

A New Post-whitening Transform Domain LMS Algorithm

Laid Chergui^{1,3*}, Saad Bouguezel^{2,3}

¹ Department of Industrial Engineering, Faculty of Sciences and Technology, University of Abbas Laghrour, Khenchela 40000, Algeria

² Department of Electronics, Faculty of Technology, University of Ferhat Abbas, Setif 1, 19000, Algeria

³ LCCNS Laboratory, ELN Department, University of Ferhat Abbas, Setif 1, 19000, Algeria

Corresponding Author Email: chergui.laid@gmail.com

<https://doi.org/10.18280/ts.360307>

ABSTRACT

Received: 5 March 2019

Accepted: 29 May 2019

Keywords:

eigen-value spread, orthogonal transforms, post-whitening, predictive decorrelation, system identification, TDMLS

This paper proposes a new post-whitening transform domain LMS (POW-TDLMS) algorithm for system identification purposes, where the post whitened and original transformed signals are used during the adaptation and filtering phases, respectively. The main idea behind the proposed algorithm is to introduce a first order adaptive post-whitening filter in the TDLMS algorithm after applying the transform to completely decorrelate the transformed signal. Linear prediction is adopted for the post-whitening and the prediction coefficients are adapted in the time domain. Furthermore, the mean convergence performance analysis of the proposed POW-TDLMS algorithm is presented. The simulation results show the superiority of the proposed POW-TDLMS algorithm compared to the conventional TDLMS algorithm in terms of the MSE convergence speed and reached steady state.

1. INTRODUCTION

The least mean square (LMS) algorithm is an adaptive filtering algorithm known for its simplicity, robustness and low computational complexity [1-3]. The convergence of this algorithm becomes slow in the case of correlated input signals. The autocorrelation matrix of such signals is said ill conditioned [4-6].

In the literature, several solutions and alternatives have been proposed to improve the performance of the LMS algorithm when it is employed with correlated input signals. In [7], Mboub et al. proposed a new whitening structure for the acoustic echo cancellation. In [8, 9], the authors devised a pre-whitening for the time domain LMS algorithm and its variants. The resulting pre-whitening techniques have allowed improvements in the convergence performance. The affine projection algorithms were developed to speed up the convergence of the gradient based algorithms in the case of correlated input signals [1, 10].

In [11], Narayan et al. proposed a new LMS algorithm operating in the transform domain, where orthogonal and unitary transforms have been used for the decorrelation of the input signal. The transformed signal is then normalised by its power, which allows the reduction of the eigenvalue spread of the autocorrelation matrix of the resulting signal, and thus offers improvements in the convergence performance of the resulting algorithm called transform domain LMS algorithm (TDLMS) [12]. The convergence performance of the TDLMS algorithm depends on the used orthogonal transforms, which are known to be limited in terms of decorrelation [13, 14]. In [13], Chergui et Bouguezel proposed a new pre-whitening (PW) of the TDLMS (PW-

TDLMS) algorithm. This pre-whitening allows the reinforcement of the decorrelation of the used orthogonal transforms, and consequently the improvements in the convergence performance of the corresponding PW-TDLMS algorithms. The PW-TDLMS algorithm was intended for applications of adaptive noise cancellation in the speech signal, where the pre-whitened and transformed signal is used during both filtering and adaptation phases. However, the PW-TDLMS algorithm cannot be applied in system identification applications, where the input signal characteristics must not be changed before the filtering phase [15, 16] and may only be modified during the adaptation phase.

In this paper, to deal with system identification applications, we propose a post-whitening (POW) of the TDLMS algorithm to ensure further decorrelation of the transformed signal and significantly reinforce the decorrelation of the used orthogonal transforms. The post-whitening process is designed using a first order adaptive decorrelation filter based on linear prediction and the prediction coefficients are adapted in the time domain. In the proposed algorithm, the post whitened and original transformed signals are used during the adaptation and filtering phases, respectively. This allows improvements in the convergence performance of the TDLMS algorithm in system identification applications.

This paper is organized as follows: Section 2 gives a review of the conventional TDLMS algorithm. Section 3 proposes a new post-whitening TDLMS (POW-TDLMS) algorithm. Section 4 presents a study of the mean convergence performance of the proposed POW-TDLMS algorithm. Section 5 discusses the simulation results. Finally, the conclusion is given in Section 6.

2. CONVENTIONAL TDLMS ALGORITHM

The conventional TDLMS algorithm presented in Figure 1 can be summarised as [17, 18]:

Let $\mathbf{x}_k^N = [x_k, x_{k-1}, \dots, x_{k-N+1}]^T$ be the input signal of length N , where $(\cdot)^T$ denotes the transposition operation. Its transformed version \mathbf{X}_k^N is also of length N and obtained using the $N \times N$ orthogonal transform matrix \mathbf{T}_N as

$$\mathbf{X}_k^N = \mathbf{T}_N \mathbf{x}_k^N = [X_k(0), X_k(1), \dots, X_k(N-1)]^T \quad (1)$$

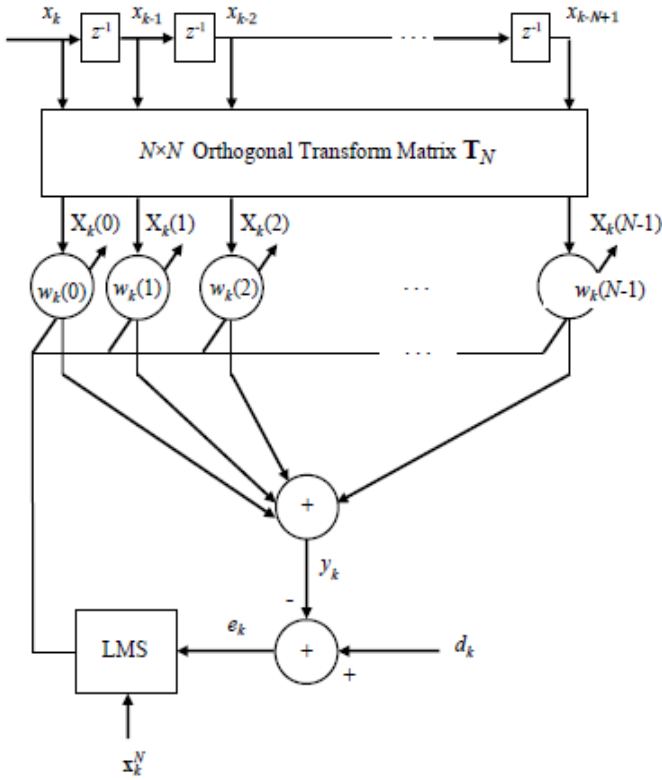


Figure 1. Block diagram of the TDLMS adaptive filter [13, 18]

