<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>A new switch-mode noise-constrained transform domain NLMS adaptive filtering algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Chan, SC; Chu, YJ; Tsui, KM; Zhang, ZG</td>
</tr>
<tr>
<td><strong>Issued Date</strong></td>
<td>2011</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10722/143311">http://hdl.handle.net/10722/143311</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>IEEE International Symposium on Circuits and Systems Proceedings. Copyright © IEEE.</td>
</tr>
</tbody>
</table>
A New Switch-Mode Noise-Constrained Transform Domain NLMS Adaptive Filtering Algorithm

S. C. Chan, Y. J. Chu, K. M. Tsui, and Z. G. Zhang
Electrical and Electronic Engineering Department, The University of Hong Kong
Hong Kong, P. R. China

Abstract—The transform domain normalized least mean squares (TDNLMS) algorithm is an efficient adaptive algorithm, which offers fast convergence speed with a reasonably low arithmetic complexity. However, its convergence speed is usually limited by the fixed step-size so as to achieve a low desired misadjustment. In this paper a new switch-mode noise-constrained TDNLMS (SNC-TDNLMS) algorithm is proposed. It employs a maximum step-size mode in initial convergence and a noise-constrained mode afterwards to improve the convergence speed and steady-state performance. The mean and mean square convergence behaviors of the proposed algorithm are studied to characterize its convergence condition and steady-state excess mean square error (EMSE). Based on the theoretical results, an automatic threshold selection scheme for mode switching is developed. Computer simulations are conducted to show the effectiveness of the proposed algorithm and verify the theoretical results.

I. INTRODUCTION

Adaptive filters are frequently used in system identification and related problems, where the statistics of the underlying signals are either unknown a priori, or slowly-varying. The adaptive filtering algorithms are usually variants of the well known LMS [1] and RLS [12] algorithms. The normalized LMS (NLMS) algorithm [2] and the transform domain NLMS (TDNLMS) are also commonly used due to their good numerical stability and computational simplicity.

In particularly, TDNLMS algorithm [3-5] is attractive due to its fast convergence speed and reasonably low arithmetic complexity. It exploits the decorrelation property of transformations, such as the discrete cosine transform (DCT) or the wavelet transform (WT), to approximately prewhiten the input signal to reduce the eigenvalue spread of the input autocorrelation matrix. Consequently, the convergence rate can be improved significantly. In conventional TDNLMS algorithms, the step-size is fixed and therefore the convergence speed is limited by the desired misadjustment. This has motivated considerable interest in designing reliable and efficient variable stepsize (VSS) algorithms to overcome this drawback [6-11]. These algorithms aim to employ large step-size to speed up the convergence rate initially and gradually decrease the step-size in order to achieve a low excess mean square error (EMSE). This is often accomplished by varying the step-size values based on a certain measure of convergence status [7-11]. In [6], the modified VSS TDNLMS (MVSS-TDNLMS) algorithm varies the step-size by estimating the noise power.

In this paper, a switch-mode noise-constrained TDNLMS (SNC-TDNLMS) algorithm is proposed. It exploits the prior knowledge of the additive noise variance as in the NCLMS approach [9] and gives rise to a VSS algorithm. Moreover, the improved performance is found to be obtained if maximum step-size is employed at initial convergence while the NC adaptation is more suitable to be used near convergence in order to reduce the steady-state misadjustment. Therefore, the proposed method is extended to include a switch-mode scheme which employs a maximum step-size mode (MSM) during initial convergence and a NC mode (NCM) afterwards so as to simultaneously improve the convergence speed and steady-state performance. The mean and mean squares convergence of the proposed SNC-TDNLMS algorithm is studied and its steady-state EMSE is characterized. Based on the theoretical results, an automatic threshold selection scheme for mode switching and recommendations for typical algorithm parameters are proposed. Simulation results show that the SNC-TDNLMS algorithm has faster convergence speed than the traditional TDNLMS algorithm. The theoretical and computer simulation results also agree well with each other. The rest of the paper is organized as follows. In Section II, the TDNLMS algorithm is briefly reviewed. This is followed by the proposed SNC-TDNLMS algorithm. Section III is devoted to the mean and mean square convergence performance of the proposed algorithm. Simulation results and comparisons with conventional methods are presented in Section IV.

