Steady-State Performance of Non-Negative Least-Mean-Square Algorithm and Its Variants

Jie Chen, *Member, IEEE*, José Carlos M. Bermudez, *Senior Member, IEEE*, and Cédric Richard, *Senior Member, IEEE*

***Abstract*—The Non-Negative Least-Mean-Square (NNLMS) algorithm and its variants have been proposed for online estimation under non-negativity constraints. The transient behavior of the NNLMS, Normalized NNLMS, Exponential NNLMS and Sign-Sign NNLMS algorithms have been studied in the literature. In this letter, we derive closed-form expressions for the steady-state excess mean-square error (EMSE) for the four algorithms. Simulation results illustrate the accuracy of the theoretical results. This work complements the understanding of the behavior of these algorithms.**

***Index Terms*—Non-negative LMS, steady-state performance, excess mean-square error, stochastic behavior.**

I. Introduction

Non-negativity is one important constraint that can be imposed on parameters to estimate. It is often imposed to avoid physically unreasonable solutions and to comply with natural physical characteristics. Non-negativity constraints appear, for example, in deconvolution problems [1]–[3], image processing [4], [5], audio processing [6], remote sensing [7]–[9], and neuroscience [10]. The Non-Negative Least-Mean-Square algorithm (NNLMS) [11] and its three variants, namely, Normalized NNLMS, Exponential NNLMS and Sign-Sign NNLMS [12], were proposed to adaptively find solutions of a typical Wiener filtering problem under non-negativity constraints. The transient behavior of these algorithms has been studied in [11], [12]. Analytical recursive models have been derived for the mean and mean-square behaviors of the adaptive weights.

This paper complements the work in [11], [12] by deriving closed form expressions for the steady-state excess mean square error of each of these algorithms. These expressions cannot be directly obtained from the transient recursions derived in [11], [12] because the weight updates include nonlinearities on the adaptive weights. Moreover, they cannot be derived following the conventional energy-conservation relations [13]. Hence, new analyses are required to understand the steady-state behavior of these algorithms.

In this paper, we derive accurate models for the steady-state behaviors of NNLMS and its variants using a common analysis framework, with clear physical interpretation of each term in the expressions. Simulations are conducted to validate the theoretical results. This work therefore complements the understanding of the behavior of these algorithms, and introduces a new methodology for the study of the steady-state performance of adaptive algorithms. We recommend that readers refer to [11], [12] for a more detailed understanding of the algorithms and their transient behavior. Readers may also refer to the associated report [14] for some detailed calculation steps.

Manuscript received January 29, 2014; revised March 18, 2014; accepted April 15, 2014. Date of publication April 29, 2014; date of current version May 07, 2014. This work was supported in part by CNPq Grant 307071/2013-8. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Peter K. Willett.

J. Chen and C. Richard are with the Lagrange Laboratory, University of Nice Sophia-Antipolis, CNRS, Côte d'Azur Observatory, 06108 Nice cedex 2, France (e-mail: jie.chen@unice.fr; cedric.richard@unice.fr).

J.-C. M. Bermudez is with the Department of Electrical Engineering, Federal University of Santa Catarina 88040-900, Florianópolis, SC, Brazil (e-mail: j.bermudez@ieee.org).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2014.2320944

II. Problem Formulation and Algorithms

Consider an unknown system with input-output relation characterized by the linear model

$$y(n) = \boldsymbol{\alpha}^{*\top}\boldsymbol{x}(n) + z(n) \tag{1}$$

with $\boldsymbol{\alpha}^* = [\alpha_1^*, \ldots, \alpha_N^*]^\top$ an unknown parameter vector, and $\boldsymbol{x}(n) = [x(n), x(n-1), \ldots, x(n-N+1)]^\top$ the regressor vector with correlation matrix $\boldsymbol{R}_x$. The input signal $x(n)$ and the reference signal $y(n)$ are assumed zero-mean stationary. The modeling error $z(n)$ is assumed stationary, independent and identically distributed (i.i.d.), with zero-mean and variance σ_z^2, and independent of any other signal. Due to inherent physical characteristics of the system, non-negativity is imposed on the estimated coefficient vector $\boldsymbol{\alpha}$. We seek to identify this system by minimizing the constrained mean-square error criterion

$$\begin{aligned} \boldsymbol{\alpha}^o = \arg\min_{\boldsymbol{\alpha}} E\left\{[y(n) - \boldsymbol{\alpha}^\top\boldsymbol{x}(n)]^2\right\} \\ \text{subject to } \alpha_i \geq 0, \quad \forall i. \end{aligned} \tag{2}$$

In order to solve this problem in an adaptive and online manner, the Non-Negative Least-Mean-Square (NNLMS) algorithm was derived in [11] with weight update relation given by

$$\boldsymbol{\alpha}(n+1) = \boldsymbol{\alpha}(n) + \eta\boldsymbol{D}_{\boldsymbol{\alpha}}(n)e(n)\boldsymbol{x}(n) \tag{3}$$

where $\boldsymbol{D}_{\boldsymbol{\alpha}}(n)$ denotes the diagonal matrix with ith diagonal entry $[\boldsymbol{D}_{\boldsymbol{\alpha}}(n)]_{ii} = \alpha_i(n)$, η denotes a fixed positive step size, and the estimation error $e(n) = y(n) - \boldsymbol{\alpha}^\top(n)\boldsymbol{x}(n)$. Several useful variants were derived to improve the NNLMS properties in some sense [12].