The sample y_k at the output of the filter is obtained as

$$y_k = (\mathbf{w}_k^N)^H \mathbf{X}_k^N \quad (2)$$

where, $(\cdot)^H$ denotes the complex conjugate transpose operation and $\mathbf{w}_k^N = [w_k(0), w_k(1), \dots, w_k(N-1)]^T$ is the filter vector of length N , which is updated at each instant sample k according to the adaptation equation given by

$$\mathbf{w}_{k+1}^N = \mathbf{w}_k^N + \mu e_k \mathbf{P}_k^{-1} \mathbf{X}_k^N \quad (3)$$

with $e_k = d_k - y_k$ being the sample of the error signal and d_k is the sample of the desired signal, $\mathbf{P}_k = \text{diag}(\sigma_k^2(0), \sigma_k^2(1), \dots, \sigma_k^2(N-1))$, the entries $\sigma_k^2(i)$, $i = 0, 1, \dots, N-1$, are obtained from the power estimate $\sigma_k^2 = [\sigma_k^2(0), \sigma_k^2(1), \dots, \sigma_k^2(N-1)]^T$ of \mathbf{X}_k^N , which can be computed recursively at each time sample k as $\sigma_k^2(i) = \beta \sigma_{k-1}^2(i) + (1 - \beta) |X_k(i)|^2$, β is a smoothing factor that takes values in the interval $]0, 1[$. The matrix \mathbf{P}_k is used for normalising the transformed signal and the constant step size μ is employed for controlling the MSE convergence of the TDLMS algorithm.

2.1 Convergence analysis

In this subsection we briefly review the convergence analysis with first order autoregressive (AR1) process. The signal x_k obtained from the AR1 process is defined at each instant sample k as [13, 20, 21]:

$$x_k = \rho x_{k-1} + \omega_k \quad (4)$$

where, ρ is the correlation coefficient that takes values in the interval $]0, 1[$ and ω_k is a stationary white Gaussian noise with variance that makes the power of x_k equal to 1.

The autocorrelation matrix of the input signal x_k is defined as:

$$\mathbf{R}_N = \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{N-1} \\ \rho & 1 & \rho & \dots & \rho^{N-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{N-1} & \rho^{N-2} & \dots & \dots & 1 \end{pmatrix} \quad (5)$$

The eigenvalue spread of \mathbf{R}_N is defined as:

$$\text{Lim}_{N \rightarrow \infty} (\text{Eigenvalue spread of } \mathbf{R}_N) = \left(\frac{1+\rho}{1-\rho} \right)^2 \quad (6)$$

A high value of ρ increases the correlation of x_k , and consequently increases the eigenvalue spread of the autocorrelation matrix \mathbf{R}_N . The orthogonal transforms namely the DCT, the DFT and the DHT are used in the TDLMS as a decorrelator, which allows the whitening of the input signal x_k before being used by the LMS algorithm and consequently improves the convergence performances of the TDLMS algorithm.

The autocorrelation matrix of the transformed signal and normalised by its power is defined as:

$$\mathbf{S}_N = (\text{diag}(\mathbf{B}_N))^{-\frac{1}{2}} \mathbf{B}_N (\text{diag}(\mathbf{B}_N))^{-\frac{1}{2}} \quad (7)$$

and $\mathbf{B}_N = \mathbf{T}_N \mathbf{R}_N \mathbf{T}_N^H$.

The eigenvalue spreads of \mathbf{S}_N are defined for various transforms as:

$$\text{Lim}_{N \rightarrow \infty} \left(\begin{array}{l} \text{Eigenvalue spread after DFT or} \\ \text{DHT and power normalisation} \end{array} \right) = \left(\frac{1+\rho}{1-\rho} \right) \quad (8)$$

$$\text{Lim}_{N \rightarrow \infty} \left(\begin{array}{l} \text{Eigenvalue spread after DCT} \\ \text{and power normalisation} \end{array} \right) = (1 + \rho) \quad (9)$$

The different transforms significantly decrease the eigenvalue spread of the autocorrelation matrix \mathbf{S}_N and the DCT is considered as suboptimal in terms of decorrelation and offers the best reduction in the eigenvalue spread of \mathbf{S}_N compared to the DFT and the DHT transforms.

3. PROPOSED POST-WHITENING TRANSFORM DOMAIN LMS ALGORITHM

3.1 Proposed post-whitening approach

The decorrelation based on the linear prediction is given in time domain as [9, 13]

$$\tilde{x}_k = x_k - (\mathbf{a}_{k-1}^{L-1})^T \mathbf{x}_{k-1}^{L-1} \quad (10)$$

where, $\mathbf{x}_{k-1}^{L-1} = [x_{k-1}, x_{k-2}, \dots, x_{k-L+1}]^T$ is the input sub-sequence of length $L-1$, which is supposed to be correlated, and $\mathbf{a}_{k-1}^{L-1} = [a_{k-1}(1), a_{k-1}(2), \dots, a_{k-1}(L-1)]^T$, which can be computed adaptively at each time sample k using LMS algorithm as

$$\mathbf{a}_k^{L-1} = \mathbf{a}_{k-1}^{L-1} + \gamma \tilde{x}_k \mathbf{x}_{k-1}^{L-1} \quad (11)$$

with γ being the adaptation step size parameter. The decorrelation filter used in Eq. (10) is of order $L-1$ and has a finite impulse response (FIR) $[1, -(a_{k-1}^{L-1})^T]$ of length L . Using Eq. (10), the preceding $N-1$ samples of \tilde{x}_k can be obtained by decorrelating the preceding $N-1$ samples of x_k as

$$\begin{aligned} \tilde{x}_{k-1} &= x_{k-1} - (\mathbf{a}_{k-2}^{L-1})^T \mathbf{x}_{k-2}^{L-1} \\ &\vdots \\ \tilde{x}_{k-N+1} &= x_{k-N+1} - (\mathbf{a}_{k-N}^{L-1})^T \mathbf{x}_{k-N}^{L-1} \end{aligned} \quad (12)$$

Therefore, Eq. (10) and Eq. (12) can be combined and

compactly formulated as

$$\tilde{\mathbf{x}}_k^N = \mathbf{x}_k^N - \left[(\mathbf{a}_{k-1}^{L-1})^T \mathbf{x}_{k-1}^{L-1}, (\mathbf{a}_{k-2}^{L-1})^T \mathbf{x}_{k-2}^{L-1}, \dots, (\mathbf{a}_{k-N}^{L-1})^T \mathbf{x}_{k-N}^{L-1} \right]^T \quad (13)$$

It is clear that the decorrelation in Eq. (13) is achieved using the FIR decorrelation filters $[1, -(a_{k-1}^{L-1})^T]$, $[1, -(a_{k-2}^{L-1})^T]$, ..., and $[1, -(a_{k-N}^{L-1})^T]$ of order $L-1$. Finally, by multiplying both sides of Eq. (13) by \mathbf{T}_N , the proposed post-whitening of the transformed signal \mathbf{X}_k^N can be obtained as

$$\tilde{\mathbf{X}}_k^N = \mathbf{X}_k^N - \mathbf{T}_N \times \left[(\mathbf{a}_{k-1}^{L-1})^T \mathbf{x}_{k-1}^{L-1}, (\mathbf{a}_{k-2}^{L-1})^T \mathbf{x}_{k-2}^{L-1}, \dots, (\mathbf{a}_{k-N}^{L-1})^T \mathbf{x}_{k-N}^{L-1} \right]^T \quad (14)$$

