

Communication

# Stochastic Model for the LMS Algorithm with Symmetric/Antisymmetric Properties

Augusto Cesar Becker <sup>1</sup>, Eduardo Vinicius Kuhn <sup>1</sup>, Marcos Vinicius Matsuo <sup>2</sup>, Jacob Benesty <sup>3</sup>,  
Constantin Paleologu <sup>4,\*</sup>, Laura-Maria Dogariu <sup>4</sup> and Silviu Ciochină <sup>4</sup>

<sup>1</sup> LAPSE—Electronics and Signal Processing Laboratory, Department of Electronics Engineering, Federal University of Technology-Paraná, Toledo 85902-490, PR, Brazil

<sup>2</sup> GEPS—Electronics and Signal Processing Group, Department of Control, Automation, and Computation, Federal University of Santa Catarina, Blumenau 89036-004, SC, Brazil

<sup>3</sup> National Institute of Scientific Research—Energy, Materials, and Telecommunications, University of Quebec, Montreal, QC H5A 1K6, Canada

<sup>4</sup> Department of Telecommunications, Faculty of Electronics, Telecommunications, and Information Technology, University Politehnica of Bucharest, 060042 Bucharest, Romania

\* Correspondence: pale@comm.pub.ro

**Abstract:** This paper presents a stochastic model for the least-mean-square algorithm with symmetric/antisymmetric properties (LMS-SAS), operating in a system identification setup with Gaussian input data. Specifically, model expressions are derived to describe the mean weight behavior of the (global and virtual) adaptive filters, learning curves, and evolution of some correlation-like matrices, which allow predicting the algorithm behavior. Simulation results are shown and discussed, confirming the accuracy of the proposed model for both transient and steady-state phases.

**Keywords:** adaptive filters; least-mean-square (LMS) algorithm; stochastic modeling; system identification; Kronecker product decomposition



**Citation:** Becker, A.C.; Kuhn, E.V.; Matsuo, M.V.; Benesty, J.; Paleologu, C.; Dogariu, L.-M.; Ciochină, S. Stochastic Model for the LMS Algorithm with Symmetric/Antisymmetric Properties. *Symmetry* **2022**, *14*, 1908. <https://doi.org/10.3390/sym14091908>

Academic Editor: Theodore E. Simos

Received: 17 August 2022

Accepted: 7 September 2022

Published: 12 September 2022

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



**Copyright:** © 2022 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/>).

## 1. Introduction

In several practical applications, adaptive filtering techniques are used to obtain an approximate representation for the input–output relationship of an unknown system in real-time, such as occurs in system identification problems, echo and noise cancellation, as well as in adaptive control [1–5]. In some of these applications, the system to be identified exhibits known special characteristics (e.g., sparseness, symmetry or antisymmetry, among others), which can be exploited by the adaptive algorithm to improve its convergence and/or reduce computational complexity. For instance, proportionate-type algorithms [6–11] exploit the sparseness of the system impulse response. In turn, algorithms operating with bilinear forms [12,13] consider a more efficient representation for characterizing multiple-input/single-output (MISO) spatiotemporal systems. While the least-mean-square with symmetric/antisymmetric properties (LMS-SAS) algorithm and its normalized version [14] exploit the intrinsic symmetric or antisymmetric characteristics of some systems to decompose their impulse responses into two smaller vectors using the Kronecker product. Therefore, considering the practical applicability, theoretical studies on the behavior of these algorithms become relevant.

In this context, stochastic models may serve as a theoretical basis for studying the behavior of adaptive algorithms without relying only on extensive Monte Carlo (MC) simulations [15–23]. Such models are useful to establish (through mathematical expressions) stability conditions for the algorithm and a suitable range of values for its parameters, as well as interesting cause-and-effect relationships between performance metrics and algorithm parameters, which can then help the designer [24–29]. Moreover, models can be used to identify undesirable behavior of the algorithm, leading thereby to the development

of enhanced algorithms [30,31]. However, despite the inherent importance, a stochastic model for the LMS-SAS algorithm [14] has not been discussed (to the best of our knowledge) in the literature so far. So, aiming to fill this gap, the present research work has the following goals:

- (i) To develop a stochastic model describing the behavior of the algorithm, considering (uncorrelated and correlated) Gaussian input data;
- (ii) To derive expressions characterizing the mean weight behavior of the adaptive filters, learning curves, and evolution of some correlation-like matrices; and
- (iii) To verify and discuss the accuracy of the model for different operating scenarios.

Note that performance comparisons with other algorithms from the literature are beyond the scope of this research work (for such, see [14]).

The remainder of the paper is organized as follows. Section 2 introduces the system model and revisits the LMS-SAS algorithm. Section 3 presents the proposed stochastic model. Section 4 provides simulation results for different operating scenarios. Section 5 presents a brief discussion about the results and highlights some algorithm characteristics. Finally, Section 6 summarizes conclusions and suggestions for future research works.

The mathematical notation adopted in this paper follows the standard practice of using lower-case boldface letters for vectors, upper-case boldface letters for matrices, and both italic Roman and Greek letters for scalar quantities. Superscripts  $T$  and  $B$  stand for the transpose and block-transpose operators [32],  $\otimes$  denotes the Kronecker product,  $\text{vec}(\cdot)$  characterizes the vectorization operator [33],  $\text{Tr}(\cdot)$  represents the trace of a matrix, and  $E(\cdot)$  is the expected value operator. Still, the notation  $[\mathbf{A}]_{(i,j)}$  represents the scalar  $(i, j)$ -entry of matrix  $\mathbf{A}$ , while  $\mathbf{A}_{i,j}$  denotes the  $i, j$ -block/submatrix of  $\mathbf{A}$ .

## 2. Problem Formulation

In a system identification setup (as depicted in Figure 1), let us consider a linear single-input/single-output (SISO) system whose output signal  $y(n)$  corrupted by an additive measurement noise  $w(n)$  at time index  $n$  yields the so-called desired signal which can be expressed as

$$\begin{aligned} d(n) &= y(n) + w(n) \\ &= \mathbf{h}^T(\mathbf{h}_1, \mathbf{h}_2)\mathbf{x}(n) + w(n) \end{aligned} \quad (1)$$

