International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:6, No:12, 2012

# Minimal Residual Method for Adaptive Filtering with Finite Termination

Noor Atinah Ahmad and Shazia Javed

*Abstract*—We present a discussion of three adaptive filtering algorithms well known for their one-step termination property, in terms of their relationship with the minimal residual method. These algorithms are the normalized least mean square (NLMS), Affine Projection algorithm (APA) and the recursive least squares algorithm (RLS). The NLMS is shown to be a result of the orthogonality condition imposed on the instantaneous approximation of the Wiener equation, while APA and RLS algorithm result from orthogonality condition in multi-dimensional minimal residual formulation. Further analysis of the minimal residual formulation for the RLS leads to a triangular system which also possesses the one-step termination property (in exact arithmetic)

*Keywords*—Adaptive filtering, minimal residual method, projection method.

## I. Introduction

ADAPTIVE filtering is commonly formulated as a stochastic adaptive least squares problem which may be solved by adaptive counterparts of standard methods for least squares problem. For standard least squares problem, solution methods may be classified into two categories: direct method and iterative method. Direct methods are able to give more accurate solution and fast convergence in general but require higher computational complexity. On the other hand, iterative methods provide a more robust environment for adaptive implementation while requiring lower computational complexity.

Standard iterative methods have been modified for application in adaptive filtering [1], [2], [3], [4], [5], [6], [7] . For example, the Least Mean Square algorithm (LMS) can be seen as an adaptive counterpart of the steepest descent method with stochastic estimation of gradient [1]. The LMS algorithm is widely used due to its simplicity, robustness and its low computational complexity. However it suffers from slow convergence rate for colored input signal such as speech. The normalized least mean square (NLMS), Affine Projection algorithm (APA) and the recursive least squares algorithm are well known adaptive filtering algorithms which have received a lot of attention due to their superior convergence compared to the LMS. The superior convergence is due to the one-step termination property shared by these algorithms and their close proximity to the Newton iteration [8], [9]. Although they are implemented as iterative methods, the one-step termination property renders the algorithms to be comparable to direct methods.

In this paper, we analyze NLMS, APA and RLS and present their relationship with the minimal residual method. After giving brief description of minimal residual method in section II, , the relationship between these algorithms with the one dimensional and multidimensional minimal residual method is presented in section III. In section IV, we show the equivalence of the RLS method to a triangular system which is solvable via forward or backward substitution. The triangular system is also shown to provide one-step termination.

Noor Atinah Ahmad is with the School of Mathematical Sciences, Universiti Sains Malaysia, 11800 Penang, Malaysia. (Email: atinah@cs.usm.my)
Shazia Javed is with the School of Mathematical Sciences, Universiti Sains Malaysia, 11800 Penang, Malaysia. (Email: shaziafateh@hotmail.com)

## II. The Minimal Residual Method (MR)

The standard MR method for solving linear system of equation $\Phi x = p$ updates the approximated solution along the current residual vector $r^{(k)} = p - \Phi x^{(k)}$, and, the stepsize is chosen so that the residual 2-norm square $\left\| p - \Phi x^{(k+1)} \right\|_2^2$ is minimized. By doing so, an orthogonality condition $r^{(k+1)T}\left(\Phi r^{(k)}\right) = 0$ is imposed where $r^{(k+1)} = p - \Phi x^{(k+1)}$. In general, a one-dimensional MR method with direction of search $d^{(k)}$, the MR update equations take the form,

$$\begin{aligned} \alpha_k &= \tfrac{r^{(k)T} b^{(k)}}{b^{(k)T} b^{(k)}} \\ x^{(k+1)} &= x^{(k)} + \alpha_k d^{(k)} \end{aligned} \tag{1}$$

where $b^{(k)} = \Phi d^{(k)}$. The iteration in (1) is a result of the orthogonality condition

$$r^{(k+1)T} b^{(k)} = \left(p - \Phi x^{(k+1)}\right)^T b^{(k)} = 0$$

i.e., (1) represents an orthogonal projection of

$$r^{(k+1)T} b^{(k)} = \left(p - \Phi x^{(k+1)}\right)^T b^{(k)} = 0$$

onto the subspace containing $b^{(k)}$.

