mathbf{H}_{\text{st}}) \underline{\mathbf{x}}(k), \qquad (38)$$

where

$$\mathbf{H}_{\rm st} = \sum_{l=1}^{L_2} \mathbf{h}_{{\rm st}1,l} \mathbf{h}_{{\rm t}2,l}^T$$
(39)

is a matrix of size $ML_1 \times L_2$ of rank equal to L_2 in general. At the same time, the spatiotemporal impulse response of the LMISO system in (22) can be decomposed as

$$\underline{\mathbf{h}} = \begin{bmatrix} \underline{\mathbf{h}}_1^T & \underline{\mathbf{h}}_2^T & \cdots & \underline{\mathbf{h}}_{L_2}^T \end{bmatrix}^T,$$
(40)

where $\underline{\mathbf{h}}_l$, $l = 1, 2, ..., L_2$ are impulse responses of length ML_1 each. Next, we can rewrite the linear term of the LMISO system as

$$y(k) = \underline{\mathbf{h}}^T \underline{\mathbf{x}}(k)$$

= $\operatorname{vec}^T(\mathbf{H}_{\mathrm{st}})\underline{\mathbf{x}}(k)$, (41)

where $\mathbf{H}_{st} = \text{ivec}(\underline{\mathbf{h}})$. It can be easily seen by comparing (38) and (41) that the LMISO and BMISO3 systems are equivalent.

4. Best Approximation

The main objective in this study is to identify the LMISO system in (13) (or, equivalently, in (15) or (22)). The LSISO system is just a particular case and has been studied before. We can achieve this goal based on what is already known about bilinear forms and how they are best approximated.

Let **a** be a real-valued vector of length *L*. The 2-norm or Euclidean norm of this vector is defined as

$$\|\mathbf{a}\|_2 = \sqrt{\mathbf{a}^T \mathbf{a}}.\tag{42}$$

Let **A** be a real-valued rectangular matrix of size $L \times C$. The Frobenius norm and the 2-norm of this matrix are, respectively,

$$\|\mathbf{A}\|_{\mathrm{F}} = \sqrt{\mathrm{tr}(\mathbf{A}^{T}\mathbf{A})} \tag{43}$$

and

$$\|\mathbf{A}\|_{2} = \max_{\|\mathbf{x}\|_{2}=1} \|\mathbf{A}\mathbf{x}\|_{2}.$$
(44)

Now, we can consider the impulse response of the BMISO3 system in (26); i.e., the matrix $\mathbf{H}_{st} = \sum_{l=1}^{L_2} \mathbf{h}_{st1,l} \mathbf{h}_{t2,l}^T$ of size $ML_1 \times L_2$ with $ML_1 \ge L_2$ (see Equation (38)). As mentioned before, this system is equivalent to the LMISO system defined by relation (22). The matrix \mathbf{H}_{st} can be factorized through the singular value decomposition (SVD):

$$\mathbf{H}_{st} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^{T}$$

$$= \sum_{l=1}^{L_{2}} \sigma_{l} \mathbf{u}_{l} \mathbf{v}_{l}^{T},$$
(45)

where **U**, of size $ML_1 \times ML_1$, and **V**, of size $L_2 \times L_2$, are orthogonal matrices and Σ is an $ML_1 \times L_2$ rectangular diagonal matrix having on the main diagonal nonnegative real numbers. The columns of **U** and **V** are known as the left-singular and right-singular vectors, respectively, of **H**_{st}, whereas the elements σ_l , $l = 1, 2, ..., L_2$ on the diagonal of Σ are called singular values of **H**_{st} with $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_{L_2} \ge 0$. Based on (39) and (45), we deduce that

$$\mathbf{h}_{\mathrm{stl},l} = \alpha(\sigma_l) \mathbf{u}_l,\tag{46}$$

$$\mathbf{h}_{\mathbf{t}2,l} = \beta(\sigma_l) \mathbf{v}_l,\tag{47}$$

with $l = 1, 2, ..., L_2$, where $\alpha(\sigma_l)\beta(\sigma_l) = \sigma_l$, \mathbf{u}_l , $l = 1, 2, ..., L_2$ are the first L_2 columns of **U**, and \mathbf{v}_l , $l = 1, 2, ..., L_2$ are the columns of **V**. It may be easily checked that $\|\mathbf{H}_{st}\|_2 = \sigma_1$ and $\|\mathbf{H}_{st}\|_F = \sqrt{\sum_{l=1}^{L_2} \sigma_l^2}$. In addition, since $\underline{\mathbf{h}} = \operatorname{vec}(\mathbf{H}_{st})$ (see (41)), the global impulse response can be decomposed as

$$\underline{\mathbf{h}} = \sum_{l=1}^{L_2} \sigma_l(\mathbf{v}_l \otimes \mathbf{u}_l)$$
$$= \sum_{l=1}^{L_2} \mathbf{h}_{t2,l} \otimes \mathbf{h}_{st1,l}.$$
(48)

However, in practical scenarios, the matrix H_{st} is never really of full rank, because of the reflections and/or sparseness in the system [57–60]. Let $P \ll L_2$ and let us define the following matrix:

$$\mathbf{H}_{\rm st}(P) = \sum_{p=1}^{P} \sigma_p \mathbf{u}_p \mathbf{v}_p^T.$$
(49)

Now, the objective is to verify whether \mathbf{H}_{st} can be well approximated by $\mathbf{H}_{st}(P)$. In the positive scenario, the LMISO system can be written as

$$d(k) = \operatorname{vec}^{T}(\mathbf{H}_{st})\underline{\mathbf{x}}(k) + w(k)$$

= $\operatorname{vec}^{T}[\mathbf{H}_{st}(P)]\underline{\mathbf{x}}(k) + b(k) + w(k),$ (50)

where

$$b(k) = \operatorname{vec}^{T}(\mathbf{Y})\underline{\mathbf{x}}(k)$$
(51)

denotes the correlated noise (considered negligible), with $\mathbf{Y} = \sum_{i=P+1}^{L_2} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$. Consequently, the goal becomes to identify the new matrix $\mathbf{H}_{st}(P)$ instead of \mathbf{H}_{st} . This new idea may have a few advantages, as is explained in the following.

Next, we state a theorem given in [61,62], which helps to prove that \mathbf{H}_{st} can be well approximated by $\mathbf{H}_{st}(P)$. Let rank $(\mathbf{H}_{st}) = R \leq L_2$ and let S be the set of $ML_1 \times L_2$ matrices of rank equal to P < R. Then, the solution to the minimization problem

$$\min_{\mathbf{H}\in\mathcal{S}} \|\mathbf{H}_{st} - \mathbf{H}\|_{2} \quad \text{or} \quad \min_{\mathbf{H}\in\mathcal{S}} \|\mathbf{H}_{st} - \mathbf{H}\|_{F}$$
(52)

is given by (49). Furthermore, we have

$$\min_{\mathbf{H}\in\mathcal{S}} \|\mathbf{H}_{st} - \mathbf{H}\|_2 = \|\mathbf{H}_{st} - \mathbf{H}_{st}(P)\|_2 = \sigma_{P+1}$$
(53)

and

$$\min_{\mathbf{H}\in\mathcal{S}} \|\mathbf{H}_{st} - \mathbf{H}\|_{F} = \|\mathbf{H}_{st} - \mathbf{H}_{st}(P)\|_{F} = \sqrt{\sum_{i=P+1}^{L_{2}} \sigma_{i}^{2}}.$$
(54)

Consequently, as long as the normalized misalignment,

$$\mathcal{M}(P) = \frac{\|\mathbf{H}_{st} - \mathbf{H}_{st}(P)\|_{F}}{\|\mathbf{H}_{st}\|_{F}},$$
(55)

remains very small, it is sufficient in practice to estimate the impulse responses $\mathbf{h}_{\text{st1},p}$ and $\mathbf{h}_{\text{t2},p}$ for p = 1, 2, ..., P.

In order to show the validity of this approach, let us consider two scenarios that will also be detailed in the simulations provided in Section 6. In the first scenario, we consider M = 4 impulse responses from the G168 Recommendation [63], which are network echo paths of length L = 500, as depicted in Figure 3. In this case, the decomposition was performed using $L_1 = 25$ and $L_2 = 20$. In the second scenario, we used two acoustic impulse responses (i.e., M = 2), each one having L = 1024 coefficients, as depicted in Figure 4. Here, we set $L_1 = L_2 = 32$ for the decomposition. In both cases, we evaluate the normalized misalignment from (55) and the evolution of the singular values σ_l ($l = 1, 2, ..., L_2$) of the matrix \mathbf{H}_{st} . As we can see in Figure 5, the normalized misalignment decreased with the value of P. This was much more apparent in the first scenario (corresponding to Figure 3), where the rank of \mathbf{H}_{st} resulted in R = 5, as shown in Figure 5a. Consequently, a good approximation was obtained for $P \ll L_2$. In case of the acoustic impulse responses (i.e., the scenario from Figure 4), the resulting matrix \mathbf{H}_{st} was closer to being full rank, so that a larger value of P was required to obtain a good approximation. Nevertheless, as we can notice in Figure 5b, a value of P significantly lower as compared to L_2 led to reasonable attenuation of the misalignment (e.g., around -20 dB). This behavior is also supported in Figure 6, where we can notice the decreasing trend of the singular values (which are normalized to the maximum value).



Figure 3. Impulse responses used in the first set of simulations from Section 6 (according to the G168 Recommendation [63]), with L = 500 and M = 4: (a) the first network echo path from [63], (b) the second network echo path from [63], (c) the fifth network echo path from [63], and (d) the sixth network echo path from from [63].



