A Stochastic Total Least Squares Solution of Adaptive Filtering Problem

Shazia Javed¹,² and Noor Atinah Ahmad¹

¹ School of Mathematical Science, Universiti Sains Malaysia, 11800 Penang, Malaysia
² Lahore College for Women University, Lahore 54000, Pakistan

Correspondence should be addressed to Shazia Javed; shaziafateh@hotmail.com

Received 30 August 2013; Accepted 5 December 2013; Published 3 February 2014

1. Introduction

Ordinary least squares methods are extensively used in many signal processing applications to extract the system parameters from input/output data [1, 2]. These methods yield an unbiased solution of adaptive least squares problem having no interference in both inputs and outputs or having interference only in the outputs of the unknown system and clean inputs. However, if interference exists in both input and output of the unknown system or adaptive filtering problem, the ordinary least squares solution gets biased [3].

Total least squares (TLS) method [4] is an efficient technique to achieve an unbiased estimate of the system parameters when both input and output are contaminated by noise. Golub and Van Loan [5] provided an analytical procedure to get an unbiased solution of the TLS problem using singular value decomposition (SVD) of data matrices. This technique is extensively used in data processing and control applications [4, 6, 7]. However, application of TLS methods in signal processing is still limited because computation of SVD requires a high complexity of $O(N^3)$ for an $N \times N$ matrix.

TLS solutions of adaptive filtering problem gained importance after the pioneer work done by Pisarenko [8]. He presented an efficient solution of adaptive TLS problem by adaptively computing the eigenvector corresponding to smallest eigenvalue of augmented input/output signal's autocorrelation matrix. Since then, several algorithms have been proposed based on the adaptive implementations of Pisarenko. The adaptive TLS algorithms proposed in [9–11] are able to achieve an unbiased TLS solution of adaptive filtering problem with a complexity of $O(N)$. However they are sensitive to the correlation properties of input signals and have a drawback of bad performance under correlated inputs.

In this paper, an iterative algorithm is presented to find an optimal TLS solution of adaptive FIR filtering problem. A stochastic technique similar to that of least mean squares (LMS) algorithm of adaptive least squares filtering is employed to develop a total least mean squares (TLMS) algorithm for adaptive total least squares problem. Instead of basing the approach on the minimum mean squares error as the LMS algorithm does, the proposed (TLMS) algorithm is based on the total mean squares, obtained by minimizing the weighted cost function for the TLS solution of adaptive filtering problem. The proposed algorithm has maintained the $O(N)$ complexity of adaptive TLS algorithms with an
2. Mathematical Formulation of Adaptive Total Least Squares Problem

Consider an unknown system to be identified by adaptive FIR filter of length \( N \) and response vector \( \mathbf{w}_n \) at time \( n \), with an assumption that both input and output are corrupted by an additive white Gaussian noise (AWGN). The noise free input vector \( \mathbf{a}_n \in \mathbb{R}^N \) is formed from the input signals \( u(n) \), such that

\[
\mathbf{a}_n = [u(n), u(n - 1), \ldots, u(n - N + 1)]^T \in \mathbb{R}^N. \tag{1}
\]

The desired output of the unknown system is then given by

\[
\tilde{s}(n) = s(n) + \Delta s(n), \tag{2}
\]

where \( s(n) = \mathbf{w}_n^T \mathbf{a}_n \) is system’s output and \( \Delta s(n) \) an added white Gaussian noise of zero mean and variance \( \sigma^2_{\Delta s} \).

The primary assumption of an adaptive least squares (ALS) problem is that perturbations occur in the output signals only and that the input signals are exactly known. This assumption is not practical because, perturbations due to sampling or modeling or measurement errors affect the input signals too. A sensible choice to overcome such situations is to introduce perturbations in input signals in addition to perturbations of output signals. A schematic diagram of an adaptive filter with perturbed input is depicted in Figure 1.

