Mean-Square Performance Analysis of Variable Step-Size l_0-NLMS Algorithm*

[1]S. Nunoo, [2]U. A. K. Chude-Okonkwo, and [3]R. Ngah
[1]University of Mines and Technology, P. O. Box 237, Tarkwa, Ghana
[2]University of Pretoria, Hatfield, Pretoria, South Africa
[3]Universiti Teknologi Malaysia, Skudai, Johor, Malaysia

Nunoo, S., Chude-Okonkwo, U. A. K. and R. Ngah, R. (2019), "Mean-Square Performance Analysis of Variable Step-Size l_0-NLMS Algorithm", *Ghana Journal of Technology*, Vol. 4, No. 1, pp. 7 - 16.

Abstract

Wireless communication systems often require accurate Channel State Information (CSI) at the receiver side. Typically, the CSI can be obtained from Channel Impulse Response (CIR). Measurements have shown that the CIR of wideband channels are often sparse. To this end, the Least Mean Square (LMS)-based algorithms have been used to estimate the CIR at the receiver side, which unfortunately is not able to accurately estimate sparse channels. In this paper, we propose a variable step-size l_0-norm Normalised LMS (NLMS) algorithm. The step-size is varied with respect to changes in the Mean Square Error (MSE), allowing the filter to track changes in the system as well as produce smaller steady-state errors. We present simulation results and compare the performance of the new algorithm with the Invariable Step-Size NLMS (ISS-NLMS), Variable Step-Size NLMS (VSS-NLMS) and the Invariable Step-Size l_0-NLMS (ISS-L0-NLMS) algorithms. The results show that the proposed algorithm performs admirably to improve the identification of sparse systems, especially at SNR of 10 dB.

Keywords: Variable Step-Size Adaptation, Normalised Least Mean Square Algorithm, Compressive Sensing

1 Introduction

The need for accurate Channel State Information (CSI) at the receiver side of many wideband communication systems is of utmost importance. Most of these channels have shown to be sparse. Adaptive Channel Estimation (ACE) is an effective approach for estimating the sparse channels. There are many ACE algorithms, such as the Least Mean Square (LMS) and the Recursive Least Squares (RLS) algorithms (Diniz, 2013; Sayed, 2008; Haykin, 2002). However, these algorithms are not able to exploit the channel sparsity due to their lack of sparse characteristics. In general, sparse channels contain very few non-zero coefficients. A typical example of sparse channel is shown in Fig. 1 with a channel length of N = 25 and K = 4 dominant coefficients.

Step-size is a critical parameter in ACE since it controls the estimation performance, convergence rate and computational cost (Nunoo *et al.*, 2013). Unfortunately, using an Invariable Step-Size (ISS) leads to performance loss and/or low convergence as well as high computational cost . For this reason, the Variable Step-Size NLMS (VSS-NLMS) was proposed by Harris *et al.* (1986) to improve estimation performance. In this case, the step-size varies as a result of the variation in the Mean Square Error (MSE) and the previous step-size estimate, thus allowing the adaptive filter to track changes in the system as well as produce a small steady-state error (Yang *et al.*, 2010).

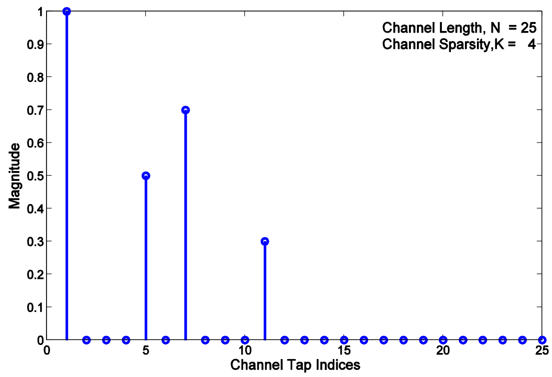

Fig. 1 A Typical Sparse Multipath Channel With A Channel Length Of 25 And 4 Non-Zero Taps

In recent times, some ACE algorithms have exploited channel sparsity to improve the identification performance (Gui *et al.*, 2013; Das *et al.*, 2011; Taheri and Vorobyov, 2014; Gui and Adachi, 2013; Li and Hamamura, 2014a; Li and Hamamura, 2014b). These algorithms adopt the Compressive Sensing (CS) approach (Donoho, 2006). To address issues related to variable sparsity, Das *et al.* (2011) proposed the use of an adaptive convex combination of LMS-based and Zero-Attracting LMS (ZA-LMS) adaptive filters. A new reweighted l_1-norm penalised LMS algorithm, proposed by Taheri and Vorobyov (2014), introduces an additional reweighting of the Channel Impulse Response (CIR) coefficient estimates to promote a sparse solution even more and approximate l_0-pseudo-norm closer. Gui and

*Manuscript received July 3, 2019
Revised version accepted September 19, 2019

Adachi (2013) proposed the L0-LMS, ZA-NLMS, reweighted ZA-NLMS (RZA-NLMS), LP-NLMS, and L0-NLMS algorithms, which all used ISS during adaptation. Li and Hamamura (2014a) also proposed the LP Proportionate NLMS (LP-PNLMS) algorithm to exploit the sparse property of broadband multipath wireless communication channels. In addition, Li and Hamamura (2014b) proposed a smooth approximation l_0-norm constrained Affine Projection Algorithm (SL0-APA) to improve the convergence speed and the steady-state error of Affine Projection Algorithm (APA) for sparse channel estimation.

Of the lot, only Gui *et al.* (2013) dealt with the issue of VSS parameter for adaptation, where the VSS-ZA-NLMS algorithm is presented. To the best of our knowledge, no other paper has reported on the use of sparse VSS-NLMS algorithm to exploit the channel sparsity. Unlike algorithms that use l_1-norm sparse penalty, that is the ZA- and RZA-based algorithms, the l_0-norm sparse penalty is a good candidate to achieve more accurate channel estimates.

In this paper, we present the VSS-based l_0-norm NLMS algorithm (Nunoo *et al.*, 2014). It presents a detailed analysis of the mean performance and the steady-state excess MSE. To conclude, we present various simulation results to confirm the effectiveness of our proposed algorithm.

The rest of the paper is organised as follows. Section 2 presents the problem formulation, which encompasses an overview of the ISS-NLMS and VSS-NLMS algorithms. It continues with a derivation of the proposed VSS-based l_0-norm NLMS algorithm. Lastly, Section 2 concludes with an analysis of the mean behaviour and the excess steady-state MSE and its computational complexity. The analysis takes into consideration the effect of the environmental noise to ascertain the robustness of the proposed algorithm as espoused by Haykin (2002). Simulation results and discussions are presented in Section 3. Finally, Section 4 presents the conclusions and recommendations for future work.

