





# On the Stochastic Modeling of the NSVR–IAF–PNLMS Algorithm for Correlated Gaussian Input Data

Eduardo Vinicius Kuhn<sup>1,2</sup>  · Fábio Luis Perez<sup>3</sup> · Francisco das Chagas de Souza<sup>4</sup> · Rui Seara<sup>1</sup> 

Received: 21 February 2019 / Accepted: 26 November 2019 / Published online: 2 January 2020  
© Brazilian Society for Automatics--SBA 2020

## Abstract

This paper presents a stochastic model for the normalized smoothed variation rate individual-activation-factor proportionate normalized least-mean-square (NSVR–IAF–PNLMS) algorithm. Specifically, taking into account correlated Gaussian input data, model expressions are derived for predicting the mean weight vector, gain distribution matrix, NSVR metric, learning curve, weight-error correlation matrix, and steady-state excess mean-square error. Such expressions are obtained by considering the time-varying characteristics of the gain distribution matrix. Simulation results are shown confirming the accuracy of the proposed model for different operating conditions.

**Keywords** Adaptive filtering · Proportionate normalized least-mean-square algorithm · Stochastic model

## 1 Introduction

Although the least-mean-square (LMS) and the normalized LMS (NLMS) algorithms have been commonly used in many real-world applications (Sayed 2009; Farhang-Boroujeny 2013; Haykin 2014), these algorithms exhibit slow conver-

gence and poor tracking capability when the plant impulse response is sparse (as compared with other algorithms designed to exploit the sparse nature of the plant) (Martin et al. 2002; Paleologu et al. 2010; Wagner and Doroslovački 2013). For instance, Duttweiler (2000) introduced the proportionate NLMS (PNLMS) algorithm, in which each adaptive coefficient is updated proportionally to its magnitude in order to improve the algorithm convergence speed. Nevertheless, it has been observed that the PNLMS algorithm exhibits slow convergence for plants with low and medium sparseness (Benesty and Gay 2002), and improved convergence characteristics are not preserved over the whole adaptation process (Deng and Doroslovački 2006). So, to overcome the aforementioned problems, several PNLMS-type algorithms have been discussed in the open literature, such as the improved PNLMS (IPNLMS) (Benesty and Gay 2002), the  $\mu$ -law PNLMS (MPNLMS) (Deng and Doroslovački 2006), the sparseness-controlled PNLMS (SC-PNLMS) (Loganathan et al. 2008), and the individual-activation-factor PNLMS (IAF–PNLMS) (de Souza et al. 2010a).

In particular, the IAF–PNLMS algorithm (de Souza et al. 2010a) presents faster convergence in comparison with other algorithms from the literature for plants with high sparseness. Such an improvement in the convergence speed is due to the use of proportional gain for both active and inactive adaptive coefficients. However, given that gains assigned to active coefficients are maintained high even after achieving the vicinity of their optimal values, inactive

✉ Rui Seara  
seara@linse.ufsc.br

Eduardo Vinicius Kuhn  
kuhn@linse.ufsc.br; kuhn@utfpr.edu.br

Fábio Luis Perez  
fabiotek@furb.br

Francisco das Chagas de Souza  
francisco.souza@ufma.br

<sup>1</sup> LINSE: Circuits and Signal Processing Laboratory, Department of Electrical and Electronics Engineering, Federal University of Santa Catarina, Florianópolis, Santa Catarina 88040-900, Brazil

<sup>2</sup> LAPSE: Signal Processing and Electronics Laboratory, Department of Electronics Engineering, Federal University of Technology - Paraná, Toledo, Paraná 85902-490, Brazil

<sup>3</sup> ASL: Adaptive Filtering and Control Systems Laboratory, Department of Telecommunications and Electrical Engineering, University of Blumenau, Blumenau, Santa Catarina 89030-080, Brazil

<sup>4</sup> LSAPS: Adaptive Systems and Signal Processing Laboratory, Department of Electrical Engineering, Federal University of Maranhão, São Luís, Maranhão 65080-805, Brazil

coefficients receive very small gain during the whole adaptation process. As a consequence, the fast initial convergence is significantly compromised after the convergence of the active coefficients (de Souza et al. 2012). So, since assigning larger gains to coefficients that reached the vicinity of their optimal values have almost no effect on the convergence of the adaptive filter, Perez et al. (2017) have devised the normalized smoothed variation rate IAF–PNLMS (NSVR–IAF–PNLMS) algorithm. In this algorithm, gain allocated to coefficients that have achieved the vicinity of their optimal values is transferred to other coefficients that have not yet converged, thus enhancing the algorithm convergence and reducing the steady-state misalignment.

A convenient way to provide a theoretical basis for the study of a given adaptive algorithm is through its stochastic model. In the stochastic modeling of adaptive algorithms, the aim is to derive expressions describing (with certain accuracy) the algorithm behavior under different operating conditions (Duttweiler 2000; Rupp 1993; Doroslovački and Deng 2006; Wagner and Doroslovački 2008a, b; Loganathan et al. 2010; de Souza et al. 2010b; Kuhn et al. 2014a; Haddad and Petraglia 2014). So, from the model expressions, cause-and-effect relationships between algorithm parameters and performance metrics can be established, resulting in useful design guidelines (Abadi and Husøy 2009; Kuhn et al. 2014b; Kuhn et al. 2015; Matsuo and Seara 2016). Moreover, model expressions can help in identifying anomalous and/or undesired behavior, allowing to modify the algorithm to circumvent such an issue (Kolodziej et al. 2009). In this context, focusing on the stochastic modeling of the NSVR–IAF–PNLMS algorithm, the present research work has the following goals:

- (i) To obtain a stochastic model for predicting the algorithm behavior considering correlated Gaussian input data and time-varying characteristics of the gain distribution matrix;
- (ii) To provide model expressions describing the mean weight vector, gain distribution matrix, NSVR metric, learning curve, and weight-error correlation matrix;
- (iii) To derive model expressions characterizing the excess mean-square error (EMSE) in steady state and misadjustment; and
- (iv) To assess the accuracy of the model for different operating conditions.

Note that the proposed stochastic model presented here extends the results given in Perez et al. (2017).

This paper is organized as follows. Section 2 describes briefly the system identification setup considered and revisits the NSVR–IAF–PNLMS algorithm. In Sect. 3, based on some assumptions and approximations, the proposed stochastic model is derived and discussed. Section 4 presents

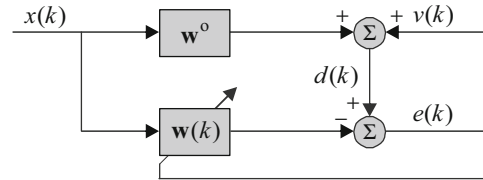


Fig. 1 Block diagram of a typical system identification setup

simulation results, aiming to assess the accuracy of the proposed model under different operating scenarios. Finally, Sect. 5 presents concluding remarks.