If \( \Delta \mathbf{a}_n = [\Delta u(n), \Delta u(n - 1), \ldots, \Delta u(n - N + 1)]^T \in \mathbb{R}^N \), denote the perturbations in input vector \( \mathbf{a}_n \), where \( \Delta u(n) \) is an additive white Gaussian noise (uncorrelated from the output noise) of zero mean and variance \( \sigma^2_{\Delta u} \), then noisy input vector is

\[
\tilde{\mathbf{a}}_n = \mathbf{a}_n + \Delta \mathbf{a}_n. \tag{3}
\]

An alternate form of \( e(n) \), in terms of augmented data vector of Definition 1, is obtained as follows:

\[
e(n) = y(n) - \tilde{s}(n) = \mathbf{w}_n^T \tilde{\mathbf{a}}_n - \tilde{s}(n) = [\mathbf{w}_n^T - 1] (\tilde{\mathbf{a}}_n - \tilde{s}(n)) = [\mathbf{w}_n^T - 1] \tilde{\mathbf{z}}_n,
\]

where \( \tilde{\mathbf{z}}_n = [\mathbf{w}_n^T - 1] \tilde{\mathbf{z}}_n \) denote the \( (N + 1) \times 1 \) extended parameter vector.

The TLS solution of adaptive filtering problem is an eigenvector associated with the smallest eigenvalue of extended autocorrelation matrix \( \tilde{\mathbf{R}}_n \):

\[
\tilde{\mathbf{R}}_n = E\left\{ \tilde{\mathbf{z}}_n \tilde{\mathbf{z}}_n^T \right\} = \begin{pmatrix} \mathbf{R}_n & \mathbf{P}_{\mathbf{a}_n} \\ \mathbf{P}_{\mathbf{a}_n}^T & \sigma^2_{\tilde{s}(n)} \end{pmatrix}, \tag{6}
\]

Instead of minimizing the mean square error \( E[e^2(n)] \), adaptive total least squares problem is concerned with minimizing the total mean square error \( E[\eta^2(n)] \) and cost function \( J(\tilde{\mathbf{w}}_n) = E[e^2(n)] \), where the total error \( \eta(n) \) is given by

\[
\eta(n) = \frac{e(n)}{\sqrt{\mathbf{w}_n^T \mathbf{w}_n}} = \frac{\tilde{\mathbf{z}}_n^T \tilde{\mathbf{w}}_n}{\sqrt{\mathbf{w}_n^T \mathbf{w}_n}} = \frac{\tilde{s}(n)}{\sqrt{\mathbf{w}_n^T \mathbf{w}_n}}. \tag{7}
\]

The TLS cost function \( J(\tilde{\mathbf{w}}_n) \) is then defined in terms of total error as

\[
\bar{J} = E\left\{ \eta^2(n) \right\} = E\left\{ \frac{\tilde{s}(n)^2}{\mathbf{w}_n^T \mathbf{w}_n} \right\} = \frac{\mathbf{R}_n}{\mathbf{w}_n^T \mathbf{w}_n} = \frac{\tilde{\mathbf{R}}_n}{\mathbf{w}_n^T \mathbf{w}_n}. \tag{8}
\]
The adaptive total least squares problem is a minimization problem of the form [10, 11]:

$$\min_{w \in \mathbb{R}^{N+1}} \overline{J} = \min_{w \in \mathbb{R}^{N+1}} \overline{\mathbf{r}}_n^T \overline{\mathbf{r}}_n.$$  (9)

Note that an optimal solution $w_{\text{opt}}$ of the TLS problem (9) is an eigenvector corresponding to the smallest eigenvalue of $\overline{\mathbf{R}}_n$. In practice SVD technique is used to solve TLS problems since it offers lower sensitivity to the computational errors; however, it is computationally expensive [5]. An alternate choice to estimate eigenvector corresponding to smallest eigenvalue is to use an adaptive algorithm [1, 2].

### 3. Derivation of Total LMS Algorithm for Adaptive Filtering Problem

In adaptive least squares problem, conventional LMS algorithm is a steepest descent method which uses an instantaneous cost function $J = \overline{e}^2(n)$ for computation of gradient vector [1]. Using a similar implementation in TLS problem, the total LMS (TLMS) algorithm is obtained by having an instantaneous estimate of the cost function (8) as $\overline{J} = \eta^2(n)$. The recursive update equation of TLMS algorithm is then given as

$$\overline{w}_{n+1} = \overline{w}_n - \mu \nabla_{\overline{w}} \overline{J},$$  (10)