2 Resources and Methods Used

2.1 Problem Formulation

Let us consider the receiver side of a typical communication system, which is represented by the system identification system like the one shown in Fig. 2. Given that, $d(k)$ is the desired signal of an adaptive filter, then:

$$d(k) = \mathbf{x}^T(k)\mathbf{h} + n(k) \tag{1}$$

where $\mathbf{x}(k) = [x_0(k)\ x_1(k)\ \ldots\ x_{L-1}(k)]^T$ is the input signal vector at iteration k, h is the sparse channel vector of the communication system that we wish to estimate, $n(k)$ is the system noise signal, which is a zero-mean uncorrelated sequence that is independent of $\mathbf{x}(k)$, $y(k)=\mathbf{x}^T(k)\mathbf{w}(k)$, $\mathbf{w}(k) = [w_0(k)\ w_1(k)\ \ldots\ w_{L-1}(k)]^T$ is the filter weight coefficient vector, and $[.]^T$ denotes vector transpose. For simplicity, the filter is assumed to have the same structure as the unknown system. Thus, the a priori estimation error $e(k)$ is also given by:

$$e(k) = d(k) - \mathbf{x}^T(k)\mathbf{w}(k) \tag{2}$$

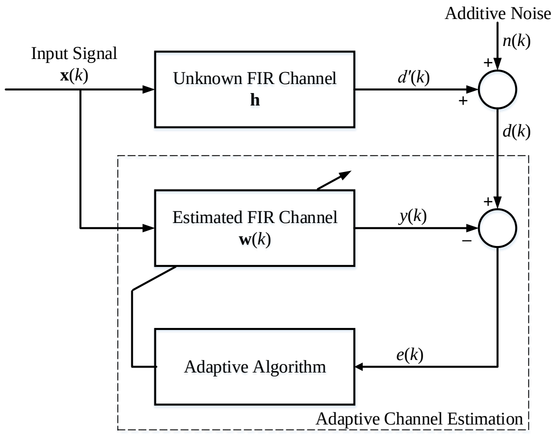

Fig. 2 A Typical System Identification Block Diagram

2.1.1 ISS-NLMS Algorithm

The LMS algorithm is sensitive to the scaling of its input. This makes it very hard, if not unfeasible, to choose a step-size that guarantees stability of the algorithm (Haykin, 2002). The NLMS algorithm solves this problem by normalising the adaptive error update section with the input power. Thus, the NLMS algorithm is described by the equation:

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu \frac{e(k)\mathbf{x}(k)}{\varepsilon + \mathbf{x}^T(k)\mathbf{x}(k)} \tag{3}$$

where ε is a regulation parameter, which is included in order to avoid large step sizes when $\mathbf{x}^T(k)\mathbf{x}(k)$ becomes small and μ is the invariable (fixed) step-size parameter. If the primary objective of the adaptation is to achieve a faster convergence, then a variable step size can be used.

2.1.2 VSS-NLMS Algorithm

The bedrock of all VSS-NLMS algorithms is to develop a means of varying the step-size parameter. The Kwong and Johnston VSS-LMS algorithm (Kwong and Johnston, 1992) makes use of the squared instantaneous estimation error to update the step-size:

$$\mu'(k+1) = \alpha\mu(k) + \gamma e^2(k) \qquad (4)$$

where 0<α<1, γ>0 and

$$\mu(k+1) = \begin{cases} \mu_{\max} & \text{if } \mu'(k+1) > \mu_{\max} \\ \mu_{\min} & \text{if } \mu'(k+1) < \mu_{\min} \\ \mu'(k+1) & \text{otherwise} \end{cases} \qquad (5)$$

To employ the use of VSS in the NLMS algorithms, from (3), the normalised version of the coefficient update equation is given as:

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu(k+1)\frac{e(k)\mathbf{x}(k)}{\varepsilon + \mathbf{x}^T(k)\mathbf{x}(k)} \qquad (6)$$

where $\mu(k+1)$ is the VSS parameter.

2.2 Proposed VSS-Based l_0-Norm NLMS Algorithm

The NLMS-based adaptive sparse channel estimation algorithm possesses the ability to mitigate the scaling interference of the training signal. This effect is due to the fact that NLMS-based methods estimate the sparse channel by normalising the power of the training signal $\mathbf{x}(k)$. The basic principle of CS-based sparse adaptive filtering is the introduction of an appropriate sparse penalty which can be generalised as follows:

$$\mathbf{w}(k+1) = \mathbf{w}(k) + [\mathit{Adaptive\ Error\ Update}] + [\mathit{Sparse\ Penalty}] \qquad (7)$$

Consider l_0-norm penalty on the LMS cost function which forces $\mathbf{w}(k)$ to approach zero. The cost function is given by:

$$\xi(k) = \frac{1}{2}e^2(k) + \lambda_{l_0}\|\mathbf{w}(k)\|_0 \qquad (8)$$

where $\lambda_{l_0} > 0$ is a regulation parameter for balancing the penalty and estimation error and $\|\mathbf{w}(k)\|_0$ is the l_0-norm sparse penalty function. Since the l_0-norm is a Non-Polynomial (NP) hard problem (Gu *et al.*, 2009), in order to reduce the computational complexity, we replace it with an approximate continuous function:

$$\|\mathbf{w}(k)\|_0 \approx \sum_{i=0}^{N-1}\left(1 - e^{-\beta|w_i(k)|}\right) \qquad (9)$$

where β is a regulation parameter. Therefore, the cost function in (8) can be rewritten as:

$$\xi(k) = \frac{1}{2}e^2(k) + \lambda_{l_0}\sum_{i=0}^{N-1}\left(1 - e^{-\beta|w_i(k)|}\right) \qquad (10)$$

From the first order Taylor series expansion of the exponential functions, the exponential function in (10) can be simplified as:

$$e^{-\beta|w_i(k)|} \approx \begin{cases} 1 - \beta|w_i(k)| & \text{when } |w_i(k)| \le \frac{1}{\beta} \\ 0 & \text{otherwise} \end{cases} \qquad (11)$$

The update equation of the l_0-norm LMS is given as:

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu e(k)\mathbf{x}(k) - \rho_{l_0}\beta\,\mathrm{sgn}[\mathbf{w}(k)]e^{-\beta\mathbf{w}(k)} \qquad (12)$$

where $\rho_{l_0} = \mu\lambda_{l_0}$ and $\mathrm{sgn}(\cdot)$ is a component-wise sign function defined as:

$$\mathrm{sgn}(x) = \begin{cases} \frac{x}{|x|} & x \ne 0 \\ 0 & \text{otherwise} \end{cases} \qquad (13)$$

Unfortunately, the exponential function in (12) will also cause high computational complexity. To further reduce the complexity, an approximation function $J[\mathbf{w}(k)]$ is introduced. Thus, the l_0-norm LMS sparse ACE is given as:

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu e(k)\mathbf{x}(k) - \rho_{l_0}J[\mathbf{w}(k)] \qquad (14)$$

with $J[\mathbf{w}(k)]$ defined as:

$$J[\mathbf{w}(k)] = \begin{cases} 2\beta^2\mathbf{w}(k) - 2\beta\,\mathrm{sgn}[\mathbf{w}(k)] & \text{when } \mathbf{w}(k) \le \frac{1}{\beta} \\ 0 & \text{otherwise} \end{cases} \qquad (15)$$

for all $i \in \{1, 2, ..., N\}$.