Figure 4. Impulse responses used in the second set of simulations from Section 6, with L = 1024 and M = 2: (a) right acoustic echo path and (b) left acoustic echo path.



Figure 5. Normalized misalignment evaluated based on (55) for different values of *P*, corresponding to (**a**) the impulse responses from Figure 3, with M = 4, L = 500, $L_1 = 25$, and $L_2 = 20$; and (**b**) the impulse responses from Figure 4, with M = 2, L = 1024, and $L_1 = L_2 = 32$. For better visualization, the representation is limited to -100 dB.



Figure 6. Singular values (normalized with respect to the maximum one) of the matrix H_{st}, corresponding to (a) the impulse responses from Figure 3, with M = 4, L = 500, $L_1 = 25$, and $L_2 = 20$; and (b) the impulse responses from Figure 4, with M = 2, L = 1024, and $L_1 = L_2 = 32$.

5. Identification with the Wiener Filter

The identification of the LMISO system in (22) involves finding a real-valued filter, \mathbf{h} , of length ML, which estimates the system **<u>h</u>**. The error signal can be defined as

$$e(k) = d(k) - \widehat{y}(k), \tag{56}$$

where $\hat{y}(k) = \hat{\mathbf{h}}^T \mathbf{x}(k)$. The optimization criterion used to find the optimal filter is the mean-squared error (MSE):

$$\mathcal{J}\left(\underline{\widehat{\mathbf{h}}}\right) = E\left[e^{2}(k)\right]$$
$$= \sigma_{d}^{2} - 2\underline{\widehat{\mathbf{h}}}^{T}\mathbf{p} + \underline{\widehat{\mathbf{h}}}^{T}\mathbf{R}\underline{\widehat{\mathbf{h}}},$$
(57)

where $E[\cdot]$ is the mathematical expectation, $\mathbf{p} = E[\underline{\mathbf{x}}(k)d(k)]$ represents the cross-correlation vector between $\underline{\mathbf{x}}(k)$ and d(k), and $\mathbf{R} = E[\underline{\mathbf{x}}(k)\underline{\mathbf{x}}^T(k)]$ denotes the covariance matrix of $\underline{\mathbf{x}}(k)$. After minimizing $\mathcal{J}(\hat{\mathbf{h}})$, the celebrated (multichannel) Wiener filter is obtained:

$$\widehat{\underline{\mathbf{h}}}_{\mathrm{W}} = \mathbf{R}^{-1}\mathbf{p}.\tag{58}$$

Since the covariance matrix in the expression above is of size $ML \times ML$, a large number of data samples (more than ML) is needed in order to obtain a reliable solution.

An alternative approach to identifying the LMISO system in (22) and estimating **h** as in the conventional case is to identify the LMISO system in (50) and estimate H_{st} . In the rest of this paper, the subscripts st and t are dropped in order to simplify the notation, and in this way $\mathbf{H}_{st} = \sum_{l=1}^{L_2} \mathbf{h}_{st1,l} \mathbf{h}_{t2,l}^T$ becomes $\mathbf{H} = \sum_{l=1}^{L_2} \mathbf{h}_{1,l} \mathbf{h}_{2,l}^T$. Next, we assume that rank $(\mathbf{H}) = P \ll L_2$. Consequently, \mathbf{h} can be decomposed as

$$\underline{\mathbf{h}} = \sum_{p=1}^{P} \mathbf{h}_{2,p} \otimes \mathbf{h}_{1,p}, \tag{59}$$

where the impulse responses $\mathbf{h}_{1,p}$ and $\mathbf{h}_{2,p}$ have lengths ML_1 and L_2 , respectively. Therefore, the filter $\hat{\mathbf{h}}$ may also be decomposed as

$$\widehat{\underline{\mathbf{h}}} = \sum_{p=1}^{p} \widehat{\mathbf{h}}_{2,p} \otimes \widehat{\mathbf{h}}_{1,p}, \tag{60}$$

where the filters $\hat{\mathbf{h}}_{1,p}$ and $\hat{\mathbf{h}}_{2,p}$ have lengths ML_1 and L_2 , respectively. With the relations

$$\widehat{\mathbf{h}}_{2,p} \otimes \widehat{\mathbf{h}}_{1,p} = \left(\widehat{\mathbf{h}}_{2,p} \times 1\right) \otimes \left(\mathbf{I}_{ML_{1}} \times \widehat{\mathbf{h}}_{1,p}\right) \\ = \left(\widehat{\mathbf{h}}_{2,p} \otimes \mathbf{I}_{ML_{1}}\right) \widehat{\mathbf{h}}_{1,p}$$
(61)

and

$$\widehat{\mathbf{h}}_{2,p} \otimes \widehat{\mathbf{h}}_{1,p} = \left(\mathbf{I}_{L_2} \times \widehat{\mathbf{h}}_{2,p}\right) \otimes \left(\widehat{\mathbf{h}}_{1,p} \times 1\right) \\ = \left(\mathbf{I}_{L_2} \otimes \widehat{\mathbf{h}}_{1,p}\right) \widehat{\mathbf{h}}_{2,p},$$
(62)

where I_{ML_1} and I_{L_2} are the identity matrices of sizes $ML_1 \times ML_1$ and $L_2 \times L_2$, respectively, (60) may be rewritten as

$$\widehat{\underline{\mathbf{h}}} = \sum_{p=1}^{p} \widehat{\mathbf{H}}_{2,p} \widehat{\mathbf{h}}_{1,p}$$
(63)

$$=\sum_{p=1}^{P}\widehat{\mathbf{H}}_{1,p}\widehat{\mathbf{h}}_{2,p},\tag{64}$$

where

$$\widehat{\mathbf{H}}_{2,p} = \widehat{\mathbf{h}}_{2,p} \otimes \mathbf{I}_{ML_1},$$

 $\widehat{\mathbf{H}}_{1,p} = \mathbf{I}_{L_2} \otimes \widehat{\mathbf{h}}_{1,p}$

are matrices of sizes $ML \times ML_1$ and $ML \times L_2$, respectively. As a result, we may express the error signal defined in (56) in two distinct ways:

$$e(k) = d(k) - \sum_{p=1}^{p} \widehat{\mathbf{h}}_{1,p}^{T} \widehat{\mathbf{H}}_{2,p}^{T} \underline{\mathbf{x}}(k)$$

$$= d(k) - \sum_{p=1}^{p} \widehat{\mathbf{h}}_{1,p}^{T} \mathbf{x}_{2,p}(k)$$

$$= d(k) - \underline{\widehat{\mathbf{h}}}_{1}^{T} \underline{\mathbf{x}}_{2}(k)$$
(65)

and

$$e(k) = d(k) - \sum_{p=1}^{P} \widehat{\mathbf{h}}_{2,p}^{T} \widehat{\mathbf{H}}_{1,p}^{T} \underline{\mathbf{x}}(k)$$

$$= d(k) - \sum_{p=1}^{P} \widehat{\mathbf{h}}_{2,p}^{T} \mathbf{x}_{1,p}(k)$$

$$= d(k) - \widehat{\mathbf{h}}_{2}^{T} \underline{\mathbf{x}}_{1}(k),$$
(66)

where

$$\begin{aligned} \mathbf{x}_{2,p}(k) &= \widehat{\mathbf{H}}_{2,p}^{T} \underline{\mathbf{x}}(k), \\ &\widehat{\mathbf{h}}_{1} = \begin{bmatrix} \widehat{\mathbf{h}}_{1,1}^{T} & \widehat{\mathbf{h}}_{1,2}^{T} & \cdots & \widehat{\mathbf{h}}_{1,P}^{T} \end{bmatrix}^{T}, \\ &\underline{\mathbf{x}}_{2}(k) &= \begin{bmatrix} \mathbf{x}_{2,1}^{T}(k) & \mathbf{x}_{2,2}^{T}(k) & \cdots & \mathbf{x}_{2,P}^{T}(k) \end{bmatrix}^{T}, \\ &\mathbf{x}_{1,p}(k) &= \widehat{\mathbf{H}}_{1,p}^{T} \underline{\mathbf{x}}(k), \\ &\widehat{\mathbf{h}}_{2} &= \begin{bmatrix} \widehat{\mathbf{h}}_{2,1}^{T} & \widehat{\mathbf{h}}_{2,2}^{T} & \cdots & \widehat{\mathbf{h}}_{2,P}^{T} \end{bmatrix}^{T}, \\ &\underline{\mathbf{x}}_{1}(k) &= \begin{bmatrix} \mathbf{x}_{1,1}^{T}(k) & \mathbf{x}_{1,2}^{T}(k) & \cdots & \mathbf{x}_{1,P}^{T}(k) \end{bmatrix}^{T}. \end{aligned}$$

