STOCHASTIC MODEL FOR THE GENERALIZED SUBBAND DECOMPOSITION εNLMS ALGORITHM WITH GAUSSIAN DATA

Javier E. Kolodziej, Orlando J. Tobias, and Rui Seara

LINSE – Circuits and Signal Processing Laboratory
Department of Electrical Engineering
Federal University of Santa Catarina
88040-900 – Florianópolis – SC – Brazil
E-mails: {javier, orlando, seara}@linse.ufsc.br

ABSTRACT

This paper proposes a stochastic model for the generalized subband decomposition normalized LMS (NLMS) algorithm. This algorithm is used as an alternative to the standard NLMS one, aiming to improve the convergence speed under correlated input data. Analytical models for the first and second moments of the filter weights are derived taking into account the time-varying nature of normalized step size. Moreover, in the model expressions a positive regularization parameter ε is added to the power estimates, preventing division by zero during the power normalization process. Through simulation results, the accuracy of the proposed analytical model can be verified.

1. INTRODUCTION

Among the various adaptive algorithms, the LMS one is the most popular and widely used due to its simplicity and robustness. However, such an algorithm suffers from slow convergence when the input signal is highly correlated. In this context, several adaptive subband structures have been proposed aiming to improve the convergence behavior of the standard LMS algorithm. One of them is the structure based on the generalized subband decomposition (GSD) of FIR filters [1]. The goal of this approach consists of implementing an N-weight FIR filter as an M-branch parallel structure with $1 \leq M \leq N$. Each branch is composed of the cascade of an interpolator and a sparse subfilter, acting on a specific frequency subband. The adaptive version of the GSD structures is obtained by adapting the subfilters while the interpolators are maintained fixed. The latter may be implemented by using an orthogonal transform such as DCT, DFT, Hadamard, etc. The fact of using frequency subband filters permits to increase the convergence speed of the overall adaptive structure, as well as to reduce the required computational burden by neglecting the frequency bands presenting low activity.

The convergence speed is improved by using the NLMS algorithm in each subfilter [1]. Thus, the GSD structure associated with the normalized-LMS results in the GSD-NLMS algorithm. It also presents the advantage of a reduced transform size when compared with the transform-domain LMS (TD-LMS) algorithm.

Regarding the stochastic analysis of the GSD-NLMS algorithm [1] a simplifying consideration is used. This analysis disregards the time-varying nature of the step-size normalizing operation. Such an assumption simplifies considerably the involved mathematics; however it leads to a statistical model that does not allow for a time-varying step-size parameter condition. Thus, the aim of this paper is to derive an appropriate statistical model to describe the GSD-NLMS algorithm behavior considering the time-varying nature of the step-size parameter. In particular, we derive analytical expressions for the first and second moments of the adaptive filter weight vector for Gaussian data and slow convergence condition. Through numerical simulations, we verify the very good agreement between the results obtained with the Monte Carlo method and the predictions from the proposed analytical model.

2. GENERALIZED SUBBAND DECOMPOSITION-LMS ALGORITHM

This section discusses the basic expressions that describe the GSD-LMS algorithm [1]. The adaptive subband structure is shown in Fig. 1. The input signal $x(n)$ is first processed by an M-point unitary transform, implemented by an M×M nonsingular matrix $\mathbf{T}$, generating the signals $u_k(n)$ which are then filtered by the sparse subfilters $W_k(z)$. All samples of the transformed signals used as inputs of the subfilters form the KM-dimensional vector $\mathbf{u}_\mathrm{a}(n)$, given by

$$\mathbf{u}_\mathrm{a}(n) = \left\{\mathbf{u}^\mathrm{T}(n) \quad \mathbf{u}^\mathrm{T}(n-L) \quad \cdots \quad \mathbf{u}^\mathrm{T}[n-(K-1)L]\right\}^\mathrm{T}, \tag{1}$$

where

$$\mathbf{u}(n) = \left[u_0(n) \quad u_1(n) \quad \cdots \quad u_{M-1}(n)\right]^\mathrm{T}. \tag{2}$$