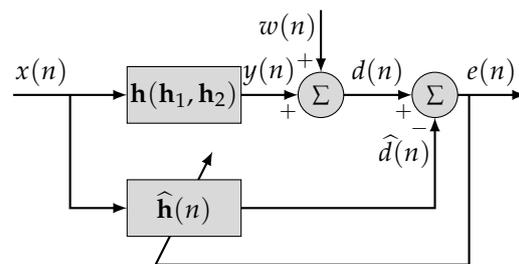
where  $\mathbf{x}(n) = [x(n) \ x(n-1) \ \dots \ x(n-L_0^2+1)]^T$  contains the  $L_0^2$  most recent samples of the input signal  $x(n)$ , while

$$\mathbf{h}(\mathbf{h}_1, \mathbf{h}_2) = \mathbf{h}_1 \otimes \mathbf{h}_2 \pm \mathbf{h}_2 \otimes \mathbf{h}_1 \quad (2)$$

denotes an  $L_0^2$ -dimensional (unknown) symmetric (+) or antisymmetric (−) system impulse response which depends on two (virtual) impulse responses  $\mathbf{h}_1$  and  $\mathbf{h}_2$  of length  $L_0$ . Note that the symmetric or antisymmetric characteristic considered here is not related to the samples of the impulse response such as in linear-phase finite impulse response filters (for details, see the discussion presented in [14]). So, the error signal can be written as

$$e(n) = d(n) - \hat{d}(n) \quad (3)$$

with  $\hat{d}(n)$  denoting the output signal of the adaptive filter  $\hat{\mathbf{h}}(n)$  whose weights are adjusted based on  $x(n)$  and  $e(n)$ .



**Figure 1.** Block diagram of a system identification setup.

In this context, instead of estimating  $\mathbf{h}(\mathbf{h}_1, \mathbf{h}_2)$  directly, the LMS-SAS algorithm [14] can be used to identify  $\mathbf{h}_1$  and  $\mathbf{h}_2$  [in (2)] through two (virtual) adaptive filters with weight vectors  $\hat{\mathbf{h}}_1(n)$  and  $\hat{\mathbf{h}}_2(n)$ , both of length  $L_0$ , in such a way that the (global) adaptive filter can be computed as

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}_1(n) \otimes \hat{\mathbf{h}}_2(n) \pm \hat{\mathbf{h}}_2(n) \otimes \hat{\mathbf{h}}_1(n). \quad (4)$$

Thus, the adjustment of  $\hat{\mathbf{h}}_1(n)$  and  $\hat{\mathbf{h}}_2(n)$  in (4) is carried out using the following update rules:

$$\hat{\mathbf{h}}_1(n) = \hat{\mathbf{h}}_1(n-1) + \mu_1 \mathbf{x}_2(n) e(n) \quad (5)$$

and

$$\hat{\mathbf{h}}_2(n) = \hat{\mathbf{h}}_2(n-1) + \mu_2 \mathbf{x}_1(n) e(n) \quad (6)$$

where  $\mu_1$  and  $\mu_2$  represent step-size parameters,

$$\mathbf{x}_1(n) = \hat{\mathbf{H}}_1^T(n-1) \mathbf{x}(n) \quad (7)$$

and

$$\mathbf{x}_2(n) = \hat{\mathbf{H}}_2^T(n-1) \mathbf{x}(n) \quad (8)$$

with

$$\hat{\mathbf{H}}_1(n) = \hat{\mathbf{h}}_1(n) \otimes \mathbf{I}_{L_0} \pm \mathbf{I}_{L_0} \otimes \hat{\mathbf{h}}_1(n) \quad (9)$$

and

$$\hat{\mathbf{H}}_2(n) = \mathbf{I}_{L_0} \otimes \hat{\mathbf{h}}_2(n) \pm \hat{\mathbf{h}}_2(n) \otimes \mathbf{I}_{L_0} \quad (10)$$

characterizing matrices of size  $L_0^2 \times L_0$ , in which  $\mathbf{I}_{L_0}$  denotes the identity matrix with dimension  $L_0 \times L_0$ , while the error signal (3) can conveniently be expressed as

$$\begin{aligned} e(n) &= d(n) - \hat{\mathbf{h}}_1^T(n-1) \mathbf{x}_2(n) \\ &= d(n) - \hat{\mathbf{h}}_2^T(n-1) \mathbf{x}_1(n). \end{aligned} \quad (11)$$

Therefore, the system identification setup and the LMS-SAS algorithm [defined through (4)–(11)] have now been characterized.

### 3. Proposed Model

In this section, the proposed stochastic model describing the behavior of the LMS-SAS algorithm is derived. Specifically, model expressions are obtained for predicting the mean weight behavior of the (global and virtual) adaptive filters, learning curves, and the evolution of some correlation-like matrices related to the adaptive weight vectors. To this end, the following assumptions and approximations are used:

**Assumption 1.** The input signal  $x(n)$  is obtained from a zero-mean (uncorrelated or correlated) Gaussian process with variance  $\sigma_x^2$  and autocorrelation matrix  $\mathbf{R} = \mathbb{E}[\mathbf{x}(n)\mathbf{x}^T(n)]$  [2,5].

**Assumption 2.** The measurement noise  $w(n)$  is obtained from a white Gaussian process with variance  $\sigma_w^2$ , which is uncorrelated with any other signal in the system [2,5].

**Assumption 3.** The adaptive weight vectors  $\hat{\mathbf{h}}_1(n)$ ,  $\hat{\mathbf{h}}_2(n)$ , and  $\hat{\mathbf{h}}(n)$  are assumed independent of each other as well as from any other variable in the system [2,5,13,29].

Note that these assumptions have been commonly used in the stochastic modeling of adaptive algorithms to make the development mathematically tractable (see [2,5]) and have led to satisfactory results (as shown later in Section 4).

### 3.1. Mean Weight Behavior

Taking into account the similar forms of (5) and (6), let us start replacing the subindexes 1 and 2 by  $\alpha$  and  $\bar{\alpha}$  (or vice-versa) such that both update rules can be expressed through

$$\hat{\mathbf{h}}_{\alpha}(n) = \hat{\mathbf{h}}_{\alpha}(n-1) + \mu_{\alpha} \mathbf{x}_{\bar{\alpha}}(n) e(n). \quad (12)$$

Then, substituting (1) and (11) into (12), taking the expected value of both sides of the resulting expression, and using Assumptions 2 and 3, the mean weight behavior of the (virtual) adaptive filters can be determined as

$$E[\hat{\mathbf{h}}_{\alpha}(n)] = [\mathbf{I}_{L_0} - \mu_{\alpha} \mathbf{S}_{\bar{\alpha}}(n)] E[\hat{\mathbf{h}}_{\alpha}(n-1)] + \mu_{\alpha} \mathbf{S}'_{\bar{\alpha}}(n) \mathbf{h} \quad (13)$$

where

$$\mathbf{S}_{\alpha}(n) = E[\mathbf{x}_{\alpha}(n) \mathbf{x}_{\alpha}^T(n)] \quad (14)$$

and

$$\mathbf{S}'_{\alpha}(n) = E[\mathbf{x}_{\alpha}(n) \mathbf{x}^T(n)]. \quad (15)$$

Notice from (13) that the convergence of the (virtual) adaptive filters depends on

- The step sizes  $\mu_{\alpha}$ , the system impulse response  $\mathbf{h}$ , the estimate of the system impulse response  $\hat{\mathbf{h}}_{\alpha}(n)$ , as well as the correlation-like matrices  $\mathbf{S}_{\alpha}(n)$  and  $\mathbf{S}'_{\alpha}(n)$ ;
- The initialization condition  $\hat{\mathbf{h}}_{\alpha}(0)$  which may impair the initial convergence or cause instability of the algorithm; and
- Each other, i.e., the behavior of one (virtual) adaptive filter is affected by the behavior of the other one and vice-versa.