The Normalized NNLMS algorithm was proposed to reduce the sensitivity of the NNLMS performance to the input power. Its weight update relation is

$$\boldsymbol{\alpha}_{\rm N}(n+1) = \boldsymbol{\alpha}_{\rm N}(n) + \frac{\eta}{\boldsymbol{x}^\top(n)\boldsymbol{x}(n)}\boldsymbol{D}_{\boldsymbol{\alpha}_{\rm N}}(n)e_{\rm N}(n)\boldsymbol{x}(n) \tag{4}$$

where a small positive value ϵ can possibly be added to the denominator in order to avoid numerical difficulties and $e_{\mathrm{N}}(n) = y(n) - \boldsymbol{\alpha}_{\mathrm{N}}^\top(n)\boldsymbol{x}(n)$. The Exponential NNLMS was proposed to better balance the convergence rates of the weights:

$$\boldsymbol{\alpha}_{\mathrm{E}}(n+1) = \boldsymbol{\alpha}_{\mathrm{E}}(n) + \eta \boldsymbol{D}_{\boldsymbol{\alpha}_{\mathrm{E}}^{(\gamma)}}(n) e_{\mathrm{E}}(n)\boldsymbol{x}(n) \tag{5}$$

where $\boldsymbol{D}_{\boldsymbol{\alpha}_{\mathrm{E}}^{(\gamma)}}(n)$ is the diagonal matrix with (i,i)th entry equal to the ith component of $\boldsymbol{\alpha}_{\mathrm{E}}^{(\gamma)}(n)$, namely $[\boldsymbol{\alpha}_{\mathrm{E}}^{(\gamma)}(n)]_i = \mathrm{sgn}\{\alpha_{\mathrm{E}_i}(n)\}|\alpha_{\mathrm{E}_i}(n)|^\gamma$, and $e_{\mathrm{E}}(n) = y(n) - \boldsymbol{\alpha}_{\mathrm{E}}^\top(n)\boldsymbol{x}(n)$. The Sign-Sign NNLMS was proposed to reduce the implementation cost in critical real-time applications. Its update relation is given by

$$\boldsymbol{\alpha}_{\mathrm{S}}(n+1) = \boldsymbol{\alpha}_{\mathrm{S}}(n) + \eta \boldsymbol{D}_{\boldsymbol{\alpha}_{\mathrm{S}}}(n)\mathrm{sgn}(\boldsymbol{x}(n)e_{\mathrm{S}}(n)) \tag{6}$$

with $e_{\mathrm{S}}(n) = y(n) - \boldsymbol{\alpha}_{\mathrm{S}}^\top(n)\boldsymbol{x}(n)$. As the errors are nonlinear functions of the weights, the theoretical analysis becomes very challenging and significantly different from those of the LMS-based algorithms employed for solving unconstrained estimation problems.

III. Steady-State Mean-Square Performance Analysis

In this introduction we use the generic notations $\boldsymbol{\alpha}(n)$, $\alpha_i(n)$ and $e(n)$ for all the algorithms. The expressions in the following subsections naturally refer to the variables for the corresponding algorithm. This simplifies the notation and conserves space without ambiguity.

Define the weight error vector $\boldsymbol{v}(n)$ as the difference between the estimated weight vector $\boldsymbol{\alpha}(n)$ and the real system coefficient vector $\boldsymbol{\alpha}^*$, namely

$$\boldsymbol{v}(n) = \boldsymbol{\alpha}(n) - \boldsymbol{\alpha}^*. \tag{7}$$

Assume that the step size of the algorithm is chosen to be sufficiently small to ensure the convergence in the mean and mean-square senses, and denote the mean weight estimate at steady-state by $E\{\boldsymbol{\alpha}(\infty)\}$. The weight error vector (7) can then be rewritten as

$$\boldsymbol{v}(n) = \underbrace{[\boldsymbol{\alpha}(n) - E\{\boldsymbol{\alpha}(\infty)\}]}_{\boldsymbol{v}'(n)} + \underbrace{[E\{\boldsymbol{\alpha}(\infty)\} - \boldsymbol{\alpha}^*]}_{E\{\boldsymbol{v}(\infty)\}}. \tag{8}$$

The first difference $\boldsymbol{v}'(n)$ on the right-hand-side (RHS) of (8) is the weight error vector with respect to the mean of the converged weights. The second difference is the mean weight error (7) at convergence, i.e., the asymptotic bias $E\{\boldsymbol{v}(\infty)\}$.