Based on the ISS-based l_0-norm LMS algorithm in (14), that of the NLMS algorithm is also given as:

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu\frac{e(k)\mathbf{x}(k)}{\varepsilon + \mathbf{x}^T(k)\mathbf{x}(k)} - \rho_{l_0}J[\mathbf{w}(k)] \qquad (16)$$

Thus to further improve the estimation performance of (16), we propose the use of variable step-size as inspirited in (Kwong and Johnston, 1992). Hence from (16):

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu(k+1)\frac{e(k)\mathbf{x}(k)}{\varepsilon + \mathbf{x}^T(k)\mathbf{x}(k)} - \rho_{l_0}J[\mathbf{w}(k)] \qquad (17)$$

where $\mu(k+1)$ is the VSS and defined in (4) and (5). The VSS-l_0-NLMS algorithm is as in Algorithm 1. A sufficient condition for mean coefficient vector convergence of the proposed algorithm is given as (Kwong and Johnston, 1992):

$$0 < \mathrm{E}\left[\mu(k)\right] < \frac{2}{\lambda_{\max}} \tag{18}$$

where $\lambda_{\max}$ is the maximum eigenvalue of the input signal autocorrelation matrix $R = \mathrm{E}\left[\mathbf{x}(k)\mathbf{x}^T(k)\right]$. The proposed algorithm performs better than the ISS-NLMS, VSS-NLMS and ISS-l_0-NLMS and other related sparsity-aware algorithms because of its utilisation of the l_0-norm.

Algorithm 1: Summary of the VSS-l_0-NLMS Algorithm

Initializations (typical values)

$\mathbf{x}(0) = \mathbf{w}(0) = [0 \cdots 0]^T$

Choose constants: α, γ, $\mu_{\min}$, $\mu_{\max}$, β, ρ_{l_0}, and ε.

Processing and adaptations

For $k = 0, 1, 2, \ldots$

$$\mathbf{x}(k) = \left[\mathbf{x}(k) \quad \mathbf{x}(k-1) \quad \cdots \quad \mathbf{x}(n-N+1)\right]$$

$$e(k) = d(k) - \mathbf{x}^T(k)\mathbf{w}(k)$$

$$\mu'(k+1) = \alpha\mu(k) + \gamma e^2(k)$$

$$\mu(k+1) = \begin{cases} \mu_{\max} & \text{if } \mu'(k+1) > \mu_{\max} \\ \mu_{\min} & \text{if } \mu'(k+1) < \mu_{\min} \\ \mu'(k+1) & \text{otherwise} \end{cases}$$

$$J\left[\mathbf{w}(k)\right] = \begin{cases} 2\beta^2\mathbf{w}(k) - 2\beta\,\mathrm{sgn}\left[\mathbf{w}(k)\right] & \text{when } \mathbf{w}(k) \le \dfrac{1}{\beta} \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu\frac{e(k)\mathbf{x}(k)}{\varepsilon + \mathbf{x}^T(k)\mathbf{x}(k)} - \rho_{l_0} J\left[\mathbf{w}(k)\right]$$

End

2.3 Performance Analysis

In this section, we present a performance analysis of the proposed VSS-l_0-NLMS algorithm. We discuss the performance analysis in three thematic areas, viz. mean behaviour analysis, steady-state excess MSE analysis and computational complexity analysis. To simplify the analysis, we adopt the following assumptions:

a. the input signal $\mathbf{x}(k)$ is a stationary ergodic process which is Gaussian distributed with zero mean and autocorrelation $\mathbf{R} = \mathrm{E}\left[\mathbf{x}^T(k)\mathbf{x}(k)\right]$,
b. the noise signal $n(k)$ is white Gaussian distributed with zero mean and variance σ_n^2, and
c. $\mathbf{w}(k)$, $\mathbf{x}(k)$ and $n(k)$ are statistically independent of each other.

2.3.1 Mean Behaviour

We adopt the energy conservation approach (Sayed, 2008) to obtain the theoretical expression for the MSE of the VSS-l_0-NLMS algorithm. From (1) and (2), we obtain:

$$e(k) = \mathbf{x}^T(k)\left[\mathbf{h} - \mathbf{w}(k)\right] + n(k) \tag{19}$$

Subtract h from both sides of the VSS-l_0-NLMS update equation given in (17), then the misalignment vector, $\Delta(k+1) = \mathbf{h} - \mathbf{w}(k+1)$, after substituting (19) in the ensuing equation, can be expressed as:

$$\begin{aligned}\Delta(k+1) = \Delta(k) \\ -\mu(k+1)\left[\frac{\left[\Delta(k)\mathbf{x}^T(k) + n(k)\right]\mathbf{x}(k)}{\varepsilon + \mathbf{x}^T(k)\mathbf{x}(k)}\right] \\ +\rho_{l_0} J\left[\mathbf{w}(k)\right]\end{aligned} \tag{20}$$

Thus,

$$\begin{aligned}\Delta(k+1) = \left\{\mathbf{I}_N - \mu(k+1)\left[\frac{\mathbf{x}^T(k)\mathbf{x}(k)}{\varepsilon + \mathbf{x}^T(k)\mathbf{x}(k)}\right]\right\}\Delta(k) \\ -\mu(k+1)\left[\frac{\mathbf{x}(k)n(k)}{\varepsilon + \mathbf{x}^T(k)\mathbf{x}(k)}\right] \\ +\rho_{l_0} J\left[\mathbf{w}(k)\right]\end{aligned} \tag{21}$$

where $\mathbf{I}_N$ is the Nth-order identity matrix. Taking the expectation of both sides of (21):

$$\begin{aligned}\mathrm{E}\left[\Delta(k+1)\right] = \mathrm{E}\left[\Delta(k)\right] \\ -\mathrm{E}\left[\mu(k+1)\right]\mathrm{E}\left[\frac{\mathbf{x}^T(k)\mathbf{x}(k)}{\varepsilon + \mathbf{x}^T(k)\mathbf{x}(k)}\right]\mathrm{E}\left[\Delta(k)\right] \\ -\mathrm{E}\left[\mu(k+1)\right]\mathrm{E}\left[\frac{\mathbf{x}(k)n(k)}{\varepsilon + \mathbf{x}^T(k)\mathbf{x}(k)}\right] \\ +\rho_{l_0}\,\mathrm{E}\left\{J\left[\mathbf{w}(k)\right]\right\}\end{aligned} \tag{22}$$