The vector $\mathbf{u}_\mathrm{a}(n)$ is related to the input vector

$$\mathbf{x}(n) = \left[x(n) \quad x(n-1) \quad \cdots \quad x(n-N+1)\right]^\mathrm{T}, \tag{3}$$

by

$$\mathbf{u}_\mathrm{a}(n) = \mathbf{T}_\mathrm{a}^\mathrm{T}\,\mathbf{x}(n). \tag{4}$$

Matrix $\mathbf{T}_\mathrm{a}^\mathrm{T}$ of dimension $KM \times N$ is related to matrix $\mathbf{T}$ by

$$\mathbf{T}_\mathrm{a}^\mathrm{T} = \underbrace{\begin{bmatrix} \underbrace{\mathbf{T}}_{M\times M} & \underbrace{\mathbf{0}}_{M\times(N-M)} & \\ \underbrace{\mathbf{0}}_{M\times L} \;\; \mathbf{T} & \underbrace{\mathbf{0}}_{M\times(N-(L+M))} & \\ & \vdots & \\ \underbrace{\mathbf{0}}_{M\times((K-1)L)} & \mathbf{T} & \end{bmatrix}}_{KM\times((K-1)L+M)}, \tag{5}$$

where $\mathbf{0}$ represent null matrixes.

By defining the generalized subband weight vector $\mathbf{w}_\mathrm{a}(n)$ as the vector containing all the weights of the subfilters as

$$\mathbf{w}_a(n) = \left[\mathbf{w}_0^T(n) \quad \mathbf{w}_1^T(n) \quad \cdots \quad \mathbf{w}_{K-1}^T(n) \right]^T, \qquad (6)$$

where

$$\mathbf{w}_l(n) = \left[w_{0,l}(n) \quad w_{1,l}(n) \quad \cdots \quad w_{M-1,l}(n) \right]^T, \quad l = 0,\ldots,K-1, \qquad (7)$$

the filter output $y(n)$ can be computed as follows:

$$y(n) = \mathbf{u}_a^T(n)\mathbf{w}_a(n). \qquad (8)$$

The corresponding error signal $e(n)$ is given by

$$e(n) = d(n) - y(n), \qquad (9)$$

where $d(n)$ represents the desired signal. By using the NLMS algorithm, the weight update equation is [1]

$$\mathbf{w}_a(n+1) = \mathbf{w}_a(n) + \tau \mathbf{D}_a^{-1}\left[d(n) - y(n)\right]\mathbf{u}_a(n), \qquad (10)$$

where

$$\mathbf{D}_a = \begin{bmatrix} \mathbf{D} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{D} \end{bmatrix}, \qquad (11)$$

and $\mathbf{D} = diag[\sigma_0^2 \ldots \sigma_{M-1}^2]$, with the operator $diag[.]$ denoting a diagonal matrix, and τ is the step size parameter.

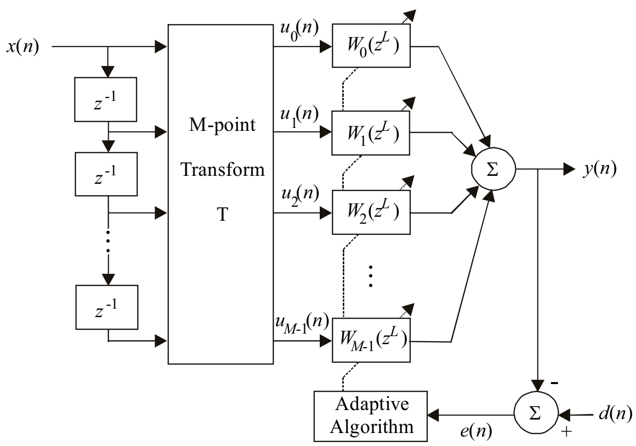


Fig. 1. Adaptive version of the GSD structure.