In the following analyses we use the conventional independence assumption, namely, that $\boldsymbol{v}(n)$ is independent of $\boldsymbol{x}(m)$ for all $m \le n$ [15].

Using (8), $e(n)$ can be written as

$$e(n) = z(n) - \boldsymbol{v}'^\top(n)\boldsymbol{x}(n) - E\{\boldsymbol{v}^\top(\infty)\}\boldsymbol{x}(n) \tag{9}$$

and the excess mean-square error (EMSE) $\zeta(n) = E\{e^2(n)\} - \sigma_z^2$ is calculated as

$$\zeta(n) = \underbrace{E\{[\boldsymbol{x}^\top(n)\boldsymbol{v}'(n)]^2\}}_{\zeta'(n)} + \underbrace{\mathrm{tr}\{\boldsymbol{R}_x E\{\boldsymbol{v}(\infty)\}E\{\boldsymbol{v}^\top(\infty)\}\}}_{\zeta^\infty} + 2E\{\boldsymbol{v}'^\top(n)\}\boldsymbol{R}_x E\{\boldsymbol{v}(\infty)\}. \tag{10}$$

The steady-state EMSE is obtained by taking the limiting value as $n \to \infty$. The second term on the RHS of (10) is deterministic. The third term vanishes as $n \to \infty$. Then, it remains to evaluate the first term $\zeta'(n)$ to determine the steady-state EMSE. The advantage of working with $\boldsymbol{v}'(n)$ instead of $\boldsymbol{v}(n)$ is that the mean value of $\boldsymbol{v}'(n)$ always converges to 0, i.e., $E\{\boldsymbol{v}'(\infty)\} = 0$, which is not true for $E\{\boldsymbol{v}(\infty)\}$ in the constrained optimization problem.

The formulation in (10) is general enough to study different non-negativity constrained optimization problems. When the algorithm solution is unbiased with respect to the unconstrained solution $\boldsymbol{\alpha}^*$, the contribution of ζ^∞ will be zero. When the algorithm solution is unbiased with respect to the constrained solution $\boldsymbol{\alpha}^o$, then ζ^∞ accounts for the error due to the constraints. Otherwise, $\boldsymbol{\alpha}^o$ cannot be analytically calculated but $E\{\boldsymbol{v}(\infty)\}$ can be determined by running the recursive models derived in [11], [12] for the mean weight behavior.

For the analyses that follow, we distinguish the weights into two sets. The set $\mathcal{S}_+$ contains the indices of the weights that converge in the mean to positive values, namely,

$$\mathcal{S}_+ = \{i : E\{\alpha_i(\infty)\} > 0\}.$$

The set $\mathcal{S}_0$ contains the indices of the weights that converge in the mean to zero, namely,

$$\mathcal{S}_0 = \{i : E\{\alpha_i(\infty)\} = 0\}.$$

Considering that the non-negativity constraint is always satisfied at steady-state, $E\{\alpha_i(\infty)\} = 0$ implies that $\alpha_i(\infty) = 0$ for $i \in \mathcal{S}_0$ for all realizations. The weight error vector $\boldsymbol{v}(\infty)$ is then deterministic and satisfies

$$v_i(\infty) = -\alpha_i^*, \quad \text{for} \quad i \in \mathcal{S}_0 \tag{11}$$

and, consequently,

$$v_i'(\infty) = 0, \quad \text{for} \quad i \in \mathcal{S}_0. \tag{12}$$

Now let $\overline{\boldsymbol{D}}_{\boldsymbol{\alpha}}^{-1}(n)$ be a diagonal matrix with entries

$$[\overline{\boldsymbol{D}}_{\boldsymbol{\alpha}}^{-1}(n)]_{ii} = \begin{cases} \frac{1}{\alpha_i(n)}, & i \in \mathcal{S}_+ \\ 0, & i \in \mathcal{S}_0 \end{cases} \tag{13}$$

and $\overline{\boldsymbol{I}}$ be the diagonal matrix such that

$$[\overline{\boldsymbol{I}}]_{ii} = \begin{cases} 1, & i \in \mathcal{S}_+ \\ 0, & i \in \mathcal{S}_0 \end{cases}. \tag{14}$$

With these matrices, we have that

$$\overline{\boldsymbol{D}}_{\boldsymbol{\alpha}}^{-1}(n)\boldsymbol{D}_{\boldsymbol{\alpha}}(n) = \overline{\boldsymbol{I}}, \tag{15}$$

and, as $n \to \infty$,

$$E\{\boldsymbol{D}_{\boldsymbol{\alpha}}(\infty)\}\overline{\boldsymbol{I}} = E\{\boldsymbol{D}_{\boldsymbol{\alpha}}(\infty)\}. \tag{16}$$

With these definitions and notations at hand, we now perform the steady-state analysis for NNLMS and its variants.