Continuing with this formulation, we can write the MSE criterion as

$$\mathcal{J}\left(\underline{\widehat{\mathbf{h}}}_{1}, \underline{\widehat{\mathbf{h}}}_{2}\right) = \sigma_{d}^{2} - 2\underline{\widehat{\mathbf{h}}}_{1}^{T}\underline{\mathbf{p}}_{2} + \underline{\widehat{\mathbf{h}}}_{1}^{T}\underline{\mathbf{R}}_{2}\underline{\widehat{\mathbf{h}}}_{1}$$

$$= \sigma_{d}^{2} - 2\underline{\widehat{\mathbf{h}}}_{2}^{T}\underline{\mathbf{p}}_{1} + \underline{\widehat{\mathbf{h}}}_{2}^{T}\underline{\mathbf{R}}_{1}\underline{\widehat{\mathbf{h}}}_{2},$$
(67)
(67)

where

$$\underline{\mathbf{p}}_{2} = \begin{bmatrix} \mathbf{p}^{T} \widehat{\mathbf{H}}_{2,1} & \mathbf{p}^{T} \widehat{\mathbf{H}}_{2,2} & \cdots & \mathbf{p}^{T} \widehat{\mathbf{H}}_{2,P} \end{bmatrix}^{T}, \\ \mathbf{R}_{2} = \begin{bmatrix} \widehat{\mathbf{H}}_{2,1}^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,1} & \widehat{\mathbf{H}}_{2,1}^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,2} & \cdots & \widehat{\mathbf{H}}_{2,1}^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,P} \\ \widehat{\mathbf{H}}_{2,2}^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,1} & \widehat{\mathbf{H}}_{2,2}^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,2} & \cdots & \widehat{\mathbf{H}}_{2,P}^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,P} \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{\mathbf{H}}_{2,P}^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,1} & \widehat{\mathbf{H}}_{2,P}^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,2} & \cdots & \widehat{\mathbf{H}}_{2,P}^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,P} \end{bmatrix}, \\ \\ \underline{\mathbf{p}}_{1} = \begin{bmatrix} \mathbf{p}^{T} \widehat{\mathbf{H}}_{1,1} & \mathbf{p}^{T} \widehat{\mathbf{H}}_{1,2} & \cdots & \mathbf{p}^{T} \widehat{\mathbf{H}}_{1,P} \end{bmatrix}^{T}, \\ \\ \mathbf{R}_{1} = \begin{bmatrix} \widehat{\mathbf{H}}_{1,1}^{T} \mathbf{R} \widehat{\mathbf{H}}_{1,1} & \widehat{\mathbf{H}}_{1,2}^{T} \mathbf{R} \widehat{\mathbf{H}}_{1,2} & \cdots & \widehat{\mathbf{H}}_{1,P}^{T} \mathbf{R} \widehat{\mathbf{H}}_{1,P} \\ \\ \widehat{\mathbf{H}}_{1,2}^{T} \mathbf{R} \widehat{\mathbf{H}}_{1,1} & \widehat{\mathbf{H}}_{1,2}^{T} \mathbf{R} \widehat{\mathbf{H}}_{1,2} & \cdots & \widehat{\mathbf{H}}_{1,P}^{T} \mathbf{R} \widehat{\mathbf{H}}_{1,P} \\ \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{\mathbf{H}}_{1,P}^{T} \mathbf{R} \widehat{\mathbf{H}}_{1,1} & \widehat{\mathbf{H}}_{1,P}^{T} \mathbf{R} \widehat{\mathbf{H}}_{1,2} & \cdots & \widehat{\mathbf{H}}_{1,P}^{T} \mathbf{R} \widehat{\mathbf{H}}_{1,P} \end{bmatrix}. \end{cases}$$

It can be noticed that the matrices $\underline{\mathbf{R}}_1$ and $\underline{\mathbf{R}}_2$ have sizes $PL_2 \times PL_2$ and $PML_1 \times PML_1$, respectively, which can be much smaller than the size of \mathbf{R} , which is $ML \times ML$. Additionally, at least PML_1 data samples are required for the estimation of the statistics in the MSE from (67) or (68), whereas in order to estimate the statistics in the conventional MSE from (57), we need at least ML data samples. When $\underline{\mathbf{\hat{h}}}_2$ is fixed, we can express (67) as

$$\mathcal{J}_{\underline{\widehat{\mathbf{h}}}_{2}}\left(\underline{\widehat{\mathbf{h}}}_{1}\right) = \sigma_{d}^{2} - 2\underline{\widehat{\mathbf{h}}}_{1}^{T}\underline{\mathbf{p}}_{2} + \underline{\widehat{\mathbf{h}}}_{1}^{T}\underline{\mathbf{R}}_{2}\underline{\widehat{\mathbf{h}}}_{1}$$

$$\tag{69}$$

and when $\underline{\hat{\mathbf{h}}}_1$ is fixed, we can write (68) as

$$\mathcal{J}_{\underline{\hat{\mathbf{h}}}_1}(\underline{\hat{\mathbf{h}}}_2) = \sigma_d^2 - 2\underline{\hat{\mathbf{h}}}_2^T \underline{\mathbf{p}}_1 + \underline{\hat{\mathbf{h}}}_2^T \underline{\mathbf{R}}_1 \underline{\hat{\mathbf{h}}}_2.$$
(70)

This represents a bilinear optimization strategy [64].

In order to obtain the optimal filters, an iterative algorithm similar to the those presented in [24,43] can be derived. At iteration 0, we may take

$$\widehat{\mathbf{h}}_{2,p}^{(0)} = \begin{bmatrix} \epsilon & 0 & \cdots & 0 \end{bmatrix}^T, \ p = 1, 2, \dots, P,$$

where $0 < \epsilon \leq 1$. Then, we may form $\widehat{\mathbf{H}}_{2,p}^{(0)} = \widehat{\mathbf{h}}_{2,p}^{(0)} \otimes \mathbf{I}_{ML_1}$ and

$$\underline{\mathbf{p}}_{2}^{(0)} = \begin{bmatrix} \mathbf{p}^{T} \widehat{\mathbf{H}}_{2,1}^{(0)} & \mathbf{p}^{T} \widehat{\mathbf{H}}_{2,2}^{(0)} & \cdots & \mathbf{p}^{T} \widehat{\mathbf{H}}_{2,P}^{(0)} \end{bmatrix}^{T}, \\ \underline{\mathbf{R}}_{2}^{(0)} = \begin{bmatrix} \left(\widehat{\mathbf{H}}_{2,1}^{(0)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,1}^{(0)} & \left(\widehat{\mathbf{H}}_{2,1}^{(0)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,2}^{(0)} & \cdots & \left(\widehat{\mathbf{H}}_{2,1}^{(0)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,P}^{(0)} \\ \left(\widehat{\mathbf{H}}_{2,2}^{(0)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,1}^{(0)} & \left(\widehat{\mathbf{H}}_{2,2}^{(0)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,2}^{(0)} & \cdots & \left(\widehat{\mathbf{H}}_{2,2}^{(0)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,P}^{(0)} \\ \vdots & \vdots & \ddots & \vdots \\ \left(\widehat{\mathbf{H}}_{2,P}^{(0)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,1}^{(0)} & \left(\widehat{\mathbf{H}}_{2,P}^{(0)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,2}^{(0)} & \cdots & \left(\widehat{\mathbf{H}}_{2,P}^{(0)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,P}^{(0)} \end{bmatrix}.$$

By substituting the quantities above into the MSE from (69), we get at iteration 1:

$$\mathcal{J}_{\underline{\widehat{\mathbf{h}}}_{2}}\left(\underline{\widehat{\mathbf{h}}}_{1}^{(1)}\right) = \sigma_{d}^{2} - 2\left(\underline{\widehat{\mathbf{h}}}_{1}^{(1)}\right)^{T} \underline{\mathbf{p}}_{2}^{(0)} + \left(\underline{\widehat{\mathbf{h}}}_{1}^{(1)}\right)^{T} \underline{\mathbf{R}}_{2}^{(0)} \underline{\widehat{\mathbf{h}}}_{1}^{(1)},\tag{71}$$

which can be minimized with respect to $\widehat{\underline{\mathbf{h}}}_1^{(1)}$, thereby yielding

$$\widehat{\underline{\mathbf{h}}}_{1}^{(1)} = \left(\underline{\mathbf{R}}_{2}^{(0)}\right)^{-1} \underline{\mathbf{p}}_{2}^{(0)}.$$
(72)

Next, using $\underline{\widehat{h}}_{1}^{(1)}$, we can construct $\widehat{H}_{1,p}^{(1)} = \mathbf{I}_{L_2} \otimes \widehat{h}_{1,p}^{(1)}$ and

$$\underline{\mathbf{p}}_{1}^{(1)} = \left[\begin{array}{ccc} \mathbf{p}^{T} \widehat{\mathbf{H}}_{1,1}^{(1)} & \mathbf{p}^{T} \widehat{\mathbf{H}}_{1,2}^{(1)} & \cdots & \mathbf{p}^{T} \widehat{\mathbf{H}}_{1,P}^{(1)} \end{array} \right]^{T}, \\ \underline{\mathbf{R}}_{1}^{(1)} = \left[\begin{array}{ccc} \left(\widehat{\mathbf{H}}_{1,1}^{(1)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{1,1}^{(1)} & \left(\widehat{\mathbf{H}}_{1,1}^{(1)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{1,2}^{(1)} & \cdots & \left(\widehat{\mathbf{H}}_{1,1}^{(1)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{1,P}^{(1)} \\ \left(\widehat{\mathbf{H}}_{1,2}^{(1)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{1,1}^{(1)} & \left(\widehat{\mathbf{H}}_{1,2}^{(1)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{1,2}^{(1)} & \cdots & \left(\widehat{\mathbf{H}}_{1,2}^{(1)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{1,P}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ \left(\widehat{\mathbf{H}}_{1,P}^{(1)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{1,1}^{(1)} & \left(\widehat{\mathbf{H}}_{1,P}^{(1)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{1,2}^{(1)} & \cdots & \left(\widehat{\mathbf{H}}_{1,P}^{(1)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{1,P}^{(1)} \end{array} \right].$$

Consequently, the MSE from (70) is

$$\mathcal{I}_{\underline{\widehat{h}}_{1}}\left(\underline{\widehat{h}}_{2}^{(1)}\right) = \sigma_{d}^{2} - 2\left(\underline{\widehat{h}}_{2}^{(1)}\right)^{T} \underline{\underline{p}}_{1}^{(1)} + \left(\underline{\widehat{h}}_{2}^{(1)}\right)^{T} \underline{\underline{R}}_{1}^{(1)} \underline{\widehat{h}}_{2}^{(1)}.$$
(73)