Employing Assumption 3 that $n(k)$ and $\mathbf{x}(k)$ are statistically independent, then $\mathrm{E}\left\{\mathbf{x}(k)n(k)/\left[\varepsilon + \mathbf{x}^T(k)\mathbf{x}(k)\right]\right\} = 0$. Thus, (22) can be simplified as:

$$\begin{aligned}\mathrm{E}\left[\Delta(k+1)\right] = \mathrm{E}\left[\Delta(k)\right] \\ -\mathrm{E}\left[\mu(k+1)\right]\mathrm{E}\left[\frac{\mathbf{x}^T(k)\mathbf{x}(k)}{\varepsilon + \mathbf{x}^T(k)\mathbf{x}(k)}\right] \\ \cdot\mathrm{E}\left[\Delta(k)\right] + \rho_{l_0}\,\mathrm{E}\left\{J\left[\mathbf{w}(k)\right]\right\}\end{aligned} \tag{23}$$

But from (15), when $\left|w_i(k)\right| \le 1/\beta$, then:

$$\begin{aligned}\mathrm{E}\left\{J\left[\mathbf{w}(k)\right]\right\} = 2\beta^2\,\mathrm{E}\left[\mathbf{w}(k)\right] \\ -2\beta\,\mathrm{E}\left\{\mathrm{sgn}\left[\mathbf{w}(k)\right]\right\}\end{aligned} \tag{24}$$

Otherwise,

$$\mathrm{E}\left\{J\left[\mathbf{w}(k)\right]\right\}=0 \tag{25}$$

Note that under steady-state conditions, previous works on sparse LMS algorithms (Chen *et al.*, 2009; Su *et al.*, 2012) have shown that:

$$\mathrm{E}\left\{\operatorname{sgn}\left[\mathbf{w}(k)\right]\right\}\approx\operatorname{sgn}(\mathbf{w}) \tag{26}$$

Therefore:

$$\mathrm{E}\left\{J\left[\mathbf{w}(k)\right]\right\}=2\beta^2\,\mathrm{E}\left[\mathbf{w}(k)\right]-2\beta\operatorname{sgn}(\mathbf{w}) \tag{27}$$

Hence, (23) can be rewritten as:

$$\mathrm{E}\left[\Delta(k+1)\right]=\left\{\mathbf{I}_N-\mathrm{E}\left[\mu(k+1)\right]\mathrm{E}\left[\frac{\mathbf{x}^T(k)\mathbf{x}(k)}{\varepsilon+\mathbf{x}^T(k)\mathbf{x}(k)}\right]\right\} \cdot\mathrm{E}\left[\Delta(k)\right]+\rho_{l_0}\left\{2\beta^2\,\mathrm{E}\left[\mathbf{w}(k)\right]-2\beta\operatorname{sgn}(\mathbf{w})\right\} \tag{28}$$

With reference to Assumption 1 and given that $\mathrm{E}\left[\mu(k+1)\right]=\mu(k+1)$, then (28) is rewritten as:

$$\mathrm{E}\left[\Delta(k+1)\right]=\left\{\mathbf{I}_N-\mu(k+1)\mathbf{R}\,\mathrm{E}\left[\frac{1}{\varepsilon+\mathbf{x}^T(k)\mathbf{x}(k)}\right]\right\} \cdot\mathrm{E}\left[\Delta(k)\right]+\rho_{l_0}\left\{2\beta^2\,\mathrm{E}\left[\mathbf{w}(k)\right]-2\beta\operatorname{sgn}(\mathbf{w})\right\} \tag{29}$$

The stability condition of the VSS-l_0-NLMS algorithm is independent of the parameter ρ_{l_0}. Given that the estimated channel vector $\mathbf{w}(k)$ converges when $n\to\infty$, then (29) is rewritten as:

$$\mathrm{E}\left[\Delta(\infty)\right]=\left\{\mathbf{I}_N-\mu(k+1)\mathbf{R}\,\mathrm{E}\left[\frac{1}{\varepsilon+\mathbf{x}^T(k)\mathbf{x}(k)}\right]\right\} \cdot\mathrm{E}\left[\Delta(\infty)\right]+\rho_{l_0}\left\{2\beta^2\,\mathrm{E}[\mathbf{h}]-2\beta\operatorname{sgn}(\mathbf{h})\right\} \tag{30}$$

Hence, from (30), assuming $\mathrm{E}\left\{1/\left[\varepsilon+\mathbf{x}^T(k)\mathbf{x}(k)\right]\right\}=\gamma_{\mathbf{xx}}$, the optimum solution of the VSS-l_0-NLMS algorithm is given as:

$$\mathrm{E}\left[\mathbf{h}(\infty)\right]=\mathbf{h}+\frac{\rho_{l_0}\mathbf{R}^{-1}}{\gamma_{\mathbf{xx}}\mu(\infty)}\left\{2\beta^2\,\mathrm{E}[\mathbf{h}]-2\beta\operatorname{sgn}(\mathbf{h})\right\} \tag{31}$$

Thus, the vector $\rho_{l_0}\,\mathrm{E}\left\{J\left[\mathbf{w}(k)\right]\right\}$ is bounded by $-\rho_{l_0}\mathbf{1}$ and $\rho_{l_0}\mathbf{1}$, where $\mathbf{1}$ is a vector of 1's. Therefore $\mathrm{E}\left[\Delta(k)\right]$ converges if the maximal eigenvalue of $\left[\mathbf{I}-\mu(k+1)\mathbf{R}\right]$ is less than 1, which is satisfied by (18). Since $\mathrm{E}\left[\mathbf{w}(k)\right]=\mathrm{E}\left[\Delta(k+1)\right]+\mathbf{h}$, then $\mathrm{E}\left[\mathbf{w}(k)\right]$ will also converge with the limiting vector given in (31).

2.3.2 Steady-State Excess MSE Analysis

Consider the linear model in (1) for the received signal, the steady-state MSE is defined as:

$$\mathrm{MSE}=\lim_{k\to\infty}\mathrm{E}\left[\left|e(k)\right|^2\right] \tag{32}$$

Firstly, multiplying both sides of (17) by $\mathbf{x}(k)$, then:

$$\mathbf{x}(k)\mathbf{w}(k+1)=\mathbf{x}^T(k)\mathbf{w}(k)+\mu(k+1) \cdot\frac{e(k)\mathbf{x}^T(k)\mathbf{x}(k)}{\varepsilon+\mathbf{x}^T(k)\mathbf{x}(k)}-\rho_{l_0}\mathbf{x}^T(k)J\left[\mathbf{w}(k)\right] \tag{33}$$

Furthermore,

$$\mathbf{x}(k)\mathbf{w}(k+1)=\mathbf{x}(k)\mathbf{w}(k)+\mu(k+1)e(k) \tag{34}$$

In addition, define the a posteriori error vector, $e_p(k)$ and the a priori error vector, $e_a(k)$ as:

$$\begin{aligned} e_p(k)&=\mathbf{x}(k)\mathbf{h}-\mathbf{x}(k)\mathbf{w}(k+1)\\ e_a(k)&=\mathbf{x}(k)\mathbf{h}-\mathbf{x}(k)\mathbf{w}(k)\end{aligned} \tag{35}$$