where $\mu$ is the stepsize parameter or convergence parameter. Note that

$$\nabla_{\overline{w}} \overline{J} = \frac{\partial}{\partial \overline{w}_n} \left( \eta(n)^2 \right)$$

$$= 2\eta(n) \frac{\partial}{\partial \overline{w}_n} \eta(n)$$

$$= 2\eta(n) \frac{\partial}{\partial \overline{w}_n} \left( \frac{\overline{w}_n^T \overline{z}_n}{\sqrt{\overline{w}_n^T \overline{w}_n}} \right)$$

$$= 2\eta(n) \left[ \frac{\overline{z}_n^T \overline{w}_n^T - \overline{w}_n^T \overline{z}_n \cdot \overline{w}_n}{\sqrt{\overline{w}_n^T \overline{w}_n}} \right]$$

$$= \frac{2\eta(n)}{(\overline{w}_n^T \overline{w}_n)^{3/2}} \left\{ \frac{\overline{z}_n \cdot \overline{w}_n^T \overline{w}_n - \overline{w}_n \cdot \overline{z}_n \cdot \overline{w}_n}{\sqrt{\overline{w}_n^T \overline{w}_n}} \right\}.$$  (11)

Using $e(n) = \overline{z}_n^T \overline{w}_n = \overline{w}_n^T \overline{z}_n$ and $\| \overline{w}_n \| = \sqrt{\overline{w}_n^T \overline{w}_n}$, then above equation becomes

$$\nabla_{\overline{w}} \overline{J} = \frac{2}{\| \overline{w}_n \|^4} \left\{ \| \overline{w}_n \|^2 \cdot \overline{z}_n^T \overline{z}_n - e(n)^2 \cdot \overline{w}_n \right\}.$$  (12)

Substituting (12) in (10), the updated equation of TLMS algorithm becomes

$$\overline{w}_{n+1} = \overline{w}_n + \frac{2e(n) \cdot \mu}{\| \overline{w}_n \|^4} \left\{ e(n) \cdot \overline{w}_n - \| \overline{w}_n \|^2 \overline{z}_n \right\}.$$  (13)

Once $\overline{w}_{n+1}$ is computed using (13), the TLS solution update $\overline{w}_{n+1}$ is obtained by the following formula:

$$\overline{w}_{n+1} = \frac{\overline{w}_{n+1}(1 : N)}{\overline{w}_{n+1}(N + 1)}.$$  (14)

The detailed TLMS algorithm is summarized in Table I. A complexity measure of the algorithms shows that it is a computationally linear algorithm, requiring a total of $6N + 9$ multiplications/divisions per iteration. This computational simplicity of adaptive TLS algorithm makes it a better choice than computationally expensive SVD based TLS algorithm, which requires $6N^3$ computations per iteration [5].

### 3.1. Convergence Analysis

In (13), inner product with $\overline{z}_n$ yields,

$$\overline{w}_{n+1}^T \overline{z}_n = \left( \overline{w}_n + \frac{2e(n) \cdot \mu}{\| \overline{w}_n \|^4} \left\{ e(n) \cdot \overline{w}_n - \| \overline{w}_n \|^2 \overline{z}_n \right\} \right)^T \overline{z}_n$$

$$= \overline{w}_n^T \overline{z}_n + \frac{2e(n) \cdot \mu}{\| \overline{w}_n \|^4} \left\{ e(n) \cdot \overline{w}_n^T \overline{z}_n - \| \overline{w}_n \|^2 \overline{z}_n^T \overline{z}_n \right\}$$

$$= e(n) + \frac{2e(n) \cdot \mu}{\| \overline{w}_n \|^4} \left\{ |e(n)|^2 - \| \overline{z}_n \|^2 \right\}.$$  (15)
Since $e(n) = \bar{w}_n^T \bar{z}_n$, Cauchy-Schwarz inequality [12] gives

$$|e(n)|^2 \leq \|\bar{w}_n\|^2 \|\bar{z}_n\|^2;$$

that is,

$$\left| e(n) \right|^2 \leq \|\bar{z}_n\|^2,$$

$$\left| \eta(n) \right|^2 \leq \|\bar{z}_n\|^2,$$

$$\left| \eta(n) \right|^2 - \|\bar{z}_n\|^2 \leq 0.$$  \hspace{1cm} (16, 17)