For simplicity, we consider in this paper the case of first order decorrelation filters, i.e. $L=2$. Thus, Eq. (14) becomes

$$\begin{aligned} \tilde{\mathbf{X}}_k^N &= \mathbf{X}_k^N - \mathbf{T}_N \times [a_{k-1}(1)x_{k-1}, a_{k-2}(1)x_{k-2}, \dots, a_{k-N}(1)x_{k-N}]^T \\ &= \mathbf{X}_k^N - \mathbf{T}_N \times \text{diag}(a_{k-1}(1), a_{k-2}(1), \dots, a_{k-N}(1)) \times \mathbf{x}_{k-1}^N \end{aligned} \quad (15)$$

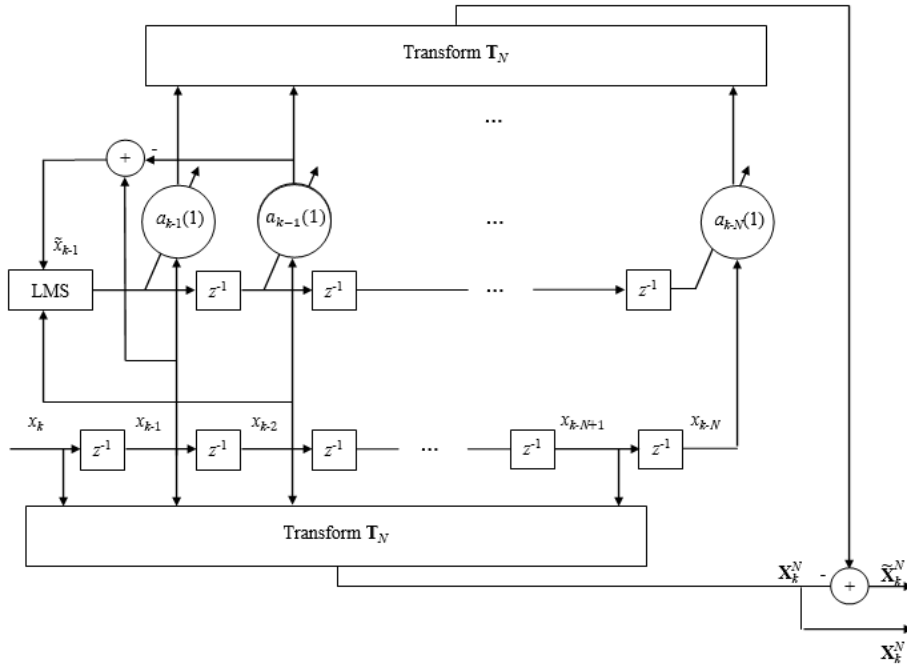


Figure 2. Proposed first order adaptive post-whitening

In this case ($L=2$), the prediction coefficients $a_{k-i}(1)$, $i=1, 2, \dots, N$, in Eq. (15) are computed using the simplified versions of Eq. (10) and Eq. (11) as

$$\tilde{x}_k = x_k - a_{k-1}(1)x_{k-1} \quad (16)$$

$$a_k(1) = a_{k-1}(1) + \gamma \tilde{x}_k x_{k-1} \quad (17)$$

Therefore, Eq. (15), Eq. (16) and Eq. (17) represent the first order adaptive post-whitening process of the transformed signal \mathbf{X}_k^N , which is illustrated by Figure 2, where the input signal $\mathbf{x}_{k-1}^N = [x_{k-1}, x_{k-2}, \dots, x_{k-N}]^T$ is weighted using the prediction coefficients $a_{k-i}(1)$, $i=1, 2, \dots, N$, before being transformed by \mathbf{T}_N and the resulting vector is subtracted from \mathbf{X}_k^N to provide the post-whitened signal $\tilde{\mathbf{X}}_k^N = [\tilde{X}_k(0),$

$$\tilde{X}_k(1), \dots, \tilde{X}_k(N-1)]^T.$$

3.2 Proposed post whitening TDLMS algorithm

In this subsection, we introduce the proposed post-whitening given by Eq. (15), Eq. (16), Eq. (17) and Figure 2 in the TDLMS to develop a new Post-whitening TDLMS (POW-TDLMS) algorithm to reinforce the decorrelation of the used orthogonal transform \mathbf{T}_N , and consequently, to decrease the eigenvalue spread of the autocorrelation matrix of the post whitened signal $\tilde{\mathbf{X}}_k^N$ normalised by its power. Therefore, the convergence performance of the TDLMS would be improved in terms of the MSE convergence behaviour.

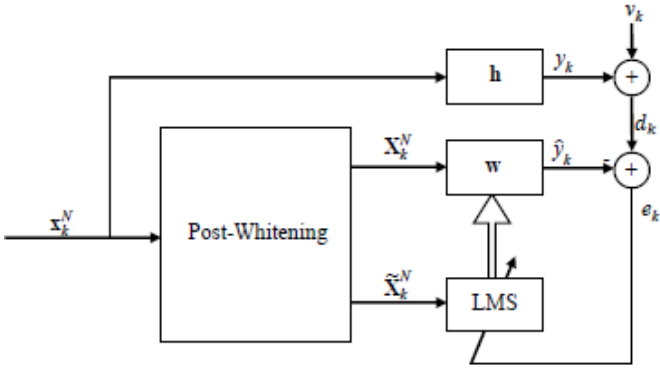


Figure 3. Proposed post-whitening TDLMS algorithm

After introducing the proposed post-whitening approach, we develop the proposed POW-TDLMS algorithm for system identification as shown in Figure 3, where v_k is the observation noise generated from a random white Gaussian process, and $\mathbf{h} = [h(0), h(1), \dots, h(N-1)]^T$ is the unknown system impulse response to be identified by the adaptive filter.

In Figure 3, the error signal at the output of the filter is given by

$$e_k = d_k - (\mathbf{w}_k^N)^H \mathbf{X}_k^N \quad (18)$$

and the desired signal is given by

$$d_k = v_k + \sum_{i=0}^{N-1} h(i) x_{k-i} \quad (19)$$

The post-whitened signal $\tilde{\mathbf{X}}_k^N$ is normalised by its power defined as

$$\tilde{\mathbf{P}}_k = \text{diag}([\tilde{\sigma}_k^2(0), \tilde{\sigma}_k^2(1), \dots, \tilde{\sigma}_k^2(N-1)]) \quad (20)$$

where, $\tilde{\sigma}_k^2(i)$ is the estimation of the power of the i^{th} input $\tilde{X}_k(i)$, which can be computed recursively as

$$\tilde{\sigma}_k^2(i) = \beta \tilde{\sigma}_{k-1}^2(i) + (1-\beta) |\tilde{X}_k(i)|^2 \quad (21)$$

The filter weight vector is updated recursively as

$$\mathbf{w}_{k+1}^N = \mathbf{w}_k^N + \mu e_k \tilde{\mathbf{P}}_k^{-1} \tilde{\mathbf{X}}_k^N \quad (22)$$

where, $\tilde{\mathbf{P}}_k^{-1} \tilde{\mathbf{X}}_k^N$ represents the normalized post-whitened signal.