The minimization of the previous expression with respect to $\widehat{\underline{h}}_2^{(1)}$ gives

í

$$\widehat{\underline{\mathbf{h}}}_{2}^{(1)} = \left(\underline{\mathbf{R}}_{1}^{(1)}\right)^{-1} \underline{\mathbf{p}}_{1}^{(1)}.$$
(74)

By iterating further, we obtain at iteration n

$$\widehat{\mathbf{h}}_{1}^{(n)} = \left(\underline{\mathbf{R}}_{2}^{(n-1)}\right)^{-1} \underline{\mathbf{p}}_{2}^{(n-1)},\tag{75}$$

$$\widehat{\underline{\mathbf{h}}}_{2}^{(n)} = \left(\underline{\mathbf{R}}_{1}^{(n)}\right)^{-1} \underline{\mathbf{p}}_{1}^{(n)},\tag{76}$$

where $\underline{\mathbf{R}}_{2}^{(n-1)}$, $\underline{\mathbf{p}}_{2}^{(n-1)}$, $\underline{\mathbf{R}}_{1}^{(n)}$, and $\underline{\mathbf{p}}_{1}^{(n)}$ are constructed similarly to $\underline{\mathbf{R}}_{2}^{(0)}$, $\underline{\mathbf{p}}_{2}^{(0)}$, $\underline{\mathbf{R}}_{1}^{(1)}$, and $\underline{\mathbf{p}}_{1}^{(1)}$, respectively. In the end, the Wiener filter at iteration *n* results as

$$\widehat{\mathbf{h}}_{\mathrm{W}}^{(n)} = \sum_{p=1}^{p} \widehat{\mathbf{h}}_{2,p}^{(n)} \otimes \widehat{\mathbf{h}}_{1,p}^{(n)}.$$
(77)

The multichannel iterative Wiener filter is summarized in Table 1. For M = 1 (i.e., single-channel case), the problem is reduced to a regular SISO scenario, and the algorithm becomes equivalent to the version developed in [43]. Additionally, if the system is perfectly

separable/decomposable, we can obtain the optimal solution for P = 1. In this case, the iterative Wiener filter for bilinear forms (proposed in [24]) is obtained.

Table 1. The multichannel iterative Wiener filter.

Data: **R**, **p** (estimated statistics based on *N* data samples), *P* (decomposition parameter), $\delta > 0$ (regularization constant) Initialization: Provide the initial coefficients of the filter $\hat{\mathbf{h}}_{2,p}^{(0)}$ (for p = 1, 2, ..., P): $\widehat{\mathbf{h}}_{2.v}^{(0)} = \begin{bmatrix} \epsilon & 0 & \cdots & 0 \end{bmatrix}^T, \ p = 1, 2, \dots, P \ (0 < \epsilon \le 1)$ Compute the initial statistics $\mathbf{p}_2^{(0)}$ and $\underline{\mathbf{R}}_2^{(0)}$: $\widehat{\mathbf{H}}_{2,p}^{(0)} = \widehat{\mathbf{h}}_{2,p}^{(0)} \otimes \mathbf{I}_{ML_1}, \ p = 1, 2, \dots, P$ $\underline{\mathbf{p}}_{2}^{(0)} = \begin{bmatrix} \mathbf{p}^{T} \widehat{\mathbf{H}}_{2,1}^{(0)} & \cdots & \mathbf{p}^{T} \widehat{\mathbf{H}}_{2,P}^{(0)} \end{bmatrix}^{T}$ $\underline{\mathbf{R}}_{2}^{(0)} = \begin{bmatrix} \left(\widehat{\mathbf{H}}_{2,1}^{(0)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,1}^{(0)} & \cdots & \left(\widehat{\mathbf{H}}_{2,1}^{(0)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,P}^{(0)} \\ \vdots & \ddots & \vdots \\ \left(\widehat{\mathbf{H}}_{2,P}^{(0)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,1}^{(0)} & \cdots & \left(\widehat{\mathbf{H}}_{2,P}^{(0)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,P}^{(0)} \end{bmatrix}$ For n = 1, 2, ...Compute the coefficients of the filter $\underline{\hat{\mathbf{h}}}_{1}^{(n)}$ based on (75): $\widehat{\mathbf{h}}_{1}^{(n)} = \left(\underline{\mathbf{R}}_{2}^{(n-1)} + \delta \mathbf{I}_{PML_{1}}\right)^{-1} \underline{\mathbf{p}}_{2}^{(n-1)} = \left[\begin{array}{ccc} \left(\widehat{\mathbf{h}}_{1,1}^{(n)}\right)^{T} & \cdots & \left(\widehat{\mathbf{h}}_{1,P}^{(n)}\right)^{T}\end{array}\right]^{T}$ Compute the statistics $\mathbf{p}_1^{(n)}$ and $\mathbf{\underline{R}}_1^{(n)}$: $\widehat{\mathbf{H}}_{1,p}^{(n)} = \mathbf{I}_{L_2} \otimes \widehat{\mathbf{h}}_{1,p}^{(n)}, \ p = 1, 2, \dots, P$ $\underline{\mathbf{p}}_{1}^{(n)} = \begin{bmatrix} \mathbf{p}^{T} \widehat{\mathbf{H}}_{1,1}^{(n)} & \cdots & \mathbf{p}^{T} \widehat{\mathbf{H}}_{1,P}^{(n)} \end{bmatrix}^{T}$ $\underline{\mathbf{R}}_{1}^{(n)} = \begin{bmatrix} \left(\widehat{\mathbf{H}}_{1,1}^{(n)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{1,1}^{(n)} & \cdots & \left(\widehat{\mathbf{H}}_{1,1}^{(1)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{1,P}^{(1)} \\ \vdots & \ddots & \vdots \\ \left(\widehat{\mathbf{H}}_{1,p}^{(n)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{1,1}^{(n)} & \cdots & \left(\widehat{\mathbf{H}}_{1,P}^{(n)} \right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{1,P}^{(n)} \end{bmatrix}$ Compute the coefficients of the filter $\underline{\hat{h}}_{2}^{(n)}$ based on (76): $\widehat{\mathbf{\underline{h}}}_{2}^{(n)} = \left(\underline{\mathbf{R}}_{1}^{(n)} + \delta \mathbf{I}_{PL_{2}}\right)^{-1} \underline{\mathbf{p}}_{1}^{(n)} = \left[\begin{array}{ccc} \left(\widehat{\mathbf{h}}_{2,1}^{(n)}\right)^{T} & \cdots & \left(\widehat{\mathbf{h}}_{2,P}^{(n)}\right)^{T}\end{array}\right]^{T}$ Compute the statistics $\mathbf{p}_2^{(n)}$ and $\mathbf{\underline{R}}_2^{(n)}$: $\widehat{\mathbf{H}}_{2,p}^{(n)} = \widehat{\mathbf{h}}_{2,p}^{(n)} \otimes \mathbf{I}_{ML_1}, \ p = 1, 2, \dots, P$ $\mathbf{p}_{2}^{(n)} = \begin{bmatrix} \mathbf{p}^{T} \widehat{\mathbf{H}}_{21}^{(n)} & \cdots & \mathbf{p}^{T} \widehat{\mathbf{H}}_{2n}^{(n)} \end{bmatrix}$ $\underline{\mathbf{R}}_{2}^{(n)} = \begin{bmatrix} \left(\widehat{\mathbf{H}}_{2,1}^{(n)}\right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,1}^{(n)} & \cdots & \left(\widehat{\mathbf{H}}_{2,1}^{(n)}\right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,P}^{(n)} \\ \vdots & \ddots & \vdots \\ \left(\widehat{\mathbf{H}}_{2,P}^{(n)}\right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,1}^{(n)} & \cdots & \left(\widehat{\mathbf{H}}_{2,P}^{(n)}\right)^{T} \mathbf{R} \widehat{\mathbf{H}}_{2,P}^{(n)} \end{bmatrix}$ Compute the coefficients of the Wiener filter $\hat{\mathbf{h}}_{W}^{(n)}$ based on (77): $\langle \rangle$

$$\widehat{\mathbf{h}}_{\mathrm{W}}^{(n)} = \sum_{p=1}^{i} \widehat{\mathbf{h}}_{2,p}^{(n)} \otimes \widehat{\mathbf{h}}_{1,p}^{(n)}$$

6. Simulation Results

In this section, we evaluate the performance of the conventional and iterative Wiener filters in two different scenarios. The first one is dedicated to the case of independent input signals, $\mathbf{x}_m(k)$, m = 1, 2, ..., M. The second scenario is more challenging, since it considers the case when the input signals are coming from the same source and they are linearly related. In both cases, the performance measure used to evaluate the overall behavior is the normalized misalignment (in dB), which is related to the spatiotemporal impulse response of the system, \mathbf{h} . In this framework, the solution provided by the conventional Wiener filter is given in (58), so that the performance measure is evaluated as

$$\mathcal{M}\left(\underline{\mathbf{h}}, \underline{\widehat{\mathbf{h}}}_{W}\right) = 20\log_{10} \frac{\left\|\underline{\mathbf{h}} - \underline{\widehat{\mathbf{h}}}_{W}\right\|_{2}}{\left\|\underline{\mathbf{h}}\right\|_{2}}.$$
(78)

Similarly, for the iterative Wiener filter from (77), the performance measure results in

$$\mathcal{M}\left(\underline{\mathbf{h}}, \underline{\widehat{\mathbf{h}}}_{W}^{(n)}\right) = 20\log_{10} \frac{\left\|\underline{\mathbf{h}} - \underline{\widehat{\mathbf{h}}}_{W}^{(n)}\right\|_{2}}{\left\|\underline{\mathbf{h}}\right\|_{2}}.$$
(79)

...