Next, substituting either (7) or (8) into (14), using the properties of the Kronecker product [33,34], as well as considering Assumption 3, we have

$$\mathbf{S}_{\alpha}(n) = \mathbf{S}_{\alpha}^{(1)}(n) \pm \mathbf{S}_{\alpha}^{(2)}(n) \pm \mathbf{S}_{\alpha}^{(2)T}(n) + \mathbf{S}_{\alpha}^{(3)}(n) \quad (16)$$

with

$$\begin{aligned} \mathbf{S}_{\alpha}^{(1)}(n) &= E\{[\mathbf{I}_{L_0} \otimes \hat{\mathbf{h}}_{\alpha}(n-1)]^T \mathbf{R} [\mathbf{I}_{L_0} \otimes \hat{\mathbf{h}}_{\alpha}(n-1)]\} \\ &= \begin{bmatrix} \text{Tr}[\mathbf{R}_{1,1} \mathbf{G}_{\alpha}(n-1)] & \cdots & \text{Tr}[\mathbf{R}_{1,L_0} \mathbf{G}_{\alpha}(n-1)] \\ \vdots & \ddots & \vdots \\ \text{Tr}[\mathbf{R}_{L_0,1} \mathbf{G}_{\alpha}(n-1)] & \cdots & \text{Tr}[\mathbf{R}_{L_0,L_0} \mathbf{G}_{\alpha}(n-1)] \end{bmatrix} \end{aligned} \quad (17)$$

$$\begin{aligned} \mathbf{S}_{\alpha}^{(2)}(n) &= E\{[\mathbf{I}_{L_0} \otimes \hat{\mathbf{h}}_{\alpha}(n-1)]^T \mathbf{R} [\hat{\mathbf{h}}_{\alpha}(n-1) \otimes \mathbf{I}_{L_0}]\} \\ &= \begin{bmatrix} \sum_{i=1}^{L_0} \mathbf{g}_{\alpha,i}^T(n-1) \mathbf{R}_{1,i} \\ \vdots \\ \sum_{i=1}^{L_0} \mathbf{g}_{\alpha,i}^T(n-1) \mathbf{R}_{L_0,i} \end{bmatrix} \end{aligned} \quad (18)$$

and

$$\begin{aligned} \mathbf{S}_\alpha^{(3)}(n) &= E\{[\widehat{\mathbf{h}}_\alpha(n-1) \otimes \mathbf{I}_{L_0}]^T \mathbf{R} [\widehat{\mathbf{h}}_\alpha(n-1) \otimes \mathbf{I}_{L_0}]\} \\ &= \sum_{i=1}^{L_0} \sum_{j=1}^{L_0} [\mathbf{G}_\alpha(n-1)]_{(i,j)} \mathbf{R}_{i,j} \end{aligned} \quad (19)$$

in which

$$\mathbf{G}_\alpha(n) = E[\widehat{\mathbf{h}}_\alpha(n) \widehat{\mathbf{h}}_\alpha^T(n)] \quad (20)$$

is the autocorrelation matrix of  $\widehat{\mathbf{h}}_\alpha(n)$ ,  $\mathbf{g}_{\alpha,i}(n)$  represents the  $i$ -th column of  $\mathbf{G}_\alpha(n)$ ,  $[\mathbf{G}_\alpha(n)]_{(i,j)}$  characterizes the  $(i, j)$ -th element of  $\mathbf{G}_\alpha(n)$ , while  $\mathbf{R}_{i,j}$  denotes an  $L_0 \times L_0$ -block/submatrix of  $\mathbf{R}$ , i.e.,

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_{1,1} & \cdots & \mathbf{R}_{1,L_0} \\ \vdots & \ddots & \vdots \\ \mathbf{R}_{L_0,1} & \cdots & \mathbf{R}_{L_0,L_0} \end{bmatrix}. \quad (21)$$

Similarly, from (7) and (8), (15) can be simplified as

$$\mathbf{S}'_\alpha(n) = E[\widehat{\mathbf{H}}_\alpha^T(n-1)] \mathbf{R} \quad (22)$$

with  $E[\widehat{\mathbf{H}}_\alpha^T(n-1)]$  being obtained by taking the expected value of both sides of either (9) or (10).

Finally, the mean weight behavior of the (global) adaptive filter  $\widehat{\mathbf{h}}(n)$  can be determined by applying the expected value to both sides of (4) and using Assumption 3. Thereby, one obtains

$$E[\widehat{\mathbf{h}}(n)] = E[\widehat{\mathbf{h}}_1(n)] \otimes E[\widehat{\mathbf{h}}_2(n)] \pm E[\widehat{\mathbf{h}}_2(n)] \otimes E[\widehat{\mathbf{h}}_1(n)] \quad (23)$$

which depends on (13); hence, the characteristics observed in the convergence of the (virtual) adaptive filters hold also for the convergence of the (global) adaptive filter. Therefore, the mean weight behavior of  $\widehat{\mathbf{h}}_1(n)$ ,  $\widehat{\mathbf{h}}_2(n)$ , and  $\widehat{\mathbf{h}}(n)$  can be predicted [from (13) and (23)] if the evolution of (20) is known.

### 3.2. Learning Curves

Aiming to characterize the learning curves, let us start by rewriting (11) using the weight-error vector

$$\mathbf{v}(n) = \mathbf{h} - \widehat{\mathbf{h}}(n) \quad (24)$$

as

$$e(n) = \mathbf{v}^T(n-1) \mathbf{x}(n) + w(n). \quad (25)$$

Then, squaring both sides of (25), taking the expected value of the resulting expression, and using Assumption 2, one obtains the following expression describing the evolution of the mean-square error (MSE):

$$J(n) = J_{\min} + J_{\text{ex}}(n) \quad (26)$$

where

$$J_{\min} = \sigma_w^2 \quad (27)$$

is the minimum MSE attainable in steady state, while

$$J_{\text{ex}}(n) = E[\mathbf{v}^T(n-1) \mathbf{x}(n) \mathbf{x}^T(n) \mathbf{v}(n-1)] \quad (28)$$

is the excess MSE (EMSE) introduced by the algorithm. Next, considering the properties of the trace operator [33] and using Assumption 3, (28) reduces to

$$J_{\text{ex}}(n) = \text{Tr}[\mathbf{R} \mathbf{K}(n-1)] \quad (29)$$

with

$$\mathbf{K}(n) = E[\mathbf{v}(n) \mathbf{v}^T(n)] \quad (30)$$

being the autocorrelation matrix of the weight-error vector. Therefore, if the evolution of (30) is known, the MSE and EMSE learning curves are completely characterized by (26), (27), and (29).