Throughout this paper, the mathematical notation adopted 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. In addition, superscript T stands for the transpose,  $\text{diag}(\cdot)$  denotes the diagonal operator, and  $E(\cdot)$  characterizes the expected value.

## 2 Problem Statement

In this section, a brief description of the environment considered in the development of the proposed model is firstly presented. Next, the general expressions describing the NSVR–IAF–PNLMS algorithm are introduced.

### 2.1 System Identification Setup

In a system identification setup (as depicted in Fig. 1), the adaptive filter is used to produce an estimate of an unknown system (plant) impulse response (Sayed 2009; Farhang-Boroujeny 2013; Haykin 2014). In this kind of setup, both the plant (system to be identified)  $\mathbf{w}^0$  and the adaptive filter  $\mathbf{w}(k)$  are driven by the same input signal  $x(k)$ . In turn, the plant output signal is corrupted by an additive measurement noise  $v(k)$ , which results in the desired signal  $d(k)$ . Thereby, the error signal can be expressed as

$$\begin{aligned} e(k) &= d(k) - \mathbf{w}^T(k) \mathbf{x}(k) \\ &= [\mathbf{w}^0 - \mathbf{w}(k)]^T \mathbf{x}(k) + v(k) \end{aligned} \quad (1)$$

where  $\mathbf{w}(k) = [w_1(k) \ w_2(k) \ \dots \ w_N(k)]^T$  characterizes the adaptive coefficient vector,  $\mathbf{w}^0$  is the plant coefficient vector, and  $\mathbf{x}(k) = [x(k) \ x(k-1) \ \dots \ x(k-N+1)]^T$  denotes the input vector with the  $N$  most recent input samples. Note that the order of the adaptive filter and the plant is assumed the same for the sake of simplicity.

### 2.2 Revisiting the NSVR–IAF–PNLMS Algorithm

For PNLMS-type algorithms, the coefficient update rule is expressed as (Duttweiler 2000; Benesty and Gay 2002)

$$\mathbf{w}(k + 1) = \mathbf{w}(k) + \frac{\beta \mathbf{G}(k) \mathbf{x}(k) e(k)}{\mathbf{x}^T(k) \mathbf{G}(k) \mathbf{x}(k) + \xi} \quad (2)$$

where  $\beta$  denotes the step-size control parameter,  $\xi > 0$  is a regularization parameter that prevents division by zero, while

$$\mathbf{G}(k) = \text{diag}[g_1(k) g_2(k) \cdots g_N(k)] \quad (3)$$

represents a diagonal gain distribution matrix. Note that the individual gain governing the step-size adjustment of the  $i$ th adaptive filter coefficient is obtained as

$$g_i(k) = \frac{\gamma_i(k)}{\sum_{j=1}^N \gamma_j(k)}, \quad i = 1, 2, 3, \dots \quad (4)$$

with  $\gamma_i(k)$  characterizing a proportionality function, which is defined in different ways depending on the particulars of the algorithm considered (e.g., see Duttweiler 2000; Benesty and Gay 2002; Deng and Doroslovački 2006; Loganathan et al. 2008; de Souza et al. 2010a).

For the NSVR–IAF–PNLMS algorithm, the proportionality function associated with the  $i$ th adaptive filter coefficient is given by (Perez et al. 2017)

$$\gamma_i(k) = \begin{cases} \frac{1}{N}, & V_i(k) \leq \varepsilon \\ \lambda_i(k), & \text{otherwise} \end{cases} \quad (5)$$

where  $\varepsilon$  is a convergence threshold parameter,

$$V_i(k) = \begin{cases} \frac{|Q_i(k) - P_i(k)|}{\min[Q_i(k), P_i(k)] + \zeta}, & k = m\Delta, \quad m = 1, 2, 3, \dots \\ V_i(k - 1), & \text{otherwise} \end{cases} \quad (6)$$

with  $\zeta > 0$  denoting a regularization parameter,

$$Q_i(k) = \sum_{j=0}^{(\Delta-2)/2} |w_i(k - j)| \quad (7)$$

and

$$P_i(k) = \sum_{j=\Delta/2}^{\Delta-1} |w_i(k - j)| \quad (8)$$

in which  $\Delta$  defines the number of samples of  $w_i(k)$  considered to assess the individual-coefficient convergence, while

$$\lambda_i(k) = \max[f_i(k), |w_i(k)|] \quad (9)$$

and

$$f_i(k) = \begin{cases} \frac{1}{2}|w_i(k)| + \frac{1}{2}\lambda_i(k - 1), & k = mN, \quad m = 1, 2, 3, \dots \\ f_i(k - 1), & \text{otherwise} \end{cases} \quad (10)$$

arise from the IAF–PNLMS algorithm given in de Souza et al. (2010a).

Notice that (6) characterizes the NSVR metric of the  $i$ th adaptive coefficient, which is used to capture the individual-coefficient convergence behavior during the adaptation process [see (7) and (8)]. In turn,  $\varepsilon$  represents the percentage value of coefficient magnitude variation for which one assumes that the coefficient has reached the vicinity of its optimal value [see (5)]. Then, the value chosen for  $\varepsilon$  must ensure that gain assigned to adaptive coefficients that have reached the vicinity of their optimal values is transferred to other coefficients that have not yet converged, thus improving the algorithm convergence. Therefore, the choice of  $\Delta$  and  $\varepsilon$  affects the change of the gain distribution policy and, hence, the algorithm performance [typical values for  $\Delta$  and  $\varepsilon$  are discussed in Perez et al. (2017)].

### 3 Proposed Stochastic Model

To derive the stochastic model of the NSVR–IAF–PNLMS algorithm [given by (1)–(10)], the following simplifying assumptions and approximations are stated:

- (A1) The input signal  $x(k)$  is obtained from a zero-mean correlated Gaussian process with variance  $\sigma_x^2$  and autocorrelation matrix  $\mathbf{R} = E[\mathbf{x}(k)\mathbf{x}^T(k)]$  (Sayed 2009; Farhang-Boroujeny 2013; Haykin 2014).
- (A2) Regularization parameters  $0 < \xi \ll 1$  and  $0 < \zeta \ll 1$  are neglected during the model development for the sake of simplicity (Kuhn et al. 2014c).
- (A3) The measurement noise  $v(k)$  is obtained from a zero-mean independent and identically distributed (i.i.d.) Gaussian process with variance  $\sigma_v^2$ , which is uncorrelated with any other signal (Sayed 2009; Farhang-Boroujeny 2013; Haykin 2014).
- (A4) For slow adaptation (small step size), the adaptive weight vector  $\mathbf{w}(k)$  and the input vector  $\mathbf{x}(k)$  are assumed statistically independent (Sayed 2009; Farhang-Boroujeny 2013; Haykin 2014).
- (A5) Due to the piecewise-constant nature of the gain distribution matrix [see (3)–(10)], variables  $\mathbf{G}(k)$ ,  $\mathbf{w}(k)$ , and  $\mathbf{x}(k)$  are assumed uncorrelated (de Souza et al. 2010b; Kuhn et al. 2014a).