Let $\delta_n = \|\bar{z}_n\|^2 - |\eta(n)|^2 \geq 0$ then (15) becomes:

$$\bar{w}_{n+1}^T \bar{z}_n = e(n) + \frac{2e(n) \cdot \mu}{\|\bar{w}_n\|^2} \{-\delta_n\}$$

$$= \left\{1 - \frac{2\mu\delta_n}{\|\bar{w}_n\|^2}\right\} e(n)$$

or,

$$\bar{w}_{n+1}^T \bar{z}_n = \left\{1 - \frac{2\mu\delta_n}{\|\bar{w}_n\|^2}\right\} \bar{w}_n^T \bar{z}_n.$$  \hspace{1cm} (18, 19)
Since \( z_n \neq 0 \),

\[
\tilde{w}_{n+1} = \left( 1 - \frac{2\mu \delta_n}{\|\tilde{w}_n\|^2} \right) \tilde{w}_n
\]  

which shows that \( \{\tilde{w}_n\} \) is a geometric progression. It would converge to an optimal solution if

\[
1 - \frac{2\mu \delta_n}{\|\tilde{w}_n\|^2} < 1
\]

or

\[
0 < \frac{\mu}{\|\tilde{w}_n\|^2} < \frac{1}{\delta_n}.
\]

This shows that the proposed algorithm is a variable stepsize algorithm, with \( \bar{\mu} = \mu/\|\tilde{w}_n\|^2 \). An appropriate way to choose \( \mu \) is to initialize the algorithm such that \( \|w_o - w_{opt}\| \) is less than \( 2\|w_{opt}\| \) [13]. According to this result for \( w_o = 0, \|\tilde{w}_n\| = 1 \) and \( \bar{\mu} = \mu \), while \( \delta_n = \|\tilde{z}_n\|^2 - \tilde{s}(n) > 0 \).

4. Application of TLMS Algorithm in System Identification

To examine the performance of proposed TLMS algorithm, an unknown system identification model, shown in Figure 2, is used.

A white Gaussian input signal of variance \( \sigma^2 = 1 \) is passed through a coloring filter with frequency response [1]:

\[
H(z) = \frac{\sqrt{1 - \alpha^2}}{1 - \alpha z^{-1}},
\]

where \( |\alpha| < 1 \), \( \alpha \) is a correlation parameter and controls the eigenvalue spread of input signals. \( \alpha = 0 \) corresponds to the case when eigenvalue spread of input signals is close to 1, and eigenvalue spread increases with an increase in the value of \( \alpha \).

A white Gaussian noise of SNR = 30 dB is added in the input signal \( u(n) \) to get noisy signal \( \tilde{u}(n) \), and an output signal \( \tilde{s}(n) \) is obtained by corrupting the output signal \( s(n) \) with an additive white Gaussian noise of SNR 30 dB. Proposed TLMS algorithm is compared with LMS and NLMS algorithms of [1] to get an FIR vector for a filter of length \( N = 10 \). Least squares misalignment \( \|w_n - w_{opt}\| \) is compared with the total least squares misalignment \( \|w_n - w_{opt}\|/\sqrt{\|w_n\|\|w_n\|} \), and simulations results are recorded for 2000 iterations with an ensemble average of 1000 independent runs.

4.1. Convergence Behavior Corresponding to Stepsize Parameter \( \mu \). Although TLMS algorithm converges for all values of \( \mu \), satisfying (22), but steady state convergence TLMS algorithm is observed when stepsize parameter \( \mu \) is a power of 2. In Figures 3(b)–3(d), four learning curves of total misalignment of TLMS algorithm are shown, corresponding to \( \mu = 0.5, 0.25, 0.125 \), and 0.0625, and it is observed that robustness increases uniformly with an increase in the value of \( \mu \). On the other hand if \( \mu \) is chosen randomly, then a change in the convergence behavior is random, though Figure 3(a) shows that algorithm still converges.

4.2. Convergence Behavior Corresponding to Correlation Parameter \( \alpha \). To check effect of changes in correlation parameter \( \alpha \) on the steady state convergence behavior of TLMS algorithm, different simulations are presented in Figure 3, each showing four learning curves of total misalignment of TLMS algorithm corresponding to \( \mu = 0.5, 0.25, 0.125 \), and 0.0625. In the first two simulations \( \alpha = 0.3 \) in Figures 3(a) and 3(b), it is 0.6 in Figure 3(c), and 0.9 in Figure 3(d). It is clear from the results of all these simulation curves that increase in correlation of data signals has not affected the steady state performance of the algorithm. Although the convergence speed seems to slow down, but all the curves converge to optimal solution.

