



On the NP-VSS-NLMS Algorithm: Model, Design Guidelines, and Numerical Results

Augusto Cesar Becker¹ · Eduardo Vinicius Kuhn¹ ·
Marcos Vinicius Matsuo² · Jacob Benesty³

Received: 6 March 2023 / Revised: 10 November 2023 / Accepted: 10 November 2023 /
Published online: 27 December 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

In this paper, a stochastic model is presented for the nonparametric variable step-size normalized least-mean-square (NP-VSS-NLMS) algorithm. This algorithm has demonstrated potential in practical applications and hence a deeper understanding of its behavior becomes crucial. In this context, model expressions are obtained for characterizing the algorithm behavior in the transient phase as well as in the steady state, considering a system identification problem and Gaussian input data. Such expressions reveal interesting algorithm characteristics that are useful for establishing design guidelines and for the advancement of more refined algorithms. Simulation results for various operating scenarios ratified both the model's accuracy and the algorithm's superior performance relative to other recent and relevant algorithms from the literature.

✉ Eduardo Vinicius Kuhn
kuhn@utfpr.edu.br

Augusto Cesar Becker
augustobecker@alunos.utfpr.edu.br

Marcos Vinicius Matsuo
marcos.matsuo@ufsc.br

Jacob Benesty
jacob.benesty@inrs.ca

¹ LAPSE–Electronics and Signal Processing Laboratory, Department of Electronics Engineering, Federal University of Technology - Paraná, Toledo 85902-490, PR, Brazil

² GEPS–Electronics and Signal Processing Group, Department of Control, Automation, and Computation Engineering, Federal University of Santa Catarina, Blumenau 89036-004, SC, Brazil

³ INRS-EMT–National Institute of Scientific Research–Energy, Materials, and Telecommunications Research Center, University of Quebec, Montreal, QC H5A-1K6, Canada

Keywords Adaptive filters · Stochastic modeling · System identification · Variable step-size algorithm

1 Introduction

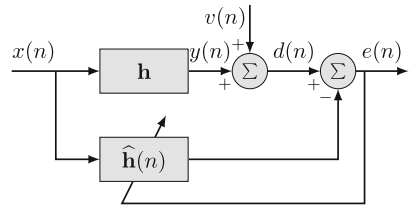
Nowadays, adaptive filtering techniques have been used in a large variety of real-world applications such as adaptive control, noise and echo cancelation, channel equalization, and adaptive beamforming [4, 6, 7, 13–15, 34]. These applications often require the use of an adaptive filter to obtain (in real-time) an approximate input–output relationship representation of an unknown system (to be identified). Such adaptive filter uses an algorithm responsible for adjusting the weights of a filtering structure in order to optimize a given performance measure [13–15, 34]. Among the existing adaptive algorithms, the least-mean-square (LMS) [40, 41] and normalized LMS (NLMS) [3, 20, 30] are well-established and have been extensively used in practical applications, due mainly to their low computational complexity and numerical robustness. Nevertheless, the adjustment of the step-size parameter in these algorithms raises a crucial tradeoff between convergence speed and steady-state performance.

Aiming to address this tradeoff, an approach from the practical point-of-view is given by variable step-size algorithms [2, 9, 11, 17, 24, 25, 28, 35, 37, 39, 42, 43]. These algorithms make use of strategies for adjusting the step-size value during adaptation, by starting with a large step size (within the stability conditions) that allows for faster initial convergence, and then gradually reducing it (slowing down the convergence speed) for achieving smaller error in steady state. Among such algorithms, the nonparametric variable step-size NLMS (NP-VSS-NLMS) [9] exhibits interesting convergence and steady-state characteristics when compared to others from the literature [10], which makes it suitable for different practical applications. However, to the best of our knowledge, the performance and robustness of this algorithm have been assessed only through Monte Carlo (MC) simulations and through model expressions derived under very strong simplifying assumptions.

In this context, a stochastic model serves as a convenient tool for supporting a more rigorous and comprehensive theoretical analysis of an adaptive algorithm. Based on such a model, the behavior of an algorithm can be predicted (through model expressions) under different operating conditions, facilitating the identification of anomalous behavior, cause-and-effect relationships between parameters and performance metrics, as well as design guidelines for parameter tuning (for further details, see [1, 21–23, 26, 27, 32, 33]). So, focusing on the NP-VSS-NLMS [9] algorithm, the goals of this research work are:

- (i) to derive a stochastic model for the algorithm, considering a system identification problem and Gaussian input data;
- (ii) to obtain expressions for predicting the mean weight behavior, the evolution of the variable step-size parameter, learning curves, as well as some correlation-like matrices;
- (iii) to investigate the algorithm behavior in steady state based on the model expressions derived;

Fig. 1 Block diagram of a system identification problem



- (iv) to assess the model’s accuracy and algorithm characteristics for different operating conditions; and
- (v) to provide performance comparisons against other recent and relevant variable step-size algorithms from the literature.

The remainder of this paper is organized as follows. Section 2 introduces the problem formulation, revisiting the NP-VSS-NLMS algorithm and the operating scenario considered. Section 3 presents the proposed model, including assumptions, mathematical development, and design guidelines. Section 4 shows simulation results assessing the model’s accuracy and the algorithm’s performance. Section 5 presents concluding remarks and suggestions for future research works.

Throughout this paper, the adopted mathematical notation 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. Superscript $(\cdot)^T$ stands for the transpose operator, $\text{Tr}(\cdot)$ characterizes the trace operator, while $E(\cdot)$ denotes the expected value, and $P(\cdot)$ represents the probability of a random event.

2 Problem Formulation

Considering a system identification problem (see the block diagram shown in Fig. 1) [13–15, 34], let us assume that the output signal $y(n)$ of the system to be identified, which is corrupted by the measurement noise $v(n)$, can be expressed as

$$\begin{aligned} d(n) &= y(n) + v(n) \\ &= \mathbf{h}^T \mathbf{x}(n) + v(n) \end{aligned} \tag{1}$$

where \mathbf{h} denotes an L -dimensional (unknown) system impulse response, and

$$\mathbf{x}(n) = [x(n) \ x(n - 1) \ \cdots \ x(n - L + 1)]^T \tag{2}$$

is a vector containing the L most recent time samples of the input signal $x(n)$. Then, consider that the NP-VSS-NLMS [9] algorithm is used to update the adaptive filter weight vector $\hat{\mathbf{h}}(n)$ (of length L), through

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n - 1) + \mu(n) \frac{\mathbf{x}(n)e(n)}{\delta + \mathbf{x}^T(n)\mathbf{x}(n)} \tag{3}$$

with the error signal given by

$$e(n) = d(n) - \hat{\mathbf{h}}^T(n-1)\mathbf{x}(n) \quad (4)$$

and the step-size adjustment rule by

$$\mu(n) = \begin{cases} 1 - \frac{\hat{\sigma}_v}{\zeta + \hat{\sigma}_e(n)}, & \hat{\sigma}_e(n) \geq \hat{\sigma}_v \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

where $\hat{\sigma}_v^2$ denotes an estimate of the measurement noise variance, while

$$\hat{\sigma}_e^2(n) = \kappa \hat{\sigma}_e^2(n-1) + (1-\kappa)e^2(n) \quad (6)$$

is an estimate of the error signal variance with parameter $0 \ll \kappa < 1$ representing a forgetting factor. Note, in (3) and (6), that δ and ζ are the regularization parameters. Therefore, based on the setup described here, the proposed stochastic model can now be developed.