### 3.3. Correlation-like Matrices

The aim is now to derive recursive expressions describing the evolution of the correlation-like matrices  $\mathbf{G}_\alpha(n)$  [given by (20)] and  $\mathbf{K}(n)$  [given by (30)]. To this end, substituting (12) into (20), and using Assumptions 2 and 3, we obtain

$$\begin{aligned} \mathbf{G}_\alpha(n) &= \mathbf{G}_\alpha(n-1) - \mu_\alpha \mathbf{S}_{\bar{\alpha}}(n) \mathbf{G}_\alpha(n-1) \\ &\quad - \mu_\alpha \mathbf{G}_\alpha(n-1) \mathbf{S}_{\bar{\alpha}}(n) + \mu_\alpha \mathbf{G}'_\alpha(n-1) \mathbf{S}'_{\bar{\alpha}}(n) \\ &\quad + \mu_\alpha \mathbf{S}'_{\bar{\alpha}}(n) \mathbf{G}'_\alpha(n-1) + \mu_\alpha^2 \mathbf{F}_{\bar{\alpha}}(n) + \mu_\alpha^2 \sigma_w^2 \mathbf{S}_{\bar{\alpha}}(n) \end{aligned} \quad (31)$$

where

$$\mathbf{G}'_\alpha(n) = \mathbf{E}[\widehat{\mathbf{h}}_\alpha(n)] \mathbf{h}^T \quad (32)$$

with  $\mathbf{E}[\widehat{\mathbf{h}}_\alpha(n)]$  being computed from (13),

$$\begin{aligned} \mathbf{F}_\alpha(n) &= \mathbf{E}[\widehat{\mathbf{H}}_\alpha^T(n-1) \mathbf{x}(n) \mathbf{x}^T(n) \mathbf{v}(n-1) \mathbf{v}^T(n-1) \mathbf{x}(n) \mathbf{x}^T(n) \widehat{\mathbf{H}}_\alpha(n-1)] \\ &\cong \mathbf{E}[\widehat{\mathbf{H}}_\alpha^T(n-1) \mathbf{E}(n) \widehat{\mathbf{H}}_\alpha(n-1)] \end{aligned} \quad (33)$$

in which

$$\begin{aligned} \mathbf{E}(n) &= \mathbf{E}[\mathbf{x}(n) \mathbf{x}^T(n) \mathbf{K}(n-1) \mathbf{x}(n) \mathbf{x}^T(n)] \\ &= 2\mathbf{R}\mathbf{K}(n-1)\mathbf{R} + \mathbf{R}\text{Tr}[\mathbf{R}\mathbf{K}(n-1)] \end{aligned} \quad (34)$$

due to the factorization theorem of Gaussian variables [2,3,5] (also known as Isserlis' theorem [35]). Thereby, such as in (16), (33) can be expressed [using (34)] as

$$\mathbf{F}_\alpha(n) = \mathbf{F}_\alpha^{(1)}(n) \pm \mathbf{F}_\alpha^{(2)}(n) \pm \mathbf{F}_\alpha^{(2)T}(n) + \mathbf{F}_\alpha^{(3)}(n) \quad (35)$$

where

$$\begin{aligned} \mathbf{F}_\alpha^{(1)}(n) &= \mathbf{E}\{[\mathbf{I}_{L_0} \otimes \widehat{\mathbf{h}}_\alpha(n-1)]^T \mathbf{E}(n) [\mathbf{I}_{L_0} \otimes \widehat{\mathbf{h}}_\alpha(n-1)]\} \\ &= \begin{bmatrix} \text{Tr}[\mathbf{E}_{1,1}(n) \mathbf{G}_\alpha(n-1)] & \cdots & \text{Tr}[\mathbf{E}_{1,L_0}(n) \mathbf{G}_\alpha(n-1)] \\ \vdots & \ddots & \vdots \\ \text{Tr}[\mathbf{E}_{L_0,1}(n) \mathbf{G}_\alpha(n-1)] & \cdots & \text{Tr}[\mathbf{E}_{L_0,L_0}(n) \mathbf{G}_\alpha(n-1)] \end{bmatrix} \end{aligned} \quad (36)$$

$$\begin{aligned} \mathbf{F}_\alpha^{(2)}(n) &= \mathbf{E}\{[\mathbf{I}_{L_0} \otimes \widehat{\mathbf{h}}_\alpha(n-1)]^T \mathbf{E}(n) [\widehat{\mathbf{h}}_\alpha(n-1) \otimes \mathbf{I}_{L_0}]\} \\ &= \begin{bmatrix} \sum_{i=1}^{L_0} \mathbf{g}_{\alpha,i}^T(n-1) \mathbf{E}_{1,i}(n) \\ \vdots \\ \sum_{i=1}^{L_0} \mathbf{g}_{\alpha,i}^T(n-1) \mathbf{E}_{L_0,i}(n) \end{bmatrix} \end{aligned} \quad (37)$$

and

$$\begin{aligned} \mathbf{F}_\alpha^{(3)}(n) &= \mathbf{E}\{[\widehat{\mathbf{h}}_\alpha(n-1) \otimes \mathbf{I}_{L_0}]\}^T \mathbf{E}(n) [\widehat{\mathbf{h}}_\alpha(n-1) \otimes \mathbf{I}_{L_0}] \\ &= \sum_{i=1}^{L_0} \sum_{j=1}^{L_0} [\mathbf{G}_\alpha(n-1)]_{(i,j)} \mathbf{E}_{i,j}(n). \end{aligned} \quad (38)$$

In turn, calculating  $\mathbf{v}(n) \mathbf{v}^T(n)$  from (24) and taking the expected value of both sides of the resulting expression, it is possible to rewrite (30) as

$$\mathbf{K}(n) = \mathbf{h} \mathbf{h}^T - \mathbf{h} \mathbf{E}[\widehat{\mathbf{h}}^T(n)] - \mathbf{E}[\widehat{\mathbf{h}}(n)] \mathbf{h}^T + \mathbf{G}(n) \quad (39)$$

where

$$\begin{aligned}\mathbf{G}(n) &= \mathbb{E}[\widehat{\mathbf{h}}(n)\widehat{\mathbf{h}}^T(n)] \\ &= \mathbf{G}_1(n) \otimes \mathbf{G}_2(n) \pm \{\text{vec}[\mathbf{G}_2(n)]\text{vec}[\mathbf{G}_1(n)]^T\}^B \\ &\quad \pm \{\text{vec}[\mathbf{G}_1(n)]\text{vec}[\mathbf{G}_2(n)]^T\}^B + \mathbf{G}_2(n) \otimes \mathbf{G}_1(n)\end{aligned}\quad (40)$$

denotes the autocorrelation matrix of  $\widehat{\mathbf{h}}(n)$  and  $(\cdot)^B$  represents the transposition of blocks [32] of dimension  $L_0 \times L_0$ . Note that (40) is obtained by calculating the outer product  $\widehat{\mathbf{h}}(n)\widehat{\mathbf{h}}^T(n)$  from (4), taking the expected value of both sides of the resulting expression, and considering Assumption 3.