A. Steady-State Performance for NNLMS

Subtracting $E\{\boldsymbol{\alpha}(\infty)\}$ from both sides of (3), we have the weight error update relation

$$\boldsymbol{v}'(n+1) = \boldsymbol{v}'(n) + \eta e(n)\boldsymbol{D}_{\boldsymbol{\alpha}}(n)\boldsymbol{x}(n). \tag{17}$$

Consider the weighted square-norm $\|\cdot\|^2_{\overline{\boldsymbol{D}}_{\alpha}^{-1}(n)}$ such that $\|\boldsymbol{x}\|^2_{\overline{\boldsymbol{D}}_{\alpha}^{-1}(n)} = \boldsymbol{x}^\top \overline{\boldsymbol{D}}_{\alpha}^{-1}(n)\boldsymbol{x}$. Then we have

$$E\left\{\|\boldsymbol{v}'(n+1)\|^2_{\overline{D}_{\alpha}^{-1}(n)}\right\} = E\left\{\|\boldsymbol{v}'(n)\|^2_{\overline{D}_{\alpha}^{-1}(n)}\right\} + 2\eta E\left\{\boldsymbol{v}'^\top(n)\overline{\boldsymbol{I}}\boldsymbol{x}(n)e(n)\right\} + \eta^2 E\left\{\boldsymbol{x}^\top(n)\overline{\boldsymbol{I}}\boldsymbol{D}_{\boldsymbol{\alpha}}(n)\boldsymbol{x}(n)e^2(n)\right\}. \tag{18}$$

Assuming convergence, we consider the following relation to be valid at steady-state:

$$\lim_{n\to\infty} E\left\{\|\boldsymbol{v}'(n+1)\|^2_{\overline{\boldsymbol{D}}_{\alpha}^{-1}(n)}\right\} = \lim_{n\to\infty} E\left\{\|\boldsymbol{v}'(n)\|^2_{\overline{\boldsymbol{D}}_{\alpha}^{-1}(n)}\right\}. \tag{19}$$

Using equation (9), the expected value of the second term on RHS of (18) with $n \to \infty$ is given by

$$\lim_{n\to\infty} E\left\{\boldsymbol{v}'^\top(n)\overline{\boldsymbol{I}}\boldsymbol{x}(n)e(n)\right\} = -\lim_{n\to\infty} E\left\{\boldsymbol{v}'^\top(n)\overline{\boldsymbol{I}}\boldsymbol{x}(n)\boldsymbol{x}^\top(n)\boldsymbol{v}'(n) + \boldsymbol{v}'^\top(n)\overline{\boldsymbol{I}}\boldsymbol{x}(n)\boldsymbol{x}^\top(n)E\{\boldsymbol{v}(\infty)\}\right\} = -\zeta'(\infty) \tag{20}$$

where we have considered that $\boldsymbol{v}'^\top(\infty)\overline{\boldsymbol{I}} = \boldsymbol{v}'^\top(\infty)$ due to (12) and (14) and that $E\{\boldsymbol{v}'(\infty)\} = 0$. For the expected value of the third term on the RHS of (18), we assume that $\|\boldsymbol{x}(n)\|^2_{\boldsymbol{D}_{\alpha}(n)}$ is independent of $e^2(n)$ at steady-state, which is similar to the approximation performed in [13]. This expected value can then be written as

$$\lim_{n\to\infty} E\{\boldsymbol{x}^\top(n)\overline{\boldsymbol{I}}\boldsymbol{D}_{\boldsymbol{\alpha}}(n)\boldsymbol{x}(n)e^2(n)\} \approx \operatorname{tr}\{E\{\boldsymbol{D}_{\boldsymbol{\alpha}}(\infty)\}\boldsymbol{R}_x\}(\sigma_z^2 + \zeta'(\infty) + \zeta^\infty). \tag{21}$$

Now, using (19) to (21) in (18) yields

$$-2\eta\zeta'(\infty) + \eta^2\operatorname{tr}\{E\{\boldsymbol{D}_{\boldsymbol{\alpha}}(\infty)\}\boldsymbol{R}_x\}(\sigma_z^2 + \zeta'(\infty) + \zeta^\infty) = 0 \tag{22}$$

which leads to

$$\zeta'(\infty) = \frac{\eta\sigma_z^2\operatorname{tr}\{E\{\boldsymbol{D}_{\boldsymbol{\alpha}}(\infty)\}\boldsymbol{R}_x\}}{2 - \eta\operatorname{tr}\{E\{\boldsymbol{D}_{\boldsymbol{\alpha}}(\infty)\}\boldsymbol{R}_x\}} + \frac{\eta\zeta^\infty\operatorname{tr}\{E\{\boldsymbol{D}_{\boldsymbol{\alpha}}(\infty)\}\boldsymbol{R}_x\}}{2 - \eta\operatorname{tr}\{E\{\boldsymbol{D}_{\boldsymbol{\alpha}}(\infty)\}\boldsymbol{R}_x\}}. \tag{23}$$