4. PERFORMANCE ANALYSIS OF THE PROPOSED POW-TDLMS ALGORITHM

4.1 Stability performance

In this sub-section, we present the mean convergence behaviour analysis of the proposed POW-TDLMS algorithm. Similar to [19], it is seen from Eq. (22) that the power normalization performed by $\tilde{\mathbf{P}}_k^{-1}$ on the step size, allowing improvement in the MSE convergence behaviour of the algorithm, increases the difficulty of analysing the stability performance. In order to simplify the analysis, we equivalently perform the power normalization on the post

whitened vector $\tilde{\mathbf{X}}_k^N$. The equivalent form is obtained by multiplying both sides of Eq. (22) by $\tilde{\mathbf{P}}_k^{1/2}$ as

$$\tilde{\mathbf{P}}_k^{1/2} \mathbf{w}_{k+1}^N = \tilde{\mathbf{P}}_k^{1/2} \mathbf{w}_k^N + \mu e_k \tilde{\mathbf{P}}_k^{-1/2} \tilde{\mathbf{X}}_k^N \quad (23)$$

We assume that $\tilde{\mathbf{X}}_k^N$ is a zero mean stationary process. Thus, it can be assumed that $\tilde{\mathbf{P}}_{k+1}^{1/2} \approx \tilde{\mathbf{P}}_k^{1/2}$ for which Eq. (23) becomes

$$\hat{\mathbf{w}}_{k+1}^N = \hat{\mathbf{w}}_k^N + \mu e_k \mathbf{V}_k^N \quad (24)$$

where, $\hat{\mathbf{w}}_{k+1}^N = \tilde{\mathbf{P}}_{k+1}^{1/2} \mathbf{w}_{k+1}^N \approx \tilde{\mathbf{P}}_k^{1/2} \mathbf{w}_{k+1}^N$, $\hat{\mathbf{w}}_k^N = \tilde{\mathbf{P}}_k^{1/2} \mathbf{w}_k^N$ and $\mathbf{V}_k^N = \tilde{\mathbf{P}}_k^{-1/2} \tilde{\mathbf{X}}_k^N$.

The weight vector \mathbf{w}_k^N of Eq. (22) converges to $\mathbf{T}_N \mathbf{w}^o$, where \mathbf{w}^o is the wiener solution. We can easily check that the weight vector $\hat{\mathbf{w}}_k^N$ of Eq. (24) converges to $\tilde{\mathbf{P}}^{1/2} \mathbf{T}_N \mathbf{w}^o$, where $\tilde{\mathbf{P}} = \mathbb{E}\{\tilde{\mathbf{P}}_k\}$. Therefore, the two forms of Eq. (22) and Eq. (24) are equivalent in stationary and mildly non stationary environments.

By subtracting the transformed and normalized wiener filter $\tilde{\mathbf{P}}^{1/2} \mathbf{T}_N \mathbf{w}^o$ from both sides of Eq. (24), we obtain

$$\tilde{\mathbf{w}}_{k+1}^N = \tilde{\mathbf{w}}_k^N + \mu e_k \mathbf{V}_k^N \quad (25)$$

where, $\tilde{\mathbf{w}}_k^N = \hat{\mathbf{w}}_k^N - \tilde{\mathbf{P}}^{1/2} \mathbf{T}_N \mathbf{w}^o$ known as the weight-error vector.

We assume that the decorrelation filter $[1, -a_{k-1}(1)]$ is converged to its optimal value, where at each time sample k after the convergence was reached $a_{k-1}(1) = a$, with a is a scalar that takes values in the interval $]0, 1[$. Thus, Eq. (16) becomes

$$\tilde{x}_k = x_k - a x_{k-1} \quad (26)$$

and Eq. (15) becomes

$$\tilde{\mathbf{X}}_k^N = \mathbf{X}_k^N - a \mathbf{X}_{k-1}^N \quad (27)$$

Let us express e_k defined in Eq. (18) in terms of $\tilde{\mathbf{X}}_k^N$. The input signal \mathbf{X}_k^N is generated from AR1 process according to Eq. (4)

Therefore,

$$\mathbf{x}_k^N = \rho \mathbf{x}_{k-1}^N + \boldsymbol{\omega}_k^N \quad (28)$$

where, $\boldsymbol{\omega}_k^N = [\omega_k, \omega_{k-1}, \dots, \omega_{k-N+1}]^T$. By multiplying both sides of Eq. (28) by \mathbf{T}_N , we obtain

$$\mathbf{X}_k^N = \rho \mathbf{X}_{k-1}^N + \boldsymbol{\omega}_k^N \quad (29)$$

where, $\boldsymbol{\omega}_k^N = \mathbf{T}_N \boldsymbol{\omega}_k^N$. Thus, Eq. (29) can be reorganized as

$$\mathbf{X}_{k-1}^N = \frac{1}{\rho} \mathbf{X}_k^N - \frac{1}{\rho} \boldsymbol{\omega}_k^N \quad (30)$$

By substituting Eq. (30) in Eq. (27) and making some rearrangement, we obtain

$$\mathbf{X}_k^N = \frac{\rho}{\rho-a} \tilde{\mathbf{X}}_k^N - \frac{a}{\rho-a} \boldsymbol{\omega}_k^N \quad (31)$$

By substituting Eq. (31) in Eq. (18), we obtain an expression for e_k as

$$e_k = d_k - (\mathbf{w}_k^N)^H \left(\alpha \tilde{\mathbf{X}}_k^N - \varphi \mathbf{w}_k^N \right) \quad (32)$$

where, $\alpha = \frac{\rho}{\rho-a}$, $\varphi = \frac{a}{\rho-a}$, $0 < a \leq 1$, $0 < \rho < 1$, ρ is the correlation coefficient and $a \neq \rho$.