The elements of $\mathbf{D}_a$ are the variances σ_i^2 of the $u_i(n)$ elements of vector $\mathbf{u}_a(n)$. Thus, to implement the GSD-NLMS algorithm it is necessary to know the variances σ_i^2 for accomplishing the step-size normalizing operation. Instead of using the instantaneous value of power, which is susceptible to overshoot during the adaptation process, it is of common practice to use an averaged power [2]. For this purpose, we use here the following estimate:

$$\sigma_i^2(n) = \frac{1}{M_w}\sum_{k=0}^{M_w-1} u_i^2(n-k), \qquad (12)$$

where $u_i(n)$ for $i = 0,1,\ldots,M-1$ is the ith output signal after the $M \times M$–transformation, and M_w is the length of the observation window. This estimation makes σ_i^2 time-varying. For that, in our approach matrix $\mathbf{D}_a$ is then replaced in (10) by its time-varying version $\mathbf{D}_a(n)$ having now as diagonal elements $\sigma_i^2(n)$. The normalization matrix $\mathbf{D}(n)$ reads then

$$\mathbf{D}(n) = diag[\sigma_0^2(n) + \varepsilon \ \ldots \ \sigma_{M-1}^2(n) + \varepsilon\ ]. \qquad (13)$$

Note that a small positive regularization parameter ε is added to each diagonal element to prevent division by zero when $\mathbf{D}_a^{-1}(n)$ is performed. Such a parameter is also considered in the proposed model.

3. ANALYSIS

3.1. Mean-weight behavior

In this section the first moment of the adaptive weight vector is obtained by taking the expectation on both sides of (10). By using the simplifying assumption that $\mathbf{w}_a(n)$ and $\mathbf{u}_a(n)$ are statistically independent [2]-[4], we can write

$$\begin{aligned} E\left[\mathbf{w}_a(n+1)\right] = E\left[\mathbf{w}_a(n)\right] + \tau E\left[\mathbf{D}_a^{-1}(n)\mathbf{u}_a(n)d(n)\right] \\ - \tau E\left[\mathbf{D}_a^{-1}(n)\mathbf{u}_a(n)\mathbf{u}_a^T(n)\right]E\left[\mathbf{w}_a(n)\right]. \end{aligned} \qquad (14)$$

The expected values of the second and third terms on the r.h.s. of (14) must now be determined. For such, we use the simplifying assumption give by the Averaging Principle [5]. It states that given $\mathbf{D}_a^{-1}(n)$ and $\mathbf{u}_a(n)\mathbf{u}_a^T(n)$, jointly stationary processes, such that $\mathbf{D}_a^{-1}(n)$ is slowly varying with respect to $\mathbf{u}_a(n)\mathbf{u}_a^T(n)$, then

$$\begin{aligned} &E[\mathbf{D}_a^{-1}(n)\mathbf{u}_a(n)\mathbf{u}_a^T(n)] \approx \\ &\qquad E[\mathbf{D}_a^{-1}(n)]E[\mathbf{u}_a(n)\mathbf{u}_a^T(n)] = E[\mathbf{D}_a^{-1}(n)]\mathbf{R}_{\mathbf{u}_a\mathbf{u}_a}. \end{aligned} \qquad (15)$$

Similarly,

$$\begin{aligned} &E[\mathbf{D}_a^{-1}(n)\mathbf{u}_a(n)d(n)] \approx \\ &\qquad E[\mathbf{D}_a^{-1}(n)]E[\mathbf{u}_a(n)d(n)] = E[\mathbf{D}_a^{-1}(n)]\mathbf{p}_{\mathbf{u}_a d}, \end{aligned} \qquad (16)$$

where $\mathbf{R}_{\mathbf{u}_a\mathbf{u}_a} = E\left[\mathbf{u}_a(n)\mathbf{u}_a^T(n)\right]$ and $\mathbf{p}_{\mathbf{u}_a d} = E\left[\mathbf{u}_a(n)d(n)\right]$ are the transformed-input autocorrelation matrix and crosscorrelation vector, respectively. By substituting (15) and (16) into (14), we obtain

$$\begin{aligned} &E[\mathbf{w}_a(n+1)] = \\ &\qquad \{\mathbf{I} - \tau E[\mathbf{D}_a^{-1}(n)]\mathbf{R}_{\mathbf{u}_a\mathbf{u}_a}\}E[\mathbf{w}_a(n)] + \tau E[\mathbf{D}_a^{-1}(n)]\mathbf{p}_{\mathbf{u}_a d}. \end{aligned} \qquad (17)$$