4.3. Comparison. Figure 4 shows the comparison of misalignment of three algorithms, that is, LMS, NLMS, and TLMS algorithms. The first two compute a least squares solution of adaptive total least squares problem, while the third one computes TLS solution of adaptive total least squares problem. Taking \( \alpha = 0.3 \), stepsize parameter for LMS algorithm is chosen as 0.015, for NLMS algorithm as 0.3, and for TLMS algorithm, it is 0.25. The results in this simulation show that the convergence of TLMS algorithm increases with an increase in the iteration, and it presents a better solution of adaptive TLS problem.

5. Conclusion

In this paper, an efficient TLMS algorithm is presented for the total least squares solution of adaptive filtering problem. The proposed algorithm is derived by using cost function of
weighted instantaneous error signals and an efficient computation of misalignment in terms of mean squares deviation. TLMS algorithm has better ability to tackle with perturbations of both input and output signals, because it is chiefly derived for the purpose. Since in real life problems, both input and output signals are contaminated by noise, therefore TLMS algorithm has great applicability. Convergence analysis shows that the proposed algorithm has global convergence, provided that the stepsize parameter is chosen appropriately. Furthermore, it is computationally simple and requires only O(N) complexity, while other algorithms for TLS problems either require higher complexity or are sensitive to correlation properties of data signals.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors are grateful to the anonymous referees for their valuable comments, appreciation, and recommendation to publish this paper. This work is supported by School of Mathematical Sciences, Universiti Sains Malaysia.

References

[1] B. Farhang-Boroujeny, *Adaptive Filters: Theory and Applications*, John Wiley & Sons, New York, NY, USA, 1998.

[2] N. A. Ahmad, “Comparative study of iterative search method for adaptive filtering problems,” in *Proceedings of the International Conference On Applied Mathematics (ICAM ’05)*, Bedong, Indonesia, August 2005.

[3] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, 2nd edition, 1991.

[4] S. Van Huffel and J. Vandewalle, *The Total Least Squares Problem: Computational Aspects and Analysis*, SIAM, Philadelphia, Pa, USA, 1991.

[5] G. H. Golub and C. F. Van Loan, *An Analysis of Total Least Squares Problem*, SIAM, Philadelphia, Pa, USA, 1980.

[6] M. A. Rahman and K.-B. Yu, “Total least squares approach for frequency estimation using linear prediction,” *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 35, no. 10, pp. 1440–1454, 1987.

[7] Y. Hua and T. K. Sarkar, “On the total least squares linear prediction method for frequency estimation,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 12, pp. 2186–2189, 1990.

[8] V. F. Pisarenko, “The retrieval of harmonics from a covariance function,” *Geophysical Journal of the Royal Astronomical Society*, vol. 33, no. 3, pp. 347–366, 1973.

[9] P. A. Thompson, “Adaptive spectral analysis technique for unbiased frequency estimation in the presence white noise,” *Proceedings of the 13th Asilomar Conference on Circuits, Systems & Computers*, pp. 529–533, 1980.

[10] D.-Z. Feng, “Efficient recursive total least squares algorithm for FIR adaptive filtering,” *IEEE Transactions on Signal Processing*, vol. 42, no. 2, pp. 268–280, 1994.

[11] D.-Z. Feng, X.-D. Zhang, D.-X. Chang, and W. X. Zheng, “A fast recursive total least squares algorithm for adaptive FIR filtering,” *IEEE Transactions on Signal Processing*, vol. 52, no. 10, pp. 2729–2737, 2004.

[12] E. Kreyszig, *Introductory Functional Analysis with Applications*, vol. 1, John Wiley & Sons, New York, NY, USA, 1989.

[13] B. E. Dunne and G. A. Williamson, “Stable simplified gradient algorithms for total least squares filtering,” in *Proceedings of the 34th Asilomar Conference on Signals, Systems, Computers*, pp. 1762–1766, November 2000.