...

Both the conventional and iterative Wiener filters rely on the estimation of the statistics, i.e., the covariance matrix \mathbf{R} and the cross-correlation vector \mathbf{p} . Considering that N data samples are available, these estimates result in

$$\widehat{\mathbf{R}} = \frac{1}{N} \sum_{k=1}^{N} \underline{\mathbf{x}}(k) \underline{\mathbf{x}}^{T}(k),$$
(80)

$$\widehat{\mathbf{p}} = \frac{1}{N} \sum_{k=1}^{N} \underline{\mathbf{x}}(k) d(k).$$
(81)

Clearly, the value of *N* influences the quality of these estimates. Nevertheless, in practice, only a small amount of data could be available, which makes the identification process more challenging. In this case, the advantages of the iterative Wiener filter (which operates with smaller data structures) become more apparent, as will be supported in the following analysis.

The additive noise may also affect the accuracy of the Wiener solution. In relation to (13) or (22), the signal-to-noise ratio (SNR) can be defined as

$$SNR = \frac{\sigma_y^2}{\sigma_w^2},$$
(82)

where σ_y^2 and σ_w^2 are the variances of the output signal and noise, respectively. In practice, the Wiener solution is satisfactory with reasonable levels of the SNR, but it is not with small values of the SNR. In our experiments, different values of the SNR were used, in order to illustrate this behavior.

All the experiments were performed using MATLAB R2018b on an Asus GL552VX device (Windows 10 OS), having an Intel Core i7-6700HQ CPU @ 2.60 GHz, with four cores, eight logical processors, and 16 GB of RAM. In the first set of experiments, we considered the case of *M* independent input signals, which were generated as AR(1) processes. These were obtained by filtering white Gaussian noise through an AR(1) model with a pole at 0.9. Of course, different other inputs can be used instead of the AR(1) model. The most common considerations are: (i) the input signal x(k) is wide-sense stationary, (ii) all of the signals (i.e., x(k), d(k), and w(k)) have zero means, and (iii) usually, the noise w(k) is not correlated with the input signal x(k).

In our scenario, the number of channels was M = 4 and their impulse responses were chosen from the G168 Recommendation [63]. They were network echo paths of length L = 500, as depicted in Figure 3.

As mentioned before, the performance of the Wiener solution is influenced by the value of N and the level of the SNR. This is supported in Figure 7, where the performance of the conventional Wiener filter from (58) is illustrated for different values of N (from 2000 to 10,000 available data samples) and SNR levels (from 0 dB to 30 dB). As we can notice, a larger value of N (i.e., $N \gg ML$) is required to obtain reasonable attenuation of the misalignment. Additionally, as expected, a more accurate solution was obtained with higher SNRs.



Figure 7. Performance of the conventional Wiener filter for different values of *N* (number of available data samples to estimate the statistics) and SNRs. The input signals were independent AR(1) processes, M = 4, and L = 500.

In this context, let us first compare the performance of the conventional and iterative Wiener filters in "favorable" conditions, using a large amount of data to estimate the statistics (i.e., N = 10,000), in a high SNR environment (i.e., SNR = 30 dB). The iterative Wiener filter from (77) uses $L_1 = 25$, $L_2 = 20$, and different values of the decomposition parameter P (from 3 to 6). These values are much lower than L_2 , which represents an important advantage, as discussed in Section 5. We should also note that for the scenario considered in this first set of experiments (using the setup from Figure 3), the rank of the matrix **H** (or \mathbf{H}_{st}) was equal to R = 5. As we can notice in Figure 8, the iterative Wiener filter was able to outperform the conventional Wiener filter for most of the values of P. Even the case of $P = 3 < R \ll L_2$ led to reasonable attenuation of the misalignment. Moreover, all these iterative solutions were obtained in only a few iterations.



Figure 8. Performances of the conventional and iterative Wiener filters, using N = 10,000. The input signals were independent AR(1) processes, M = 4, L = 500, and SNR = 30 dB.

The advantages of the iterative Wiener filter became more apparent when less data were available to estimate the statistics. The previous simulation was repeated, but using N = 2500 (Figure 9). According to the results shown in Figure 7, the performance of the conventional Wiener filter was affected in this case (even for SNR = 30 dB); its misalignment level was close to -15 dB. This result is also confirmed in Figure 9. Most importantly, the iterative Wiener filter outperformed the conventional solution for all the values of *P*, thereby being more robust in this case due to the low-dimensional data structures used in its development.



Figure 9. Performances of the conventional and iterative Wiener filters, using N = 2500. The input signals were independent AR(1) processes, M = 4, L = 500, and SNR = 30 dB.

For the scenario considered in this first set of experiments, the length of the spatiotemporal impulse response was $ML = 4 \times 500 = 2000$. Therefore, the case N = 2000 represents a limit in terms of the available amount of data. As we can notice in Figure 7, the conventional Wiener filter could not cope with this limit. On the other hand, the iterative Wiener filter was still able to obtain good performances, for $P \ll L_2$, as supported in Figure 10.



Figure 10. Performance of the iterative Wiener filter with different values of *P*, using N = 2000. The input signals were independent AR(1) processes, M = 4, L = 500, and SNR = 30 dB.

Furthermore, using N < ML is a significant challenge in terms of system identification, since apparently we deal with an "incomplete" scenario, when trying to estimate ML coefficients using less data. Clearly, the conventional Wiener filter cannot be used in this case. However, the iterative Wiener filter reformulates the original system identification problem (of size ML) as a combination of low-dimension solutions of size PML_1 and PL_2 , with $P \ll L_2$. Hence, it could overcome this limit of N < ML. This is supported in Figure 11: only N = 1000 data samples were available for the estimation of the statistics. Even in this challenging case, the iterative Wiener filter was able to provide a good attenuation of the misalignment, in a relatively small number of iterations. This represents an important feature and a significant advantage when dealing with small amounts of data. In other words, the iterative Wiener filter exploiting the decomposition-based approach can be used to solve system identification problems with highly incomplete information, which is a condition imposed in many important applications.



Figure 11. Performance of the iterative Wiener filter with different values of *P*, using N = 1000. The input signals were independent AR(1) processes, M = 4, L = 500, and SNR = 30 dB.

The SNR is also a critical factor in system identification problems. As shown in Figure 7, the performance of the conventional Wiener filter is highly influenced by the SNR level. In the following simulations from this first set of experiments, we considered a more challenging environment, by setting SNR = 10 dB. In this case, even for a large amount of data (i.e., N = 10,000), the conventional Wiener filter was outperformed by the iterative version, as supported in Figure 12.



Figure 12. Performances of the conventional and iterative Wiener filters, using N = 10,000. The input signals were independent AR(1) processes, M = 4, L = 500, and SNR = 10 dB.

The gain becomes more apparent when fewer data are available to estimate the statistics. Such a case is considered in Figure 13, where SNR = 10 dB and N = 2500. As we can see, the conventional Wiener filter could not provide an accurate solution (as also indicated in Figure 7), whereas its iterative counterpart still attenuated the misalignment to an acceptable degree.



Figure 13. Performances of the conventional and iterative Wiener filters, using N = 2500. The input signals were independent AR(1) processes, M = 4, L = 500, and SNR = 10 dB.

As previously explained (related to Figures 10 and 11), using $N \le ML$ represents a critical scenario for the conventional Wiener filter. In this context, a lower SNR level made the situation even more challenging. Nevertheless, the iterative Wiener filter is reasonably robust even in these adverse conditions. This behavior is supported in Figures 14 and 15, where N = 2000 and 1500, respectively. The same noisy conditions were considered, with SNR = 10 dB. As we can see in these figures, the best behavior was obtained for P = 3, which is significantly lower than $L_2 = 20$.



Figure 14. Performance of the iterative Wiener filter with different values of *P*, using N = 2000. The input signals were independent AR(1) processes, M = 4, L = 500, and SNR = 10 dB.



Figure 15. Performance of the iterative Wiener filter with different values of *P*, using N = 1500. The input signals were independent AR(1) processes, M = 4, L = 500, and SNR = 10 dB.

Finally, the last simulation of the first set of experiments was performed in "extreme" SNR conditions, using SNR = 0 dB. As we already know from Figure 7, the conventional Wiener filter cannot provide an accurate solution in this case, despite the value of *N*. In Figure 16, N = 10,000, and SNR = 0 dB. As expected, the conventional Wiener filter failed to provide an accurate estimate. On the other hand, the iterative Wiener filter was able to reach a much lower misalignment level, using $P \ll L_2$. Therefore, it is much more robust in noisy conditions, which are frequent in practice.



Figure 16. Performances of the conventional and iterative Wiener filters, using N = 10,000. The input signals were independent AR(1) processes, M = 4, L = 500, and SNR = 0 dB.

The second set of experiments was performed in a more challenging situation that appeared in the context of stereophonic acoustic echo cancellation (SAEC) [65–68]. There were two acoustic echo paths to identify (for each microphone); i.e., M = 2. Consequently, the reference (or microphone) signal resulted in

$$d(k) = \mathbf{h}_{\mathrm{R}}^{T} \mathbf{x}_{\mathrm{R}}(k) + \mathbf{h}_{\mathrm{L}}^{T} \mathbf{x}_{\mathrm{L}}(k) + w(k)$$

= $\begin{bmatrix} \mathbf{h}_{\mathrm{R}}^{T} & \mathbf{h}_{\mathrm{L}}^{T} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{\mathrm{R}}(t) \\ \mathbf{x}_{\mathrm{L}}(k) \end{bmatrix} + w(k),$ (83)

where \mathbf{h}_{R} and \mathbf{h}_{L} correspond to the loudspeaker-to-microphone acoustic impulse responses (right and left, respectively), and $\mathbf{x}_{R}(k)$ and $\mathbf{x}_{L}(k)$ comprise the loudspeaker signal samples (right and left, respectively).