3 Proposed Model

Here, the proposed model describing the behavior of the NP-VSS-NLMS [9] algorithm is derived, which includes expressions for predicting the mean weight behavior, the evolution of the variable step-size parameter, learning curves, some correlation-like matrices related to the input data and weight-error vector, as well as analytical expressions characterizing the variable step size and MSE in steady state. Based on such expressions, some interesting characteristics of the algorithm behavior are further discussed, aiming to provide useful design guidelines. To this end, one considers the following commonly used assumptions and approximations to make the mathematical development tractable [13–15, 22, 23, 26, 27, 33, 34]:

- (A1) The input data are obtained through a real-valued Gaussian process with zero mean, variance σ_x^2 , and autocorrelation matrix $\mathbf{R} = \mathbf{E}[\mathbf{x}(n)\mathbf{x}^T(n)]$.
- (A2) The measurement noise $v(n)$ is obtained through a white Gaussian process of zero mean and variance σ_v^2 .
- (A3) The adaptive weight vector $\hat{\mathbf{h}}(n)$ is considered independent of any other variable in the system.
- (A4) The variable step-size parameter $\mu(n)$ is considered independent of any other variable in the system.
- (A5) The parameters δ and ζ are considered small, so their effect can be neglected.

These assumptions and approximations yield satisfactory results as shown later in Sect. 4.

3.1 Mean Weight Behavior

Substituting (1) and (4) into (3), applying the expected value, and using Assumptions (A2)–(A5) to simplify the resulting expression, we get

$$E[\widehat{\mathbf{h}}(n)] = \{\mathbf{I} - E[\mu(n)]\mathbf{R}_1\}E[\widehat{\mathbf{h}}(n-1)] + E[\mu(n)]\mathbf{R}_1\mathbf{h} \quad (7)$$

where \mathbf{I} is an $L \times L$ identity matrix, $E[\mu(n)]$ characterizes the mean behavior of the variable step size, and [32]

$$\mathbf{R}_1 = E \begin{bmatrix} \mathbf{x}(n)\mathbf{x}^T(n) \\ \mathbf{x}^T(n)\mathbf{x}(n) \end{bmatrix}. \quad (8)$$

So, using the results presented in [23, Appendix A] for computing the normalized correlation-like matrix \mathbf{R}_1 vis-à-vis (uncorrelated and correlated) Gaussian input data, the mean weight behavior of the adaptive filter can be predicted through (7) if $E[\mu(n)]$ is known.

3.2 Learning Curves

Rewriting (4) in terms of the weight-error vector:

$$\mathbf{v}(n) = \mathbf{h} - \widehat{\mathbf{h}}(n) \quad (9)$$

as

$$e(n) = \mathbf{v}^T(n-1)\mathbf{x}(n) + v(n) \quad (10)$$

calculating $e^2(n)$, applying the expected value, and using Assumption (A3), the following expression is obtained for predicting the evolution of the mean-square error (MSE) [13–15, 34]:

$$J(n) = J_{\min} + J_{\text{ex}}(n) \quad (11)$$

with the minimum MSE given by

$$J_{\min} = \sigma_v^2 \quad (12)$$

and the excess MSE (EMSE) by

$$\begin{aligned} J_{\text{ex}}(n) &= \text{Tr}[\mathbf{R}\mathbf{K}(n-1)] \\ &= \boldsymbol{\lambda}^T \mathbf{k}'(n-1) \end{aligned} \quad (13)$$

where vector λ contains the eigenvalues arising from the eigendecomposition $\mathbf{R} = \mathbf{Q}\Lambda\mathbf{Q}^T$ [13–15, 34],

$$\mathbf{K}(n) = E[\mathbf{v}(n)\mathbf{v}^T(n)] \quad (14)$$

denotes the autocorrelation matrix of the weight-error vector, while $\mathbf{k}'(n)$ represents a vector containing the diagonal elements of $\mathbf{K}'(n) = \mathbf{Q}^T\mathbf{K}(n)\mathbf{Q}$. Note that the mean-square deviation (MSD) [13–15, 34] can be straightforwardly determined [from (9)] as

$$m(n) = \text{Tr}[\mathbf{K}(n)]. \quad (15)$$

Therefore, based on (11)–(13) and (15), the (MSE, EMSE, and MSD) learning curves can be predicted if $\mathbf{k}'(n)$ is known.

3.3 Weight-Error Autocorrelation Matrix

Subtracting \mathbf{h} from both sides of (3), considering (9) and (10), determining $\mathbf{v}(n)\mathbf{v}^T(n)$, taking the expected value, using (14), and applying Assumptions (A2)–(A5) to simplify both sides of the resulting expression, one gets

$$\begin{aligned} \mathbf{K}(n) &= \mathbf{K}(n-1) - E[\mu(n)][\mathbf{K}(n-1)\mathbf{R}_1 + \mathbf{R}_1\mathbf{K}(n-1)] \\ &\quad + E[\mu^2(n)]\mathbf{R}_2(n) + E[\mu^2(n)]\mathbf{R}_3\sigma_v^2 \end{aligned} \quad (16)$$

with the normalized correlation-like matrices $\mathbf{R}_2(n)$ and $\mathbf{R}_3(n)$ [32] given as

$$\mathbf{R}_2(n) = E \left\{ \frac{\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{K}(n-1)\mathbf{x}(n)\mathbf{x}^T(n)}{[\mathbf{x}^T(n)\mathbf{x}(n)]^2} \right\} \quad (17)$$

and

$$\mathbf{R}_3 = E \left\{ \frac{\mathbf{x}(n)\mathbf{x}^T(n)}{[\mathbf{x}^T(n)\mathbf{x}(n)]^2} \right\}. \quad (18)$$

Then, pre- and post-multiplying (16) by \mathbf{Q}^T and \mathbf{Q} , respectively, taking the diagonal elements, and considering the results presented in [22] for computing \mathbf{R}_1 , \mathbf{R}_2 , and \mathbf{R}_3 , we obtain

$$\mathbf{k}'(n) = \{\mathbf{I} - 2E[\mu(n)]\mathbf{H} + E[\mu^2(n)](2\mathbf{T} + \mathbf{P})\}\mathbf{k}'(n-1) + E[\mu^2(n)]\mathbf{s}\sigma_v^2 \quad (19)$$

where the diagonal matrices \mathbf{H} and \mathbf{T} , the full matrix \mathbf{P} , as well as the vector \mathbf{s} are determined as in [23, Appendix A] for (uncorrelated and correlated) Gaussian input data. Therefore, the evolution of $\mathbf{k}'(n)$ can be computed if $E[\mu(n)]$ and $E[\mu^2(n)]$ are known, thus making it possible to predict (in a recursive way) the mean-square behavior of the algorithm.