In the above expression, the first term accounts for the EMSE contribution associated with unbiased components, which is equivalent to EMSE of the LMS algorithm with component-wise step sizes $\eta E\{\alpha_i(\infty)\}$. This result is reasonable when observing the weight update equation (3). The second term accounts for EMSE introduced in the adaptive process by the bias with respect to the unconstrained solution. Finally considering the relation (10), i.e., adding the direct bias contribution, the steady-state EMSE is given by

$$\zeta(\infty) = \frac{\eta\operatorname{tr}\{E\{\boldsymbol{D}_{\boldsymbol{\alpha}}(\infty)\}\boldsymbol{R}_x\}(\sigma_z^2 + \zeta^\infty)}{2 - \eta\operatorname{tr}\{E\{\boldsymbol{D}_{\boldsymbol{\alpha}}(\infty)\}\boldsymbol{R}_x\}} + \zeta^\infty. \tag{24}$$

B. Steady-State Performance for Normalized NNLMS

It is common to neglect the correlation between the denominator $\boldsymbol{x}^\top(n)\boldsymbol{x}(n)$ and the numerator of the weight update in equation (4) for large filter lengths, as the former tends to vary much slower [16], [17]. Moreover, for sufficiently large N, $E\{\boldsymbol{x}^\top(n)\boldsymbol{x}(n)\} \approx N\sigma_x^2$ and the Normalized NNLMS performs as the NNLMS algorithm with the equivalent step size $\tilde{\eta} = \eta/(N\sigma_x^2)$. Based on this approximation, the Normalized NNLMS steady-state EMSE is directly obtained by using $\tilde{\eta}$ in equation (24):

$$\zeta(\infty) = \frac{\tilde{\eta}\operatorname{tr}\{E\{\boldsymbol{D}_{\boldsymbol{\alpha}}(\infty)\}\boldsymbol{R}_x\}(\sigma_z^2 + \zeta^\infty)}{2 - \tilde{\eta}\operatorname{tr}\{E\{\boldsymbol{D}_{\boldsymbol{\alpha}}(\infty)\}\boldsymbol{R}_x\}} + \zeta^\infty. \tag{25}$$

C. Steady-State Performance for Exponential NNLMS

Let $\overline{\boldsymbol{D}}^{-1}_{\boldsymbol{\alpha}^{(\gamma)}}(n)$ be a matrix defined as in equation (13), with entries $[\overline{\boldsymbol{D}}^{-1}_{\boldsymbol{\alpha}^{(\gamma)}}(n)]_{ii} = 1/[\alpha_i^\gamma(n)]$ for $i \in \mathcal{S}_+$, $[\overline{\boldsymbol{D}}^{-1}_{\boldsymbol{\alpha}^{(\gamma)}}(n)]_{ii} = 0$ otherwise. Following the same steps that led to the EMSE for the NNLMS algorithm, except by taking the weighted square-norm $\|\boldsymbol{x}\|^2_{\overline{\boldsymbol{D}}^{-1}_{\boldsymbol{\alpha}^{(\gamma)}}(n)}$ when writing the norm equality (18), yields the following steady-state EMSE for the Exponential NNLMS algorithm:

$$\zeta(\infty) = \frac{\eta\operatorname{tr}\{E\{\boldsymbol{D}_{\boldsymbol{\alpha}^{(\gamma)}}(\infty)\}\boldsymbol{R}_x\}(\sigma_z^2 + \zeta^\infty)}{2 - \eta\operatorname{tr}\{E\{\boldsymbol{D}_{\boldsymbol{\alpha}^{(\gamma)}}(\infty)\}\boldsymbol{R}_x\}} + \zeta^\infty. \tag{26}$$

D. Steady-State Performance for Sign-Sign NNLMS

In this subsection, we shall derive the EMSE for Sign-Sign NNLMS in detail due to the particular nonlinearity introduced by sgn function. Subtracting $E\{\boldsymbol{\alpha}(\infty)\}$ from both sides of the weight update relation (6), we have the relation:

$$\boldsymbol{v}'(n+1) = \boldsymbol{v}'(n) + \eta\boldsymbol{D}_{\boldsymbol{\alpha}}(n)\operatorname{sgn}(\boldsymbol{x}(n)e(n)). \tag{27}$$