Therefore, since the evolution of the correlation-like matrices  $\mathbf{G}_\alpha(n)$ ,  $\mathbf{K}(n)$ ,  $\mathbf{G}(n)$  has been properly characterized, the behavior of the LMS-SAS algorithm can now be predicted.

#### 4. Simulation Results

This section aims to assess the accuracy of the proposed model by comparing results obtained from MC simulations (average of 200 independent runs) with model predictions. To this end, three examples are presented, covering uncorrelated and correlated Gaussian input data, several signal-to-noise ratio (SNR) values, as well as distinct system impulse responses and initialization conditions for the adaptive filters. Specifically, the input signal  $x(n)$  is obtained from an autoregressive (AR) process [5], given by

$$x(n) = -a_1x(n-1) - a_2x(n-2) + v(n) \quad (41)$$

in which  $a_1$  and  $a_2$  denote the AR(2) coefficients, while  $v(n)$  is a white Gaussian noise whose variance is determined as

$$\sigma_v^2 = \sigma_x^2 \left( \frac{1-a_2}{1+a_2} \right) [(1+a_2)^2 - a_1^2] \quad (42)$$

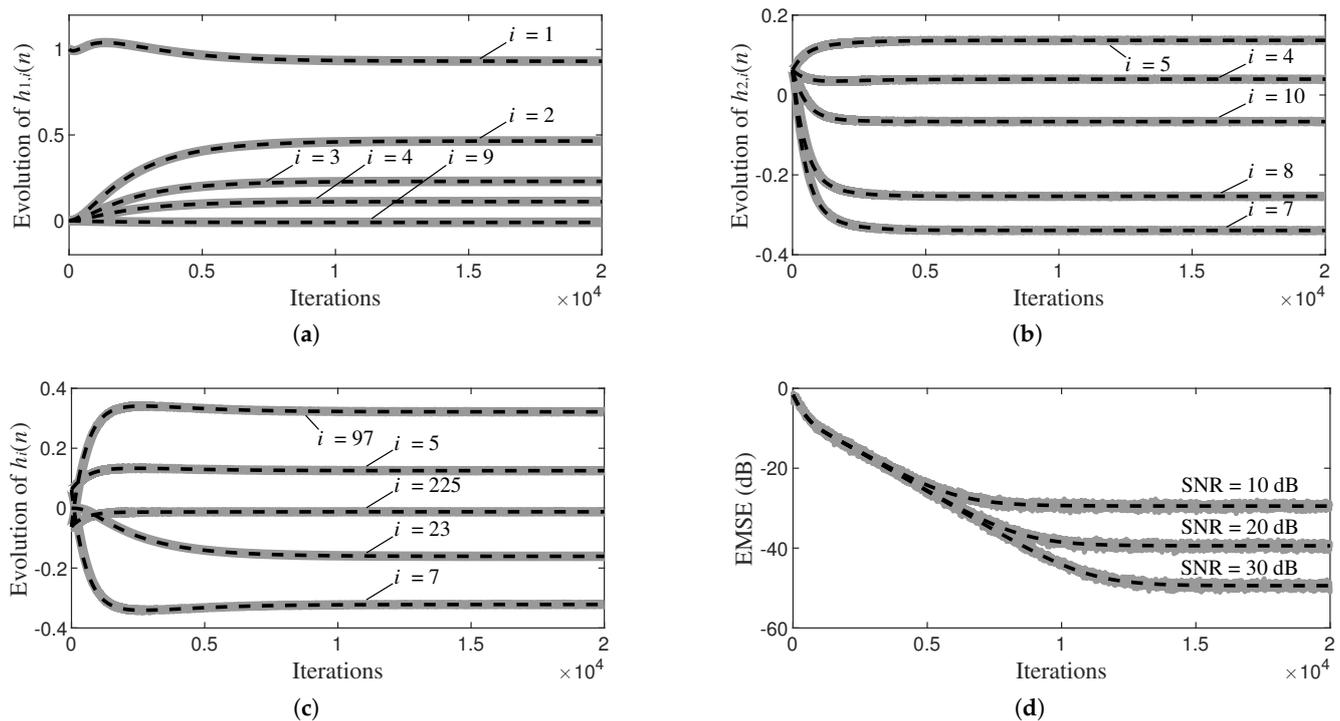
such that  $\sigma_x^2 = 1$ . The SNR is defined (in dB) as [14,29]

$$\text{SNR} = 10 \log_{10} \left( \frac{\sigma_y^2}{\sigma_w^2} \right) \quad (43)$$

with  $\sigma_y^2 = \mathbf{h}^T(\mathbf{h}_1, \mathbf{h}_2)\mathbf{R}\mathbf{h}(\mathbf{h}_1, \mathbf{h}_2)$  characterizing the variance of the system output signal; in particular, three values of SNR are considered here, i.e., 10, 20, and 30 dB. The (virtual) adaptive filters 1 and 2 are initialized, respectively, as  $\widehat{\mathbf{h}}_1(0) = [1 \ 0 \ \dots \ 0]^T$  and  $\widehat{\mathbf{h}}_2(0) = L_0^{-1}[1 \ 1 \ \dots \ 1]^T$  unless otherwise stated to prevent them from stalling at the beginning of the adaptation process [14], while the step sizes are chosen as  $\mu_1 = \mu_2 = 10^{-3}$ .

##### 4.1. Example 1

Here, the proposed model is verified for uncorrelated Gaussian input data, different SNR values, and a system impulse response with antisymmetric characteristic. Specifically, the input signal is obtained from (41) by making  $a_1 = a_2 = 0$ , which results in an eigenvalue spread [2,3,5] for the input autocorrelation matrix of  $\chi = 1$ . The antisymmetric system impulse response  $\mathbf{h}(\mathbf{h}_1, \mathbf{h}_2)$  is obtained from (2), with  $\mathbf{h}_1$  containing the first  $L_0 = 16$  weights of the echo path model 1 given in the ITU-T G.168 Recommendation [36] while  $[\mathbf{h}_2]_l = (-0.9)^{l-1}$  for  $l = 1, 2, \dots, L_0$ ; consequently, the length of  $\mathbf{h}(\mathbf{h}_1, \mathbf{h}_2)$  is  $L_0^2 = 256$ . The results obtained for this operating scenario are presented in Figure 2.



