<table>
<thead>
<tr>
<th>Title</th>
<th>A new switch-mode noise-constrained transform domain NLMS adaptive filtering algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Chan, SC; Chu, YJ; Tsui, KM; Zhang, ZG</td>
</tr>
<tr>
<td>Issued Date</td>
<td>2011</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/10722/143311">http://hdl.handle.net/10722/143311</a></td>
</tr>
<tr>
<td>Rights</td>
<td>IEEE International Symposium on Circuits and Systems Proceedings. Copyright © IEEE.; This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.; ©2011 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the 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 analyzed 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 particular, 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 (NMC) 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 \( \mathbf{w}^* \) 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 \( \eta(n) \), so that:

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

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

The update equations for the TDNLMS algorithm are:
\[ e(n) = d(n) - W^T(n)X_C(n), \]

\[ W(n+1) = W(n) + \mu(n)X_C(n)e(n), \]

where \( W(n) = \text{Cw}(n) = [W_{c1}(n), W_{c2}(n), \ldots, W_{cn}(n)]^T \) and \( X_C(n) = [X_{c1}(n), X_{c2}(n), \ldots, X_{cn}(n)]^T \) 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. \( \lambda_i \) is the independence assumption, which is a good approximation commonly used in the analysis of Gaussian sequence with zero-mean and covariance matrix \( \Sigma \). 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.

\[ \alpha, \beta, \gamma \text{ are constant parameters and } \hat{\lambda}(n) = e^T(n) - \sigma^2. \]

\[ \mu(n) = \left\{ \begin{array}{ll}
\mu_{\text{max}}, & \lambda(n) \geq T, (\text{MSM}), \\
\max[s(1 + \gamma(n)), s(1 + \delta)], & \lambda(n) < T, (\text{NCM}).
\end{array} \right. \]

\[ \lambda(n+1) = (1 - \beta)\lambda(n) + \beta \hat{\lambda}(n)/2, \]

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 \( \hat{\beta} \), is used to estimate a short-term EMSE \( \hat{\lambda}(n) \) 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 \( \hat{\lambda}(n) \).

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

1. \( \mu(n) \) is independent of the input and error sequence; \( W(n) \), \( x(n) \) and \( \{\eta(n)\} \) are statistically independent;
2. \( \{\eta(n)\} \) is an independent identically distributed (i.i.d) Gaussian sequence with zero-mean and covariance matrix \( \Sigma \).

(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 \( W^* = R_{\text{ave}}P_{\alpha}, \) 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 \( w_{\text{opt}} = R_{\text{ave}}P_{\alpha} = CW^* \).

### Mean Convergence Analysis

First, let the weight error vector at time \( n \) be \( v(n) = W^* - W(n) \). 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):

\[ E[v(n+1)] = \lambda(n) - \lambda(n)]E_{\Sigma_{X_c}, v}[e_i(n)X_c(n)](n) \]

\[ = (I - E[\mu(n)]D_nX_{c1})E[v(n)], \]

\[ E[\mu(n)] = \alpha(1 + \gamma \lambda(n)), \]

\[ E[\lambda(n) + 1] = (1 - \beta)E[\lambda(n)] + \beta \lambda(n)/2, \]

where \( E_{\Sigma_{X_c}, v}[] \) denotes the expectation over \( \{v, X_c(n), \eta(n)\} \). \( D_n = \text{diag}(\alpha_1, \ldots, \alpha_n) \) is a diagonal matrix with \( \alpha_i = \ldots, \alpha_i \ldots \)
\[ \int \exp(-ue)(g(\tilde{\beta}))^{2} du, \quad g(\tilde{\beta}) = (1 + 2\beta R_{X_{c},X_{c}}), \quad \tilde{\beta} = \alpha u \text{ and } R_{X_{c},X_{c},i,j} \text{ being the } (i,j)\text{-th element of } R_{X_{c},X_{c}}. \]