Combining (34) and (35), then:

$$e_p(k)=e_a(k)-\mu(k+1)e(k) \tag{36}$$

In addition, using (1), (2) and (35), it can easily be shown that:

$$e(k)=e_a(k)+n(k) \tag{37}$$

Substituting (37) into (36), then:

$$e_p(k)=\left[\mathbf{I}_N-\mu(k+1)\right]e(k)-n(k) \tag{38}$$

From (36), $e(k)$ is given as:

$$e(k)=\frac{1}{\mu(k+1)}\left[e_a(k)-e_p(k)\right] \tag{39}$$

Substituting (39) into (17), then:

$$\mathbf{w}(k+1)=\mathbf{w}(k)+\frac{\left[e_a(k)-e_p(k)\right]\mathbf{x}(k)}{\varepsilon+\mathbf{x}^T(k)\mathbf{x}(k)} -\rho_{l_0}J\left[\mathbf{w}(k)\right] \tag{40}$$

Taking the expectation of both sides of (40), then:

$$\mathrm{E}\left[\mathbf{w}(k+1)\right]=\mathrm{E}\left[\mathbf{w}(k)\right]+\mathrm{E}\left[\frac{e_a(k)\mathbf{x}(k)}{\varepsilon+\mathbf{x}^T(k)\mathbf{x}(k)}\right]$$
$$-\mathrm{E}\left[\frac{e_p(k)\mathbf{x}(k)}{\varepsilon+\mathbf{x}^T(k)\mathbf{x}(k)}\right]$$
$$-\rho_{l_0}\,\mathrm{E}\left\{J\left[w(k)\right]\right\} \qquad (41)$$

Given that at steady-state, as $k\to\infty$:

$$\mathrm{E}\left[\left\|\mathbf{w}(k+1)\right\|^2\right]\approx\mathrm{E}\left[\left\|\mathbf{w}(k)\right\|^2\right] \qquad (42)$$

and assuming that $e_a(k)$, $e_p(k)$ and $\mathbf{w}(k)$ are statistically independent of $\mathbf{x}(k)$, then:

$$\mathrm{E}\left[\frac{e_p^T(k)\mathbf{x}(k)\mathbf{x}^T(k)e_p(k)}{\left\|\varepsilon+\mathbf{x}^T(k)\mathbf{x}(k)\right\|^2}\right]=\mathrm{E}\left[\frac{e_a^T(k)\mathbf{x}(k)\mathbf{x}^T(k)e_a(k)}{\left\|\varepsilon+\mathbf{x}^T(k)\mathbf{x}(k)\right\|^2}\right]$$
$$-\rho_{l_0}^2\,\mathrm{E}\left\{\left\|J\left[\mathbf{w}(k)\right]\right\|^2\right\} \qquad (43)$$

Substituting (38) into the Left-Hand Side (LHS) of (43), then:

$$\mathrm{LHS}=\mathrm{E}\left\{\frac{\left[\mathbf{I}_N-\mu(k+1)\right]^2e^T(k)\mathbf{x}(k)\mathbf{x}^T(k)e(k)}{\left\|\varepsilon+\mathbf{x}^T(k)\mathbf{x}(k)\right\|^2}\right\}$$
$$-\mathrm{E}\left\{\frac{\left[\mathbf{I}_N-\mu(k+1)\right]n^T(k)\mathbf{x}(k)\mathbf{x}^T(k)e(k)}{\left\|\varepsilon+\mathbf{x}^T(k)\mathbf{x}(k)\right\|^2}\right\}$$
$$-\mathrm{E}\left\{\frac{\left[\mathbf{I}_N-\mu(k+1)\right]e^T(k)\mathbf{x}(k)\mathbf{x}^T(k)n(k)}{\left\|\varepsilon+\mathbf{x}^T(k)\mathbf{x}(k)\right\|^2}\right\}$$
$$+\mathrm{E}\left\{\frac{n^T(k)\mathbf{x}(k)\mathbf{x}^T(k)n(k)}{\left\|\varepsilon+\mathbf{x}^T(k)\mathbf{x}(k)\right\|^2}\right\} \qquad (44)$$

Similarly, substituting (37) into the Right-Hand Side (RHS) of (43), then:

$$\mathrm{RHS}=\mathrm{E}\left\{\frac{e^T(k)\mathbf{x}(k)\mathbf{x}^T(k)e(k)}{\left\|\varepsilon+\mathbf{x}^T(k)\mathbf{x}(k)\right\|^2}\right\}$$
$$-\mathrm{E}\left\{\frac{n^T(k)\mathbf{x}(k)\mathbf{x}^T(k)e(k)}{\left\|\varepsilon+\mathbf{x}^T(k)\mathbf{x}(k)\right\|^2}\right\}$$
$$-\mathrm{E}\left\{\frac{e^T(k)\mathbf{x}(k)\mathbf{x}^T(k)n(k)}{\left\|\varepsilon+\mathbf{x}^T(k)\mathbf{x}(k)\right\|^2}\right\}$$
$$+\mathrm{E}\left\{\frac{n^T(k)\mathbf{x}(k)\mathbf{x}^T(k)n(k)}{\left\|\varepsilon+\mathbf{x}^T(k)\mathbf{x}(k)\right\|^2}\right\}$$
$$-\rho_{l_0}^2\,\mathrm{E}\left\{\left\|J\left[\mathbf{w}(k)\right]\right\|^2\right\} \qquad (45)$$

Combine (44) and (45):

$$\mathrm{E}\left\{\left[2\mu(k+1)-\left\|\mu(k+1)\right\|^2\right]\frac{e^T(k)\mathbf{x}(k)\mathbf{x}^T(k)e(k)}{\left\|\varepsilon+\mathbf{x}^T(k)\mathbf{x}(k)\right\|^2}\right\}$$
$$=\mathrm{E}\left\{\mu(k+1)\frac{n^T(k)\mathbf{x}(k)\mathbf{x}^T(k)e(k)}{\left\|\varepsilon+\mathbf{x}^T(k)\mathbf{x}(k)\right\|^2}\right\}$$
$$+\mathrm{E}\left\{\mu(k+1)\frac{e^T(k)\mathbf{x}(k)\mathbf{x}^T(k)n(k)}{\left\|\varepsilon+\mathbf{x}^T(k)\mathbf{x}(k)\right\|^2}\right\}$$
$$+\rho_{l_0}^2\,\mathrm{E}\left\{\left\|J\left[\mathbf{w}(k)\right]\right\|^2\right\} \qquad (46)$$

With reference to Assumption 3 that the additive noise $n(k)$ is statistically independent of the input signal $\mathbf{x}(k)$, then (46) can be simplified as:

$$\mathrm{E}\left\{\frac{e^T(k)\mathbf{x}(k)\mathbf{x}^T(k)e(k)}{\left\|\varepsilon+\mathbf{x}^T(k)\mathbf{x}(k)\right\|^2}\right\}=\frac{1}{\mathrm{E}\left[2\mathbf{I}_N-\mu(k+1)\right]}$$
$$\cdot\mathrm{E}\left\{\frac{n^T(k)\mathbf{x}(k)\mathbf{x}^T(k)n(k)}{\left\|\varepsilon+\mathbf{x}^T(k)\mathbf{x}(k)\right\|^2}\right\}$$
$$+\frac{1}{\mathrm{E}\left[2\mu(k+1)-\left\|\mu(k+1)\right\|^2\right]}$$
$$\cdot\rho_{l_0}^2\,\mathrm{E}\left\{\left\|J\left[\mathbf{w}(k)\right]\right\|^2\right\} \qquad (47)$$