At first glance, from a system identification perspective, we need to identify the global impulse response $\begin{bmatrix} \mathbf{h}_{R}^{T} & \mathbf{h}_{L}^{T} \end{bmatrix}^{T}$. Nevertheless, in an SAEC scenario, the difficulty is manyfold. One of the main challenges is the so-called nonuniqueness problem [66,67], which comes from the fact that the loudspeaker (input) signals are linearly related. This issue can be addressed by manipulating the signals transmitted to the receiving room, e.g., using a preprocessor on the loudspeaker signals to make them less coherent, without affecting the stereo perception and the signal quality much. A simple but efficient nonlinear method uses positive and negative half-wave rectifiers on each channel, respectively [67]. In this case, the nonlinearly transformed signals become

$$\mathbf{x}_{\mathrm{L}}'(k) = \mathbf{x}_{\mathrm{L}}(k) + \alpha \frac{\mathbf{x}_{\mathrm{L}}(k) + |\mathbf{x}_{\mathrm{L}}(k)|}{2},\tag{84}$$

$$\mathbf{x}_{\mathrm{R}}'(k) = \mathbf{x}_{\mathrm{R}}(k) + \alpha \frac{\mathbf{x}_{\mathrm{R}}(k) - |\mathbf{x}_{\mathrm{R}}(k)|}{2},\tag{85}$$

where α is a parameter used to control the amount of nonlinearity; the recommended interval for this parameter is $0 < \alpha \le 0.5$ [67]. The distortion parameter α (which controls the amount of nonlinearity) is provided a priori. Clearly, this distortion must be performed in such a way that the quality of the signals and the stereo effect are not degraded. Experiments reported in [67] (and also in many subsequent works) show that stereo perception is not affected even with an α as large as 0.5. Additionally, the audible distortion is small because of the psychoacoustic masking effects [69].

However, we should note that other methods can be used to address the nonuniqueness problem; e.g., see [70,71] and the references therein. An analysis of their influences on the overall performance of the decomposition-based approach is beyond the scope of this paper.

In our simulations, the source signal (in the transmission room) was white Gaussian noise. The acoustic impulse responses in the transmission room had 2048 coefficients. The acoustic impulse responses in the receiving room had L = 1024 coefficients, as depicted in Figure 4. The background noise w(k) (in the receiving room) was white and Gaussian, with SNR = 30 dB.

The influence of the preprocessing technique from (84) and (85) on the loudspeaker signals can be seen in Figure 17, where the performance of the conventional Wiener filter is evaluated for different values of N (i.e., the available amount of data used to estimate the statistics) and using different values of α , which acted as a distortion parameter. First, we can notice that the performance was clearly improved when preprocessing the input signals using positive and negative half-wave rectifiers with larger values of α . Additionally, even with preprocessing, a large amount of data (i.e., $N \gg ML$) is required for the conventional Wiener filter, in order to obtain reasonable misalignment attenuation.



Figure 17. Performance of the conventional Wiener filter for different values of *N* (number of available data samples to estimate the statistics) and different values of α (the distortion parameter). The source signal (white Gaussian noise) was preprocessed with positive and negative half-wave rectifiers. The numbers of channels were M = 2 (stereophonic scenario), L = 1024, and SNR = 30 dB.

In the following simulations, the identification problem was addressed using the conventional and iterative Wiener filters. The length of the global impulse response was $L = L_1L_2 = 1024$, so that we set $L_1 = L_2 = 32$. Due to the nature of acoustic impulse responses, the matrix **H** (or **H**_{st}) was closer to full rank, so a higher value of *P* was required. In this context, the iterative Wiener filter involved in the experiments used P = 6, 8, 10, and 12. Nevertheless, these values are lower than $L_2 = 32$.

For the results shown in Figure 18, N = 10,000 data samples were used to estimate the statistics, and the input signals were preprocessed using positive and negative half-wave rectifiers, with $\alpha = 0.5$. These represent favorable conditions for the identification, so the conventional Wiener filter had good accuracy (as also indicated in Figure 17). The performance of the iterative Wiener filter was improved for a larger value of *P* and even outperformed the conventional filter in the case of P = 12.



Figure 18. Performances of the conventional and iterative Wiener filters, using N = 10,000. The source signal (white Gaussian noise) was preprocessed with positive and negative half-wave rectifiers, using $\alpha = 0.5$. The numbers of channels were M = 2 (stereophonic scenario), L = 1024, and SNR = 30 dB.

Reducing the value of the distortion parameter influenced the performance for the conventional and the iterative Wiener filters, as shown in Figure 19, where N = 10,000 and $\alpha = 0.3$. However, the iterative Wiener filter is more robust to this modification, since for P = 12 it outperformed the conventional benchmark, and for P = 8 it reached a misalignment level close to the conventional Wiener solution.



Figure 19. Performances of the conventional and iterative Wiener filters, using N = 10,000. The source signal (white Gaussian noise) was preprocessed with positive and negative half-wave rectifiers, using $\alpha = 0.3$. The numbers of channels were M = 2 (stereophonic scenario), L = 1024, and SNR = 30 dB.

Next, for the simulations shown in Figures 20 and 21, less data were used to estimate the statistics—i.e., N = 2500. The other conditions were the same as in Figures 18 and 19, respectively. As we can notice, the conventional Wiener filter could not obtain an accurate solution in these cases, despite the value of the distortion parameter. On the other hand, the iterative Wiener filter was still able to achieve reasonable results (for $P < L_2$), thereby far outperforming the conventional solution. The difference was even more apparent for a lower value of α , as supported in Figure 21.



Figure 20. Performances of the conventional and iterative Wiener filters, using N = 2500. The source signal (white Gaussian noise) was preprocessed with positive and negative half-wave rectifiers, using $\alpha = 0.5$. The numbers of channels were M = 2 (stereophonic scenario), L = 1024, and SNR = 30 dB.



Figure 21. Performances of the conventional and iterative Wiener filters, using N = 2500. The source signal (white Gaussian noise) was preprocessed with positive and negative half-wave rectifiers, using $\alpha = 0.3$. The numbers of channels were M = 2 (stereophonic scenario), L = 1024, and SNR = 30 dB.

A challenging case N < ML was considered in the following two simulations, where we set N = 2000 (while ML = 2048). Two values of the distortion parameter were considered, i.e., $\alpha = 0.5$ and 0.3, and the results are depicted in Figures 22 and 23, respectively. The conventional Wiener filter was not included in these experiments, since it could not provide an accurate solution (as indicated in Figure 17). As we can notice in both cases, despite the adverse conditions, the iterative Wiener filter was still able to attenuate the misalignment to a reasonable extent and provide a robust solution, for a value of *P* smaller than L_2 .



Figure 22. Performance of the iterative Wiener filter with different values of *P*, using N = 2000. The source signal (white Gaussian noise) was preprocessed with positive and negative half-wave rectifiers, using $\alpha = 0.5$. The numbers of channels were M = 2 (stereophonic scenario), L = 1024, and SNR = 30 dB.



Figure 23. Performance of the iterative Wiener filter with different values of *P*, using N = 2000. The source signal (white Gaussian noise) was preprocessed with positive and negative half-wave rectifiers, using $\alpha = 0.3$. The numbers of channels were M = 2 (stereophonic scenario), L = 1024, and SNR = 30 dB.

As shown in Figure 17, the conventional Wiener filter could not obtain an accurate solution for small values of the distortion parameter α (i.e., closer to 0). For example, when using $\alpha = 0.1$, its misalignment could not reach -5 dB even with a large amount of data (i.e., N = 10,000); and for less data (e.g., N = 2500), it provided a far from accurate solution. These cases are considered in Figures 24 and 25, using $\alpha = 0.1$; N = 10,000 and 2500, respectively. For N = 10,000 (Figure 24), the iterative Wiener filter outperformed the conventional solution for P = 10 and 12. The decomposition using $P = 8 \ll L_2$ reached the misalignment level provided by the conventional Wiener filter. However, the differences (in favor of the iterative version, for all the values of P) were significant when N = 2500, as supported in Figure 25.



Figure 24. Performances of the conventional and iterative Wiener filters, using N = 10,000. The source signal (white Gaussian noise) was preprocessed with positive and negative half-wave rectifiers, using $\alpha = 0.1$. The numbers of channels were M = 2 (stereophonic scenario), L = 1024, and SNR = 30 dB.



Figure 25. Performances of the conventional and iterative Wiener filters, using N = 2500. The source signal (white Gaussian noise) was preprocessed with positive and negative half-wave rectifiers, using $\alpha = 0.1$. The numbers of channels were M = 2 (stereophonic scenario), L = 1024, and SNR = 30 dB.

Finally, as shown in Figure 26, N = 10,000 data samples are used to estimate the statistics, but the input signals were not preprocessed (i.e., $\alpha = 0$). Even if $N \gg ML$, the conventional Wiener filter could not obtain an accurate solution in this case. On the other hand, the iterative Wiener filter with $P \ll L_2$ was able to outperform the conventional Wiener solution, even in this extremely difficult scenario. In other words, the influence of the nonuniqueness problem is less significant for the iterative Wiener filter. This is probably due to the fact that the matrices within the iterative Wiener filter are smaller as compared to the full matrix of size $ML \times ML$ within the conventional Wiener filter.