Taking the expected value of $\|\cdot\|^2_{\overline{\boldsymbol{D}}_{\alpha}^{-1}(n)}$, we have

$$E\left\{\|\boldsymbol{v}'(n+1)\|^2_{\overline{\boldsymbol{D}}_{\alpha}^{-1}(n)}\right\} = E\left\{\|\boldsymbol{v}'(n)\|^2_{\overline{\boldsymbol{D}}_{\alpha}^{-1}(n)}\right\} + 2\eta E\left\{\boldsymbol{v}'^\top(n)\overline{\boldsymbol{I}}\operatorname{sgn}(\boldsymbol{x}(n)e(n))\right\} + \eta^2 E\left\{\operatorname{sgn}(\boldsymbol{x}^\top(n)e(n))\overline{\boldsymbol{I}}\boldsymbol{D}_{\boldsymbol{\alpha}}(n)\operatorname{sgn}(\boldsymbol{x}(n)e(n))\right\}. \tag{28}$$

Assuming convergence, we consider the following relation to be valid at steady-state:

$$\lim_{n\to\infty} E\left\{\|\boldsymbol{v}'(n+1)\|^2_{\overline{\boldsymbol{D}}_{\alpha}^{-1}(n)}\right\} = \lim_{n\to\infty} E\left\{\|\boldsymbol{v}'(n)\|^2_{\overline{\boldsymbol{D}}_{\alpha}^{-1}(n)}\right\}. \tag{29}$$

The expected value of the second term on RHS of equation (28) with $n \to \infty$ is given by

$$\lim_{n\to\infty} E\left\{\boldsymbol{v}'^\top(n)\overline{\boldsymbol{I}}\operatorname{sgn}(\boldsymbol{x}(n)e(n))\right\} = \frac{2}{\pi}\lim_{n\to\infty} E\left\{\boldsymbol{v}'^\top(n)\overline{\boldsymbol{I}}\sin^{-1}\left(-\frac{\boldsymbol{R}_x\boldsymbol{v}'(n)}{\sigma_x\sigma_{e|\boldsymbol{v}'(n)}}\right)\right\} \tag{30}$$

where we used Price's theorem to obtain this result since $x_i(n)$ and $e(n)$ are jointly Gaussian when conditioned on $\boldsymbol{v}'(n)$ [12]. The conditional variance of $e(n)$ is given by

$$\sigma^2_{e|\boldsymbol{v}'(n)} = \sigma_z^2 + \operatorname{tr}\left\{\boldsymbol{R}_x \boldsymbol{v}'(n)\boldsymbol{v}'^\top(n)\right\} + \operatorname{tr}\left\{\boldsymbol{R}_x E\{\boldsymbol{v}(\infty)\}\, E\{\boldsymbol{v}^\top(\infty)\}\right\}. \tag{31}$$

The term within the expectation in (30) is highly nonlinear due to function $\sin^{-1}(\cdot)$. It is reasonable to approximate the nonlinear function $\sin^{-1}(\cdot)$ using the linear expansion about the point $E\{\boldsymbol{v}'(\infty)\}$, since the weight errors fluctuate about $E\{\boldsymbol{v}'(\infty)\}$ at steady-state. As $E\{\boldsymbol{v}'(\infty)\} = 0$, we have

$$\begin{aligned}(30) &\approx -\frac{2}{\pi}\lim_{n\to\infty} E\{\boldsymbol{v}'^\top(n)\bar{\boldsymbol{I}}\frac{\boldsymbol{R}_x}{\sigma_x\sigma_{e|E\{\boldsymbol{v}'(\infty)\}}}\boldsymbol{v}'(n)\} \\ &= -\frac{2}{\pi\sigma_x\sigma_{e|E\{\boldsymbol{v}'(\infty)\}}}\zeta'(\infty) \end{aligned} \tag{32}$$

with $\sigma^2_{e|E\{\boldsymbol{v}'(\infty)\}} = \sigma_z^2 + \zeta^\infty$. The expected value of the third term on the RHS of (28) for $n \to \infty$ is given by

$$\begin{aligned}&\lim_{n\to\infty} E\left\{\operatorname{sgn}(\boldsymbol{x}^\top(n)e(n))\bar{\boldsymbol{I}}\boldsymbol{D}_{\boldsymbol{\alpha}}(n)\operatorname{sgn}(\boldsymbol{x}(n)e(n))\right\} \\ &= \lim_{n\to\infty} E\left\{\sum_{i=1}^{N}\operatorname{sgn}(x^2(n-i+1)e^2(n))\alpha_i(n)\right\} \\ &= \operatorname{tr}\{E\{\boldsymbol{D}_{\boldsymbol{\alpha}}(\infty)\}\}. \end{aligned} \tag{33}$$