The derivation of (17) is concluded by determining the expected value $E[\mathbf{D}_a^{-1}(n)]$. For such, we assume that the process $\{u_i^2(n)\}$ has a Chi-square distribution with M_w degrees of freedom [6]. Such an assumption is valid for independent Gaussian processes [6]. In our case, $\{u_i^2(n)\}$ does not fulfill completely such a requirement. However, we have verified through simulations that the obtained model leads to very good results even for highly correlated signals $u_i(n)$. Finally, the sought expected value is given by [2]

$$\begin{aligned} E[\mathbf{D}_a^{-1}(n)] = \frac{M_w}{(M_w-2)}[\mathrm{diag}(\mathbf{R}_{\mathbf{u}_a\mathbf{u}_a})]^{-1} \\ - \varepsilon\frac{M_w^2}{(M_w-2)(M_w-4)}[\mathrm{diag}(\mathbf{R}_{\mathbf{u}_a\mathbf{u}_a}^2)]^{-1}. \end{aligned} \qquad (18)$$

3.2. Steady-state weight vector

By substituting (4) into (17), we obtain

$$E[\mathbf{w}_\mathrm{a}(n+1)] = \left\{\mathbf{I} - \tau E[\mathbf{D}_\mathrm{a}^{-1}(n)]\mathbf{T}_\mathrm{a}^\mathrm{T}\mathbf{R}_{\mathbf{xx}}\mathbf{T}_\mathrm{a}\right\}E[\mathbf{w}_\mathrm{a}(n)] + \tau E[\mathbf{D}_\mathrm{a}^{-1}(n)]\mathbf{T}_\mathrm{a}^\mathrm{T}\mathbf{p}_{\mathbf{x}d} \quad (19)$$

Now, by applying a rotation on the axis as

$$\mathbf{v}(n) = \mathbf{B}E[\mathbf{w}_\mathrm{a}(n)] = \begin{bmatrix}\mathbf{B}_1\\ \mathbf{B}_2\end{bmatrix}E[\mathbf{w}_\mathrm{a}(n)] \quad (20)$$

where matrix $\mathbf{B}$ of dimension $(KM)\times(KM)$ has the first N rows termed $\mathbf{B}_1$, spanning the row space of $\mathbf{T}_\mathrm{a}E[\mathbf{D}_\mathrm{a}^{-1}(n)]$, and the remainder $(KM-N)$ rows, termed $\mathbf{B}_2$, spanning the null space of $\mathbf{T}_\mathrm{a}E[\mathbf{D}_\mathrm{a}^{-1}(n)]$. Since matrix $\mathbf{B}$ must be orthogonal, then the first N elements of $\mathbf{v}(n)$ are the components $\mathbf{v}_{//}(n)$ included in the space spanned by the rows of $\mathbf{B}_1$, and the other $KM-N$ components called $\mathbf{v}_\perp(n)$ are perpendicular to this space. Thus,

$$\mathbf{v}(n) = \begin{bmatrix}\mathbf{v}_{//}(n)\\ \mathbf{v}_\perp(n)\end{bmatrix}. \quad (21)$$