3.4 Variable Step-Size Behavior

Using Assumption (A5) and introducing the quantity:

$$\delta_e(n) = \hat{\sigma}_e^2(n) - \mathbb{E}[\hat{\sigma}_e^2(n)] \quad (20)$$

to rewrite the nonzero part of (5) as

$$\tilde{\mu}(n) \cong 1 - \frac{\hat{\sigma}_v}{\sqrt{\mathbb{E}[\hat{\sigma}_e^2(n)] + \delta_e(n)}} \quad (21)$$

one can write the expected value of the variable step size and its squared version as [31, 36]

$$\begin{aligned} \mathbb{E}[\mu^i(n) | \mu(n) > 0] &= \int_{-\infty}^{\infty} \mu^i(n) f[\delta_e(n)] d\delta_e(n) \\ &= \int_{\hat{\sigma}_v^2 - \mathbb{E}[\hat{\sigma}_e^2(n)]}^{\infty} \tilde{\mu}^i(n) f[\delta_e(n)] d\delta_e(n) \end{aligned} \quad (22)$$

for $i = 1, 2$ where the probability density function (PDF) of $\delta_e(n)$ is assumed to follow a Gaussian distribution, i.e.,

$$f[\delta_e(n)] \cong \frac{1}{\sqrt{2\pi\mathbb{E}[\delta_e^2(n)]}} \exp\left\{\frac{-\delta_e^2(n)}{2\mathbb{E}[\delta_e^2(n)]}\right\} \quad (23)$$

due to the central limit theorem [31, 36]. Next, we resort to second-order Taylor polynomials [38] to approximate $\tilde{\mu}^i(n)$ for $i = 1, 2$ [given by (21)] as

$$\tilde{\mu}(n) \cong 1 - \frac{\hat{\sigma}_v}{\sqrt{\mathbb{E}[\hat{\sigma}_e^2(n)]}} \left\{ 1 - \frac{1}{2} \frac{\delta_e(n)}{\mathbb{E}[\hat{\sigma}_e^2(n)]} + \frac{3}{8} \frac{\delta_e^2(n)}{\mathbb{E}[\hat{\sigma}_e^2(n)]^2} \right\} \quad (24)$$

and

$$\begin{aligned} \tilde{\mu}^2(n) &\cong 1 - \frac{2\hat{\sigma}_v}{\sqrt{\mathbb{E}[\hat{\sigma}_e^2(n)]}} \left\{ 1 - \frac{1}{2} \frac{\delta_e(n)}{\mathbb{E}[\hat{\sigma}_e^2(n)]} + \frac{3}{8} \frac{\delta_e^2(n)}{\mathbb{E}[\hat{\sigma}_e^2(n)]^2} \right\} \\ &\quad + \frac{\hat{\sigma}_v^2}{\mathbb{E}[\hat{\sigma}_e^2(n)]} \left\{ 1 - \frac{\delta_e(n)}{\mathbb{E}[\hat{\sigma}_e^2(n)]} + \frac{\delta_e^2(n)}{\mathbb{E}[\hat{\sigma}_e^2(n)]^2} \right\}. \end{aligned} \quad (25)$$

Then, substituting (23) and (24) into (22) and solving the resulting expression, we have

$$\begin{aligned} \mathbb{E}[\mu(n) | \mu(n) > 0] &= \mathbb{P}[\mu(n) > 0] - \frac{\hat{\sigma}_v}{\sqrt{\mathbb{E}[\hat{\sigma}_e^2(n)]}} \times \\ &\quad \left\{ \mathbb{P}[\mu(n) > 0] - \frac{1}{2} \frac{\mathbb{E}[\delta_e(n) | \mu(n) > 0]}{\mathbb{E}[\hat{\sigma}_e^2(n)]} + \frac{3}{8} \frac{\mathbb{E}[\delta_e^2(n) | \mu(n) > 0]}{\mathbb{E}[\hat{\sigma}_e^2(n)]^2} \right\} \end{aligned} \quad (26)$$

while, from (25),

$$\begin{aligned}
 E[\mu^2(n)|\mu(n) > 0] &= P[\mu(n) > 0] - \frac{2\hat{\sigma}_v}{\sqrt{E[\hat{\sigma}_e^2(n)]}} \times \\
 &\left\{ P[\mu(n) > 0] - \frac{1}{2} \frac{E[\delta_e(n)|\mu(n) > 0]}{E[\hat{\sigma}_e^2(n)]} + \frac{3}{8} \frac{E[\delta_e^2(n)|\mu(n) > 0]}{E[\hat{\sigma}_e^2(n)]^2} \right\} \\
 &+ \frac{\hat{\sigma}_v^2}{E[\hat{\sigma}_e^2(n)]} \left\{ P[\mu(n) > 0] - \frac{E[\delta_e(n)|\mu(n) > 0]}{E[\hat{\sigma}_e^2(n)]} + \frac{E[\delta_e^2(n)|\mu(n) > 0]}{E[\hat{\sigma}_e^2(n)]^2} \right\}
 \end{aligned} \tag{27}$$

with

$$\begin{aligned}
 P[\mu(n) > 0] &= \frac{1}{2} \operatorname{erfc} \left\{ \frac{\hat{\sigma}_v^2 - E[\hat{\sigma}_e^2(n)]}{\sqrt{2E[\delta_e^2(n)]}} \right\} \\
 E[\delta_e(n)|\mu(n) > 0] &= \sqrt{\frac{E[\delta_e^2(n)]}{2\pi}} \exp \left(-\frac{\{\hat{\sigma}_v^2 - E[\hat{\sigma}_e^2(n)]\}^2}{2E[\delta_e^2(n)]} \right)
 \end{aligned} \tag{28}$$

$$\tag{29}$$

and

$$E[\delta_e^2(n)|\mu(n) > 0] = \{\hat{\sigma}_v^2 - E[\hat{\sigma}_e^2(n)]\}E[\delta_e(n)|\mu(n) > 0] + E[\delta_e^2(n)]P[\mu(n) > 0]. \tag{30}$$

Finally, notice that (26)–(30) require knowledge of $E[\hat{\sigma}_e^2(n)]$ and $E[\delta_e^2(n)]$. So, taking the expected value, one gets from (6) that

$$E[\hat{\sigma}_e^2(n)] = \kappa E[\hat{\sigma}_e^2(n - 1)] + (1 - \kappa)J(n). \tag{31}$$

In turn, substituting (6) and (31) into (20), squaring both sides of the resulting expression, taking the expected value, and approximating $E[\delta_e(n - 1)e^2(n)] \cong E[\delta_e(n - 1)]E[e^2(n)] = 0$, it is possible to show that

$$E[\delta_e^2(n)] = \kappa^2 E[\delta_e^2(n - 1)] + (1 - \kappa)^2 \{E[e^4(n)] - J^2(n)\} \tag{32}$$

where [from (10)]

$$\begin{aligned}
 E[e^4(n)] &\cong 2\boldsymbol{\lambda}^T \mathbf{K}^{\prime 2} (n - 1) \boldsymbol{\lambda} + [\boldsymbol{\lambda}^T \mathbf{k}'(n - 1)]^2 \\
 &+ 6\sigma_v^2 \boldsymbol{\lambda} \mathbf{k}'(n - 1) + 3\sigma_v^4
 \end{aligned} \tag{33}$$

due to the factorization theorem of a fourth-order moment of Gaussian variables [31, 36]. Still, by substituting (33) into (32) and approximating $\boldsymbol{\lambda}^T \mathbf{K}^{\prime 2} (n - 1) \boldsymbol{\lambda} \cong J_{\text{ex}}^2(n)$, (32) can be further simplified to

$$E[\delta_e^2(n)] \cong \kappa^2 E[\delta_e^2(n - 1)] + 2(1 - \kappa)^2 J^2(n). \tag{34}$$

So, given that $E[\mu(n)]$ and $E[\mu^2(n)]$ can be computed recursively through (26)–(31) and (34), the behavior of the algorithm is completely characterized now during the transient phase.