For notational convenience, let \( b = a - 1 \), then \( J_{s}(n) = E[v^{2}(n)] - \alpha^{2} \sigma_{v}^{2} = J(n) - b \sigma_{v}^{2}. \)

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 \( V(n) = D_{v}^{(i)}v(n) \), we get
\[ E[V(n+1)] = (I - E[\mu(n)]D_{v}^{(i)} )E[V(n)], \]
where \( R_{X_{c},X_{c},i,j} = D_{v}^{(i)} D_{v}^{(j)} \) is the correlation matrix of a scaled input vector \( X_{o} = D_{v}^{(i)} X_{c} \). Since it is symmetric, it can be written as the following eigenvalue decomposition (EVD):
\[ R_{X_{c},X_{c}} = U_{X_{c}} \Lambda_{X_{c}} U_{X_{c}}^{T}, \quad \Lambda_{X_{c}} = \text{diag}(\lambda_{1},\lambda_{2},\cdots,\lambda_{n}) \]
contains corresponding eigenvalues. Since Eqn. (12) is identical to the difference equation of the weight error covariance matrix:
\[ \sigma_{v}^{2}(n) = \frac{1}{n_{E}} \sum_{n_{E}} v^{2}(n) \]
and \( \lambda_{n} \) is the maximum eigenvalue of \( R_{X_{c},X_{c}} \).

### 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:
\[ \mathcal{E}(n+1) = \mathcal{E}(n) - E[\mu(n)]D_{v}^{(i)} R_{X_{c},X_{c},i,j} \mathcal{E}(n) - E[\mu(n)]D_{v}^{(i)} E[v^{2}(n)], \]
where \( \mathcal{E}(n) = E[v(n)v^{T}(n)] \), \( s_{i,j} = E[X_{c,i}X_{c,j}^{T}] \) and \( \lambda_{n} \), \( \lambda_{n}^{T} \), \( \lambda_{n}^{T} \) are diagonal matrices with the \( n \)-th row of \( \lambda_{n} \) and \( \lambda_{n}^{T} \) is the maximum eigenvalue of \( R_{X_{c},X_{c}} \).

The \( (i,j)\)-th element of \( s_{i,j} \) is evaluated to be [5]
\[ s_{i,j}^{(k)} = \sum_{n=2}^{\infty} p_{n} (-4R_{X_{c},X_{c},i,j})^{n} e_{t_{n}}^{2} (\lambda_{n}^{T})^{2}, \]
where \( k = 2, 3, 4 \). Using (14), the last term in (13) can be simplified to
\[ E[v^{2}(n)]S_{i,j}^{(i)} = (R_{X_{c},X_{c}} \mathcal{E}(n)R_{X_{c},X_{c}}^{T})^{(i)} D_{v}^{(i)} + D_{v}^{(i)} S_{i,j}^{(i)}  \]
\[ + 4\sum_{n=0}^{\infty} \alpha^{2}_{(n)} \Lambda_{X_{c}}(n) \mathcal{E}(n)D_{v}^{(i)} + \sigma_{v}^{2}(n) \Gamma_{v}(n), \]
where \( \alpha^{2}_{(n)} = \sum_{n=0}^{\infty} \alpha^{2}_{(n)} \Lambda_{X_{c}}(n) \mathcal{E}(n)D_{v}^{(i)} \) and \( \Gamma_{v}(n) = \sum_{n=0}^{\infty} \alpha^{2}_{(n)} \Lambda_{X_{c}}(n) \mathcal{E}(n)D_{v}^{(i)} \).

\[ \Lambda_{X_{c}} = \text{diag}(\lambda_{1},\lambda_{2},\cdots,\lambda_{n}) \]

It can be shown that a sufficient condition for mean square convergence of (15) is given by
\[ \frac{2}{\mu} \leq \frac{\operatorname{Tr}(D_{v}^{(i)} D_{v}^{(j)} + \operatorname{Tr}(D_{v}^{(i)} \Gamma_{v}(n)))}{\mu_{\text{max}}}, \]
where \( D_{v}^{(i)} = (D_{v}^{(i)} + D_{v}^{(j)} + D_{v}^{(j)} + D_{v}^{(j)}) \) and \( D_{v}^{(i)} \) are diagonal matrices with the \( i \)-th diagonal element equal to that of \( S_{i,j}^{(i)} \).