Note that these assumptions and approximations have been commonly used in the stochastic modeling of adaptive algorithms and are required to make the development mathematically tractable.

### 3.1 Mean Weight Vector

To derive an expression that describes the mean weight vector of the algorithm, let us substitute (1) into (2) and take the expected value of both sides of the resulting expression, which yields

$$E[\mathbf{w}(k+1)] = E[\mathbf{w}(k)] + \beta E \left[ \frac{\mathbf{G}(k)\mathbf{x}(k)\mathbf{x}^T(k)}{\mathbf{x}^T(k)\mathbf{G}(k)\mathbf{x}(k) + \xi} \right] \mathbf{w}^0 - \beta E \left[ \frac{\mathbf{G}(k)\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{w}(k)}{\mathbf{x}^T(k)\mathbf{G}(k)\mathbf{x}(k) + \xi} \right] + \beta E \left[ \frac{\mathbf{G}(k)\mathbf{x}(k)v(k)}{\mathbf{x}^T(k)\mathbf{G}(k)\mathbf{x}(k) + \xi} \right]. \quad (11)$$

Then, taking into account Assumptions (A2)–(A5), (11) can be simplified to

$$E[\mathbf{w}(k+1)] \cong [\mathbf{I} - \beta \mathbf{R}_1(k)]E[\mathbf{w}(k)] + \beta \mathbf{R}_1(k)\mathbf{w}^0 \quad (12)$$

where  $\mathbf{I}$  denotes the identity matrix and

$$\mathbf{R}_1(k) \cong E \left[ \frac{\bar{\mathbf{G}}(k)\mathbf{x}(k)\mathbf{x}^T(k)}{\mathbf{x}^T(k)\bar{\mathbf{G}}(k)\mathbf{x}(k)} \right] \quad (13)$$

in which  $\bar{\mathbf{G}}(k)$  characterizes the mean gain distribution matrix. Therefore, using the solution presented in Kuhn et al. (2014a) for computing  $\mathbf{R}_1(k)$  [see Assumption (A1)], the behavior of the mean weight vector can be predicted through (12) if the mean gain distribution matrix is known.

### 3.2 Mean Gain Distribution Matrix

An expression that describes the mean gain distribution matrix can be obtained by taking the expected value of both sides of (3)–(10) and assuming that  $E[|w_i(k)|] \cong |E[w_i(k)]|$  (Kuhn et al. 2014a). Thereby,

$$\bar{\mathbf{G}}(k) = \text{diag}[\bar{g}_1(k) \bar{g}_2(k) \cdots \bar{g}_N(k)] \quad (14)$$

in which

$$\bar{g}_i(k) \cong \frac{E[\gamma_i(k)]}{\sum_{j=1}^N E[\gamma_j(k)]} \quad (15)$$

with

$$E[\gamma_i(k)] \cong \begin{cases} \frac{1}{N}, & E[V_i(k)] \leq \varepsilon \\ E[\lambda_i(k)], & \text{otherwise} \end{cases} \quad (16)$$

where

$$E[\lambda_i(k)] \cong \max\{E[f_i(k)], |E[w_i(k)]|\} \quad (17)$$

and

$$E[f_i(k)] \cong \begin{cases} \frac{1}{2}|E[w_i(k)]| + \frac{1}{2}E[\lambda_i(k-1)], & k = mN, \quad m = 1, 2, 3, \dots \\ E[f_i(k-1)], & \text{otherwise.} \end{cases} \quad (18)$$

So, considering (12), the evolution of the mean gain distribution matrix can be predicted through (14)–(18) if the mean behavior of the NSVR metric is known.

### 3.3 Mean Behavior of the NSVR Metric

Aiming to determine an expression that describes the evolution of the NSVR metric  $V_i(k)$  for active coefficients, let us start by taking the expected value of both sides of (6) along with Assumption (A2), i.e.,

$$E[V_i(k)] \cong \begin{cases} E \left[ \frac{|Q_i(k) - P_i(k)|}{\min[Q_i(k), P_i(k)]} \right], & k = m\Delta, \quad m = 1, 2, \dots \\ E[V_i(k-1)], & \text{otherwise.} \end{cases} \quad (19)$$

Next, considering that the numerator varies slowly with respect to the denominator in such a way that the averaging principle (AP) (Samson and Reddy 1983) can be invoked, one gets

$$E \left[ \frac{|Q_i(k) - P_i(k)|}{\min[Q_i(k), P_i(k)]} \right] \cong \frac{E[|Q_i(k) - P_i(k)|]}{E\{\min[Q_i(k), P_i(k)]\}}. \quad (20)$$

Also, assuming that coefficients hold the same algebraic sign during the whole adaptation process, (20) becomes

$$E \left[ \frac{|Q_i(k) - P_i(k)|}{\min[Q_i(k), P_i(k)]} \right] \cong \frac{|E[Q_i(k)] - E[P_i(k)]|}{\min\{E[Q_i(k)], E[P_i(k)]\}}. \quad (21)$$

Finally, substituting (21) into (19), the mean behavior of the NSVR for the  $i$ th (active) adaptive coefficient can be approximated to

$$E[V_i(k)] \cong \begin{cases} \frac{|E[Q_i(k)] - E[P_i(k)]|}{\min\{E[Q_i(k)], E[P_i(k)]\}}, & k = m\Delta, \quad m = 1, 2, 3, \dots \\ E[V_i(k-1)], & \text{otherwise} \end{cases} \tag{22}$$

with

$$E[Q_i(k)] \cong \sum_{j=0}^{(\Delta-2)/2} |E[w_i(k-j)]| \tag{23}$$

and

$$E[P_i(k)] \cong \sum_{j=\Delta/2}^{\Delta-1} |E[w_i(k-j)]| \tag{24}$$

obtained by taking the expected value of both sides of (7) and (8), respectively. Thereby, using (12), the mean behavior of the NSVR metric can be predicted from (22) to (24).