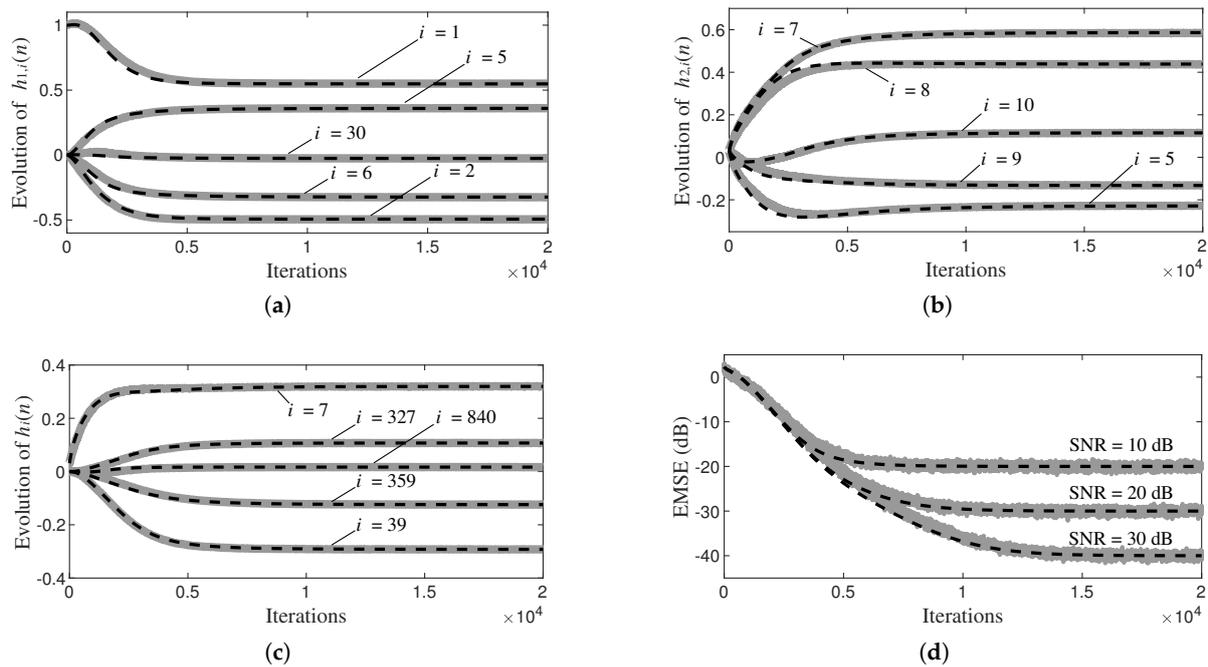
**Figure 2.** Example 1. Results obtained from MC simulations (gray-ragged lines) and predicted from the proposed model (dark-dashed lines): (a) Evolution of (five) weights of the (virtual) adaptive filter 1; (b) Evolution of (five) weights of the (virtual) adaptive filter 2; (c) Evolution of (five) weights of the (global) adaptive filter; (d) Evolution of the EMSE learning curves.

#### 4.2. Example 2

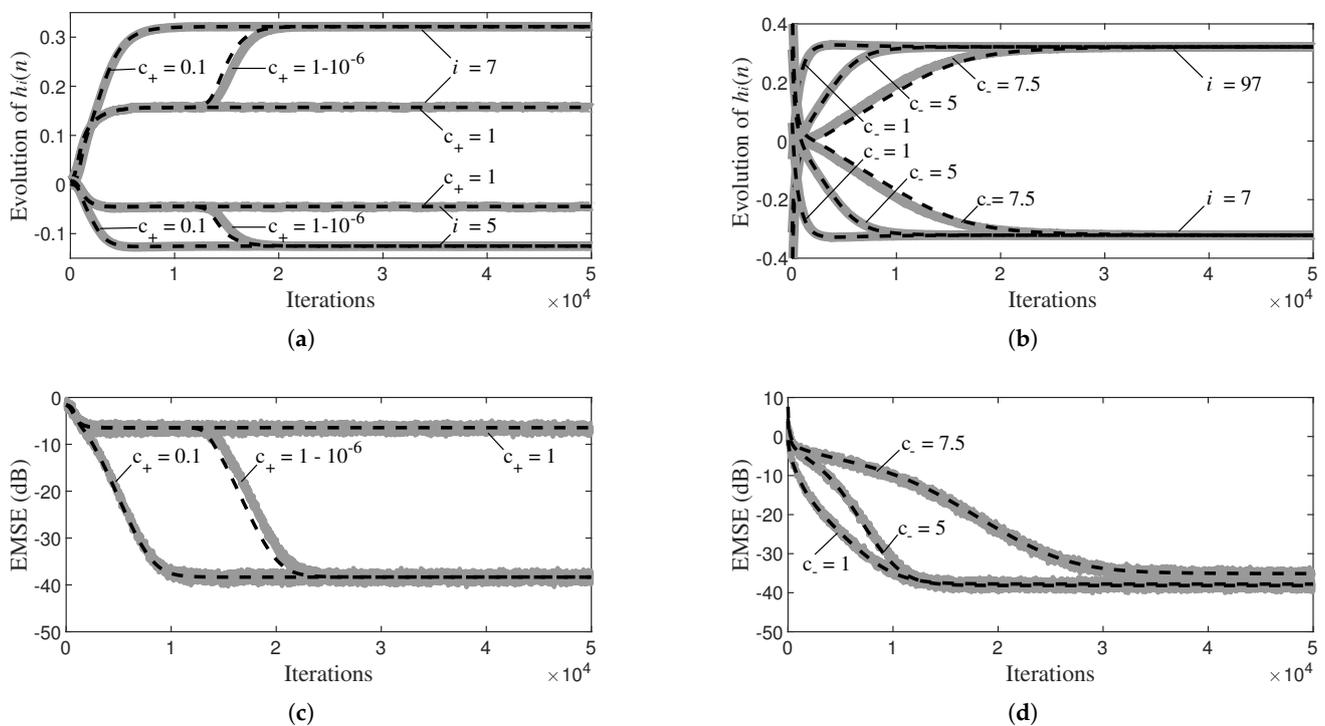
This example aims now to assess the proposed model for correlated Gaussian input data, different SNR values, and a system impulse response with symmetric characteristic. Particularly, the input signal is taken from (41) with  $a_1 = -0.6$  and  $a_2 = 0.8$ , which results in an eigenvalue spread of  $\chi = 162.13$  for the input autocorrelation matrix. The symmetric system impulse response  $\mathbf{h}(\mathbf{h}_1, \mathbf{h}_2)$  is obtained from (2), with  $\mathbf{h}_1$  containing the first  $L_0 = 32$  weights of the echo path model 1 given in the ITU-T G.168 Recommendation [36] while  $[\mathbf{h}_2]_l = 0.5^{l-1}$  for  $l = 1, 2, \dots, L_0$ ; consequently, the length of  $\mathbf{h}(\mathbf{h}_1, \mathbf{h}_2)$  is  $L_0^2 = 1024$ . The results obtained for this operating scenario are presented in Figure 3.