3.5 Variable Step Size in Steady State

Assuming convergence, letting $n \rightarrow \infty$, considering from (31) and (34) that

$$E[\widehat{\sigma}_e^2(\infty)] = J(\infty) \quad (35)$$

and

$$E[\delta_e^2(\infty)] = 2\rho J^2(\infty) \quad (36)$$

hold, and substituting into (28), (29), and (30), we have [from (26)] that

$$E[\mu(\infty)|\mu(\infty) > 0] = \frac{1}{2} \operatorname{erfc}\left(\frac{\tau}{2\sqrt{\rho}}\right) \left(1 - \frac{\widehat{\sigma}_v}{\sqrt{J(\infty)}} - \frac{3\rho\widehat{\sigma}_v}{4\sqrt{J(\infty)}}\right) - \frac{\widehat{\sigma}_v}{2\sqrt{J(\infty)}} \sqrt{\frac{\rho}{2\pi}} \exp\left(-\frac{\tau^2}{4\rho}\right) \left(\frac{3}{4}\tau - 1\right) \quad (37)$$

while, from (27),

$$E[\mu^2(\infty)|\mu(\infty) > 0] = \frac{1}{2} \operatorname{erfc}\left(\frac{\tau}{2\sqrt{\rho}}\right) \left[1 - \frac{2\widehat{\sigma}_v}{\sqrt{J(\infty)}} \left(1 + \frac{3\rho}{4}\right) + \frac{\widehat{\sigma}_v^2}{J(\infty)} (1 + 2\rho)\right] + \sqrt{\frac{\rho}{2\pi}} \exp\left(-\frac{\tau^2}{4\rho}\right) \left[\frac{\widehat{\sigma}_v^2}{J(\infty)} (\tau - 1) - \frac{\widehat{\sigma}_v}{\sqrt{J(\infty)}} \left(\frac{3}{4}\tau - 1\right)\right] \quad (38)$$

in which

$$\rho = \frac{1 - \kappa}{1 + \kappa} \quad (39)$$

and

$$\tau = \frac{\widehat{\sigma}_v^2 - J(\infty)}{J(\infty)}. \quad (40)$$

So, the variable step size $E[\mu(\infty)]$ and its squared version $E[\mu^2(\infty)]$ in steady state can be predicted from (37)–(40) if $J(\infty)$ is known.

3.6 MSE in Steady State

Using (9) and (10) to express (3) in terms of $\mathbf{v}(n)$, determining $\mathbf{v}^T(n)\mathbf{v}(n)$, applying the expected value, making $n \rightarrow \infty$ on both sides, considering Assumptions (A2)–(A5) to simplify the resulting expression, and approximating [12]

$$\mathbf{R}_1 \cong E \left[\frac{1}{\mathbf{x}^T(n)\mathbf{x}(n)} \right] \mathbf{R} \tag{41}$$

one obtains

$$- 2E[\mu(\infty)|\mu(\infty) > 0][J(\infty) - \sigma_v^2] + E[\mu^2(\infty)|\mu(\infty) > 0]J(\infty) = 0 \tag{42}$$

with $E[\mu(\infty)|\mu(\infty) > 0]$ and $E[\mu^2(\infty)|\mu(\infty) > 0]$ depending intrinsically on $J(\infty)$. Now, substituting (37) and (38) into (42), we get

$$J^{3/2}(\infty) - \sigma_v^2[2 + \alpha\gamma(\epsilon)]J^{1/2}(\infty) + 2\sqrt{\alpha}\beta(\epsilon)\sigma_v^3 = 0 \tag{43}$$

in which ϵ denotes an auxiliary variable included to make it possible to obtain an approximate solution through a perturbation method [16],

$$\alpha = \frac{\hat{\sigma}_v^2}{\sigma_v^2} \tag{44}$$

defines the ratio between the estimate of the measurement noise variance and its true value, while

$$\beta(\epsilon) = 1 + \frac{3}{4}\rho + \epsilon\sqrt{\frac{\rho}{\pi}} \left(\frac{3}{4}\tau - 1 \right) g \left(\frac{\tau}{2\sqrt{\rho}} \right) \tag{45}$$

$$\gamma(\epsilon) = 1 + 2\rho + 2\epsilon\sqrt{\frac{\rho}{\pi}} (\tau - 1) g \left(\frac{\tau}{2\sqrt{\rho}} \right) \tag{46}$$

and

$$g(z) = \frac{\exp(-z^2)}{\operatorname{erfc}(z)}. \tag{47}$$

Note that (43) leads to the original problem for $\epsilon = 1$ and to a cubic equation for $\epsilon = 0$, whose appropriate solution is given by [19]

$$\begin{aligned} \sqrt{J_\infty(0)} &= \sigma_v \left(\frac{1}{2} + i \frac{\sqrt{3}}{2} \right)^3 \sqrt[3]{\sqrt{\alpha}\beta(0) - \sqrt{\alpha\beta^2(0) - \frac{[2 + \alpha\gamma(0)]^3}{27}}} \\ &+ \sigma_v \left(\frac{1}{2} - i \frac{\sqrt{3}}{2} \right)^3 \sqrt[3]{\sqrt{\alpha}\beta(0) + \sqrt{\alpha\beta^2(0) - \frac{[2 + \alpha\gamma(0)]^3}{27}}} \end{aligned} \tag{48}$$

with $J_\infty(\epsilon)$ being a differentiable function. So, taking the implicit derivative with respect to ϵ in (43), making $\epsilon = 0$, and solving the resulting expression, we have

$$\begin{aligned} \left. \frac{dJ_\infty(\epsilon)}{d\epsilon} \right|_{\epsilon=0} &= \frac{2\sigma_v^2 J_\infty^{3/2}(0)}{J_\infty^{3/2}(0) - \sqrt{\alpha}\sigma_v^3} \sqrt{\frac{\rho}{\pi}} g\left(\frac{\tau_0}{2\sqrt{\rho}}\right) \\ &\times \left\{ \tau_0 \left[\alpha - \frac{3}{4} \frac{\sqrt{\alpha}\sigma_v}{J_\infty^{1/2}(0)} \right] - \left[\alpha - \frac{\sqrt{\alpha}\sigma_v}{J_\infty^{1/2}(0)} \right] \right\} \end{aligned} \tag{49}$$

with

$$\tau_0 = \frac{\widehat{\sigma}_v^2 - J_\infty(0)}{J_\infty(0)}. \tag{50}$$

Based on (48) and (49), the steady-state MSE can be approximated as

$$J(\infty) \cong J_\infty(0) + \left. \frac{dJ_\infty(\epsilon)}{d\epsilon} \right|_{\epsilon=0} \tag{51}$$

which holds well for

$$\alpha \leq \frac{4 - 3\sqrt{\rho\pi}}{4(1 - \sqrt{\rho\pi})}. \tag{52}$$

Therefore, based on (37)–(40) and (44)–(51), the algorithm behavior in steady state is completely characterized now.