II. THE SNC-TDNLMS ALGORITHM

A. Review of the TDNLMS Algorithm

Consider the identification of a linear time-invariant (LTI) finite impulse response (FIR) system by an adaptive filter with the same length. The impulse response coefficient vector of the system is assumed to be \( w^* \) and it is of \( L \) taps. The unknown system and adaptive filter are both excited by an input \( x(n) \). The measured output of the system is \( d(n) \), which is assumed to be corrupted by an additive noise \( \eta(n) \),

\[
d(n) = (w^*)^T x(n) + \eta(n),
\]

where \( x(n) = [x(n), \ldots, x(n-L+1)]^T \) is the input vector.

The update equations for the TDNLMS algorithm are:
\[ e(n) = d(n) - W^T(n)X_C(n), \]
\[ W(n+1) = W(n) + \mu e(n)X(n), \]
where \( W(n) = Cw(n) = [w_{c,1}(n), w_{c,2}(n), \ldots, w_{c,L}(n)]^T \) and \( X_C(n) = [X_{c,1}(n), X_{c,2}(n), \ldots, X_{c,L}(n)]^T \) are the transformed adaptive weight vector and signal vector. \( C \) is an \( L \times L \) transformation matrix such as DFT or DCT. \( \mu \) is the step-size. \( X_i^T = \text{diag}(\varepsilon_i(n), \varepsilon_i(n), \ldots, \varepsilon_i(n)) \) is an element-wise normalization matrix with \( \varepsilon_i(n) \) being the estimated power of the \( i \)-th signal component after transformation. In this paper, \( \varepsilon_i(n) = \sigma_i + \alpha_i \) is considered, where \( \alpha_i \) is a positive forgetting factor smaller than one. \( \sigma_i \) is a small positive value which can be chosen as certain prior power estimate of the corresponding component.

### B. The NC-TDNLMS Algorithm

In [10], a transformation approach was proposed to derive the NC-NLMS algorithm from its LMS counterpart. This method is also applicable to the TDNLMS algorithm. Hence, the NC-based TDNLMS algorithm can be updated as

\[ \mu(n) = \alpha + \gamma \hat{\epsilon}(n), \]
\[ \lambda(n+1) = (1 - \beta)\lambda(n) + \epsilon \hat{J}(n), \]
where \( \alpha, \beta, \gamma \) are constant parameters and \( \hat{J}(n) = e^2(n) - \sigma^2_i \) is the instantaneous estimate of the EMSE. It can be seen that the convergence measure \( \lambda(n) \) is comparatively large during initial convergence. Hence, a larger value of \( \mu(n) \) will be chosen in order to speed up the convergence rate. As the EMSE decreases, \( \mu(n) \) is then gradually decreased to achieve a lower steady-state EMSE.

As suggested in [9], after fixing the nominal step-size \( \alpha \), \( \gamma \) should be chosen as a value large as possible to obtain a fast convergence speed, while \( \beta \) should be chosen as a small value to achieve a desired EMSE. However, the values of \( \gamma \) and \( \alpha \) are still constrained so that the step-size and hence the convergence speed will be significantly limited. From the mean convergence analysis, to be presented in Section III, we found that the mean weight error vector will converge faster if a maximum possible step-size is employed. On the other hand, the NC adaptation should be used when the adaptive filter is nearly converged in order to achieve the desired steady-state EMSE.

Because of the above observations and possible advantages, we propose below a novel switch-mode scheme for the variable step-size. It employs

1) the maximum step-size mode (MSM), where a designed maximum step-size \( \mu_{\text{max}} \) is employed to achieve a faster convergence speed during initial convergence, and
2) the noise constrained mode (NCM), where the step-size is adjusted as in the NC algorithms according to (4) and (5). Thus, the desired EMSE can be achieved after the maximum step-size mode is nearly converged.