### 3.4 Learning Curve

An expression that describes the algorithm learning curve [mean-square error (MSE)] can be derived by rewriting (1) in terms of the weight-error vector

$$\mathbf{z}(k) = \mathbf{w}(k) - \mathbf{w}^o \tag{25}$$

as

$$e(k) = -\mathbf{z}^T(k) \mathbf{x}(k) + v(k) \tag{26}$$

determining  $e^2(k)$ , taking the expected value of both sides of the resulting expression, and considering Assumptions (A3) and (A4) (Sayed 2009; Farhang-Boroujeny 2013; Haykin 2014). Thus,

$$J(k) = E[e^2(k)] = J_{\min} + J_{\text{ex}}(k) \tag{27}$$

where

$$J_{\min} = \sigma_v^2 \tag{28}$$

denotes the minimum MSE, while

$$J_{\text{ex}}(k) = \text{tr}[\mathbf{R}\mathbf{K}(k)] \tag{29}$$

characterizes the excess MSE (EMSE). Note in (29) that the weight-error correlation matrix is defined as

$$\mathbf{K}(k) = E[\mathbf{z}(k)\mathbf{z}^T(k)]. \tag{30}$$

Therefore, the learning curve (MSE) can be predicted through (27)–(29) if the weight-error correlation matrix is known.

### 3.5 Weight-Error Correlation Matrix

Since the expression that describes the behavior of the algorithm learning curve requires knowledge of the weight-error covariance matrix, a recursive expression characterizing the evolution of  $\mathbf{K}(k)$  is derived now. To this end, using (2) and (26), the weight update equation is first rewritten in terms of the weight-error vector  $\mathbf{z}(k)$  [defined in (25)] as

$$\mathbf{z}(k+1) = \mathbf{z}(k) - \beta \frac{\mathbf{G}(k)\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{z}(k)}{\mathbf{x}^T(k)\mathbf{G}(k)\mathbf{x}(k) + \xi} + \beta \frac{\mathbf{G}(k)\mathbf{x}(k)v(k)}{\mathbf{x}^T(k)\mathbf{G}(k)\mathbf{x}(k) + \xi}. \tag{31}$$

Then, determining  $\mathbf{z}(k+1)\mathbf{z}^T(k+1)$ , taking the expected value of both sides of the resulting expression, and considering (30), one obtains

$$\begin{aligned} \mathbf{K}(k+1) &= \mathbf{K}(k) - \beta E \left[ \frac{\mathbf{z}(k)\mathbf{z}^T(k)\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{G}(k)}{\mathbf{x}^T(k)\mathbf{G}(k)\mathbf{x}(k) + \xi} \right] \\ &\quad - \beta E \left[ \frac{\mathbf{G}(k)\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{z}(k)\mathbf{z}^T(k)}{\mathbf{x}^T(k)\mathbf{G}(k)\mathbf{x}(k) + \xi} \right] \\ &\quad - \beta^2 E \left\{ \frac{v(k)\mathbf{G}(k)\mathbf{x}(k)\mathbf{z}^T(k)\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{G}(k)}{[\mathbf{x}^T(k)\mathbf{G}(k)\mathbf{x}(k) + \xi]^2} \right\} \\ &\quad + \beta^2 E \left[ \frac{v(k)\mathbf{G}(k)\mathbf{x}(k)\mathbf{z}^T(k)}{\mathbf{x}^T(k)\mathbf{G}(k)\mathbf{x}(k) + \xi} \right] \\ &\quad - \beta^2 E \left\{ \frac{v(k)\mathbf{G}(k)\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{z}(k)\mathbf{z}^T(k)\mathbf{G}(k)}{[\mathbf{x}^T(k)\mathbf{G}(k)\mathbf{x}(k) + \xi]^2} \right\} \\ &\quad + \beta^2 E \left\{ \frac{v^2(k)\mathbf{G}(k)\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{G}(k)}{[\mathbf{x}^T(k)\mathbf{G}(k)\mathbf{x}(k) + \xi]^2} \right\} \\ &\quad + \beta^2 E \left\{ \frac{\mathbf{G}(k)\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{z}(k)\mathbf{z}^T(k)\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{G}(k)}{[\mathbf{x}^T(k)\mathbf{G}(k)\mathbf{x}(k) + \xi]^2} \right\} \\ &\quad + \beta^2 E \left[ \frac{v(k)\mathbf{z}(k)\mathbf{x}^T(k)\mathbf{G}(k)}{\mathbf{x}^T(k)\mathbf{G}(k)\mathbf{x}(k) + \xi} \right]. \end{aligned} \tag{32}$$

Finally, due to the statistical characteristics of the measurement noise [see Assumption (A3)], all terms involving  $v(k)$  in (32) are equal to zero, except that with  $v^2(k)$ . Additionally, using Assumptions (A2)–(A5), (32) reduces to

$$\mathbf{K}(k+1) \cong \mathbf{K}(k) - \beta\mathbf{K}(k)\mathbf{R}_1^T(k) - \beta\mathbf{R}_1(k)\mathbf{K}(k) + \beta^2 J_{\min}\mathbf{R}_2(k) + \beta^2\mathbf{R}_3(k) \tag{33}$$

with

$$\mathbf{R}_2(k) \cong E \left[ \frac{\bar{\mathbf{G}}(k)\mathbf{x}(k)\mathbf{x}^T(k)\bar{\mathbf{G}}(k)}{[\mathbf{x}^T(k)\bar{\mathbf{G}}(k)\mathbf{x}(k)]^2} \right] \quad (34)$$

and

$$\mathbf{R}_3(k) \cong E \left[ \frac{\bar{\mathbf{G}}(k)\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{K}(k)\mathbf{x}(k)\mathbf{x}^T(k)\bar{\mathbf{G}}(k)}{[\mathbf{x}^T(k)\bar{\mathbf{G}}(k)\mathbf{x}(k)]^2} \right]. \quad (35)$$

So, using the solutions presented in Kuhn et al. (2014a) for computing  $\mathbf{R}_1(k)$ ,  $\mathbf{R}_2(k)$ , and  $\mathbf{R}_3(k)$  [see Assumption (A1)], along with (14)–(24), the evolution of the weight-error correlation matrix can now be predicted through (33); consequently, the algorithm learning curve can also be predicted from (27).