#### 4.3. Example 3

In this example, the accuracy of the proposed model is verified considering different initialization conditions for the adaptive weight vectors. To this end, the input signal is obtained from (41) by making  $a_1 = -0.6$  and  $a_2 = 0.8$ , yielding an eigenvalue spread of  $\chi = 160.55$  for the input autocorrelation matrix. Symmetric and antisymmetric system impulse responses are used, which are determined from (2) with  $\mathbf{h}_1$  and  $\mathbf{h}_2$  given as in Example 2 but now with  $L_0 = 16$  weights; hence, the length of  $\mathbf{h}(\mathbf{h}_1, \mathbf{h}_2)$  becomes  $L_0^2 = 256$ . In turn, the adaptive weight vectors are initialized as  $\hat{\mathbf{h}}_1(0) = L_0^{-1}[1 \ 1 \ \dots \ 1]^T$  and  $\hat{\mathbf{h}}_2(0) = c_+ L_0^{-1}[1 \ 1 \ \dots \ 1]^T$  with  $c_+ = \{10^{-1}, 1 - 10^{-6}, 1\}$  when the system impulse response is symmetric, while  $\hat{\mathbf{h}}_1(0) = [1 \ 0 \ \dots \ 0]^T$  and  $\hat{\mathbf{h}}_2(0) = c_- L_0^{-1}[1 \ 1 \ \dots \ 1]^T$  with  $c_- = \{1, 5, 7.5\}$  when the antisymmetric impulse response is used. The obtained results are presented in Figure 4, considering only SNR = 20 dB for simplicity.



**Figure 3.** Example 2. Results obtained from MC simulations (gray-ragged lines) and predicted from the proposed model (dark-dashed lines): (a) Evolution of (five) weights of the (virtual) adaptive filter 1; (b) Evolution of (five) weights of the (virtual) adaptive filter 2; (c) Evolution of (five) weights of the (global) adaptive filter; (d) Evolution of the EMSE learning curves.



**Figure 4.** Example 3. Results obtained from MC simulations (gray-ragged lines) and predicted from the proposed model (dark-dashed lines): (a) Evolution of (two) weights of the (global) adaptive filter for the symmetric impulse response; (b) Evolution of (two) weights of the (global) adaptive filter for the antisymmetric impulse response; (c) Evolution of the EMSE learning curves for the symmetric impulse response; (d) Evolution of the EMSE learning curves for the antisymmetric impulse response.

## 5. Discussion

Figures 2–4 present the results obtained from MC simulations and model predictions for the operating scenarios described in Examples 1, 2, and 3, respectively. Specifically, Figures 2a–c and 3a–c depict the evolution of (five) adaptive weights of  $\hat{\mathbf{h}}_1(n)$ ,  $\hat{\mathbf{h}}_2(n)$ , and  $\hat{\mathbf{h}}(n)$ , while Figures 2d and 3d show EMSE learning curves. [Results obtained for SNR values of 10 and 30 dB have been omitted in Figures 2a–c and 3a–c since they are very similar.] In turn, Figure 4a,b illustrate the impact of the initialization of the adaptive weight vectors on the evolution of (two) adaptive weights of  $\hat{\mathbf{h}}(n)$ , while Figure 4c,d depict the effect on the EMSE learning curves, from which one verifies that the convergence speed and the steady-state EMSE are affected by the choice of the initialization vectors of the adaptive filters; in this particular case, when  $\hat{\mathbf{h}}_1(0)$  and  $\hat{\mathbf{h}}_2(0)$  assume small and distinct values, improved convergence and lower steady-state EMSE are achieved. Despite that, one observes from these figures that the behavior predicted from the proposed model matches very well the one obtained from MC simulations, during both transient and steady-state phases. Therefore, the accuracy of the model is confirmed for uncorrelated and correlated Gaussian input data, different SNR values and initialization conditions for the adaptive weight vectors, as well as for system impulse responses with symmetric and antisymmetric characteristics.

## 6. Conclusions

A stochastic model for the LMS-SAS algorithm was derived in this paper. The proposed model allows predicting the mean weight behavior of the (global and virtual) adaptive filters, learning curves, and evolution of some correlation-like matrices. Simulation results confirmed the accuracy of the model irrespective of the input data correlation level, SNR value, initialization condition of the adaptive weight vectors, and system impulse response. Based on the proposed model, further research could address the derivation of stability bounds for the step size, analytical expressions characterizing the algorithm behavior in steady state, and the development of models for the normalized version of the LMS-SAS algorithm.

**Author Contributions:** Conceptualization, A.C.B. and E.V.K.; methodology, A.C.B. and M.V.M.; software, A.C.B.; validation, M.V.M.; investigation, A.C.B.; writing—original draft preparation, A.C.B., E.V.K., and M.V.M.; writing—review and editing, J.B., C.P., and L.-M.D.; supervision, J.B. and S.C.; funding acquisition, L.-M.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by a grant from the Romanian Ministry of Education and Research, CNCS-UEFISCDI, project no. PN-III-P1-1.1-PD-2019-0340.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data and source codes related to the present research work can be made available upon request to the corresponding author.

**Acknowledgments:** The authors thank the Academic Editor and the anonymous reviewers for their valuable and constructive comments.

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

## References

1. Åström, K.J.; Wittenmark, B. *Adaptive Control*, 2nd ed.; Dover Publications: Mineola, NY, USA, 2008.
2. Sayed, A.H. *Adaptive Filters*; John Wiley & Sons: Hoboken, NJ, USA, 2008.
3. Diniz, P.S.R. *Adaptive Filtering: Algorithms and Practical Implementation*, 4th ed.; Springer-Verlag: New York, NY, USA, 2013.
4. Farhang-Boroujeny, B. *Adaptive Filters: Theory and Applications*, 2nd ed.; John Wiley & Sons: Chichester, UK, 2013.
5. Haykin, S. *Adaptive Filter Theory*, 5th ed.; Prentice Hall: Upper Saddle River, NJ, USA, 2014.
6. Duttweiler, D.L. Proportionate normalized least-mean-squares adaptation in echo cancelers. *IEEE Trans. Speech Audio Process.* **2000**, *8*, 508–518. [[CrossRef](#)]