Using these results in the equality (28), we have the equation

$$\eta^2\operatorname{tr}\{E\{\boldsymbol{D}_{\boldsymbol{\alpha}}(\infty)\}\} - 2\eta\frac{2}{\pi\sigma_x\sigma_{e|E\{\boldsymbol{v}'(\infty)\}}}\zeta'(\infty) = 0 \tag{34}$$

which yields

$$\zeta'(\infty) = \frac{\eta\pi}{4}\operatorname{tr}\{E\{\boldsymbol{D}_{\boldsymbol{\alpha}}(\infty)\}\}\sigma_x\sigma_{e|E\{\boldsymbol{v}'(\infty)\}}. \tag{35}$$

Finally, from equation (10) the the steady-state EMSE of the Sign-Sign NNLMS algorithm is given by

$$\zeta(\infty) = \frac{\eta\pi}{4}\operatorname{tr}\{E\{\boldsymbol{D}_{\boldsymbol{\alpha}}(\infty)\}\}\,\sigma_x\sqrt{\sigma_z^2 + \zeta^\infty} + \zeta^\infty. \tag{36}$$

IV. Experiment Validation

In this section, we present examples to illustrate the correspondence between theoretical steady-state EMSE and simulated results for NNLMS and its variants. Consider an unknown system of order $N = 15$ and weights defined by

$$\begin{aligned}\boldsymbol{\alpha}^* = [&0.8, 0.6, 0.5, -0.05, 0.4, -0.04, 0.3, -0.03, \\ &0.2, -0.02, 0.1, -0.01, 0, 0, 0]^\top, \end{aligned} \tag{37}$$

where negative coefficients were explicitly included to activate the non-negativity constraint. The input signal was the first-order AR progress given by $x(n) = 0.5x(n-1) + w(n)$, where $w(n)$ is an i.i.d. zero-mean Gaussian sequence with variance $\sigma_w^2 = 0.75$ (so that $\sigma_x^2 = 1$) and independent of any other signal. The additive independent noise $z(n)$ was zero-mean i.i.d. Gaussian with variance $\sigma_z^2 = 0.01$. The adaptive weights were initialized with $\alpha_i(0) = 0.1$ for $i = 1, \ldots, N$. The step sizes were equal to $\eta = 0.01N\sigma_x^2$ for NNLMS and $\eta = 0.01$ for the NNLMS, Exponential NNLMS and Sign-Sign NNLMS algorithms. Monte Carlo simulation results were obtained by averaging over 100 runs. Fig. 1 shows the simulation results and the behavior predicted by the analytical models. The theoretical transient EMSE behaviors were obtained using results in [11], [12], and the theoretical steady-state EMSE (horizontal dashed lines) were calculated by the expressions derived in this letter. Fig. 2 shows the relative bias introduced by the assumptions used in the analysis. The bias is specially low for $\eta < \eta_{\max}/10$, which is the most used step size range in practical application. These figures clearly validate the proposed theoretical results.

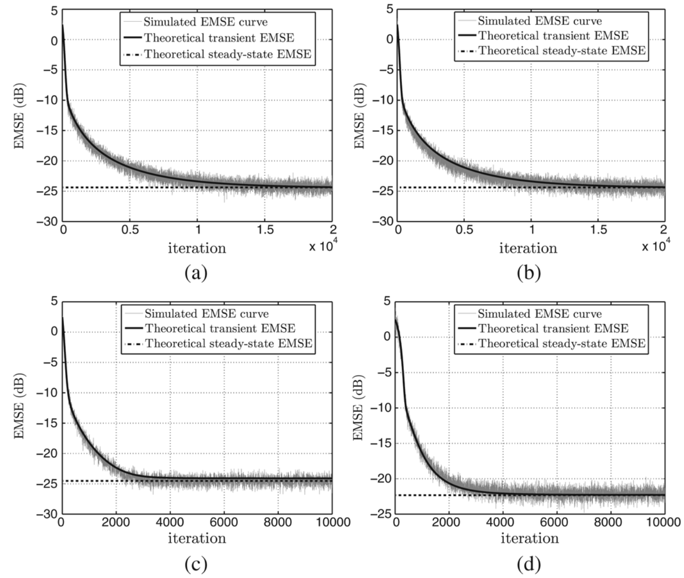

Fig. 1. Steady-state EMSE model validation for NNLMS and its variants (a) Original NNLMS (b) Normalized NNLMS (c) Exponential NNLMS (d) Sign-Sign NNLMS.