4 Simulation Results

In this section, the model’s accuracy and the algorithm’s performance are assessed through MC simulations (average of 200 independent runs), highlighting some interesting characteristics. To this end, four examples are presented and discussed, considering systems with different lengths, distinct input data correlation levels, several values of signal-to-noise ratio (SNR), as well as different values for parameters $\widehat{\sigma}_v^2$ and κ [required in (5) and (6)]. In particular, the system impulse responses \mathbf{h}_1 (with $L = 64$ weights) and \mathbf{h}_2 (with $L = 128$ weights) are obtained from echo path models for testing of speech echo cancelers given in the ITU-T G.168 Recommendation [18, Models 1 and 4] (as depicted in Fig. 2). In turn, the input signal $x(n)$ is generated through [15]

$$x(n) = -a_1x(n - 1) - a_2x(n - 2) + w(n) \tag{53}$$

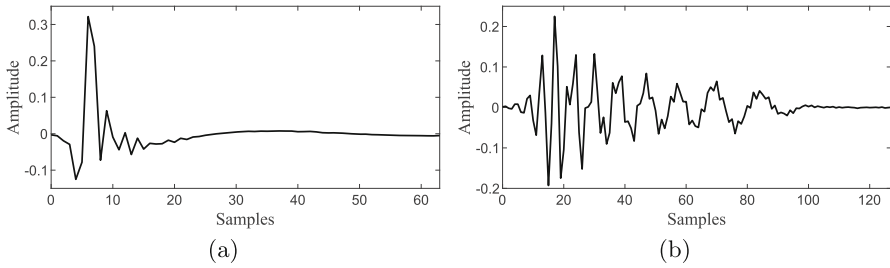


Fig. 2 System impulse responses obtained from echo path models for testing of speech echo cancelers given in the ITU-T G.168 Recommendation [18]. **a** System impulse response \mathbf{h}_1 with $L = 64$ weights (based on [18, Model 1]). **b** System impulse response \mathbf{h}_2 with $L = 128$ weights (based on [18, Model 4])

with a_1 and a_2 being the autoregressive coefficients, and $w(n)$ a white Gaussian noise whose variance

$$\sigma_w^2 = \sigma_x^2 \left(\frac{1 - a_2}{1 + a_2} \right) [(1 + a_2)^2 - a_1^2] \quad (54)$$

is adjusted such that $\sigma_x^2 = 1$. The SNR is defined (in dB) as [5, 8]

$$\text{SNR} = 10 \log_{10} \left(\frac{\sigma_y^2}{\sigma_v^2} \right) \quad (55)$$

with $\sigma_y^2 = \mathbf{h}^T \mathbf{R} \mathbf{h}$ characterizing the variance of $y(n)$ (see Fig. 1). Lastly, the algorithm variables are initialized as $\hat{\mathbf{h}}(0) = [1 \ 0 \ \dots \ 0]^T$ and $\hat{\sigma}_e^2(0) = 0$, while regularization parameters are set as $\zeta = \hat{\sigma}_v/1000$ and $\delta = 10^{-3}$.

4.1 Example 1

Here, the proposed model's accuracy is assessed for both uncorrelated and correlated input data as well as different SNR values. In particular, we consider two eigenvalue spread values for the input data autocorrelation matrix, i.e., $\chi = 1$ [obtained from (53) for $a_1 = a_2 = 0$] and $\chi = 454.09$ [obtained from (53) for $a_1 = -0.5$ and $a_2 = 0.9$]. Three SNR values are used, i.e., 20, 30, and 40 dB. The system impulse response \mathbf{h}_1 with $L = 64$ weights (see Fig. 2a) is used, while the remaining algorithm parameters are $\kappa = 0.95$ and $\hat{\sigma}_v^2 = \sigma_v^2$ (i.e., a perfect estimate of the measurement noise variance).

Figure 3 shows the results obtained for this operating scenario. In particular, Figs. 3a and b present the mean behavior of four adaptive weights, Fig. 3c and d depict the evolution of the variable step size, while Fig. 3e and f illustrate the MSE learning curves. These figures show a very good match between MC simulations and model predictions, during the transient phase as well as in the steady state, irrespective of the input data correlation level and SNR considered. Moreover, one observes, from such figures, that $\mu(n)$ plays an important role on the evolution of the adaptive weights and MSE learning curves. Specifically, $\mu(n) \rightarrow 1$ at the beginning of the adaptation

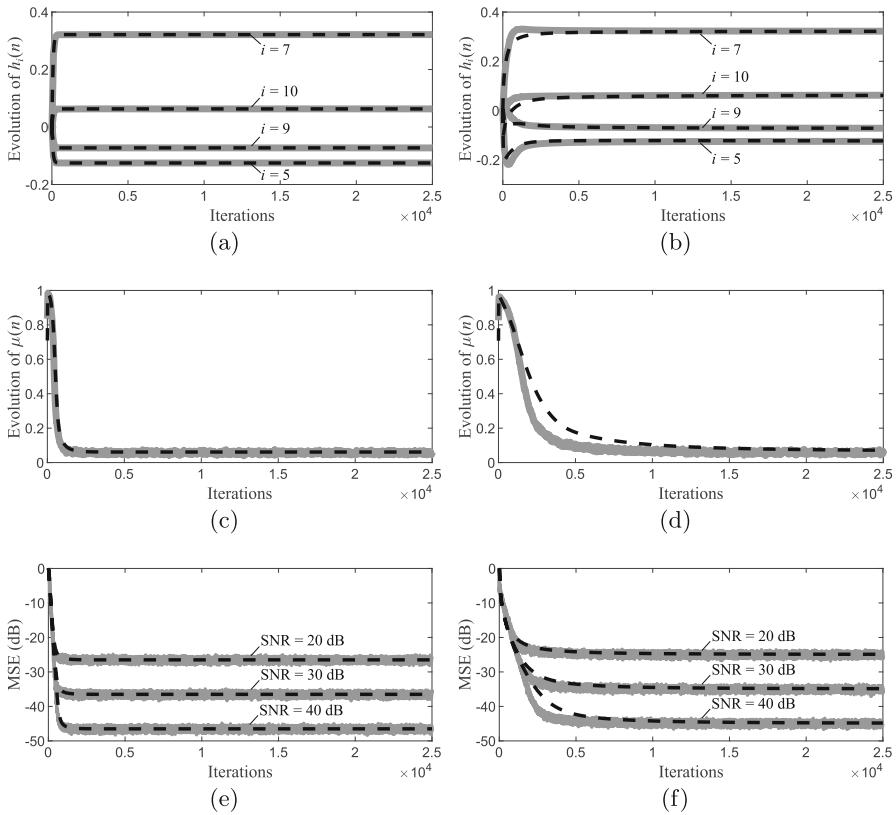


Fig. 3 Example 1. Results obtained from MC simulations (gray-ragged lines) and predicted from the proposed model (dark-dashed lines), considering uncorrelated $\chi = 1$ (left) and correlated $\chi = 454.09$ (right) Gaussian input data. **a, b** Evolution of (four) weights of the adaptive filter. **c, d** Evolution of the variable step size. **e, f** Evolution of the MSE learning curve

process, speeding up the algorithm convergence; in turn, $\mu(n) \rightarrow 0$ as the algorithm converges, aiming to reduce the EMSE in steady state. Nevertheless, as opposed to [9, Sec. II-C], we verify that $\mu(n)$ tends to a small positive value as $n \rightarrow \infty$, preventing any further reduction in the EMSE in steady state. Such effect is captured by the presence of $\delta_e(n)$ in the model expressions; so, finding ways to suppress the effect of $\delta_e(n)$ can result in improved versions of the algorithm. Therefore, one concludes that the proposed model may be successfully used to gain insights into the algorithm behavior without resorting exclusively to extensive simulations and to support the development of improved algorithms.