Hence, from (47), it can be shown that the MSE of the VSS-l_0-NLMS algorithm is given as:

$$\mathrm{E}\left[e^T(k)e(k)\right]=\frac{\sigma_n^2}{\mathrm{E}\left[2\mathbf{I}_N-\mu(k+1)\right]}$$
$$+\frac{1}{\kappa_{\mathbf{xx}}\,\mathrm{E}\left[2\mu(k+1)-\left\|\mu(k+1)\right\|^2\right]}$$
$$\cdot\rho_{l_0}^2\,\mathrm{E}\left\{\left\|J\left[\mathbf{w}(k)\right]\right\|^2\right\} \qquad (48)$$

where $\kappa_{\mathbf{xx}}=\mathrm{E}\left\{\left[\mathbf{x}^T(k)\mathbf{x}(k)\right]/\left\|\varepsilon+\mathbf{x}^T(k)\mathbf{x}(k)\right\|^2\right\}$.

2.3.3 Computational Complexity

Considering the computational complexity of the VSS-l_0-NLMS algorithm in terms of the number of additions and multiplications required relative to the ISS-l_0-NLMS algorithm, an additional complexity is introduced by the VSS-l_0-NLMS algorithm as compared to the ISS-l_0-NLMS algorithm. This arises from the computation of the VSS parameter given in (4). Therefore, the additional computations that (4) requires are $2N+1$ multiplications and $N-1$ additions.

3 Results and Discussion

This section discusses computer simulation results performed to verify the theory presented in the previous section as well as experimentally compare the performance of the VSS-l_0-NLMS algorithm with the ISS-NLMS, VSS-NLMS and ISS-l_0-NLMS algorithms. Each simulation result is the steady-state statistical average of 200 runs, with 3000 iterations in each run. The received Signal-to-Noise Ratio (SNR) is defined as $10\log\left(E_0/\sigma_n^2\right)$, where $E_0=1$ is the received signal power and the noise power is given by $\sigma_n^2=10^{-\mathrm{SNR}/10}$.

The channel estimators are evaluated by averaging the mean square error (MSE) which is defined as:

$$\mathrm{MSE}\left[\mathbf{w}(k)\right]=\mathrm{E}\left[\left\|\mathbf{w}-\hat{\mathbf{w}}(k)\right\|_2^2\right] \qquad (49)$$

where w and $\hat{\mathbf{w}}(k)$ are the actual and the kth iterative channel update, respectively, and $\|\cdot\|_2$ is the Euclidean norm operator.

3.1 Experiment 1

In this experiment, a channel with length of 25 and with the number of dominant taps set to K = 1 and K = 4 is used. We compared the performance of the algorithms for three separate SNR values: 10, 20, and 30 dB. Other simulation parameters of importance are given in Tables 1 and 2. The algorithm variables, for this experiment when K = 1, having varying values, as indicated in Table 1, at each simulated SNR investigated is given in Table 2. In the case of K = 4, only the regulation parameter ρ_{l_0} varies as the SNR changes. The values used in this case are, 5×10^{-4} when SNR is 10 dB and 5×10^{-5} when the SNR is 20 dB or 30 dB. It is worthy of note that the sparsity-aware adaptive algorithms provide optimal performance at higher SNR values.

Table 1 Parameters Values Used For Simulating all the SNR Values Investigated

Algorithm		Sparsity Level	
		K = 1	*K = 4*
ISS-NLMS	μ	0.8	0.8
	ε	5×10^{-5}	5×10^{-5}
VSS-NLMS	$\mu(0)$	Varying	0.3
	μ_{min}	Varying	0
	μ_{max}	Varying	0.3
	ε	5×10^{-5}	5×10^{-5}
	α	0.99808	0.99808
	γ	Varying	0.9
ISS-l_0-NLMS	μ	0.8	0.8
	ε	5×10^{-5}	5×10^{-5}
	β	Varying	0.9999
	ρ	Varying	Varying
VSS-l_0-NLMS	$\mu(0)$	Varying	0.3
	μ_{min}	0	0
	μ_{max}	Varying	0.3
	ε	5×10^{-5}	5×10^{-5}
	α	0.99808	0.99808
	γ	Varying	0.9
	β	Varying	0.9999
	ρ	Varying	Varying

Table 2 Simulation Parameters of the Proposed Vss-l0-Nlms Algorithm and other Comparing Algorithms when K = 1

Algorithm		SNR Values (dB)		
		10	*20*	*30*
VSS-NLMS	$\mu(0)$	0.3	0.3	0.7
	μ_{min}	0	0	0.5
	μ_{max}	0.3	0.3	0.7
	γ	0.01	0.9	0.9
ISS-l_0-NLMS	β	0.09	0.09	0.9999
	ρ	0.01	0.01	5×10^{-4}
VSS-l_0-NLMS	$\mu(0)$	0.5	0.2	1
	μ_{max}	0.5	0.2	1
	γ	0.01	0.9	0.9
	β	0.09	0.09	0.9999
	ρ	0.01	0.004	5×10^{-4}

Figs. 3-8 show the average MSE to number of iterations for sparsity of K = 1 and K = 4 and different SNR regimes. At an SNR of 10 dB, in Fig. 3 and 4, the VSS algorithms provided the best performance at when K = 1 and K = 4. There is significant reduction in the convergence when K = 4. When the SNR is increased to 20 dB, again the VSS algorithms performed better with the sparsity-aware VSS algorithm providing the best performance which confirms the theory that such algorithms should provide optimal performance at higher SNR values as seen in Fig. 5 and 6. Here again, the convergence is suboptimal compared to the ISS algorithms when K = 4 with the VSS-l_0-NLMS algorithm providing the best convergence result when K = 1. In Fig. 7 and 8, when the SNR is 30 dB, the VSS-l_0-NLMS algorithm performs better than the other algorithms with an average MSE of about -38 dB in both scenarios. Significantly, the ISS-l_0-NLMS algorithm performed better than the ISS-NLMS algorithm when K = 1.