To evaluate the steady-state EMSE, \( E[\mu^{2}(n)] \) and \( E[X_{c}^{2}(n)] \) are evaluated from (4) and (5) as follows
\[ E[\mu^{2}(n)] = \alpha^{2}_{\mu} + 2\beta E[\lambda_{(n)}] + E[X_{c}^{2}(n)], \]
where \( \lambda_{(n)} \) is the limiting value of \( \mu^{2}(n) \) is obtained by using (17)(18),
\[ E[X_{c}^{2}(n)] = \lambda_{m}^{2}, \quad \lambda_{m}, \quad \lambda_{m} = \lambda_{m}^{2}, \quad \lambda_{m} = \lambda_{m}^{2}. \]
Here we have dropped the 1st term in (15) since from numerical results it is observed small compared with the 2nd term.

Assuming \( J \) to be small, the terms involving \( J^{2} \) and \( J_{s} \) can be dropped to obtain a good approximation of \( J \), as
\[ J_{s} = \frac{c_{s} \phi_{\text{TDNLMS}}(\alpha)}{\sum_{n_{E}} (1 + \delta)} - \frac{1}{2} \frac{c_{s} \phi_{\text{TDNLMS}}(\alpha)}{\sum_{n_{E}} (1 + \delta)} \]
where \( \delta = (A_{s} + \phi_{\text{TDNLMS}}(\alpha) - 1)/(1 - \phi_{\text{TDNLMS}}(\alpha)) \).

To prevent (22) from being unbounded if the denominator is zero, the following gives an approximated condition on the maximum nominal step-size for mean squares convergence
\[ \alpha < \frac{2 - \phi_{\text{TDNLMS}}(\alpha)}{2} \approx \mu_{\text{max}}, \]
\[ \text{(C) Switching Threshold and Parameter Selection} \]
1) Selection of \( T \): From (22), the steady-state EMSE at a fixed step-size is lower bounded by \( \mu_{\text{max}} \phi_{\text{TDNLMS}}(\alpha) \).

Based on (11) and (18), \[ \text{var}[\tilde{X}(\infty)] = \text{E}[\tilde{X}(\infty)^{2}] = (1 - \phi_{\text{TDNLMS}}(\alpha)) \mu_{\text{max}} c_{r} + 1 \]
\[ \text{for } \alpha = 0 \text{ and } \phi_{\text{TDNLMS}}. \]

Assuming \( \tilde{X}(\infty) \) is Gaussian distributed, \( T \) can be chosen as \( \kappa \text{ } \sigma_{v}^{2} \) upper bound of \( \tilde{X}(\infty), \) i.e., \( T = (\mu_{\text{max}} c_{r} \text{ } \phi_{\text{TDNLMS}} + \kappa \text{ } \sigma_{v}^{2}) \).

If \( \mu_{\text{max}} \) as in (16) is used, then \( \mu_{\text{max}} c_{r} = 2 \). On the other hand, \( \kappa \) can be adjusted.
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 of $\alpha$ and $(1+\delta)$ 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 $\delta$ in (22). If $\sigma^2_i$ is not exactly known, we recommend to use the upper bound of $\sigma^2_{\text{max}}$ 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 $v_i(n) = \sigma_i + \alpha X_{i\infty}(n)$, where $\sigma_i$ is the input power and $\alpha = 0.1$. The results for the estimated power in place of $\sigma^2_{\text{max}}$ 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 III. The maximum step-size is $\mu_{\text{max}} = 0.13$ and $\kappa = 4$. Since the TD algorithms are usually employed when the input is colored, the first order autoregressive 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 validness 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.