### 3.6 Steady-State Analysis

Based on energy-conservation arguments (Sayed 2009), an expression characterizing the steady-state EMSE  $J_{\text{ex}}(\infty)$  can be derived by rewriting (2) in terms of the weight-error vector  $\mathbf{z}(k)$  [defined in (25)], which results in

$$\mathbf{z}(k+1) = \mathbf{z}(k) + \beta \frac{\mathbf{G}(k)\mathbf{x}(k)e(k)}{\mathbf{x}^T(k)\mathbf{G}(k)\mathbf{x}(k) + \xi}. \quad (36)$$

Then, calculating  $\mathbf{z}^T(k+1)\mathbf{G}^{-1}(k)\mathbf{z}(k+1)$  from (36), i.e.,

$$\begin{aligned} \mathbf{z}^T(k+1)\mathbf{G}^{-1}(k)\mathbf{z}(k+1) &= \mathbf{z}^T(k)\mathbf{G}^{-1}(k)\mathbf{z}(k) \\ &+ 2\beta \frac{\mathbf{z}^T(k)\mathbf{x}(k)e(k)}{\mathbf{x}^T(k)\mathbf{G}(k)\mathbf{x}(k) + \xi} + \beta^2 \frac{\mathbf{x}^T(k)\mathbf{G}(k)\mathbf{x}(k)e^2(k)}{[\mathbf{x}^T(k)\mathbf{G}(k)\mathbf{x}(k) + \xi]^2} \end{aligned} \quad (37)$$

taking the expected value of both sides, letting  $k \rightarrow \infty$ , assuming convergence, considering (26), as well as Assumptions (A2) and (A3), the following variance relation is obtained:

$$2 \lim_{k \rightarrow \infty} E \left\{ \frac{[\mathbf{z}^T(k)\mathbf{x}(k)]^2}{\mathbf{x}^T(k)\mathbf{G}(k)\mathbf{x}(k)} \right\} = \beta \lim_{k \rightarrow \infty} E \left[ \frac{e^2(k)}{\mathbf{x}^T(k)\mathbf{G}(k)\mathbf{x}(k)} \right]. \quad (38)$$

Now, recalling the definition of the EMSE [given in (29)], taking into account Assumptions (A4) and (A5), as well as assuming that  $M$  is large in such a way that the AP (Samson and Reddy 1983) can be used, the left-hand side term of (38) is approximated by

$$\lim_{k \rightarrow \infty} E \left\{ \frac{[\mathbf{z}^T(k)\mathbf{x}(k)]^2}{\mathbf{x}^T(k)\mathbf{G}(k)\mathbf{x}(k)} \right\} \cong \frac{J_{\text{ex}}(\infty)}{\text{tr}[\mathbf{R}\bar{\mathbf{G}}(\infty)]} \quad (39)$$

with

$$\bar{\mathbf{G}}(\infty) = \lim_{k \rightarrow \infty} E[\mathbf{G}(k)] \quad (40)$$

denoting the mean gain distribution matrix in steady state. Similarly, using (26), (28), (29), and (40), along with Assumptions (A3)–(A5), the right-hand side term of (38) is simplified to

$$\lim_{k \rightarrow \infty} E \left[ \frac{e^2(k)}{\mathbf{x}^T(k)\mathbf{G}(k)\mathbf{x}(k)} \right] \cong \frac{J_{\text{ex}}(\infty)}{\text{tr}[\mathbf{R}\bar{\mathbf{G}}(\infty)]} + J_{\text{min}}E_1(\infty) \quad (41)$$

where

$$E_1(\infty) = E \left[ \frac{1}{\mathbf{x}^T(k)\bar{\mathbf{G}}(\infty)\mathbf{x}(k)} \right]. \quad (42)$$

So, substituting (39) and (41) into (38), and solving the resulting expression for  $J_{\text{ex}}(\infty)$ , the steady-state EMSE is obtained according to

$$J_{\text{ex}}(\infty) \cong \frac{\beta}{(2-\beta)} J_{\text{min}} \text{tr}[\mathbf{R}\bar{\mathbf{G}}(\infty)] E_1(\infty). \quad (43)$$

Finally, to determine the steady-state behavior of the mean gain distribution matrix  $\bar{\mathbf{G}}(\infty)$  [required in (42) and (43)], let  $k \rightarrow \infty$  of both sides of (15) in such a way that

$$\bar{g}_i(\infty) \cong \frac{E[\gamma_i(\infty)]}{\sum_{j=1}^N E[\gamma_j(\infty)]} \quad (44)$$

with

$$E[\gamma_i(\infty)] = \lim_{k \rightarrow \infty} E[\gamma_i(k)]. \quad (45)$$

Note from (45) that knowledge of the mean value of the proportionality function in steady state is required. Then, recalling from (22) to (24) that

$$\lim_{k \rightarrow \infty} E[V_i(k)] \rightarrow 0 \quad (46)$$

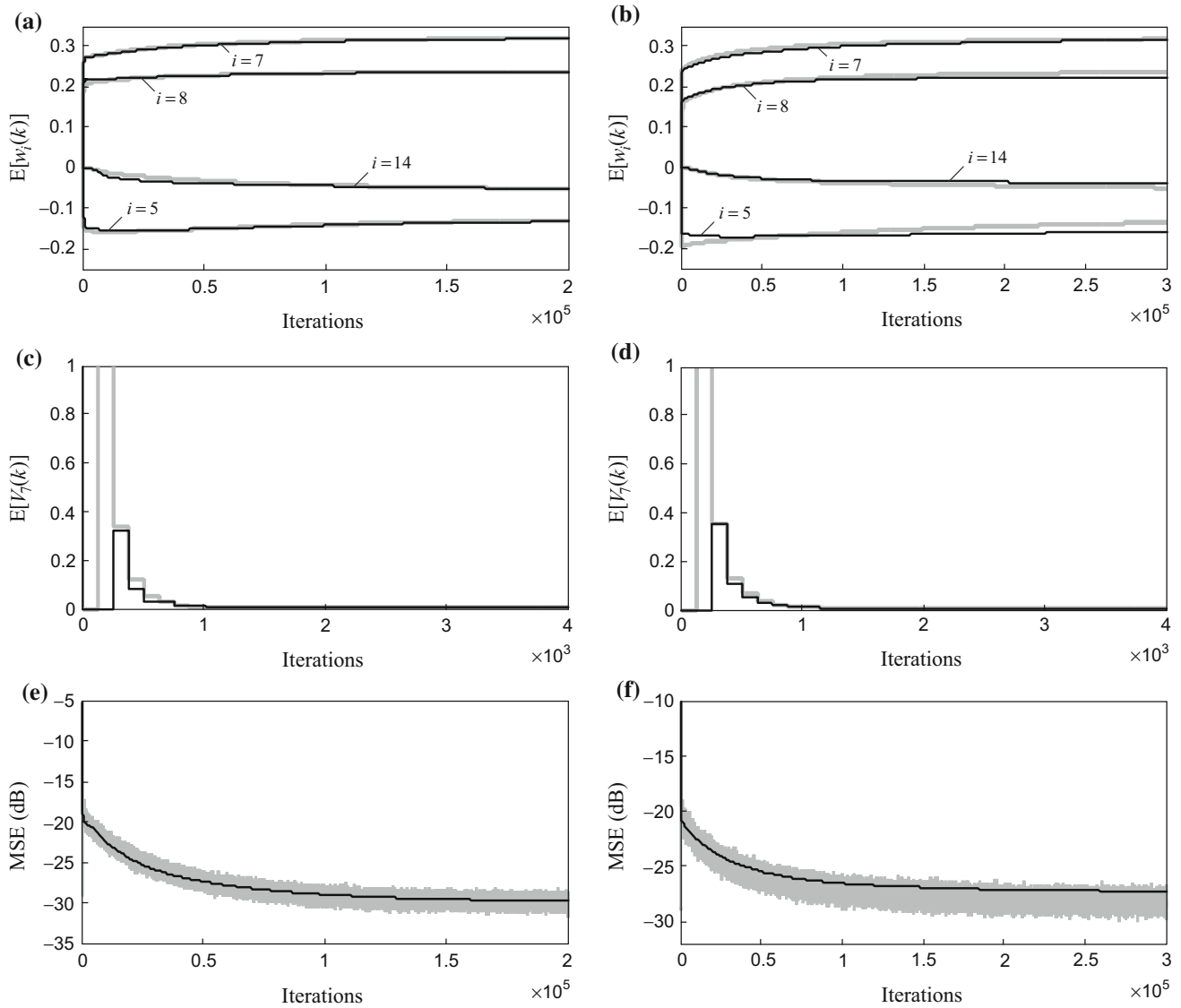
since the adaptive coefficients do not change significantly as the algorithm tends to the steady state, it can be verified [from (16)] that