By substituting (20) and (21) into (19), we get

$$\begin{bmatrix}\mathbf{v}_{//}(n+1)\\ \mathbf{v}_\perp(n+1)\end{bmatrix} = \begin{bmatrix}\mathbf{v}_{//}(n)\\ \mathbf{v}_\perp(n)\end{bmatrix} + 2\tau\begin{bmatrix}\mathbf{\Lambda}\mathbf{p}_{\mathbf{x}d}\\ \mathbf{0}\end{bmatrix} - 2\tau\begin{bmatrix}\mathbf{\Lambda}\mathbf{R}_{\mathbf{xx}}\mathbf{T}_\mathrm{a}\mathbf{B}_1^\mathrm{T}\mathbf{v}_{//}(n) + \mathbf{\Lambda}\mathbf{R}_{\mathbf{xx}}\mathbf{T}_\mathrm{a}\mathbf{B}_2^\mathrm{T}\mathbf{v}_\perp(n)\\ \mathbf{0}\end{bmatrix} \quad (22)$$

where $\mathbf{\Lambda} = \mathbf{B}_1E[\mathbf{D}_\mathrm{a}^{-1}(n)]\mathbf{T}_\mathrm{a}^\mathrm{T}$. From (22) it can be seen that only the $\mathbf{v}_{//}(n)$ components are updated. In other words, vector $\mathbf{v}_\perp(n)$ remains with its given initial value. Thus, we can write

$$\mathbf{v}_{//}(\infty) = (\mathbf{T}_\mathrm{a}\mathbf{B}_1^\mathrm{T})^{-1}[\mathbf{R}_{\mathbf{xx}}^{-1}\mathbf{p}_{\mathbf{x}d} - \mathbf{T}_\mathrm{a}\mathbf{B}_2^\mathrm{T}\mathbf{v}_\perp(0)], \quad (23)$$

and

$$\mathbf{v}_\perp(\infty) = \mathbf{v}_\perp(0) = \mathbf{B}_2E[\mathbf{w}_\mathrm{a}(0)]. \quad (24)$$

The steady-state value $E[\mathbf{w}_\mathrm{a}(\infty)]$ can now be calculated by using (23) and (20). Thus,

$$E[\mathbf{w}_\mathrm{a}(\infty)] = \mathbf{B}^\mathrm{T}\begin{bmatrix}\mathbf{v}_{//}(\infty)\\ \mathbf{v}_\perp(\infty)\end{bmatrix} = \mathbf{B}^\mathrm{T}\begin{bmatrix}(\mathbf{T}_\mathrm{a}\mathbf{B}_1^\mathrm{T})^{-1}[\mathbf{R}_{\mathbf{xx}}^{-1}\mathbf{p}_{\mathbf{x}d} - \mathbf{T}_\mathrm{a}\mathbf{B}_2^\mathrm{T}\mathbf{v}_\perp(0)]\\ \mathbf{v}_\perp(0)\end{bmatrix} \quad (25)$$

Substituting (25) into the optimum weight vector expression, given by [1]

$$\mathbf{R}_{\mathbf{u}_\mathrm{a}\mathbf{u}_\mathrm{a}}\mathbf{w}_\mathrm{opt} = \mathbf{p}_{\mathbf{u}_\mathrm{a}d}, \quad (26)$$

it can be demonstrated that (25) satisfies relation (26).

3.3. Learning curve and second moment of the weight-error matrix

Now, let us consider the weight-error vector given by $\mathbf{v}_\mathrm{a}(n) = \mathbf{w}_\mathrm{a}(n) - \mathbf{w}_\mathrm{opt}$, where $\mathbf{w}_\mathrm{opt}$ is obtained from (25). By expressing the error signal as a function of both the transformed signals and weight-error vector, we have

$$\begin{aligned} e(n) &= d(n) - \mathbf{u}_\mathrm{a}^\mathrm{T}(n)\mathbf{v}_\mathrm{a}(n) - \mathbf{u}_\mathrm{a}^\mathrm{T}(n)\mathbf{w}_\mathrm{opt} + z(n)\\ &= e_\mathrm{o}(n) - \mathbf{u}_\mathrm{a}^\mathrm{T}(n)\mathbf{v}_\mathrm{a}(n), \end{aligned} \quad (27)$$

where $z(n)$ is a measurement noise, i.i.d., zero-mean with variance σ_z^2 and uncorrelated with any other signal in the system; $e_\mathrm{o}(n)$ represents the estimation error, given by

$$e_\mathrm{o}(n) = d(n) - \mathbf{u}_\mathrm{a}^\mathrm{T}(n)\mathbf{w}_\mathrm{opt} + z(n). \quad (28)$$