4.2 Example 2

Now, the accuracy of the proposed model is verified (via EMSE learning curve) considering different lengths L for the system impulse response and distinct values for the smoothing parameter κ . For such, both system impulse responses \mathbf{h}_1 with $L = 64$

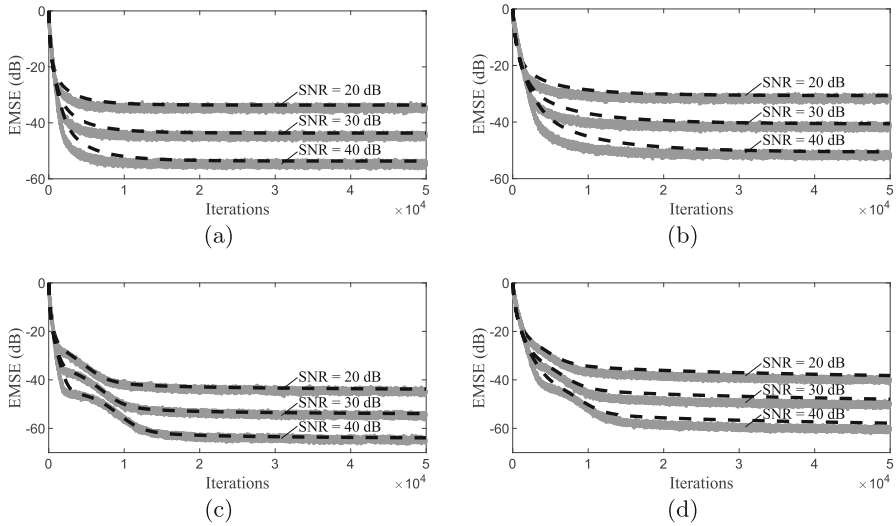


Fig. 4 Example 2. Results obtained from MC simulations (gray-ragged lines) and predicted from the proposed model (dark-dashed lines), considering a system impulse response with $L = 64$ (left) and $L = 128$ (right) weights. **a, b** Smoothing parameter $\kappa = 0.9$. **c, d** Smoothing parameter $\kappa = 0.999$

and \mathbf{h}_2 with $L = 128$ (depicted in Fig. 2) and two values for the smoothing parameter $\kappa = \{0.9, 0.999\}$ are used. Also, we consider correlated input data [obtained from (53) with $a_1 = -0.6$ and $a_2 = 0.8$], yielding an eigenvalue spread of $\chi = 144.78$ (for $L = 64$) and $\chi = 156.40$ (for $L = 128$) for the input autocorrelation matrix. Still, it is assumed perfect estimate of the measurement noise variance (i.e., $\hat{\sigma}_v^2 = \sigma_v^2$) and SNR values of 20, 30, and 40 dB.

Figure 4 depicts the results obtained for this operating scenario. Specifically, EMSE learning curves obtained for different SNR values are presented in Fig. 4a assuming $L = 64$ and $\kappa = 0.9$, Fig. 4b with $L = 128$ and $\kappa = 0.9$, Fig. 4c for $L = 64$ and $\kappa = 0.999$, while Fig. 4d considering $L = 128$ and $\kappa = 0.999$. Notice, comparing either Fig. 4a and c or Fig. 4b and d, that smaller values of steady-state EMSE are achieved when κ is increased. Nevertheless, values of κ close to 1 slow down the update of $\hat{\sigma}_e^2(n)$ [see (6)], maintaining $\mu(n)$ [see (5)] high for too long; as a consequence, one verifies a plateau in the EMSE learning curves after the initial transient phase. Despite these characteristics of the algorithm behavior, one observes that the model predictions exhibit a very good match with MC simulations, during the transient phase and in steady state, irrespective of the length L of the system impulse response, the value selected for the smoothing parameter κ , the correlation level of the input data, and/or the SNR considered. So, one concludes that the proposed model may be considered to investigate the impact of the algorithm parameters on its performance without relying only on trial-and-error procedures.

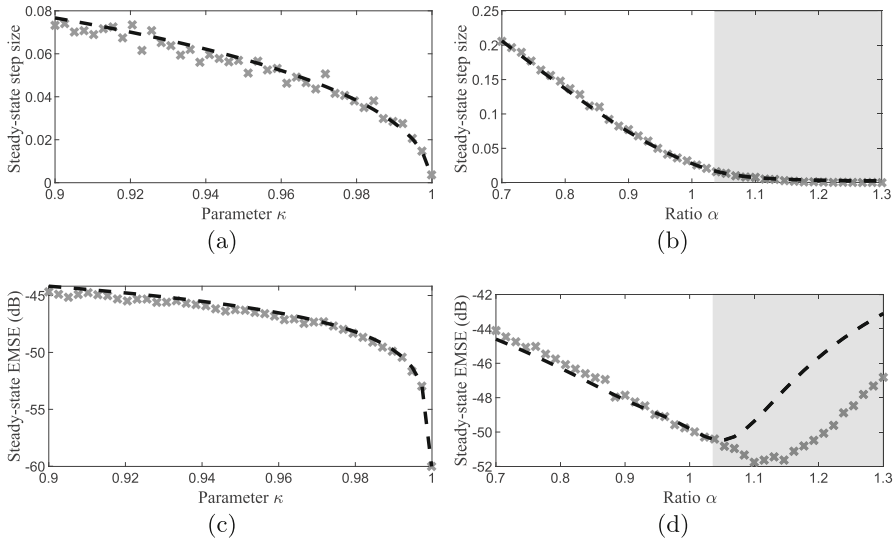


Fig. 5 Example 3. Results obtained from MC simulations (gray-cross markers) and predicted from the proposed model (dark-dashed lines), considering different values of κ with $\alpha = 1$ (left) and different values of α with $\kappa = 0.99$ (right). **a, b** Steady-state value of the variable step size. **c, d** Steady-state EMSE. [Condition (52) is depicted (as a shaded area) in **b** and **d**.]

4.3 Example 3

This example aims to verify the accuracy of expressions describing the steady-state algorithm behavior (i.e., variable step size and EMSE in steady state), as a function of the smoothing parameter κ and the ratio α between the estimate of the measurement noise variance and its true value [see (44)]. To this end, we assume either (i) different values for parameter κ ranging from 0.9 to 0.9999 while $\alpha = 1$ is kept fixed; or (ii) the ratio α ranging from 0.7 to 1.3 while $\kappa = 0.99$ is fixed. Note that, following the approach described in [34, pp. 250], we have averaged the last 100 values in steady state for each variable of interest to visualize better the experimental results. The remaining parameter values are the same as in Example 2 with $L = 128$, except for the SNR which is assumed here equal to 30 dB for simplicity.

Figure 5 exhibits the results obtained for this operating scenario. Particularly, Fig. 5a and c present curves characterizing the step size and EMSE in steady state as a function of κ with $\alpha = 1$ (fixed), while Fig. 5b and d depict curves assuming $\kappa = 0.99$ (fixed) and varying α . Notice, from Fig. 5a and c, that model predictions match satisfactorily the experimental results over a wide range of values of κ , when the measurement noise variance is perfectly estimated (i.e., when $\hat{\sigma}_v^2 = \sigma_v^2$ which implies $\alpha = 1$). In turn, although Fig. 5b confirms that the steady-state step-size value can be satisfactorily predicted, observe from Fig. 5d that the model expression describing the steady-state EMSE fails as α increases above the condition given in (52) (i.e., when $\hat{\sigma}_v^2$ overestimates σ_v^2). Despite these aspects, one

verifies from Fig. 5c that the EMSE achieved in steady state decreases as $\kappa \rightarrow 1$. In addition, Fig. 5d highlights how an imperfect estimate of the measurement noise variance affects the steady-state EMSE achieved; in other words, the steady-state EMSE increases, whereas $\hat{\sigma}_v^2$ moves away from σ_v^2 , thus affecting the algorithm performance. Hence, one concludes that parameters κ and $\hat{\sigma}_v^2$ must be determined carefully.