Consequently, the corresponding updates for the step-size can be summarized by the following equations

\[ \mu(n) = \begin{cases} \mu_{\text{max}}, & \text{if } \lambda(n) \geq T, \text{(MSM)}, \\ \max[\sigma(1 + \gamma \hat{\epsilon}(n)), \sigma(1 + \delta)], & \text{if } \lambda(n) < T, \text{(NCM)}, \end{cases} \]
\[ \lambda(n+1) = (1 - \beta)\lambda(n) + \beta \hat{J}(n)/2, \]
where \( \mu_{\text{max}} \) is the designed maximum step-size. To switch between the two modes, we employ the noise power estimate to measure the convergence status. However, to achieve a fast switching response, a large value of \( \beta \), denoted as \( \beta \), is used to estimate a short-term EMSE \( \lambda(n) \) as shown in (7).

When \( \lambda(n) \) is larger than a certain threshold \( T \), the MSM is invoked. When \( \lambda(n) \) is smaller than \( T \), the NCM is invoked, where a small \( \beta \) is used to estimate a long-term EMSE \( \lambda(n) \) as shown in (8). The value of \( \lambda(n) \) immediately after mode switching is obtained from \( \lambda(n) \). In cases of noise variance mismatch, the true noise variance in (7) and (8) should be replaced by \( \sigma^2_i \), which is assumed to be \( a \) times of \( \sigma^2_i \).

The key issue with the switch-mode approach is the proper selection of the switching threshold \( T \) and the other related parameters. In this paper, a novel threshold selection scheme is proposed based on the performance analysis in the next section. The selection of parameters will also be discussed.

### III. PERFORMANCE ANALYSIS

In this section, we analyze the convergence performance of the proposed SNC-TDNLMS algorithm. The following commonly used assumptions are made:

(A1) \( \mu(n) \) is independent of the input and error sequence;

(A2) \( W(n), \{ x(n) \} \) and \( \{ \eta(n) \} \) are statistically independent;

(A3) \( \{ x(n) \} \) is an independent identically distributed (i.i.d) Gaussian sequence with zero-mean and covariance matrix \( R_x \).

(A1) is an approximation commonly used in the analysis of VSS LMS algorithms to make it mathematically tractable. (A2) is the independence assumption, which is a good approximation for large value of \( L \) and for small to medium step-size to simplify the convergence analysis. Moreover, we denote \( W^* = R_\text{xx}^{-1}P_{\alpha}, \) where \( P_{\alpha} = E[d(n)X_c(n)] \) is the ensemble averaged cross-correlation vector between \( X_c(n) \) and \( d(n) \).

\( W^* \) is related to the Wiener solution by \( w_{\text{opt}} = R_\text{xx}^{-1}P_{\alpha} = CW^* \).

#### A. Mean Convergence Analysis

First, let the weight error vector at time \( n \) be \( \{ v(n) = W(n) - W(n) \} \). By using (2), (3), (6)-(8) and the assumptions above, the difference equations of the mean weight error vector, mean step-size and mean multiplier \( E[\lambda(n)] \) can be derived using the results in [5]:

\[ E[v(n+1)] = E[v(n)] - E[\mu(n)]E_{\varepsilon,x_c}E[\varepsilon_i(n)X_i(n)] 
= (I - E[\mu(n)]D_{\alpha})E[v(n)], \]
\[ E[\mu(n)] = \alpha(1 + \gamma E[\lambda(n)]), \]
where \( E_{\varepsilon,x_c}E[\varepsilon_i(n)X_i(n)] \) denotes the expectation over \( \{ v, X_i(n) \}, \eta(n) \). \( D_{\alpha} = \text{diag}(\alpha_i, \ldots, \alpha_i) \) is a diagonal matrix with \( \alpha_i = \ldots, \alpha_i \).