Figure 26. Performances of the conventional and iterative Wiener filters, using N = 10,000. The source signal (white Gaussian noise) was not preprocessed (no distortion, i.e., $\alpha = 0$). The numbers of channels were M = 2 (stereophonic scenario), L = 1024, and SNR = 30 dB.

Summarizing, the main feature of the multichannel iterative Wiener filter is that it operates with smaller data structures, i.e., of size PML_1 and PL_2 , with P values far smaller than L_2 . On the other hand, the conventional Wiener filter addresses a system identification problem of size $L = L_1L_2$. In our experiments, we tried to cover a wide range of scenarios,

in order to properly assess the performance of the multichannel iterative Wiener filter, as compared to its conventional counterpart. Consequently, we used different values of N (the available data to estimate the statistics), different noise levels (SNR), and different values of the decomposition parameter P. Moreover, in the SAEC scenario, we also used different values of the distortion parameter α . As an overall conclusion, the iterative version significantly outperformed the conventional Wiener filter, especially in difficult conditions and environments, e.g., when using small amounts of data or low SNRs. These scenarios are of great importance in practice, as in real-world applications, only small amounts of data may be available (to estimate the statistics) or the filters may need to operate in noisy environments.

7. Conclusions and Perspectives

In this review paper, we have addressed the system identification problem from an efficient decomposition-based perspective. The contributions are threefold. First, we have shown how the main categories of linear SISO and MISO systems can be interpreted and related in a unified framework, taking advantage of the bilinear form representation. Second, we have demonstrated that the resulting spatiotemporal impulse response of the MISO system can be efficiently identified using the nearest Kronecker product decomposition, followed by low-rank approximations. Third, we have developed an iterative Wiener filter based on these techniques which outperforms the conventional Wiener filter in terms of accuracy and robustness of the solution. The main feature of the overall approach consists of an efficient (re)formulation of a high-dimension system identification problem (e.g., identification of a long spatiotemporal impulse response) as a combination of low-dimension solutions, which result from the optimization of shorter component filters.

In this study, we have illustrated the performance of the decomposition-based approach only in terms of the Wiener filter, which represents a benchmark tool for system identification problems. Simulation results have indicated that the iterative version is able to outperform the conventional Wiener filter, especially when a small amount of data is available for the estimation of the statistics. This represents an important performance feature, taking into account that in many real-world applications we deal with incomplete information related to the inputs/outputs of the system.

In perspective, improved solutions based on adaptive filtering should also be developed, which could further extend the applicability of the decomposition-based approach. For example, several preliminaries toward this goal can be found in [44,45]. Furthermore, a rigorous convergence analysis of these algorithms could reveal the influence of the decomposition parameters, which could be further exploited in order to improve the overall behavior. In addition, finding a practical method to determine the optimal value of the decomposition parameter P represents one of our main tasks for future works. Nevertheless, this is not a straightforward task, since the decomposition parameter depends on the nature of the system to be identified (which is unknown in practice). However, we can take advantage on some a priori knowledge of the system. For example, in the case of network echo paths (which are usually very sparse), the value of P is much smaller than L_2 . For acoustic impulse responses, the value of P should be increased, but it must still be considerably lower than L_2 . A preliminary study on the influence of the decomposition parameter can be found in [43].

Author Contributions: Conceptualization, J.B.; methodology, C.P.; validation, L.-M.D.; formal analysis, S.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by a grant of the Romanian Ministry of Research and Innovation, CNCS—UEFISCDI, project number PN-III-P1-1.1-PD-2019-0340, within PNCDI III.

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

References

- 1. Ljung, L. *System Identification: Theory for the User*, 2nd ed.; Prentice-Hall: Upper Saddle River, NJ, USA, 1999.
- 2. Benesty, J.; Huang, Y. (Eds.) Adaptive Signal Processing-Applications to Real-World Problems; Springer: Berlin, Germany, 2003.
- 3. Mohler, R.R.; Kolodziej, W.J. An overview of bilinear system theory and applications. *IEEE Trans. Syst. Man Cybern.* **1980**, *10*, 683–688.
- Halawani, T.U.; Mohler, R.R.; Kolodziej, W.J. A two-step bilinear filtering approximation. *IEEE Trans. Acoust. Speech Signal Process.* 1984, 32, 344–352. [CrossRef]
- 5. Inagaki, M.; Mochizuki, H. Bilinear system identification by Volterra kernels estimation. *IEEE Trans. Autom. Control* **1984**, *29*, 746–749. [CrossRef]
- 6. Baik, H.K.; Mathews, V.J. Adaptive lattice bilinear filters. *IEEE Trans. Signal Process.* **1993**, *41*, 2033–2046. [CrossRef]
- Forssén, U. Adaptive bilinear digital filters. *IEEE Trans. Circuits Syst. II Analog. Digit. Signal Process.* 1993, 40, 729–735. [CrossRef]
 Ma, G.-K.; Lee, J.; Mathews, V.J. A RLS bilinear filter for channel equalization. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Adelaide, SA, Australia, 19–22 April 1994; pp. III-257–III-260.
- 9. Lee, J.; Mathews, V.J. Adaptive bilinear predictors. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Adelaide, SA, Australia, 19–22 April 1994; pp. III-489–III-492.
- 10. Hu, R.; Hassan, H.M. Echo cancellation in high speed data transmission systems using adaptive layered bilinear filters. *IEEE Trans. Commun.* **1994**, *42*, 655–663.
- 11. Bose, T.; Chen, M.-Q. Conjugate gradient method in adaptive bilinear filtering. *IEEE Trans. Signal Process.* **1995**, *43*, 1503–1508. [CrossRef]
- 12. Lee, J.; Mathews, V.J. Output-error LMS bilinear filters with stability monitoring. In Proceedings of the 1995 International Conference on Acoustics, Speech, and Signal Processing, Detroit, MI, USA, 9–12 May 1995; pp. 965–968.
- 13. Gesbert, D.; Duhamel, P. Robust blind joint data/channel estimation based on bilinear optimization. In Proceedings of the 8th Workshop on Statistical Signal and Array Processing, Corfu, Greece, 24–26 June 1996; pp. 168–171.
- 14. Stenger, A.; Kellermann, W.; Rabenstein, R. Adaptation of acoustic echo cancellers incorporating a memoryless nonlinearity. In Proceedings of the Proceedings of 8th Workshop on Statistical Signal and Array Processing, Corfu, Greece, 24–26 June 1996; 4p.
- 15. Stenger, A.; Kellermann, W. Adaptation of a memoryless preprocessor for nonlinear acoustic echo cancelling. *Signal Process.* **2000**, *80*, 1747–1760. [CrossRef]
- 16. Zhu, Z.; Leung, H. Adaptive identification of nonlinear systems with application to chaotic communications. *IEEE Trans. Circuits Syst. I Fundam. Theory Appl.* 2000, 47, 1072–1080. [CrossRef]
- 17. Kuo, S.M.; Wu, H.-T. Nonlinear adaptive bilinear filters for active noise control systems. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2005**, *52*, 617–624. [CrossRef]
- 18. Abrahamsson, R.; Kay, S.M.; Stoica, P. Estimation of the parameters of a bilinear model with applications to submarine detection and system identification. *Digit. Signal Process.* **2007**, *17*, 756–773. [CrossRef]
- 19. Lopes dos Santos, P.; Ramos, J.A.; Martins de Carvalho, J.L. Identification of bilinear systems with white noise inputs: An iterative deterministic-stochastic subspace approach. *IEEE Trans. Control. Syst. Technol.* **2009**, *17*, 1145–1153.. [CrossRef]
- 20. Zhao, H.; Zeng, X.; He, Z. Low-complexity nonlinear adaptive filter based on a pipelined bilinear recurrent neural network. *IEEE Trans. Neural Netw.* **2011**, *22*, 1494–1507. [CrossRef]
- Tan, L.; Jiang, J. Nonlinear active noise control using diagonal-channel LMS and RLS bilinear filters. In Proceedings of the 2014 IEEE 57th International Midwest Symposium on Circuits and Systems (MWSCAS), College Station, TX, USA, 3–6 August 2014; pp. 789–792.
- 22. Huang, Y.; Skoglund, J.; Luebs, A. Practically efficient nonlinear acoustic echo cancellers using cascaded block RLS and FLMS adaptive filters. In Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, USA, 5–9 March 2017; pp. 596–600.
- 23. Bai, E.-W.; Li, D. Convergence of the iterative Hammerstein system identification algorithm. *IEEE Trans. Autom. Control* 2004, 49, 1929–1940. [CrossRef]
- 24. Benesty, J.; Paleologu, C.; Ciochina, S. On the identification of bilinear forms with the Wiener filter. *IEEE Signal Process. Lett.* **2017**, 24, 653–657. [CrossRef]
- 25. Haykin, S. Adaptive Filter Theory, 4th ed.; Prentice-Hall: Upper Saddle River, NJ, USA, 2002.
- 26. Diniz, P.S.R. Adaptive Filtering: Algorithms and Practical Implementation, 4th ed.; Springer: New York, NY, USA, 2013.
- 27. Rupp, M.; Schwarz, S. A tensor LMS algorithm. In Proceedings of the 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), South Brisbane, QLD, Australia, 19–24 April 2015; pp. 3347–3351.
- 28. Paleologu, C.; Benesty, J.; Ciochină, S. Adaptive filtering for the identification of bilinear forms. *Digit. Signal Process.* **2018**, *75*, 153–167. [CrossRef]
- 29. Elisei-Iliescu, C.; Stanciu, C.; Paleologu, C.; Benesty, J.; Anghel, C.; Ciochină, S. Efficient recursive least-squares algorithms for the identification of bilinear forms. *Digit. Signal Process.* **2018**, *83*, 280–296. [CrossRef]
- Dogariu, L.; Paleologu, C.; Ciochină, S.; Benesty, J.; Piantanida, P. Identification of bilinear forms with the Kalman filter. In Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 15–20 April 2018; pp. 4134–4138.