Let us now express e_k given by Eq. (32) in terms of $\tilde{\mathbf{w}}_k^N$ and \mathbf{V}_k^N as

$$\begin{aligned} e_k &= d_k - (\mathbf{w}_k^N)^H \tilde{\mathbf{P}}_k^{1/2} \tilde{\mathbf{P}}_k^{-1/2} \left(\alpha \tilde{\mathbf{X}}_k^N - \varphi \mathbf{w}_k^N \right) \\ &= d_k - (\mathbf{w}_k^N)^H \tilde{\mathbf{P}}_k^{1/2} \left(\alpha \tilde{\mathbf{P}}_k^{-1/2} \tilde{\mathbf{X}}_k^N - \varphi \tilde{\mathbf{P}}_k^{-1/2} \mathbf{w}_k^N \right) \\ &= d_k - (\tilde{\mathbf{w}}_k^N)^H \left(\alpha \mathbf{V}_k^N - \varphi \tilde{\mathbf{P}}_k^{-1/2} \mathbf{w}_k^N \right) \\ &= d_k - \left(\alpha \mathbf{V}_k^N - \varphi \tilde{\mathbf{P}}_k^{-1/2} \mathbf{w}_k^N \right)^H \tilde{\mathbf{w}}_k^N \\ &= d_k - \left(\alpha \mathbf{V}_k^N - \varphi \tilde{\mathbf{P}}_k^{-1/2} \mathbf{w}_k^N \right)^H \tilde{\mathbf{w}}_k^N \\ &\quad + \left(\alpha \mathbf{V}_k^N - \varphi \tilde{\mathbf{P}}_k^{-1/2} \mathbf{w}_k^N \right)^H \tilde{\mathbf{P}}_k^{1/2} \mathbf{T}_N \mathbf{w}^o \\ &\quad - \left(\alpha \mathbf{V}_k^N - \varphi \tilde{\mathbf{P}}_k^{-1/2} \mathbf{w}_k^N \right)^H \tilde{\mathbf{P}}_k^{1/2} \mathbf{T}_N \mathbf{w}^o \\ &= e^o - \left(\alpha \mathbf{V}_k^N - \varphi \tilde{\mathbf{P}}_k^{-1/2} \mathbf{w}_k^N \right)^H \tilde{\mathbf{w}}_k^N \\ &\quad + \left(\alpha \mathbf{V}_k^N - \varphi \tilde{\mathbf{P}}_k^{-1/2} \mathbf{w}_k^N \right)^H \tilde{\mathbf{P}}_k^{1/2} \mathbf{T}_N \mathbf{w}^o \\ &= e^o - \left(\alpha \mathbf{V}_k^N - \varphi \tilde{\mathbf{P}}_k^{-1/2} \mathbf{w}_k^N \right)^H \left(\tilde{\mathbf{w}}_k^N - \tilde{\mathbf{P}}_k^{1/2} \mathbf{T}_N \mathbf{w}^o \right) \\ &= e^o - \left(\alpha \mathbf{V}_k^N - \varphi \tilde{\mathbf{P}}_k^{-1/2} \mathbf{w}_k^N \right)^H \tilde{\mathbf{w}}_k^N \end{aligned} \quad (33)$$

where, $e^o = d_k - \left(\alpha \mathbf{V}_k^N - \varphi \tilde{\mathbf{P}}_k^{-1/2} \mathbf{w}_k^N \right)^H \tilde{\mathbf{P}}_k^{1/2} \mathbf{T}_N \mathbf{w}^o$ is the error when the filter is optimum.

By substituting Eq. (33) in Eq. (25), we obtain,

$$\begin{aligned} \tilde{\mathbf{w}}_{k+1}^N &= \tilde{\mathbf{w}}_k^N + \mu \mathbf{V}_k^N \left[e^o - \left(\alpha \mathbf{V}_k^N - \varphi \tilde{\mathbf{P}}_k^{-1/2} \mathbf{w}_k^N \right)^H \tilde{\mathbf{w}}_k^N \right] = \tilde{\mathbf{w}}_k^N + \\ &\quad \mu \left[\mathbf{V}_k^N e^o - \left(\alpha \mathbf{V}_k^N (\mathbf{V}_k^N)^H - \varphi \mathbf{V}_k^N (\mathbf{w}_k^N)^H \tilde{\mathbf{P}}_k^{-1/2} \right) \tilde{\mathbf{w}}_k^N \right] \end{aligned} \quad (34)$$

Assuming that at time k the coefficients of $\tilde{\mathbf{w}}_k^N$ are independent of \mathbf{V}_k^N , \mathbf{w}_k^N and $\tilde{\mathbf{P}}_k^{-1/2}$. This assumption can be justified by taking small values of μ .

By taking the expectation of both sides of Eq. (34), we obtain,

$$\mathbb{E}\{\tilde{\mathbf{w}}_{k+1}^N\} = \mathbb{E}\{\tilde{\mathbf{w}}_k^N\} + \mu \times$$

$$\begin{aligned} &\left[\mathbb{E}\{\mathbf{V}_k^N e^o\} - \left(\alpha \mathbb{E}\{\mathbf{V}_k^N (\mathbf{V}_k^N)^H\} - \right. \right. \\ &\quad \left. \left. \mathbb{E}\{\varphi \mathbf{V}_k^N (\mathbf{w}_k^N)^H \tilde{\mathbf{P}}_k^{-1/2}\} \right) \mathbb{E}\{\tilde{\mathbf{w}}_k^N\} \right] \\ &= [\mathbf{I} - \mu \alpha \tilde{\mathbf{S}}_N] \mathbb{E}\{\tilde{\mathbf{w}}_k^N\} + \\ &\quad \mu \varphi \mathbb{E}\{\mathbf{V}_k^N (\mathbf{w}_k^N)^H \tilde{\mathbf{P}}_k^{-1/2}\} \mathbb{E}\{\tilde{\mathbf{w}}_k^N\} \end{aligned} \quad (35)$$

where, $\mathbb{E}\{\mathbf{V}_k^N e^o\} = 0$, $\tilde{\mathbf{S}}_N = \mathbb{E}\{\mathbf{V}_k^N (\mathbf{V}_k^N)^H\}$.

We assume that the proposed algorithm is convergent. Consequently, $\mathbf{V}_k^N \approx \tilde{\mathbf{P}}_k^{-1/2} \mathbf{w}_k^N$, $\mathbb{E}\{\mathbf{V}_k^N (\mathbf{w}_k^N)^H \tilde{\mathbf{P}}_k^{-1/2}\} \approx \tilde{\mathbf{S}}_N$ and Eq. (35) becomes

$$\begin{aligned} \mathbb{E}\{\tilde{\mathbf{w}}_{k+1}^N\} &= \left[\mathbf{I} - \mu \left(\frac{\rho}{\rho-a} \right) \tilde{\mathbf{S}}_N \right] \mathbb{E}\{\tilde{\mathbf{w}}_k^N\} + \mu \left(\frac{a}{\rho-a} \right) \tilde{\mathbf{S}}_N \mathbb{E}\{\tilde{\mathbf{w}}_k^N\} \\ &= [\mathbf{I} - \mu \tilde{\mathbf{S}}_N] \mathbb{E}\{\tilde{\mathbf{w}}_k^N\} \end{aligned} \quad (36)$$