$$\lim_{k \rightarrow \infty} E[\gamma_i(k)] \rightarrow \frac{1}{N}. \quad (47)$$

Consequently, the mean gain distribution matrix in steady state reduces to

$$\bar{\mathbf{G}}(\infty) \cong \frac{1}{N} \mathbf{I}. \quad (48)$$

Therefore, using the solution presented in Kuhn et al. (2014b) for computing  $E_1(\infty)$  [see Assumption (A1)], along with the expression characterizing the mean gain distribution matrix in steady state  $\bar{\mathbf{G}}(\infty)$  [given by (48)], the steady-state EMSE can be predicted through (43); likewise, the misadjustment can be straightforwardly determined.



**Fig. 2** Example 1: Algorithm behavior obtained by MC simulations (gray-ragged lines) and through the proposed model (dark-solid lines) considering  $\chi = 71.71$  (left) and  $\chi = 338.23$  (right). **a, b** Mean weight

behavior of four adaptive coefficients. **c, d** Evolution of the NSVR metric associated with coefficient  $w_7(k)$ . **e, f** Learning curve

### 4 Simulation Results

To verify the accuracy of the proposed model, three examples are presented considering distinct operating scenarios in a system identification setup. In these examples, the results obtained from Monte Carlo (MC) simulations (average of 200 independent runs) are compared with those predicted through the proposed stochastic model. The zero-mean and unit variance Gaussian input signal  $x(k)$  [see Assumption (A1)] is obtained from (Haykin 2014)

$$x(k) = -a_1x(k - 1) - a_2x(k - 2) + \eta(k) \tag{49}$$

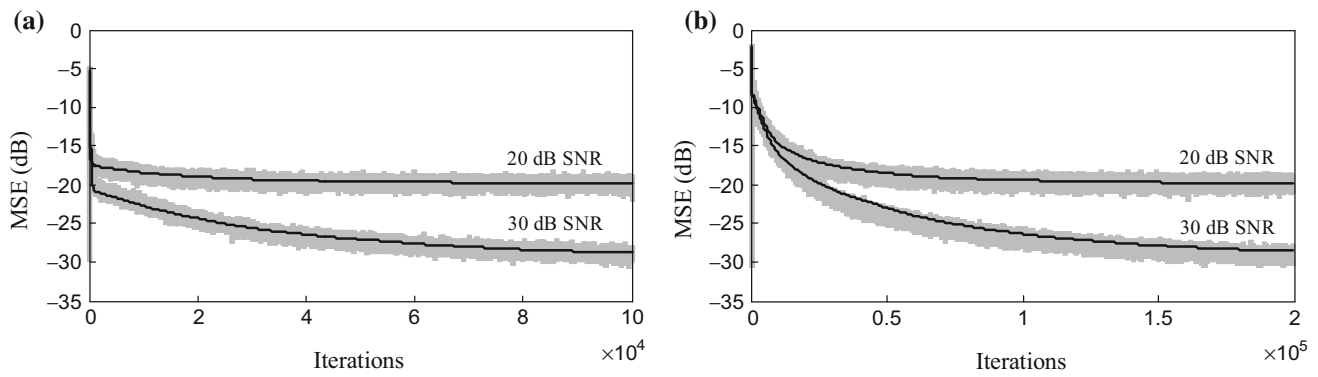
where  $a_1$  and  $a_2$  denote the AR(2) coefficients, and  $\eta(k)$  is a white Gaussian noise with variance

$$\sigma_\eta^2 = \left( \frac{1 - a_2}{1 + a_2} \right) [(1 + a_2)^2 - a_1^2]. \tag{50}$$

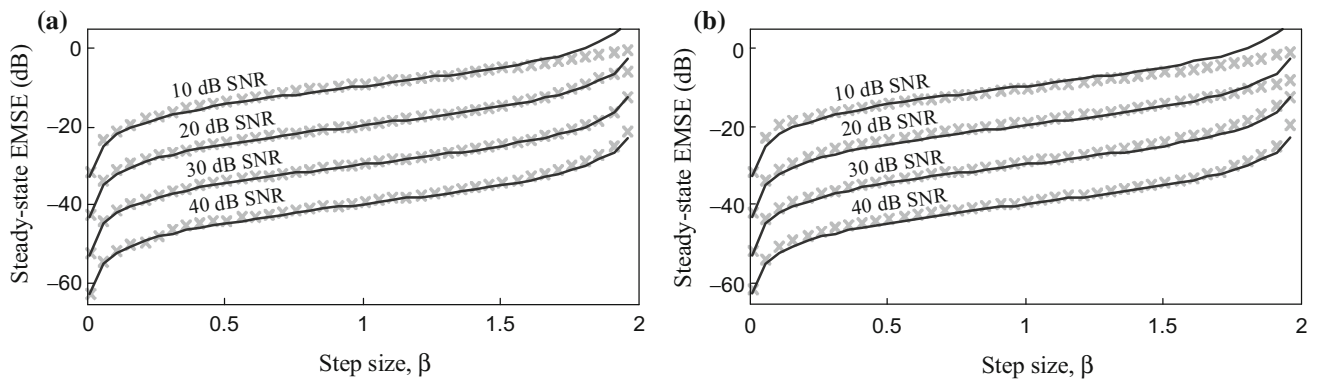
The signal-to-noise ratio (SNR) is defined here as

$$\text{SNR} = 10 \log_{10} \left( \frac{\sigma_x^2}{\sigma_v^2} \right). \tag{51}$$