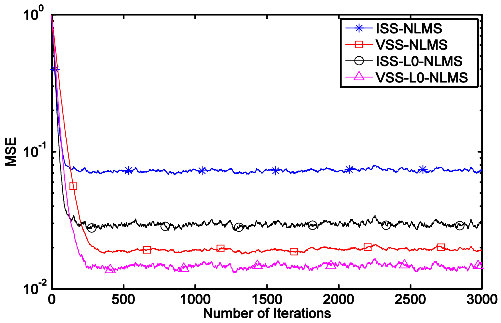

Fig. 3 Mse Performance at an SNR of 10 Db For When K = 1

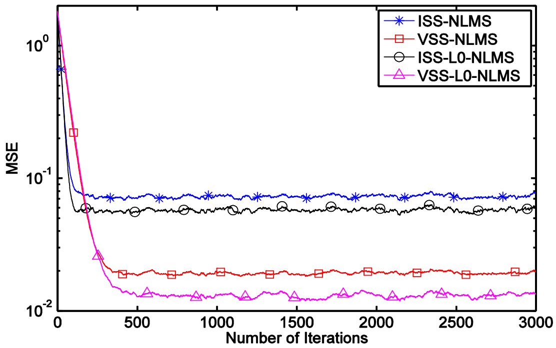

Fig. 4 MSE performance at an SNR of 10 dB for when K = 4

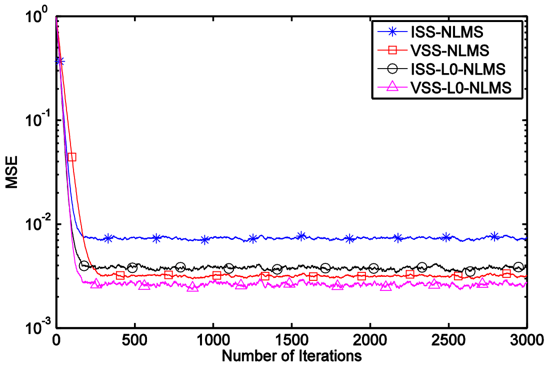

Fig. 5 MSE performance at an SNR of 20 dB for when K = 1

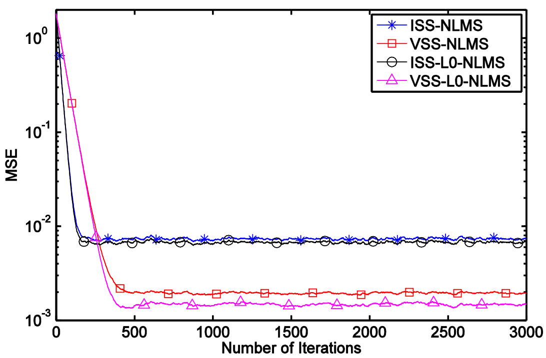

Fig. 6 MSE Performance at an SNR of 20 dB for when K = 4

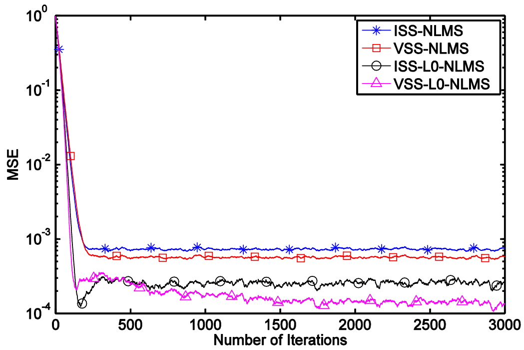

Fig. 7 MSE performance at an SNR of 30 dB for when K = 1

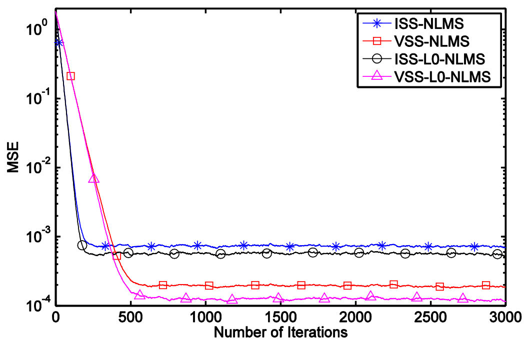

Fig. 8 Mse Performance at An SNR Of 30 Db For When K = 4

3.2 Experiment 2

In this experiment, we consider a sparse multipath communication channel, typical of UWB channels, adopted from (Carbonelli *et al.*, 2007), with a channel length of 30 having 8 dominant taps as shown in Fig. 9. Here also, several simulations are conducted for this analysis we compared the performance of the algorithms for three separate SNR values: 10, 20 and 30 dB. The simulation parameters used in this experiment are listed as follows: $\mu = 0.8$, $\mu(0) = 1$, $\mu_{\min} = 0$, $\mu_{\max} = 0.5$, $\alpha = 0.99808$, $\gamma = 0.9$, $\varepsilon = 5 \times 10^{-5}$, $\beta = 0.9999$ and $\rho = 5 \times 10^{-5}$. Other than these settings, others are same as described in Experiment 1.

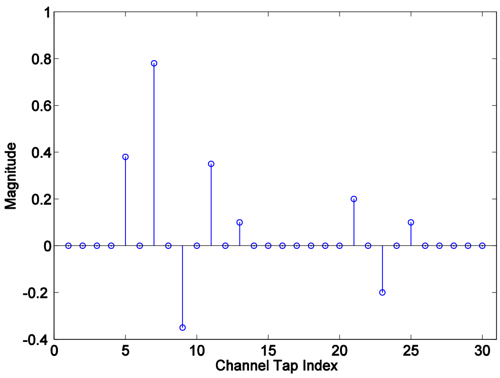

Fig. 9 a Typical Sparse Multipath Channel Impulse Response

Figures 10-12 show the average MSE to number of iterations for different SNR regimes. In all three scenarios under consideration, the VSS algorithms provided the best performance. The results show significant reduction in the convergence as the SNR increases.

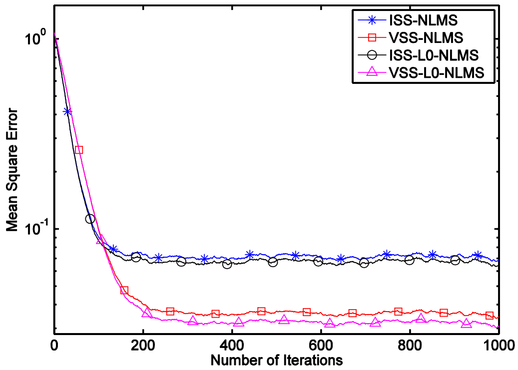

Fig. 10 MSE Performance when SNR = 10 dB

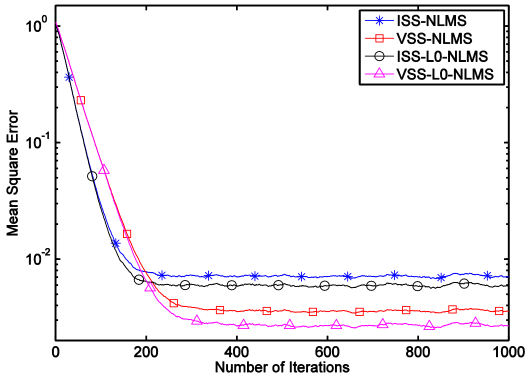

Fig. 11 MSE Performance when SNR = 20 dB.