It is clear from Eq. (36) that the equivalent forms of Eq. (22) and Eq. (24) are mean-square stabilized with the following sufficient condition

$$|1 - \mu 3 \text{tr}\{\tilde{\mathbf{S}}_N\}| < 1 \quad (37)$$

So, we can write

$$0 < \mu < \frac{2}{3 \text{tr}\{\tilde{\mathbf{S}}_N\}} \quad (38)$$

where, $\text{Tr}(\cdot)$ denotes the trace of matrix. The autocorrelation matrix $\tilde{\mathbf{S}}_N$ is defined as

$$\tilde{\mathbf{S}}_N = \mathbb{E}\{\mathbf{V}_k^N (\mathbf{V}_k^N)^H\} = \left(\text{diag}\{\tilde{\mathbf{B}}_N\} \right)^{-1/2} \tilde{\mathbf{B}}_N \left(\text{diag}\{\tilde{\mathbf{B}}_N\} \right)^{-1/2} \quad (39)$$

where, $\tilde{\mathbf{B}}_N = \mathbb{E}\{\tilde{\mathbf{X}}_k^N (\tilde{\mathbf{X}}_k^N)^H\}$, and $\left(\text{diag}\{\tilde{\mathbf{B}}_N\} \right)^{-1/2}$ allows the power normalisation of $\tilde{\mathbf{X}}_k^N$. Then, the trace of $\tilde{\mathbf{S}}_N$ can be obtained as

$$\begin{aligned} \text{tr}\{\tilde{\mathbf{S}}_N\} &= \text{tr}\left\{ \mathbb{E}\{\mathbf{V}_k^N (\mathbf{V}_k^N)^H\} \right\} = \text{tr}\left\{ \left(\text{diag}\{\tilde{\mathbf{B}}_N\} \right)^{-1/2} \tilde{\mathbf{B}}_N \left(\text{diag}\{\tilde{\mathbf{B}}_N\} \right)^{-1/2} \right\} \\ &= \text{tr}\left\{ \left(\text{diag}\{\tilde{\mathbf{B}}_N\} \right)^{-1/2} \left(\text{diag}\{\tilde{\mathbf{B}}_N\} \right)^{-1/2} \tilde{\mathbf{B}}_N \right\} = N \end{aligned}$$

where, N is the filter length. Thus, the stability condition given by Eq. (38) becomes

$$0 < \mu < \frac{2}{3N} \quad (40)$$

which is the same as that of the conventional TDLMS algorithm.

Table 1. Eigenvalue spreads obtained by different algorithms for $N = 128$ and various values of ρ

Correlation coefficient ρ	Eigenvalue spread of \mathbf{R}_N	Eigenvalue spread of \mathbf{S}_N			Eigenvalue spread of $\tilde{\mathbf{S}}_N$		
		DCT-LMS	DFT-LMS	DHT-LMS	POW-DCT-LMS	POW-DFT-LMS	POW-DHT-LMS
0.9	341	2.06	18.24	17.85	1.12	1.10	1.12
0.8	83	1.94	8.57	8.46	1.13	1.10	1.12
0.7	34	1.84	5.39	5.46	1.13	1.10	1.11
0.5	16	1.7	3.90	4.02	1.12	1.11	1.12
0.6	9	1.65	3.06	2.95	1.14	1.10	1.11

4.2 Steady state performance of the proposed POW-TDLMS algorithm

The misadjustment can be computed as [19, 13]

$$M = \frac{\text{MSE}(\infty) - \text{MSE}_{\min}}{\text{MSE}_{\min}} \quad (41)$$

where, $\text{MSE}(\infty)$ denotes the steady state mean square error and MSE_{\min} denotes the minimum mean square error that can be achieved by an optimal filter.

5. SIMULATIONS AND RESULTS

In this section, we evaluate the performances of the proposed POW-TDLMS algorithm by comparing it to that of the conventional TDLMS algorithm for the DCT, DFT and DHT transforms. The corresponding algorithms are the DCT-LMS, DFT-LMS, DHT-LMS for the conventional TDLMS and the POW-DCT-LMS, POW-DFT-LMS, POW-DHT-LMS for the proposed POW-TDLMS. The step size parameter of the adaptive post whitening is set $\gamma = 0.001$ for all the experiments and all the simulations are averaged over 100 independent runs.

5.1 Experiment 01

In this experiment, we compare the Eigenvalue spreads of \mathbf{S}_N and $\tilde{\mathbf{S}}_N$ achieved by the conventional TDLMS algorithm and the proposed POW-TDLMS, respectively. We use the input signals \mathbf{x}_k^N generated from AR1 process with different correlation coefficients values $\rho = 0.9, 0.8, 0.7, 0.6$ and 0.5 to ensure different eigenvalue spread levels of the autocorrelation matrix of the input signal.

It is clear from Table 1 that the proposed POW-TDLMS algorithm significantly decreases the eigenvalue spreads, which approaches to 1 for various transforms, independently of the correlation level of the input signal. Consequently, the post whitened signal becomes nearly white and the proposed POW-TDLMS algorithm should provide better MSE convergence behavior compared to the conventional TDLMS algorithm.

5.2 Experiment 02

In this experiment, the proposed POW-TDLMS algorithm and the conventional TDLMS algorithm are implemented in the context of system identification. The system \mathbf{h} to be identified is a low pass filter with a cutoff frequency equal to 0.5 and different lengths $N = 16, N = 64, N = 128$, its impulse response is presented in Figure 4. The tests are performed using input signals \mathbf{x}_k^N generated from AR1 process with the correlation coefficients values $\rho = 0.5$ and $\rho = 0.9$ to ensure different eigenvalue spread levels of the autocorrelation matrix \mathbf{R}_N of the input signal as shown in table 1. For the different algorithms, we use the weight vector adaptation step sizes $\mu = 0.001$ for the case of the filter lengths $N = 16$ and $N = 64$ and $\mu = 0.003$ for the case of the filter length $N = 128$, the smoothing factor $\beta = 0.95$ and the regularisation parameter $\varepsilon = 0.025$. A white Gaussian noise is added to the signal at the output of the system \mathbf{h} with an SNR = 30 dB.

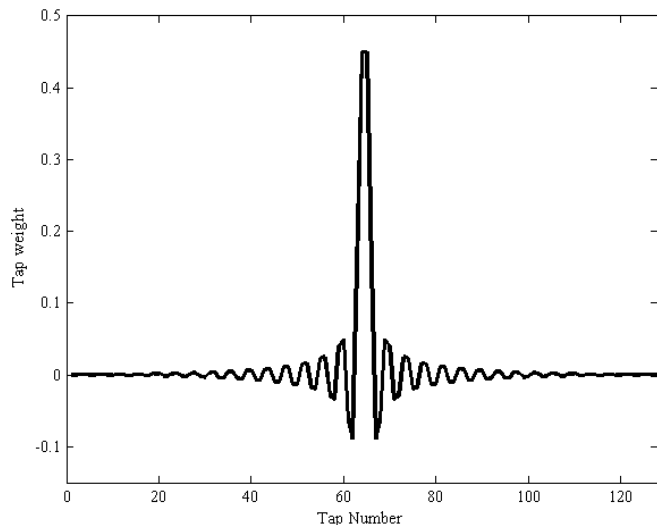


Figure 4. Impulse response of the system \mathbf{h} to be identified with $N = 128$