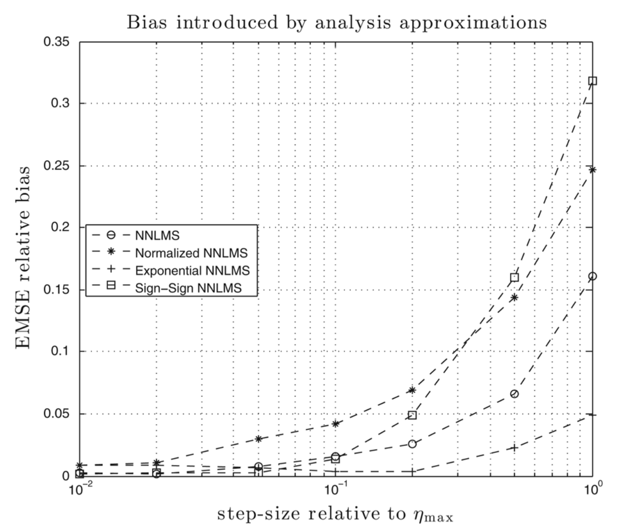

Fig. 2. Bias introduced by the assumptions made in the analysis. The bias is calculated as the relative difference of the EMSE obtained from simulations and predicted by the models for step sizes relative to the stability limit of each algorithm. The relative bias is calculated as $|\mathrm{EMSE}_{\mathrm{theory}} - \mathrm{EMSE}_{\mathrm{simul}}|/\mathrm{EMSE}_{\mathrm{simul}}$.

V. Conclusion

In this letter, we derived closed-form expressions for the steady-state excess mean-square errors of the Non-Negative LMS algorithm and its variants. Experiments illustrated the accuracy of the derived results. Future work may include the derivation of other useful variants of NNLMS and the study of their stochastic performance.

References

[1] M. D. Plumbley, "Algorithms for nonnegative independent component analysis," *IEEE Trans. Neural Netw.*, vol. 14, no. 3, pp. 534–543, Mar. 2003.

[2] S. Moussaoui, D. Brie, A. Mohammad-Djafari, and C. Carteret, "Separation of non-negative mixture of non-negative sources using a bayesian approach and MCMC sampling," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4133–4145, Nov. 2006.

[3] Y. Lin and D. D. Lee, "Bayesian regularization and nonnegative deconvolution for room impulse response estimation," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 839–847, Mar. 2006.

[4] F. Benvenuto, R. Zanella, L. Zanni, and M. Bertero, "Nonnegative least-squares image deblurring: Improved gradient projection approaches," *Inv. Probl.*, vol. 26, no. 1, p. 025004, Feb. 2010.

[5] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.

[6] A. Cont and S. Dubinov, "Realtime multiple pitch and multiple-instrument recognition for music signals using sparse non-negative constraints," in *Proc. 10th Int. Conf. Digital Audio Effects (DAFx-07)*, Bordeaux, France, Sep. 2007, pp. 85–92.

[7] J. Chen, C. Richard, and P. Honeine, "Nonlinear unmixing of hyperspectral data based on a linear-mixture/nonlinear-fluctuation model," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 480–492, 2013.

[8] J. Chen, C. Richard, H. Lantéri, C. Theys, and P. Honeine, "A gradient based method for fully constrained least-squares unmixing of hyperspectral images," in *Proc. IEEE Statistical Signal Processing Workshop (SSP)*, Nice, France, Jun. 2011, pp. 301–304.

[9] P. Honeine and C. Richard, "Geometric unmixing of large hyperspectral images: A barycentric coordinate approach," *IEEE Trans. Geosci. Remote Sensing*, vol. 50, no. 6, pp. 2185–2195, 2012.

[10] A. Cichocki, R. Zdunek, and A. H. Phan, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*. Hoboken, NJ, USA: Wiley, 2009.

[11] J. Chen, C. Richard, J.-C. M. Bermudez, and P. Honeine, "Nonnegative least-mean-square algorithm," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5225–5235, Nov. 2011.

[12] J. Chen, C. Richard, J.-C. M. Bermudez, and P. Honeine, Variants of non-negative least-mean-square algorithm and convergence analysis Univ. Nice Sophia-Antipolis, France, Tech. Rep., 2014 [Online]. Available: http://www.cedric-richard.fr/Articles/chen2013variants.pdf

[13] A. H. Sayed, *Adaptive Filters*. Hoboken, NJ, USA: Wiley, 2008.

[14] J. Chen, J.-C. M. Bermudez, and C. Richard, Steady-state performance of non-negative steady-state performance of non-negative least-mean-square algorithm and its variants Univ. Nice Sophia-Antipolis, France, Tech. Rep., Jan. 2014 [Online]. Available: http://arxiv.org/pdf/1401.6376v1.pdf

[15] S. Haykin, *Adaptive Filter Theory*, 4th Ed. ed. Delhi, India: Pearson Education India, 2005.

[16] C. Samson and V. U. Reddy, "Fixed point error analysis of the normalized ladder algorithms," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-31, no. 10, pp. 1177–1191, Oct. 1983.

[17] S. J. M. Almeida, J.-C. M. Bermudez, and N. J. Bershad, "A statistical analysis of the affine projection algorithm for unity step size and autoregressive inputs," *IEEE Trans. Circuits Syst. I: Fund. Theory Applicat.*, vol. 52, no. 7, pp. 1394–1405, Jul. 2005.