4.4 Example 4

In this last example, performance comparisons are conducted between the NP-VSS-NLMS algorithm and the algorithms given in [11, 17, 28, 37, 39, 42]. For such, we employ the system impulse response \mathbf{h}_2 with $L = 128$ (see Fig. 2b), correlated input data [obtained from (53) with $a_1 = -0.6$ and $a_2 = 0.8$] with eigenvalue spread of $\chi = 156.40$ for the input autocorrelation matrix, and SNR values of 20 and 40 dB. Parameters of the algorithms are adjusted manually for achieving the same steady-state EMSE, in order to obtain fair comparisons [29]. In particular, the parameter values are: $\kappa = 0.99$ for the NP-VSS-NLMS; $\alpha = 0.99$, $\gamma = 0.01$, $a = 0.99$, $b = 0.9993$, $A_0 = 0$, and $B_0 = 0.01$ for [42]; $\gamma = 0.994$, $\eta_0 = 0$, and $\rho_0 = 0.1$ for [28]; $\alpha = 0.99$, $\beta = 15$, and $\zeta_{th} = 0.35$, $\mu_0 = 1$, and $\mu_{max} = 1$ for [17]; $m_0 = 2\hat{\sigma}_v^2$ and $\sigma_{w,0} = 0$ for [11]; $\mu_{ref} = 1$, $\gamma = 8$, $\alpha^+ = 4$, and $\alpha_0 = 0$ for [37]; while $\lambda = 0.8$ for 20 dB and 0.99992 for 40 dB SNR, $\epsilon = 10^{-12}$, and $\sigma_{e,0}^2 = 0$ for [39]. Perfect estimate of the measurement noise variance (i.e., $\hat{\sigma}_v^2 = \sigma_v^2$) is assumed. Tracking performance is assessed by multiplying \mathbf{h}_2 by -1 after the algorithms have reached convergence.

Figure 6 presents performance comparisons involving the NP-VSS-NLMS algorithm and other recent and relevant algorithms from the literature. Specifically, EMSE learning curves obtained from the NP-VSS-NLMS and from the algorithms given in [11, 17, 28, 37, 39, 42] are depicted in Fig. 6a–f, respectively. Notice that the NP-VSS-NLMS algorithm outperforms the algorithms given in [17, 28, 39, 42] in terms of convergence speed when algorithms are properly adjusted to reach the same steady-state EMSE. Although the NP-VSS-NLMS exhibits similar transient characteristics in comparison to the algorithm given in [11], it is observed that the EMSE achieved by this latter continues to reduce as $n \rightarrow \infty$. Lastly, one verifies that both the NP-VSS-NLMS and the algorithm given in [37] demonstrate comparable performance; nevertheless, the latter requires the fine tuning of more parameters. Therefore, even when compared to other variable step-size algorithms from the literature, we conclude that the NP-VSS-NLMS is suitable in practice, requiring the tuning of fewer parameters for proper operation.

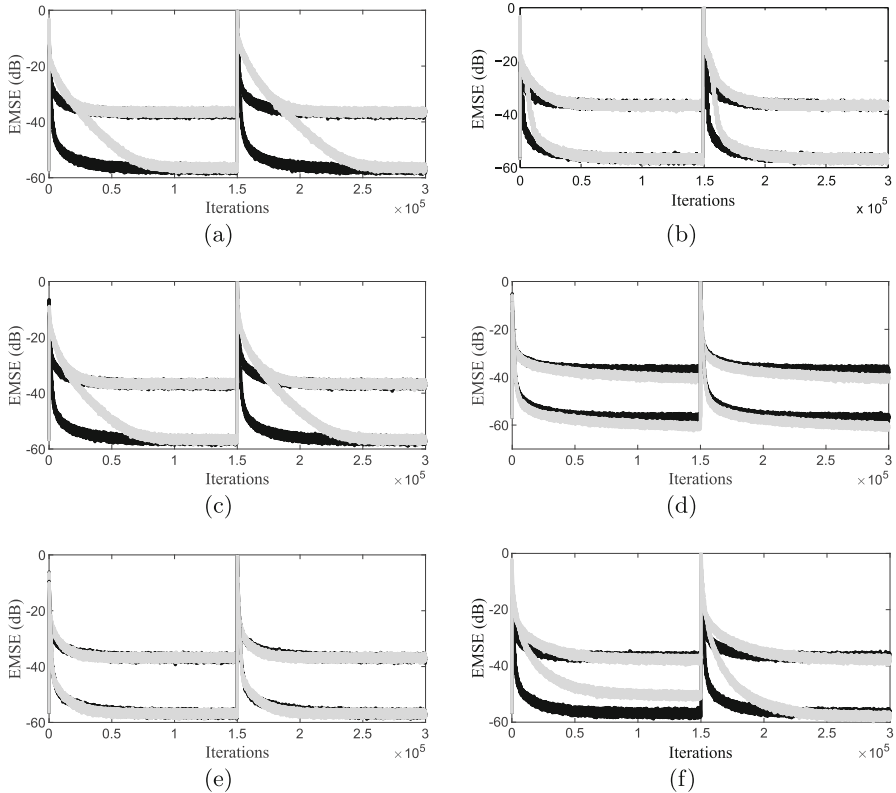


Fig. 6 Example 4. Results obtained from MC simulations comparing the performance of the NP-VSS-NLMS algorithm (dark-ragged lines) with important algorithms from the literature (gray-ragged lines) for the same steady-state EMSE. **a** Algorithm given in [42]. **b** Algorithm given in [28]. **c** Algorithm given in [17]. **d** Algorithm given in [11]. **e** Algorithm given in [37]. **f** Algorithm given in [39]

5 Conclusions

In this paper, a stochastic model for the NP-VSS-NLMS algorithm was presented. Specifically, model expressions were derived for predicting the algorithm behavior in the transient phase and the steady state, considering a system identification problem and (uncorrelated and correlated) Gaussian input data. These expressions revealed interesting algorithm characteristics and provided some design guidelines. For instance, it was observed that small steady-state EMSE values are achieved by making the smoothing parameter close to 1. Also, it was found that an imperfect estimate of the measurement noise variance significantly affects the steady-state EMSE achieved by the algorithm. Simulation results for various operating scenarios attested both the model's accuracy and the algorithm's superior performance relative to other recent and relevant algorithms from the literature. Note that, based on the obtained results, further research works could address the derivation of models for other algo-

rithms which incorporate rules for estimating the measurement noise variance as well as the development of improved variable step-size algorithms.

Acknowledgements The authors are thankful to the Associate Editor and the anonymous reviewers whose valuable comments and constructive suggestions have significantly benefited this paper.

Author Contributions ACB was involved in formal analysis, investigation, methodology, software, writing – original draft, and writing – review & editing. EVK was involved in conceptualization, supervision, methodology, software, writing – original draft, and writing – review & editing. MVM was involved in investigation, methodology, validation, writing – original draft, and writing – review & editing. JB was involved in conceptualization, supervision, project administration, and writing – review & editing.

Data and Code Availability Data and/or code generated during the development of the current study are available from the corresponding author on request.