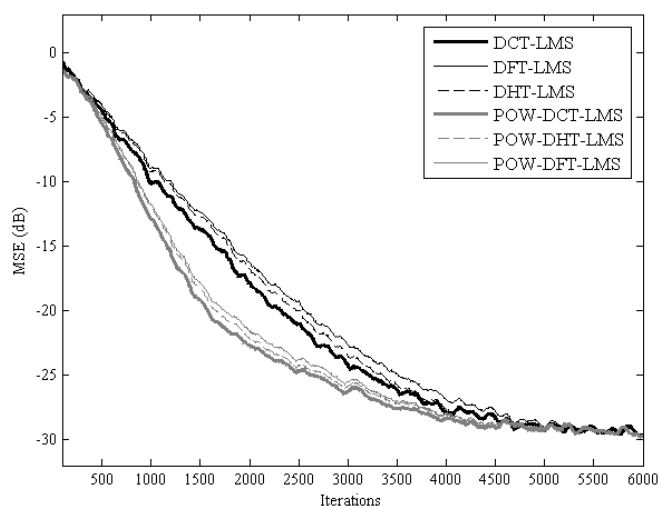


Figure 5. MSE convergence of the TDLMS and the proposed POW-TDLMS for various transforms, $N = 16$, for AR(1) input signal with $\rho = 0.9$

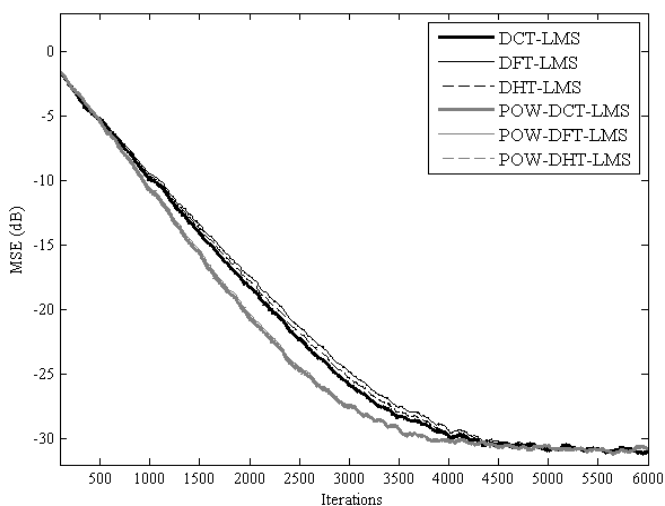


Figure 6. MSE convergence of the TDLMS and the proposed POW-TDLMS for various transforms, $N = 16$, for AR(1) input signal with $\rho = 0.5$

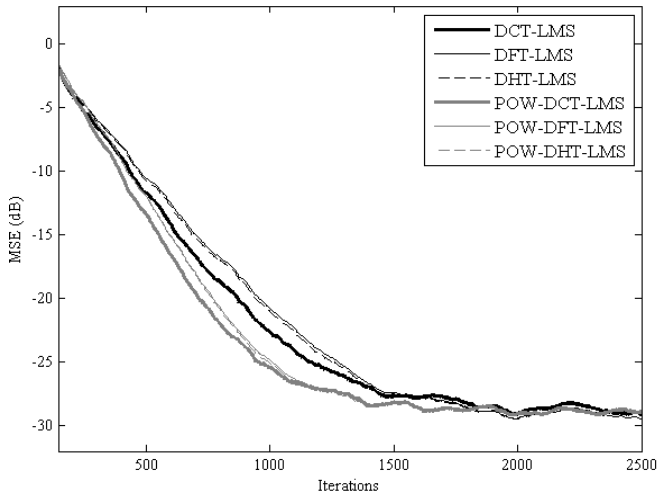


Figure 7. MSE convergence of the TDLMS and the proposed POW-TDLMS for various transforms, $N = 64$, for AR(1) input signal with $\rho = 0.9$

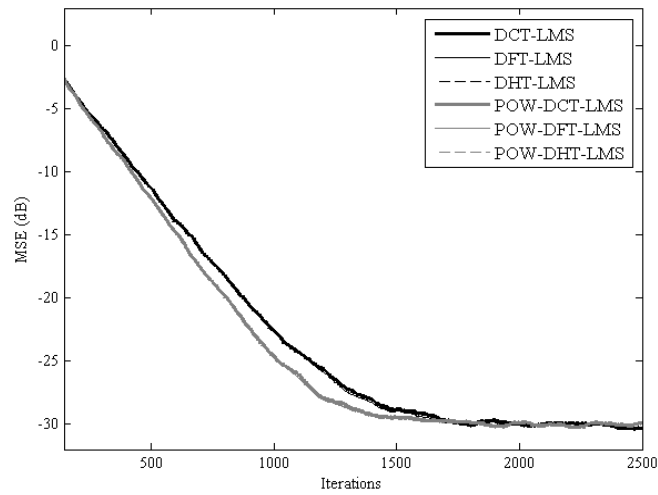


Figure 8. MSE convergence of the TDLMS and the proposed POW-TDLMS for various transforms, $N = 64$, for AR(1) input signal with $\rho = 0.5$

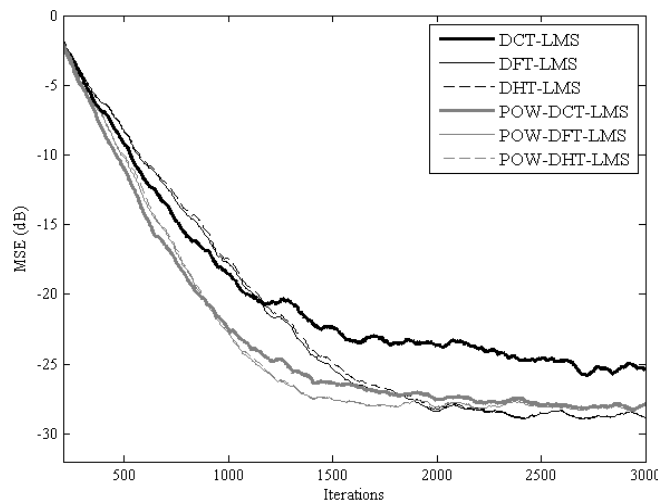


Figure 9. MSE convergence of the TDLMS and the proposed POW-TDLMS for various transforms, $N = 128$, for AR(1) input signal with $\rho = 0.9$

The simulation results for the MSE convergence

performance are presented in Figure 5, Figure 6, Figure 7, Figure 8, Figure 9 and Figure 10. It is clear from these figures that for all the considered cases, the proposed POW-TDLMS algorithm outperforms the conventional TDLMS algorithm for various transforms in terms of the MSE convergence speed. It is seen from figures 9 and 10 which corresponds to the cases of relatively long filter that the proposed POW-TDLMS algorithm outperforms the conventional TDLMS algorithm for various transforms, especially for the DCT, in terms of the reached steady state.