By squaring both sides of (27) and calculating the expected value of the resulting expression, we get

$$\begin{aligned} E[e^2(n)] &= E[e_\mathrm{o}^2(n)] - 2E[e_\mathrm{o}(n)\mathbf{u}_\mathrm{a}^\mathrm{T}(n)\mathbf{v}_\mathrm{a}(n)]\\ &\quad + E[\mathbf{v}_\mathrm{a}^\mathrm{T}(n)\mathbf{u}_\mathrm{a}(n)\mathbf{u}_\mathrm{a}^\mathrm{T}(n)\mathbf{v}_\mathrm{a}(n)]. \end{aligned} \quad (29)$$

According to the Principle of Orthogonality $E[e_\mathrm{o}(n)\mathbf{u}_\mathrm{a}(n)] = 0$, and by considering the analysis assumptions mentioned previously, the learning curve is given by

$$\begin{aligned} E[e^2(n)] &= e_\mathrm{min} + E\{\mathbf{v}_\mathrm{a}^\mathrm{T}(n)E[\mathbf{u}_\mathrm{a}(n)\mathbf{u}_\mathrm{a}^\mathrm{T}(n)]\mathbf{v}_\mathrm{a}(n)\}\\ &= e_\mathrm{min} + \mathrm{tr}\{\mathbf{R}_{\mathbf{u}_\mathrm{a}\mathbf{u}_\mathrm{a}}E[\mathbf{v}_\mathrm{a}(n)\mathbf{v}_\mathrm{a}^\mathrm{T}(n)]\}, \end{aligned} \quad (30)$$

where $e_\mathrm{min} = E[e_\mathrm{o}^2(n)]$ is the minimum error in the steady-state condition. Note that (30) is completely determined if the weight-error covariance matrix $\mathbf{K}(n) = E[\mathbf{v}_\mathrm{a}(n)\mathbf{v}_\mathrm{a}^\mathrm{T}(n)]$ is known. Then, by subtracting $\mathbf{w}_\mathrm{opt}$ from both sides of (19), determining the outer product $\mathbf{v}_\mathrm{a}(n)\mathbf{v}_\mathrm{a}^\mathrm{T}(n)$, and taking the expectation on both sides of the resulting expression according to the simplifying assumptions, we obtain

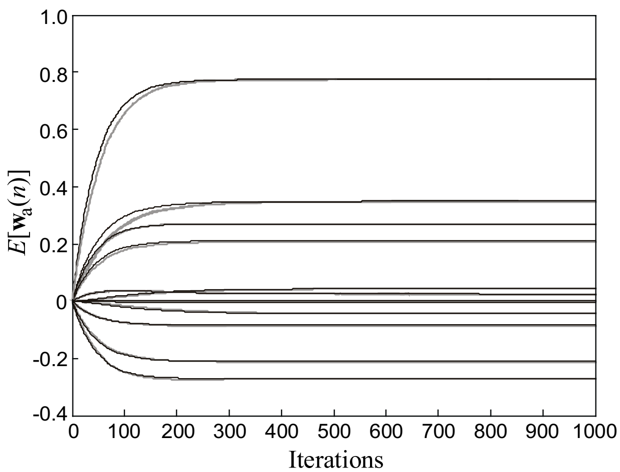
$$\begin{aligned} \mathbf{K}(n+1) &= \mathbf{K}(n) - \tau\mathbf{K}(n)\mathbf{R}_{\mathbf{u}_\mathrm{a}\mathbf{u}_\mathrm{a}}E[\mathbf{D}_\mathrm{a}^{-1}(n)]\\ &\quad - \tau E[\mathbf{D}_\mathrm{a}^{-1}(n)]\mathbf{R}_{\mathbf{u}_\mathrm{a}\mathbf{u}_\mathrm{a}}\mathbf{K}(n)\\ &\quad + 2\tau^2E[\mathbf{D}_\mathrm{a}^{-1}(n)]\{2\mathbf{R}_{\mathbf{u}_\mathrm{a}\mathbf{u}_\mathrm{a}}\mathbf{K}(n)\mathbf{R}_{\mathbf{u}_\mathrm{a}\mathbf{u}_\mathrm{a}}\\ &\quad + \mathbf{R}_{\mathbf{u}_\mathrm{a}\mathbf{u}_\mathrm{a}}\mathrm{tr}[\mathbf{R}_{\mathbf{u}_\mathrm{a}\mathbf{u}_\mathrm{a}}\mathbf{K}(n)]\}E[\mathbf{D}_\mathrm{a}^{-1}(n)]\\ &\quad + 2\tau^2E[\mathbf{D}_\mathrm{a}^{-1}(n)]\mathbf{R}_{\mathbf{u}_\mathrm{a}\mathbf{u}_\mathrm{a}}E[\mathbf{D}_\mathrm{a}^{-1}(n)]e_\mathrm{min}. \end{aligned} \quad (31)$$