Two plants  $w_A^0$  and  $w_B^0$  (with  $N = 128$  weights) obtained from echo-path models for testing of speech echo cancellers defined in the (ITU-T Recommendation G.168 2009, Models 1 and 4) are used. (Note that  $w_A^0$  has been formed by zero



**Fig. 3** Example 2: Learning curves (MSE) obtained by MC simulations (gray-ragged lines) and through the proposed model (dark-solid lines). **a** Plant  $w_A^o$  [ $S(w_A^o) = 0.78$ ]. **b** Plant  $w_B^o$  [ $S(w_B^o) = 0.42$ ]



**Fig. 4** Example 3: Steady-state EMSE obtained by MC simulations (gray-cross marker) and through the proposed model (dark-solid lines). **a** Plant  $w_A^o$  [ $S(w_A^o) = 0.78$ ]. **b** Plant  $w_B^o$  [ $S(w_B^o) = 0.42$ ]

padding the echo-path model.) Sparseness degrees (Hoyer 2004) of such plants are  $S(w_A^o) = 0.78$  and  $S(w_B^o) = 0.42$ , respectively. Unless otherwise stated, parameter values of the algorithm are  $\beta = 0.01$ ,  $\xi = 10^{-6}$ ,  $f_i(0) = \lambda_i(0) = 10^{-4}$ ,  $\Delta = N$ ,  $\varepsilon = 10^{-2}$ , and  $\zeta = 10^{-4}$ , while  $w(0) = \mathbf{0}$  and  $V_i(0) = 0$ . Still, to ensure the stability in the computation of  $R_1(k)$ ,  $R_2(k)$ ,  $R_3(k)$ , and  $E_1(\infty)$ , the threshold value used for discarding less significant eigenvalues is  $5 \times 10^{-4}$  (Kuhn et al. 2014a).

**4.1 Example 1**

Here, the accuracy of the model expressions characterizing the mean weight vector (12), NSVR metric (22), and the algorithm learning curve (27) is assessed for different input data correlation levels. To this end, we consider plant  $w_A^o$  [ $S(w_A^o) = 0.78$ ] and 30 dB SNR. The eigenvalue spread values<sup>1</sup> of the input autocorrelation matrix are  $\chi = 71.71$  [obtained from (49) for  $a_1 = -0.75$  and  $a_2 = 0.68$ ] and  $\chi =$

<sup>1</sup> Since the eigenvalue spread is defined as a ratio between the maximum and minimum eigenvalues of the input data autocorrelation matrix, the larger the eigenvalue spread, the higher the correlation level of the input data (Haykin 2014).

338.23 [obtained from (49) by changing  $a_2$  to 0.85]. Figure 2 shows the results obtained by MC simulations and predicted through the proposed model. Specifically, Fig. 2a, b depicts the mean behavior of four adaptive weights (for the sake of clarity). Figure 2c, d illustrates the evolution of the NSVR metric for one active coefficient. Figure 2e, f presents the algorithm learning curve (MSE). Notice from these figures that the proposed model satisfactorily predicts the algorithm behavior for both transient and steady-state phases, irrespective of the input data correlation level considered. This result also confirms the accuracy of the methodology used for computing the normalized autocorrelation-like matrices  $R_1(k)$ ,  $R_2(k)$ , and  $R_3(k)$ .

**4.2 Example 2**

This example verifies the accuracy of the proposed model (via learning curve) for plants with different sparseness degrees and two distinct SNR values. For this operating scenario, plants  $w_A^o$  [ $S(w_A^o) = 0.78$ ] and  $w_B^o$  [ $S(w_B^o) = 0.42$ ] are used, considering 20 and 30 dB SNR. The input signal is obtained from (49) with  $a_1 = -0.75$  and  $a_2 = 0.75$ , which yields an eigenvalue spread of  $\chi = 120.25$  for the input autocorrela-



tion matrix. Figure 3 depicts the results obtained through MC simulations and from model predictions. Specifically, Fig. 3a shows the results obtained for plant  $\mathbf{w}_A^0$  [ $S(\mathbf{w}_A^0) = 0.78$ ], while Fig. 3b for plant  $\mathbf{w}_B^0$  [ $S(\mathbf{w}_B^0) = 0.42$ ]. From these figures, very good match between MC simulations and model predictions can also be verified for both transient and steady-state phases, irrespective of the plant sparseness degrees and SNR values considered.

### 4.3 Example 3

In this last example, considering plants with different sparseness degrees and a wide range of SNR values, the accuracy of the model expressions characterizing the steady-state EMSE (43) is assessed for step-size values ranging from 0.01 to 1.96. To this end, we consider plants  $\mathbf{w}_A^0$  [ $S(\mathbf{w}_A^0) = 0.78$ ] and  $\mathbf{w}_B^0$  [ $S(\mathbf{w}_B^0) = 0.42$ ], as well as four different SNR values (i.e., 10, 20, 30, and 40 dB SNR). The remaining parameter values are the same as in Example 2. [Note also that the last 100 EMSE values in steady state have been averaged to provide a better visualization of the experimental results obtained from MC simulations (Sayed 2009, pp. 250).] Figure 4 shows the steady-state EMSE curves obtained by MC simulations along with those predicted from the model expressions. Specifically, Fig. 4a illustrates the results obtained for plant  $\mathbf{w}_A^0$  [ $S(\mathbf{w}_A^0) = 0.78$ ], while Fig. 4b for plant  $\mathbf{w}_B^0$  [ $S(\mathbf{w}_B^0) = 0.42$ ]. Observe from such figures that the proposed model predicts very well the steady-state algorithm behavior for a wide range of step-size values, irrespective of the plant sparseness degrees and SNR values considered.