118
\[ \exp(-ue^i(g(\hat{b}))^{1/2})du, \quad g(\hat{b}) = (1 + 2\hat{\beta}R_{\hat{X}_k, \hat{X}_k})^{-1} \]

\[ \hat{b} = \alpha_u \text{ and } R_{X_k, X_{k-1}} \text{ being the (i,j)-th element of } R_{X_k, X_k}. \]

For notational convenience, let \( b = a - 1 \), then \( J(s) = E[e^2(n)] \)

\[-a_s^2 J(n) - b \sigma_s^2.\]

We shall only focus on the NC adaptation mode, as the
MSM mode is equivalent to the TDNLMS algorithm with a maximum step-size. The latter can be obtained by assuming
\( \mu(n) \) to be a constant and the details can be found in [5].

Based on (9) and expressing the weight error \( v(n) \) as
\[ V(n) = D_{\mu}^{1/2}v(n), \]
we get
\[ E[V(n+1)] = (I - E[\mu(n)]D_{\mu}R_{X_k, X_k}E[V(n)], \]

(12)

where \( R_{X_k, X_k} = D_{\mu}^{1/2} R_{X_k, X_k} D_{\mu}^{1/2} \) is the correlation matrix of a scaled input vector \( X_o = D_{\mu}^{1/2} X_k. \) Since it is symmetric, it can be written as the following eigenvalue decomposition (EVD):

\[ R_{X_k, X_k} = U_{\lambda} \Lambda_{\lambda} U_{\lambda}^T \]

and \( \Lambda_{\lambda} = \text{diag}(\lambda_1', \lambda_2', \ldots, \lambda_L') \) contains corresponding eigenvalues. Since Eqn. (12) is identical to the difference equation of the LMS algorithm, the classical result of the maximum possible step-size can be obtained as

\[ \mu_{\text{max}} = 2 / \lambda_m', \]
where \( \lambda_m' \) is the maximum eigenvalue of \( R_{X_k, X_k}. \)

**B. Mean Square Convergence Analysis**

To evaluate the mean square behavior, multiplying \( v(n) \) by its transpose and taking expectation on both sides, one gets a difference equation of the weight error covariance matrix:

\[ \Xi(n+1) = \Xi(n) - E[\mu(n)]D_{\mu}R_{X_k, X_k}E[\nu(n)], \]

(13)

where \( E[\nu(n)v^T(n)] = 0, \text{and } E[\nu(n)v^T(n)] = \Lambda_m \text{, where } \Lambda_m \text{ is the maximum eigenvalue of } R_{X_k, X_k}. \)

**Switching Threshold and Parameter Selection**

To switch to state 1, let the mean square error (MSE) of the weights be bounded and the following equation be solved:

\[ \xi(n) = \sqrt{J(n)}(1 + \delta \lambda(n)) \]

(14)

Therefore, the maximum mean square error (MSE) in state 1 is

\[ \Xi(n) = \sum_{m=1}^n \Lambda_m \text{, where } \Lambda_m \text{ is the maximum eigenvalue of } R_{X_k, X_k}. \]

The threshold for switching is chosen as

\[ J_s(n) = \frac{1}{\lambda_s(n)} \]

(15)

where \( \lambda_s(n) \) is the maximum eigenvalue of the correlation matrix of the input vector.