- 31. Cichocki, A.; Zdunek, R.; Pan, A.H.; Amari, S. Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multiway Data Analysis and Blind Source Separation; Wiley: Chichester, UK, 2009.
- 32. Boussé, M.; Debals, O.; de Lathauwer, L. A tensor-based method for large-scale blind source separation using segmentation. *IEEE Trans. Signal Process.* 2017, 65, 346–358. [CrossRef]
- 33. Benesty, J.; Cohen, I.; Chen, J. Array Processing-Kronecker Product Beamforming; Springer: Cham, Switzerland, 2019.
- Ribeiro, L.N.; de Almeida, A.L.F.; Mota, J.C.M. Separable linearly constrained minimum variance beamformers. *Signal Process*. 2019, 158, 15–25. [CrossRef]
- Vasilescu, M.A.O.; Kim, E. Compositional hierarchical tensor factorization: Representing hierarchical intrinsic and extrinsic causal factors. In Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), Anchorage, AK, USA, 4–8 August 2019; 9p.
- Vasilescu, M.A.O.; Kim, E.; Zeng, X.S. CausalX: Causal eXplanations and block multilinear factor analysis. In Proceedings of the 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 10–15 January 2021; 8p.
- 37. Vervliet, N.; Debals, O.; Sorber, L.; de Lathauwer, L. Breaking the curse of dimensionality using decompositions of incomplete tensors: Tensor-based scientific computing in big data analysis. *IEEE Signal Process. Mag.* **2014**, *31*, 71–79. [CrossRef]
- Cichocki, A.; Mandic, D.; de Lathauwer, L.; Zhou, G.; Zhao, Q.; Caiafa, C.; Phan, A.H. Tensor decompositions for signal processing applications: From two-way to multiway component analysis. *IEEE Signal Process. Mag.* 2015, 32, 145–163. [CrossRef]
- Sidiropoulos, N.; de Lathauwer, L.; Fu, X.; Huang, K.; Papalexakis, E.; Faloutsos, C. Tensor decomposition for signal processing and machine learning. *IEEE Trans. Signal Process.* 2017, 65, 3551–3582. [CrossRef]
- 40. da Costa, M.N.; Favier, G.; Romano, J.M.T. Tensor modelling of MIMO communication systems with performance analysis and Kronecker receivers. *Signal Process.* **2018**, *145*, 304–316. [CrossRef]
- 41. Dogariu, L.-M.; Stanciu, C.L.; Elisei-Iliescu, C.; Paleologu, C.; Benesty, J.; Ciochină, S. Tensor-based adaptive filtering algorithms. *Symmetry* **2021**, *13*, 481. [CrossRef]
- 42. Dogariu, L.-M.; Paleologu, C.; Benesty, J.; Stanciu, C.L.; Oprea, C.C.; Ciochină, S. A Kalman filter for multilinear forms and its connection with tensorial adaptive filters. *Sensors* **2021**, *21*, 3555. [CrossRef] [PubMed]
- 43. Paleologu, C.; Benesty, J.; Ciochină, S. Linear system identification based on a Kronecker product decomposition. *IEEE/ACM Trans. Audio Speech Lang. Process.* 2018, 26, 1793–1808. [CrossRef]
- 44. Elisei-Iliescu, C.; Paleologu, C.; Benesty, J.; Stanciu, C.; Anghel, C.; Ciochină, S. Recursive least-squares algorithms for the identification of low-rank systems. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2019**, 27, 903–918. [CrossRef]
- 45. Dogariu, L.-M.; Paleologu, C.; Benesty, J.; Ciochină, S. An efficient Kalman filter for the identification of low-rank systems. *Signal Process.* **2020**, *166*, 107239. [CrossRef]
- Benesty, J.; Paleologu, C.; Oprea, C.C.; Ciochină, S. An iterative multichannel Wiener filter based on a Kronecker product decomposition. In Proceedings of the 2020 28th European Signal Processing Conference (EUSIPCO), Amsterdam, The Netherlands, 18–21 January 2021; pp. 211–215.
- Elisei-Iliescu, C.; Paleologu, C.; Benesty, J.; Stanciu, C.; Anghel, C.; Ciochină, S. A multichannel recursive least-squares algorithm based on a Kronecker product decomposition. In Proceedings of the 2020 43rd International Conference on Telecommunications and Signal Processing (TSP), Milan, Italy, 7–9 July 2020; pp. 14–18.
- Cohen, I.; Benesty, J.; Chen, J. Differential Kronecker product beamforming. *IEEE/ACM Trans. Audio Speech Lang. Process.* 2019, 27, 892–902. [CrossRef]
- 49. Bhattacharjee, S.S.; Kumar, K.; George, N.V. Nearest Kronecker product decomposition based generalized maximum correntropy and generalized hyperbolic secant robust adaptive filters. *IEEE Signal Process. Lett.* **2020**, *27*, 1525–1529. [CrossRef]
- 50. Bhattacharjee, S.S.; George, N.V. Nonlinear system identification using exact and approximate improved adaptive exponential functional link networks. *IEEE Trans. Circuits Syst. II Express Briefs* **2020**, *67*, 3542–3546. [CrossRef]
- 51. Bhattacharjee, S.S.; George, N.V. Fast and efficient acoustic feedback cancellation based on low rank approximation. *Signal Process.* **2021**, *182*, 107984. [CrossRef]
- 52. Bhattacharjee, S.S.; George, N.V. Nearest Kronecker product decomposition based linear-in-the-parameters nonlinear filters. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2021**, accepted for publication.
- 53. Yang, W.; Huang, G.; Chen, J.; Benesty, J.; Cohen, I.; Kellermann, W. Robust dereverberation with Kronecker product based multichannel linear prediction. *IEEE Signal Process. Lett.* **2021**, *28*, 101–105. [CrossRef]
- 54. Kuhn, E.V.; Pitz, C.A.; Matsuo, M.V.; Bakri, K.J.; Seara, R.; Benesty, J. A Kronecker product CLMS algorithm for adaptive beamforming. *Digit. Signal Process.* **2021**, *111*, 102968. [CrossRef]
- He, H.; Chen, J.; Benesty, J.; Yu, Y. Robust recursive least M-estimate adaptive filter for the identification of low-rank acoustic systems. In Proceedings of the ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, ON, Canada, 6–11 June 2021; pp. 940–944.
- 56. Van Loan, C.F. The ubiquitous Kronecker product. J. Comput. Appl. Math. 2000, 123, 85-100. . [CrossRef]
- 57. Gay, S.L.; Benesty, J. (Eds.) Acoustic Signal Processing for Telecommunication; Kluwer Academic Publisher: Boston, MA, USA, 2000.
- 58. Benesty, J.; Gänsler, T.; Morgan, D.R.; Sondhi, M.M.; Gay, S.L. *Advances in Network and Acoustic Echo Cancellation*; Springer: Berlin, Germany, 2001.
- 59. Paleologu, C.; Benesty, J.; Ciochină, S. *Sparse Adaptive Filters for Echo Cancellation*; Morgan & Claypool Publishers: Williston, VT, USA, 2010.

- 60. Liu, J.; Grant, S.L. Proportionate adaptive filtering for block-sparse system identification. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2016**, 24, 623–630. [CrossRef]
- 61. Golub, G.H.; van Loan, C.F. Matrix Computations, 3rd ed.; The John Hopkins University Press: Baltimore, MD, USA, 1996.
- 62. Gander, W.; Gander, M.J.; Kwok, F. Scientific Computing–An Introduction Using Maple and MATLAB; Springer: Berlin, Germany, 2014.
- 63. Digital Network Echo Cancellers; ITU-T Recommendations G.168; ITU: Geneva, Switzerland, 2002.
- 64. Bertsekas, D.P. Nonlinear Programming, 2nd ed.; Athena Scientific: Belmont, MA, USA, 1999.
- 65. Sondhi, M.M.; Morgan, D.R. Acoustic echo cancellation for stereophonic teleconferencing. In Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), New Paltz, NY, USA, 20–23 October 1991.
- 66. Sondhi, M.M.; Morgan, D.R.; Hall, J.L. Stereophonic acoustic echo cancellation–An overview of the fundamental problem. *IEEE Signal Process. Lett.* **1995**, *2*, 148–151. [CrossRef]
- 67. Benesty, J.; Morgan, D.R.; Sondhi, M.M. A better understanding and an improved solution to the specific problems of stereophonic acoustic echo cancellation. *IEEE Trans. Speech Audio Process.* **1998**, *6*, 156–165. [CrossRef]
- 68. Benesty, J.; Paleologu, C.; Gänsler, T.; Ciochină, S. *A Perspective on Stereophonic Acoustic Echo Cancellation*; Springer: Berlin, Germany, 2011.
- 69. Moore, B.C.J. An Introduction to the Psychology of Hearing; Academic Press: London, UK, 1989.
- 70. Romoli, L.; Cecchi, S.; Peretti, P.; Piazza, F. A mixed decorrelation approach for stereo acoustic echo cancellation based on the estimation of the fundamental frequency. *IEEE Trans. Audio Speech Lang. Process.* **2012**, *20*, 690–698. [CrossRef]
- 71. Schneider, M.; Kellermann, W. Multichannel acoustic echo cancellation in the wave domain with increased robustness to nonuniqueness. *IEEE/ACM Trans. Audio Speech Lang. Process.* 2016, 24, 518–529. [CrossRef]