## 5 Concluding Remarks

In this paper, considering correlated Gaussian input data, a stochastic model was obtained for the NSVR–IAF–PNLMS algorithm. Specifically, we derived expressions describing the mean weight vector, gain distribution matrix, NSVR metric, learning curve, weight-error correlation matrix, and steady-state EMSE. To this end, a reduced number of assumptions were used, leading to expressions that predict satisfactorily both the algorithm transient and steady-state phases. Simulation results confirmed the accuracy of the model for different operating conditions, i.e., covering different plants, SNR and step-size values, and several eigenvalue spreads. Note that the methodology discussed here can be extended to other PNLMS-type algorithms from the literature.

**Acknowledgements** The authors would like to thank the Editor and the Reviewers for the positive comments on this paper. Also, we are thanked to the Brazilian National Council for Scientific and Technological Development (CNPq) for supporting in part the development of this research work.

## References

- Abadi, M. S. E., & Husøy, J. H. (2009). On the application of a unified adaptive filter theory in the performance prediction of adaptive filter algorithms. *Digital Signal Processing*, 19(3), 410–432.
- Benesty, J., & Gay, S. L. (2002). An improved PNLMS algorithm. In *IEEE international conference of acoustics, speech, and signal processing (ICASSP)* (pp. 1881–1884).
- de Souza, F. C., Seara, R., & Morgan, D. R. (2012). An enhanced IAF-PNLMS adaptive algorithm for sparse impulse response identification. *IEEE Transactions on Signal Processing*, 60(6), 3301–3307.
- de Souza, F. C., Tobias, O. J., Seara, R., & Morgan, D. R. (2010a). A PNLMS algorithm with individual activation factors. *IEEE Transactions on Signal Processing*, 58(4), 2036–2047.
- de Souza, F. C., Tobias, O. J., Seara, R., & Morgan, D. R. (2010b). Stochastic model for the mean weight evolution of the IAF-PNLMS algorithm. *IEEE Transactions on Signal Processing*, 58(11), 5895–5901.
- Doroslovački, M. I., & Deng, H. (2006). On convergence of proportionate-type NLMS adaptive algorithms. In *IEEE international conference of acoustics, speech, and signal processing (ICASSP)* (pp. 105–108).
- Deng, H., & Doroslovački, M. (2006). Proportionate adaptive algorithms for network echo cancellation. *IEEE Transactions on Signal Processing*, 54(5), 1794–1803.
- Duttweiler, D. L. (2000). Proportionate normalized least-mean-squares adaptation in echo cancellers. *IEEE Transactions on Speech and Audio Processing*, 8(5), 508–518.
- Farhang-Boroujeny, B. (2013). *Adaptive filters: Theory and applications* (2nd ed.). West Sussex: Wiley.
- Haddad, D. B., & Petraglia, M. R. (2014). Transient and steady-state MSE analysis of the IMPNLMS algorithm. *Digital Signal Processing*, 33, 50–59.
- Haykin, S. (2014). *Adaptive filter theory* (5th ed.). Upper Saddle River, NJ: Pearson Education.
- Hoyer, P. O. (2004). Non-negative matrix factorization with sparseness constraints. *Journal of Machine Learning Research*, 5, 1457–1469.
- ITU-T Recommendation G.168. (2015). *Digital network echo cancellers*. Geneva: International Telecommunications Union.
- Kolodziej, J. E., Tobias, O. J., Seara, R., & Morgan, D. R. (2009). On the constrained stochastic gradient algorithm: Model, performance, and improved version. *IEEE Transactions on Signal Processing*, 57(4), 1304–1315.
- Kuhn, E. V., de Souza, F. C., Seara, R., & Morgan, D. R. (2014a). On the stochastic modeling of the IAF-PNLMS algorithm for complex and real correlated Gaussian input data. *Signal Processing*, 99, 103–115.
- Kuhn, E. V., de Souza, F. C., Seara, R., & Morgan, D. R. (2014b). On the steady-state analysis of PNLMS-type algorithms for correlated Gaussian input data. *IEEE Signal Processing Letters*, 21(11), 1433–1437.
- Kuhn, E. V., Kolodziej, J. E., & Seara, R. (2014c). Stochastic modeling of the NLMS algorithm for complex Gaussian input data and nonstationary environment. *Digital Signal Processing*, 30, 55–66.
- Kuhn, E. V., Kolodziej, J. E., & Seara, R. (2015). Analysis of the TDLMS algorithm operating in a nonstationary environment. *Digital Signal Processing*, 45, 69–83.
- Loganathan, P., Habets, E. A. P., & Naylor, P. A. (2010). Performance analysis of IPNLMS for identification of time-varying systems. In *IEEE international conference of acoustics, speech, and signal processing (ICASSP)* (pp. 317–320).

- Loganathan, P., Khong, A. W. H., & Naylor, P. A. (2008). A sparseness controlled proportionate algorithm for acoustic echo cancellation. In *European signal processing conference (EUSIPCO)* (pp. 1–5).
- Martin, R. K., Sethares, W., Williamson, R. C., & Johnson, C. R. (2002). Exploiting sparsity in adaptive filters. *IEEE Transactions on Signal Processing*, 50(8), 1883–1894.
- Matsuo, M. V., & Seara, R. (2016). On the stochastic analysis of the NLMS algorithm for white and correlated Gaussian inputs in time-varying environments. *Signal Processing*, 128, 291–302.
- Paleologu, C., Benesty, J., & Ciochina, S. (2010). *Sparse adaptive filters for echo cancellation*. San Rafael, CA: Morgan and Claypool Publishers.
- Perez, F. L., Kuhn, E. V., de Souza, F. C., & Seara, R. (2017). A novel gain distribution policy based on individual-coefficient convergence for PNLMS-type algorithms. *Signal Processing*, 138, 294–306.
- Rupp, M. (1993). The behavior of LMS and NLMS algorithms in the presence of spherically invariant processes. *IEEE Transactions on Signal Processing*, 41(3), 1149–1160.
- Samson, C., & Reddy, V. (1983). Fixed point error analysis of the normalized ladder algorithm. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 31(5), 1177–1191.
- Sayed, A. H. (2009). *Adaptive filters*. Hoboken, NJ: Wiley.
- Wagner, K. T., & Doroslovački, M. I. (2008). Analytical analysis of transient and steady-state properties of the proportionate NLMS algorithm. In *Asilomar conference on signals, systems, and computers* (pp. 256–260).
- Wagner, K. T., & Doroslovački, M. I. (2008). Towards analytical convergence analysis of proportionate-type NLMS algorithms. In *IEEE international conference of acoustics, speech, and signal processing (ICASSP)* (pp. 3825–3828).
- Wagner, K., & Doroslovački, M. I. (2013). *Proportionate-type normalized least-mean-square algorithms* (1st ed.). Croydon: ISTE and Wiley.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.