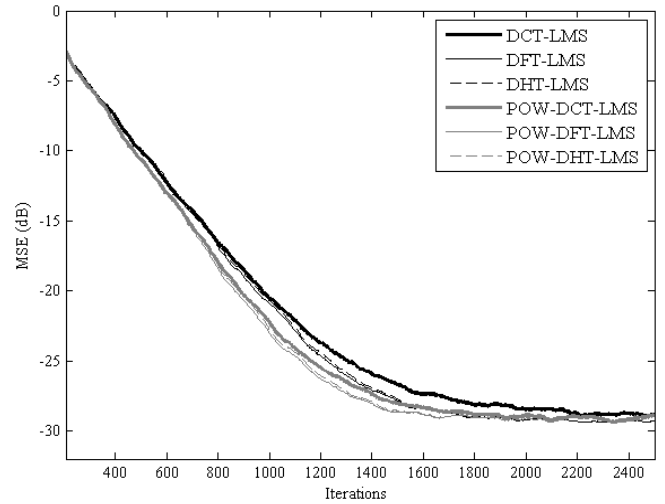


Figure 10. MSE convergence of the TDLMS and the proposed POW-TDLMS for various transforms, $N = 128$, for AR(1) input signal with $\rho = 0.5$

6. CONCLUSION

In this paper, a new post-whitening transform domain LMS (POW-TDLMS) algorithm has been developed by introducing a first order adaptive post whitening after applying the transform to reinforce its decorrelation. The mean convergence performance analysis of the proposed POW-TDLMS algorithm has also been presented. It has been shown that the proposed POW-TDLMS algorithm significantly reduces the eigenvalue spread of the autocorrelation matrix of the post whitened signal normalized by its power compared to the conventional TDLMS algorithm for the DCT, DFT and DHT transforms. The performance of the proposed POW-TDLMS algorithm has also been evaluated in the context of system identification and in terms of the MSE convergence speed and reached steady state compared to that of the conventional TDLMS algorithm. The obtained results confirm the efficiency and the superiority of the proposed POW-TDLMS algorithm. The proposed algorithm has shown its efficiency in system identification application. Therefore, as a perspective, it would be interesting to investigate its usefulness in adaptive acoustic echo cancellation.

REFERENCES

- [1] Simon, H. (2002). Adaptive Filter Theory. Prentice-hall, 4rd ed.
- [2] Alexander, D.P., Zayed, M.R. (2006). Adaptive filtering primer with matlab. (CRC) Taylor & Francis.

- [3] Widrow, B., Hoff, M.E. (1960). Adaptive switching circuits. Tech. Rep. 1553-1, Stanford Electron. Labs., Stanford, CA.
- [4] Kim, D.I., Wilde, P.D. (2000). Performance analysis of signed self-orthogonalizing adaptive lattice filter. *IEEE transactions on circuits and systems. Analogue and Digital Signal Processing*, 47(11): 1227-1237. <https://doi.org/10.1109/82.885130>
- [5] Dimitris, G.M., Vinay, K.I., Stephen, M.K. (2005). *Statistical and adaptive signal processing: Spectral estimation, signal modeling. Adaptive Filtering and Array Processing*. Artech House.
- [6] Homer, J.(2000). Quantifying the convergence speed of LMS adaptive FIR filter with autoregressive inputs. In *Electronics Letters*, 36(6): 585-586. <https://doi.org/10.1049/el:20000469>
- [7] Mboup, M., Bonnet, M., Bershada, N. (1994). LMS coupled adaptive prediction and system identification: A statistical model and transient mean analysis. *IEEE Transactions on Signal Processing*, 42(10): 2607-2615. <https://doi.org/10.1109/78.324727>
- [8] Douglas, S.C., Cichocki, A., Amari, S. (1999). Self-whitening algorithms for adaptive equalization and deconvolution. *IEEE Transactions on Signal Processing*, 47(4): 1161-1165. <https://doi.org/10.1109/78.752617>
- [9] Gazor, S., Lieu, T.(2005). Adaptive filtering with decorrelation for coloured AR environments. *IEE Proceedings - Vision, Image and Signal Processing*, 152(6): 806-818. <https://doi.org/10.1049/ip-vis:20045260>
- [10] Gonzalez, A., Ferrer, M., Albu, F., Diego, M.D. (2012). Affine projection algorithms: evolution to smart and fast multichannel algorithms and applications. In *Proc. of EUSIPCO*.
- [11] Narayan, S., Peterson, A., Narasimha, M. (1983). Transform domain LMS algorithm. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 31(3): 609-615. <https://doi.org/10.1109/TASSP.1983.1164121>
- [12] Lu, J., Qiu X., Zou, H. (2014). A modified frequency-domain block LMS algorithm with guaranteed optimal steady-state performance. *Signal Processing*, 104: 27-32. <https://doi.org/10.1016/j.sigpro.2014.03.029>
- [13] Chergui, L., Bouguezal, S. (2017). A new pre-whitening transform domain LMS algorithm and its application to speech denoising. *Signal Processing*, 130: 118-128. <http://doi.org/10.1016/j.sigpro.2016.06.021>
- [14] Boroujeuny, B.F., Gazor, S. (1992). Selection of orthonormal transforms for improving the performance of the transform domain normalised LMS algorithm. *IEEE Proceeding F in Radar and Signal Process*, 139(5): 327-335. <https://doi.org/10.1049/ip-f-2.1992.0046>
- [15] Lee, K.A., Gan, W.S., Kuo, S.M. (2009). *Subband adaptive filtering, theory and implementation*. Willey. 1st ed.
- [16] Mayyas, K. (2003). New transform domain adaptive algorithms for acoustic echo cancellation. *Digital Signal Processing*, 13(3): 415-432. [https://doi.org/10.1016/S1051-2004\(03\)00004-6](https://doi.org/10.1016/S1051-2004(03)00004-6)
- [17] Shynk, J.J. (1992). Frequency domain and multirate adaptive filtering. *IEEE Signal Processing Magazine*, 9(1): 14-37. <https://doi.org/10.1109/79.109205>
- [18] Kuhn, E.V., Kolodziej, J.E., Seara, R. (2015). Analysis of the TDLMS algorithm operating in a nonstationary environment. *Digital Signal Processing*, 45: 69-83. <https://doi.org/10.1016/j.dsp.2015.05.013>
- [19] Zhao, S., Man, Z., Khoo, S., Wu, H.R. (2009). Stability and convergence analysis of transform-domain lms adaptive filters with second-order autoregressive process. *IEEE Transactions on Signal Processing*, 57(1): 119-130. <https://doi.org/10.1109/TSP.2008.2007618>
- [20] Beaufays, F. (1995). Transform domain adaptive filters: an analytical approach. *IEEE Transactions on Signal Processing*, 43(2): 422-431. <https://doi.org/10.1109/78.348125>
- [21] Murmu, G., Nath, R. (2011). Convergence performance comparison of transform domain LMS adaptive filters for correlated signal. *International Conference on Devices and Communications (ICDeCom)*, Mesra. <https://doi.org/10.1109/ICDECOM.2011.5738534>