4. SIMULATION RESULTS

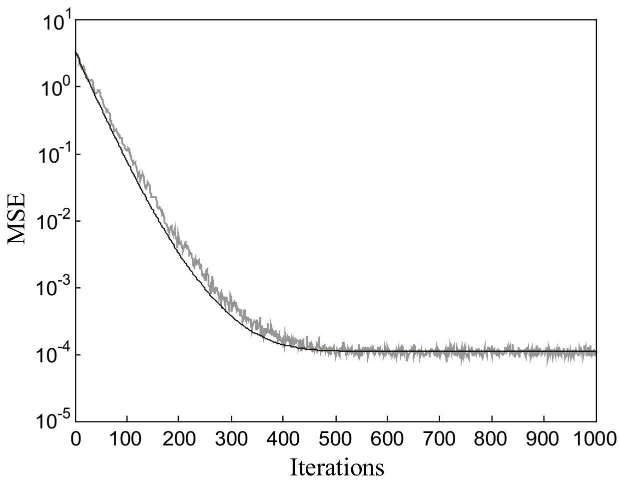
To assess the accuracy of the proposed model some examples are presented considering a system identification problem, using white and colored real input signals. The latter is obtained from an AR(2) process given by $x(n) = a_1x(n-1) + a_2x(n-2) + v(n)$, where $v(n)$ is a white noise signal with variance σ_v^2. The AR coefficients are $a_1 = 1.3214$ and $a_2 = -0.8500$, resulting in an eigenvalue dispersion of the input autocorrelation matrix of 305.89. The measurement noise $z(n)$ has a variance $\sigma_z^2 = 10^{-4}$ (SNR = 40 dB). The step-size values are $\tau = 0.005$ and $\tau = 0.01$ for white and colored signals, respectively. The other parameters are $M = 4$, $K = 4$, $L = 2$, observation window length $M_\mathrm{w} = 32$, and the plant is a length-10 vector Hanning window. The transformation used is DCT.

The curves of Fig. 2 are obtained for a white input signal, where Figs. 2(a) and (b) illustrate the first and second moments of the adaptive filter weights, respectively, obtained from Monte Carlo (MC) simulations (average of 200 independent runs) and the proposed model (17) (first moment) and (30) (learning curve). From that figure, a very good match between numerical simulations and models is verified.

Fig. 3 illustrates the case for a colored input signal. Again, we can observe a satisfactory accuracy of the predictions obtained with the proposed model.