**Switching Threshold and Parameter Selection**

1) Selection of \( T \): From (15), the steady-state EMSE can be calculated as

\[ J_s(n) = \frac{1}{\lambda_s(n)} \]

(16)

where \( T = (\mu_{\text{max}} c_f + \sqrt{2\rho^2(1 - \mu_{\text{max}} c_f)} \mu_{\text{max}} c_f + 1) \sigma_s^2 \text{ and } \mu_{\text{max}} \text{ as in (16) is used, then } \mu_{\text{max}} c_f = 2. \) On the other hand, \( \kappa \) can be adjusted

\[ \lambda_s(n) = \sum_{m=1}^n \Lambda_m \text{, where } \Lambda_m \text{ is the maximum eigenvalue of } R_{X_k, X_k}. \]

The threshold for switching is chosen as

\[ J_s(n) = \frac{1}{\lambda_s(n)} \]

(17)

Therefore, the maximum mean square error (MSE) in state 1 is

\[ \Xi(n) = \sum_{m=1}^n \Lambda_m \text{, where } \Lambda_m \text{ is the maximum eigenvalue of } R_{X_k, X_k}. \]
experimentally and appropriate values are around 3 to 5.

2) Choice of $\beta$ and $\overline{\beta}$: Generally, we observe that the parameter $(1 - \beta)$ (or $(1 - \overline{\beta})$) acts as a forgetting factor and controls the averaging process of the instantaneous MSE. The best value of $\beta$ (or $\overline{\beta}$) depends mildly on the convergence speed. The recommended value for $\beta$ is around 0.01. For $\overline{\beta}$, a larger value around 0.1 can be used because the algorithm is converging at the fastest speed under the MSM mode.

3) Choice of $\alpha$, $\delta$ and $\gamma$: According to (22), the product $\alpha (1 + \delta)$ is fixed for a desired EMSE. Since $\alpha$ contributes more to the convergence speed during the NC mode, it is advantageous to increase $\alpha$ and decrease $\delta$. A typical value of $\delta$ is 0.1. Finally, $\gamma$ can be computed from the definition of $\beta_2$ in (22). If $\sigma_v^2$ is not exactly known, we recommend to use the upper bound of $\sigma_v^2\text{max}$ in (22).

IV. SIMULATION RESULTS

In this section, computer simulations are conducted to evaluate the convergence behavior of the proposed algorithm and verify the analytical results obtained in Section III. As a comparison, we also consider the conventional TDNLMS and MVSS-TDNLMS algorithms [6]. These algorithms use a DCT transformation due to its wide usage and efficiency in practice. To simplify the comparison with the other algorithms, the estimated power of input element is chosen as $e_i(n) = \sigma_i + \alpha_i X^2_{ii}(n)$, where $\sigma_i$ is the input power and $\alpha_i = 0.1$. The results for the estimated power in place of $\sigma_v$ are similar. The simulations are performed using the system identification model and the unknown system to be estimated is an L-order ($L=8$) FIR filter. Different signal-to-noise ratios (SNRs) are used to examine the performance of the parameter selection scheme proposed in Section IIIIC. The maximum step-size $\mu_{\text{max}} = 0.13$ and $\kappa = 4$. Since the TD algorithms are usually employed when the input is colored, the first order auto-regressive process is considered: $x(n) = 0.9x(n-1) + g(n)$, where $g(n)$ is a zero-mean white Gaussian noise. For fair comparison, the algorithm parameters are chosen such that all the algorithms achieve the same steady-state EMSE. The step-size for the TDNLMS algorithm is 0.007; $\gamma$ in MVSS-TDNLMS is 0.996. For SNC-TDNLMS $\delta$ is chosen to be 0.1 and $\alpha$ is determined to be 0.0064. Thus, $\gamma$ is calculated from (22) as 0.2, 2 and 20 for SNR=0 dB, 10 dB and 20 dB, respectively. The recommended values $\beta = 0.01$ and $\overline{\beta} = 0.1$ are used. The learning curves of EMSE are shown in Figs. 1(a), (b), (c). It can be seen that the SNC-TDNLMS algorithm generally converges at the fastest speed. The improvement is more significant as the SNR increases. The theoretical and simulation results agree well with each other, especially when the algorithms are near convergence. The deviation at initial convergence at high SNR is caused by the maximum step-size used, where the validity of the independent assumption in (A2) becomes less accurate. Simulations for white Gaussian input can be found in the supplementary document [13]. And results for longer filter length are similar.

REFERENCES