7. de Souza, F.C.; Tobias, O.J.; Seara, R.; Morgan, D.R. A PNLMS algorithm with individual activation factors. *IEEE Trans. Signal Process.* **2010**, *58*, 2036–2047. [[CrossRef](#)]
8. Wagner, K.T.; Doroslovacki, M.I. *Proportionate-Type Normalized Least Mean Square Algorithms*; ISTE Ltd.: London, UK, 2013.
9. Perez, F.L.; de Souza, F.C.; Seara, R. An improved mean-square weight deviation-proportionate gain algorithm based on error autocorrelation. *Signal Process.* **2014**, *94*, 503–513. [[CrossRef](#)]
10. Beck, E.; Batista, E.L.O.; Seara, R. Norm-constrained adaptive algorithms for sparse system identification based on projections onto intersections of hyperplanes. *Signal Process.* **2016**, *118*, 259–271. [[CrossRef](#)]
11. Perez, F.L.; Kuhn, E.V.; de Souza, F.C.; Seara, R. A novel gain distribution policy based on individual-coefficient convergence for PNLMS-type algorithms. *Signal Process.* **2017**, *138*, 294–306. [[CrossRef](#)]
12. Benesty, J.; Paleologu, C.; Ciochină, S. On the identification of bilinear forms with the Wiener filter. *IEEE Signal Process. Lett.* **2017**, *24*, 653–657. [[CrossRef](#)]
13. Paleologu, C.; Benesty, J.; Ciochină, S. Adaptive filtering for the identification of bilinear forms. *Digit. Signal Process.* **2018**, *75*, 153–167. [[CrossRef](#)]
14. Benesty, J.; Paleologu, C.; Ciochină, S.; Kuhn, E.V.; Bakri, K.J.; Seara, R. LMS and NLMS algorithms for the identification of impulse responses with intrinsic symmetric or antisymmetric properties. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Singapore, 23–27 May 2022; pp. 5662–5666.
15. Bershad, N. Analysis of the normalized LMS algorithm with Gaussian inputs. *IEEE Trans. Acoust. Speech Signal Process.* **1986**, *34*, 793–806. [[CrossRef](#)]
16. Rupp, M. The behavior of LMS and NLMS algorithms in the presence of spherically invariant processes. *IEEE Trans. Signal Process.* **1993**, *41*, 1149–1160. [[CrossRef](#)]
17. Lopes, C.G.; Sayed, A.H. Diffusion least-mean squares over adaptive networks: Formulation and performance analysis. *IEEE Trans. Signal Process.* **2008**, *56*, 3122–3136. [[CrossRef](#)]
18. Abadi, M.; Husøy, J. On the application of a unified adaptive filter theory in the performance prediction of adaptive filter algorithms. *Digit. Signal Process.* **2009**, *19*, 410–432. [[CrossRef](#)]
19. Takahashi, N.; Yamada, I.; Sayed, A.H. Diffusion least-mean squares with adaptive combiners: Formulation and performance analysis. *IEEE Trans. Signal Process.* **2010**, *58*, 4795–4810. [[CrossRef](#)]
20. Parreira, W.D.; Bermudez, J.C.M.; Richard, C.; Tourneret, J.Y. Stochastic behavior analysis of the Gaussian kernel least-mean-square algorithm. *IEEE Trans. Signal Process.* **2012**, *60*, 2208–2222. [[CrossRef](#)]
21. Nosrati, H.; Shamsi, M.; Taheri, S.M.; Sedaaghi, M.H. Adaptive networks under non-stationary conditions: Formulation, performance analysis, and application. *IEEE Trans. Signal Process.* **2015**, *63*, 4300–4314. [[CrossRef](#)]
22. Yang, F.; Enzner, G.; Yang, J. A unified approach to the statistical convergence analysis of frequency-domain adaptive filters. *IEEE Trans. Signal Process.* **2019**, *67*, 1785–1796. [[CrossRef](#)]
23. Eweda, E.; Bershad, N.J.; Bermudez, J.C.M. Stochastic analysis of the diffusion least mean square and normalized least mean square algorithms for cyclostationary white Gaussian and non-Gaussian inputs. *Int. J. Adapt. Control. Signal Process.* **2021**, *35*, 2466–2486. [[CrossRef](#)]
24. de Almeida, S.; Bermudez, J.; Bershad, N.; Costa, M. A statistical analysis of the affine projection algorithm for unity step size and autoregressive inputs. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2005**, *52*, 1394–1405. [[CrossRef](#)]
25. Eweda, E. A new approach for analyzing the limiting behavior of the normalized LMS algorithm under weak assumptions. *Signal Process.* **2009**, *89*, 2143–2151. [[CrossRef](#)]
26. Batista, E.; Seara, R. On the performance of adaptive pruned Volterra filters. *Signal Process.* **2013**, *93*, 1909–1920. [[CrossRef](#)]
27. Matsuo, M.V.; Kuhn, E.V.; Seara, R. Stochastic analysis of the NLMS algorithm for nonstationary environment and deficient length adaptive filter. *Signal Process.* **2019**, *160*, 190–201. [[CrossRef](#)]
28. Matsuo, M.V.; Kuhn, E.V.; Seara, R. On the diffusion NLMS algorithm applied to adaptive networks: Stochastic modeling and performance comparisons. *Digit. Signal Process.* **2021**, *113*, 103018. [[CrossRef](#)]
29. Bakri, K.J.; Kuhn, E.V.; Seara, R.; Benesty, J.; Paleologu, C.; Ciochină, S. On the stochastic modeling of the LMS algorithm operating with bilinear forms. *Digit. Signal Process.* **2022**, *122*, 103359. [[CrossRef](#)]
30. Kolodziej, J.E.; Tobias, O.J.; Seara, R.; Morgan, D.R. On the constrained stochastic gradient algorithm: Model, performance, and improved version. *IEEE Trans. Signal Process.* **2009**, *57*, 1304–1315. [[CrossRef](#)]
31. Yu, Y.; Zhao, H.; Lu, L. Steady-state behavior of the improved normalized subband adaptive filter algorithm and its improvement in under-modeling. *Signal Image Video Process.* **2018**, *12*, 617–624. [[CrossRef](#)]
32. Mackey, D.S. Structured Linearizations for Matrix Polynomials. Ph.D. Thesis, University of Manchester, Manchester, UK, 2006.
33. Bernstein, D.S. *Matrix Mathematics: Theory, Facts, and Formulas*; Princeton University Press: Princeton, NJ, USA, 2009.
34. Van Loan, C.F. The ubiquitous Kronecker product. *J. Comput. Appl. Math.* **2000**, *123*, 85–100. [[CrossRef](#)]
35. Isserlis, L. On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika* **1918**, *12*, 134–139. [[CrossRef](#)]
36. ITU-T Recommendation G.168—Digital Network Echo Cancellers; Switzerland, International Telecommunications Union—Telecommunication Standardization Sector: Geneva, Switzerland, April 2015.