(a)

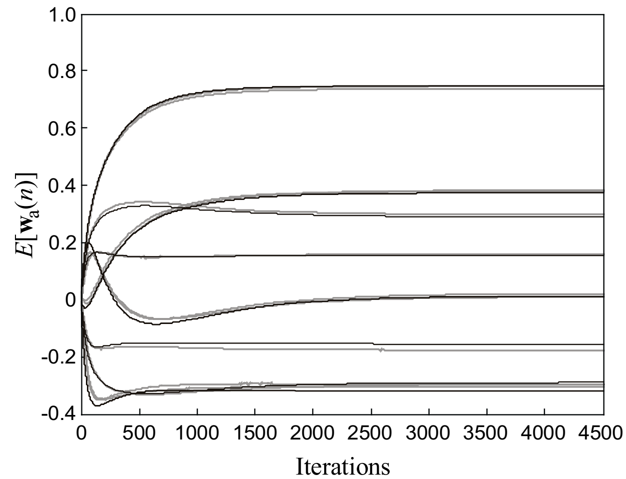


(b)

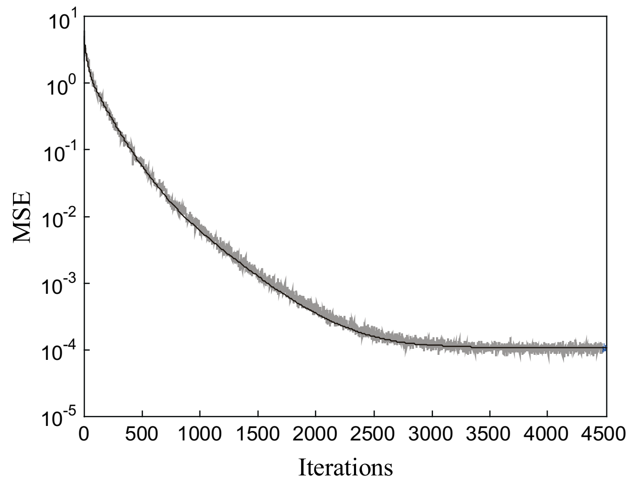
Fig. 2. Model performance for white input signal with $\tau = 0.01$. (a) Mean-weight behavior. (Gray lines) MC simulation (average of 200 independent runs). (Dark lines) proposed model. (b) MSE curves. (Gray-ragged line) MC simulation. (Dark-solid line) proposed model.

5. CONCLUSIONS

This paper has presented a stochastic model for the GSD-NLMS algorithm. The proposed model is independent of the order of the adaptive filter as well as of the type of orthogonal transform used. The proposed analytical model is derived for a slow adaptation condition and takes into account a regularization parameter (added to the power estimate) which prevents division by zero in the power normalization operation, being possible to assess its effect on the algorithm behavior. Under the analysis conditions, the proposed model exhibits a very good matching with numerical simulations.



(a)



(b)

Fig. 3. Model performance for colored input signal with $\tau = 0.005$. (a) Mean-weight behavior. (Gray lines) MC simulation (average of 300 independent runs). (Dark lines) proposed model. (b) MSE curves. (Gray-ragged line) MC simulation. (Dark-solid line) proposed model.

6. REFERENCES

[1] M. R. Petraglia and S. K. Mitra, "Adaptive FIR filter structure based on the generalized subband decomposition of FIR filters," *IEEE Trans. Circuits and Systems II*, vol. 40, no. 6, pp. 354-362, June 1993.

[2] E. M Lobato, O. J. Tobias, and R. Seara, "A stochastic model for the transform-domain LMS algorithm", *XII European Signal Processing Conf.*, Vienna, Austria, Sep. 2004, pp. 1833-1836.

[3] S. Haykin, *Adaptive Filter Theory*, 4th ed., Upper Saddler River, NJ: Prentice-Hall, 2002.

[4] B. Farhang-Boroujeny, *Adaptive Filters: Theory and Applications*, John Wiley & Sons, 1998.

[5] C. G. Samson and U. Reddy, "Fixed point error analysis of the normalized ladder algorithm," *IEEE Trans. Acoust., Speech, and Signal Processing*, vol. 31, no. 5, pp. 1177-1191, Oct. 1983.

[6] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, 3ed., McGraw-Hill, 1991.