Declarations

Competing Interests and Funding The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. M.S.E. Abadi, J.H. Husøy, On the application of a unified adaptive filter theory in the performance prediction of adaptive filter algorithms. *Digit. Signal Process.* **19**(3), 410–432 (2009)
2. T. Aboulnasr, K. Mayyas, A robust variable step-size LMS-type algorithm: analysis and simulations. *IEEE Trans. Signal Process.* **45**(3), 631–639 (1997)
3. A.E. Albert, L.S. Gardner Jr., *Stochastic Approximation and Nonlinear Regression*, 1st edn. (MIT Press, Cambridge, MA, 1967)
4. K.J. Aström, B. Wittenmark, *Adaptive Control*, 2nd edn. (Dover Publications, Mineola, NY, 2008)
5. K.J. Bakri, E.V. Kuhn, R. Seara, J. Benesty, C. Paleologu, S. Ciochină, On the stochastic modeling of the LMS algorithm operating with bilinear forms. *Digit. Signal Process.* **122**, 103359 (2022)
6. J. Benesty, I. Cohen, J. Chen, *Array Processing-Kronecker Product Beamforming* (Springer, Cham, 2019)
7. J. Benesty, T. Gänslér, D.R. Morgan, M.M. Sondhi, S.L. Gay, *Advances in Network and Acoustic Echo Cancellation* (Springer, Berlin, 2001)
8. J. Benesty, C. Paleologu, S. Ciochină, E.V. Kuhn, K.J. Bakri, R. Seara, LMS and NLMS algorithms for the identification of impulse responses with intrinsic symmetric or antisymmetric properties. In *Proceedings of the IEEE International Conference on Acoustic, Speech, Signal Process (ICASSP)*, pp. 5662–5666, Singapore, Singapore, May (2022)
9. J. Benesty, H. Rey, L.R. Vega, S. Tressens, A nonparametric VSS-NLMS algorithm. *IEEE Signal Process. Lett.* **13**, 581–584 (2006)
10. D. Bismor, K. Czyz, Z. Ogonowski, Review and comparison of variable step-size LMS algorithms. *Int. J. Acoust. Vib.* **21**(1), 24–39 (2016)
11. S. Ciochina, C. Paleologu, J. Benesty, An optimized NLMS algorithm for system identification. *Signal Process.* **118**, 115–121 (2016)
12. M.H. Costa, J.C.M. Bermudez, An improved model for the normalized LMS algorithm with Gaussian inputs and large number of coefficients. In *Proceedings of the IEEE International Conference on Acoustic, Speech, Signal Processing (ICASSP)*, volume 2, pp. 1385–1388, Orlando, FL, May (2002)
13. P.S.R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*, 4th edn. (Springer, New York, NY, 2013)
14. B. Farhang-Boroujeny, *Adaptive Filters: Theory and Applications*, 2nd edn. (Wiley, Chichester, 2013)
15. S. Haykin, *Adaptive Filter Theory*, 5th edn. (Prentice Hall, Upper Saddle River, NJ, 2014)
16. M.H. Holmes, *Introduction to Perturbation Methods*, 2nd edn. (Springer, New York, NY, 2013)
17. H.-C. Huang, J. Lee, A new variable step-size NLMS algorithm and its performance analysis. *IEEE Trans. Signal Process.* **60**(4), 2055–2060 (2012)

18. ITU-T Recommendation G.168 – Digital Network Echo Cancellers. Switzerland, International Telecommunications Union - Telecommunication Standardization Sector, Geneva (2015)
19. A. Jeffrey, H.-H. Dai, *Handbook of Mathematical Formulas and Integrals*, 4th edn. (Academic Press, Burlington, MA, 2008)
20. S. Kaczmarz, Angenäherte auflösung von systemen linearer gleichungen. *Bull. Int. Acad. Pol. Sci. Lettres A*, **35**(III), 355–357 (1937)
21. J.E. Kolodziej, O.J. Tobias, R. Seara, D.R. Morgan, On the constrained stochastic gradient algorithm: Model, performance, and improved version. *IEEE Trans. Signal Process.* **57**(4), 1304–1315 (2009)
22. E.V. Kuhn, J.E. Kolodziej, R. Seara, Stochastic modeling of the NLMS algorithm for complex Gaussian input data and nonstationary environment. *Digit. Signal Process.* **30**, 55–66 (2014)
23. E.V. Kuhn, J.G.F. Zipf, R. Seara, On the stochastic modeling of a VSS-NLMS algorithm with high immunity against measurement noise. *Signal Process.* **147**, 120–132 (2018)
24. R.H. Kwong, E.W. Johnston, A variable step size LMS algorithm. *IEEE Trans. Signal Process.* **40**(7), 1633–1642 (1992)
25. A. Mader, H. Puder, G.U. Schmidt, Step-size control for acoustic echo cancellation filters: An overview. *Signal Process.* **80**(9), 1697–1719 (2000)
26. M.V. Matsuo, E.V. Kuhn, R. Seara, Stochastic analysis of the NLMS algorithm for nonstationary environment and deficient length adaptive filter. *Signal Process.* **160**, 190–201 (2019)
27. M.V. Matsuo, R. Seara, On the stochastic analysis of the NLMS algorithm for white and correlated Gaussian inputs in time-varying environments. *Signal Process.* **128**, 291–302 (2016)
28. K. Mayyas, F. Momani, An LMS adaptive algorithm with a new step-size control equation. *J. Franklin Inst.* **348**(4), 589–605 (2011)
29. D.R. Morgan, Comments on ‘Convergence and performance analysis of the normalized LMS algorithm with uncorrelated Gaussian data’. *IEEE Trans. Inform. Theory* **35**(6), 1299 (1989)
30. J.I. Nagumo, A. Noda, A learning method for system identification. *IEEE Trans. Autom. Control* **12**(3), 282–287 (1967)
31. A. Papoulis, S.U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th edn. (McGraw-Hill, New York, NY, 2002)
32. M. Rupp, The behavior of LMS and NLMS algorithms in the presence of spherically invariant processes. *IEEE Trans. Signal Process.* **41**(3), 1149–1160 (1993)
33. M.O.B. Saeed, LMS-based variable step-size algorithms: A unified analysis approach. *Arab. J. Sci. Eng.* **42**(7), 2809–2816 (2017)
34. A.H. Sayed, *Adaptive Filters* (Wiley, Hoboken, NJ, 2008)
35. H.-C. Shin, A.H. Sayed, W.-J. Song, Variable step-size NLMS and affine projection algorithms. *IEEE Signal Process. Lett.* **11**(2), 132–135 (2004)
36. C.W. Therrien, *Discrete Random Signals and Statistical Signal Processing* (Prentice Hall, Englewood Cliffs, NJ, 1992)
37. D.G. Tiglea, R. Candido, M.T.M. Silva, A variable step size adaptive algorithm with simple parameter selection. *IEEE Signal Process. Lett.* **29**, 1774–1778 (2022)
38. J.M. VerHoef, Who invented the delta method? *Am Stat.* **66**(2), 124–127 (2012)
39. W. Wang, H. Zhang, A new and effective nonparametric variable step-size normalized least-mean-square algorithm and its performance analysis. *Signal Process.* **210**, 109060 (2023)
40. B. Widrow, M.E. Hoff, Adaptive switching circuits. *Proc. IRE WESCON Conv. Rec.* **4**, 96–104 (1960)
41. B. Widrow, J. McCool, M. Ball, The complex LMS algorithm. *Proc. IEEE* **63**(4), 719–720 (1975)
42. S. Zhao, Z. Man, S. Khoo, H.R. Wu, Variable step-size LMS algorithm with a quotient form. *Signal Process.* **89**(1), 67–76 (2009)
43. J.G.F. Zipf, O.J. Tobias, R. Seara, Non-parametric VSS-NLMS algorithm with control parameter based on the error correlation. In *IEEE International Telecommunication Symposium (ITS)*, pp. 1–5, Manaus, AM, Brazil (2010)

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.