4 Conclusions and Recommendation

In this paper, we proposed the VSS-l_0-NLMS algorithm for sparse multipath channel estimation. This research is motivated by the properties of the step-size parameter in controlling the estimation accuracy, convergence speed and computational complexity. The algorithm estimates the sparseness of the impulse response. Computer simulations show that in high sparsity scenarios, the proposed algorithm exhibits faster convergence and better performance. As future work, we will consider the method of partial updating coefficients to help reduce the computational complexity, which in effect will reduce the computational load as well.

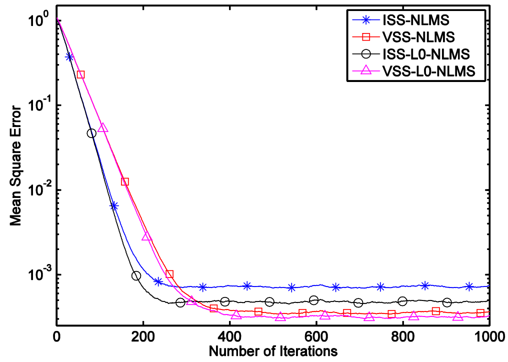

Fig. 12 MSE Performance when SNR = 30 dB

References

Carbonelli, C., Vedantam, S. and Mitra, U. (2007), "Sparse Channel Estimation with Zero Tap Detection", *IEEE Transactions on Wireless Communications*, Vol. 6, No. 5, pp. 1743–1763.

Chen, Y., Gu, Y. and Hero, A. O. (2009), "Sparse LMS for System Identification", *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, April 2009 IEEE, Taipei, pp. 3125–3128.

Das, B. K., Chakraborty, M. and Banerjee, S. (2011), "Adaptive Identification of Sparse Systems with Variable Sparsity", *2011 IEEE International Symposium of Circuits and Systems (ISCAS)*, May 2011 IEEE, Rio de Janeiro, Brazil, pp. 1267–1270.

Diniz, P. S. R. (2013), *Adaptive Filtering*, Fourth Edi., Springer US, Boston, MA.

Donoho, D. L. (2006), "Compressed Sensing", *IEEE Transactions on Information Theory*, Vol. 52, No. 4, pp.1289–1306.

Gu, Y., Jin, J. and Mei, S. (2009), "l_0 Norm Constraint LMS Algorithm for Sparse System Identification", *IEEE Signal Processing Letters*, Vol. 16, No. 9, pp. 774–777.

Gui, G. and Adachi, F. (2013), "Improved Least Mean Square Algorithm with Application to Adaptive Sparse Channel Estimation", *EURASIP Journal on Wireless Communications and Networking*, Vol. 2013, No. 1, pp. 204.

Gui, G., Kumagai, S., Mehbodniya, A. and Adachi, F. (2013), "Variable is Good: Adaptive Sparse Channel Estimation Using VSS-ZA-NLMS Algorithm", *2013 International Conference on Wireless Communications and Signal Processing*. October 2013 IEEE, pp. 1–5.

Harris, R. W., Chabries, D. M. and Bishop, F. A. (1986), "A Variable Step (VS) Adaptive Filter Algorithm", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-34, No. 2, pp. 306–316.

Haykin, S. (2002), *Adaptive Filter Theory*, Fourth Edi., Prentice-Hall, Inc., New Jersey.

Kwong, R. H. and Johnston, E. W. (1992), "A Variable Step Size LMS Algorithm", *IEEE*

Transactions on Signal Processing, Vol. 40, No. 7, pp. 1633–1642.

Li, Y. and Hamamura, M. (2014a), "An Improved Proportionate Normalized Least-Mean-Square Algorithm for Broadband Multipath Channel Estimation", *The Scientific World Journal*, Vol. 2014, pp. 1–9.

Li, Y. and Hamamura, M. (2014b), "Smooth Approximation l(0)-Norm Constrained Affine Projection Algorithm and its Applications in Sparse Channel Estimation", *The Scientific World Journal*, Vol. 2014, pp.1–15.

Nunoo, S., Chude-Okonkwo, U. A. K. and Ngah, R. (2013), "Performance of LMS, NLMS and LMF algorithms in tracking time-varying UWB channels", *2013 IEEE International Conference on Signal and Image Processing Applications*. October 2013 IEEE, Melaka, Malaysia, pp. 312–316.

Nunoo, S., Chude-Okonkwo, U. A. K., Ngah, R. and Zahedi, Y. K. (2014), "Variable Step-Size l_0-Norm NLMS Algorithm for Sparse Channel Estimation", *2014 IEEE Asia Pacific Conference on Wireless and Mobile*. August 2014 IEEE, Bali, Indonesia, pp. 88–91.

Sayed, A. H. (2008), *Adaptive Filters*, John Wiley & Sons, Inc., Hoboken, NJ, USA.

Su, G., Jin, J., Gu, Y. and Wang, J. (2012), "Performance Analysis of l_0 Norm Constraint Least Mean Square Algorithm", *IEEE Transactions on Signal Processing*, Vol. 60, No. 5, pp. 2223–2235.

Taheri, O. and Vorobyov, S. A. (2014), "Reweighted l_1-Norm Penalized LMS for Sparse Channel Estimation and its Analysis", *Signal Processing*, Vol. 104, pp. 70–79.

Yang, Y., Zhao, J., Wang, Z. and Yan, Y. (2010), "An Novel Variable Step Size LMS Adaptive Filtering Algorithm Based on Hyperbolic Tangent Function", *2010 International Conference on Computer Application and System Modeling (ICCASM 2010)*. 2010 IEEE, Taiyuan, pp. V14-233-V14-236.

Authors

Solomon Nunoo is currently a Senior Lecturer at the Department of Electrical and Electronic Engineering, UMaT He holds BSc degree in electrical engineering from Western University College of Kwame Nkrumah University of Science and Technology, now University of Mines and Technology (UMaT), the MPhil degree in electrical and electronic engineering, and the PhD in electrical engineering from Universiti Teknologi Malaysia. He. His research interest is in signal processing for wireless communications with emphasis on adaptive filtering and compressive sampling.

Uche A. K. Chude-Okonkwo is currently a Senior Research Fellow with the Department of Electrical, Electronic and Computer Engineering, University of Pretoria, South Africa. He holds PhD degree in Electrical Engineering from Universiti Teknologi Malaysia, His current research interests include molecular communication applied to advanced healthcare delivery, signal processing, wireless sensor network, wireless communication, and systems biology.

Razali Ngah is currently an Associate Professor at Wireless Communication Centre (WCC), Faculty of Electrical Engineering, Universiti Teknologi Malaysia (UTM) Skudai. He holds BSc degree in Electrical Engineering (Communication) from Universiti Teknologi Malaysia, Skudai, MSc in RF Communication Engineering from University of Bradford, UK and PhD in Photonics from University of Northumbria, UK. His research interests are Mobile Radio Propagation, Antenna and RF design, Photonics Network, Wireless Communication Systems and Radio over Fiber (RoF).