A general framework for multidimensional projection method as described in [10] extracts an approximation to the solution from a subspace of $R^N$ which is spanned by $P$ linearly independent directions (with $P \le N$) so that

$$\begin{aligned} x^{(k+1)} &= x^{(k)} + \alpha_1^{(k)} d_1^{(k)} + \alpha_2^{(k)} d_2^{(k)} + \cdots + \alpha_P^{(k)} d_P^{(k)} \\ &= x^{(k)} + V^{(k)} y^{(k)} \end{aligned} \tag{2}$$

where

$$V_P^{(k)} = \begin{pmatrix} | & | & & | \\ d_1^{(k)} & d_2^{(k)} & \cdots & d_P^{(k)} \\ | & | & & | \end{pmatrix}, \quad y^{(k)} = \begin{pmatrix} \alpha_1^{(k)} \\ \alpha_2^{(k)} \\ \vdots \\ \alpha_P^{(k)} \end{pmatrix}$$

Each of the stepsizes

$$\alpha_i^{(k)} = \frac{r^{(k)T}\left(\Phi d_i^{(k)}\right)}{\left(\Phi d_i^{(k)}\right)^T \left(\Phi d_i^{(k)}\right)}$$

International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:6, No:12, 2012

for $i = 1, \ldots, P$, satisfies the orthogonality condition

$$r^{(k+1)T} b_i^{(k)} = \left(p - \Phi x^{(k+1)}\right)^T b_i^{(k)} = 0$$

where $b_i^{(k)} = \Phi d_i^{(k)}$, therefore, minimizing the residual 2-norm along the direction $d_i^{(k)}$respectively.

## III. Minimal Residual Properties of Several Existing Adaptive Filtering Algorithm

Consider a typical setup for a transversal finite impulse response (FIR) adaptive filter with adjustable coefficient vector $x \in R^N$ where $N$ is the filter order. The filter output at the $n$th instant is $y(n) = a^{(n)T} x$, where vectors $a^{(n)} = \left[\begin{array}{cccc} u(n) & u(n-1) & \cdots & u(n-N+1) \end{array}\right]^T$ are formed by the input signal $u(n)$.

We seek an estimate of the filter coefficient vector such that the output signal is a good estimate of the desired (measured) signal $s(n)$. The mean-squared error (MSE) is a common criterion for determining how good $y(n)$ approximates $d(n)$ and by this criterion, the optimal solution is the solution of the Wiener-Hopf equation

$$\Phi x = p$$

where $\Phi = E\left\{a^{(n)} a^{(n)T}\right\}$ is the autocorrelation of the filter input signal ($E\{\cdot\}$ denotes the expectation operator), and, $p = E\left\{a^{(n)} s(n)\right\}$ is the cross correlation vector.

We shall now discuss the minimal residual properties of several adaptive filtering algorithm for estimating the Wiener solution.

### *A. The Normalized Least Mean Square (NLMS) Algorithm as a One-Dimensional Minimal Residual Method*

The Normalized Least Mean Square (NLMS) update equation is given by

$$x^{(k+1)} = x^{(k)} + \mu \frac{1}{\varepsilon + \left\|a^{(k)}\right\|_2^2} a^{(k)} e^{(k)} \quad (3)$$

where $e^{(k)} = d(k) - a^{(k)T} x^{(k)}$ is the instantaneous error at the $k$th instant. This equation is obtained by updating the approximate solution along the direction of the instantaneous residual vector $a^{(k)} e^{(k)}$ with a normalized stepsize of $\frac{\mu}{\varepsilon + \left\|a^{(k)}\right\|_2^2}$, and

$$b^{(k)} = \Phi^{(k)} d^{(k)} = (a^{(k)} a^{(k)T}) a^{(k)} e^{(k)} = \left\|a^{(k)}\right\|_2^2 a^{(k)} e^{(k)}$$

It is straightforward to show that, when $\mu = 1$ and $\varepsilon = 0$,

$$\left(p^{(k)} - \Phi^{(k)} x^{(k+1)}\right)^T b^{(k)} = 0$$

which is the orthogonality condition that minimizes the instantaneous residual 2-norm

$$\left\|a^{(k)} e^{(k)}\right\|_2^2 = \left\|a^{(k)} s(k) - \left(a^{(k)} a^{(k)T}\right) x\right\|_2^2$$

Another interesting observation is that, when $\mu = 1$ and $\varepsilon = 0$, $x^{(k+1)}$ is the exact solution to the instantaneous normal equation $\left(a^{(k)} a^{(k)T}\right) x = a^{(k)} s(k)$. Moreover,

$$a^{(k)} \left(a^{(k)T} x^{(k+1)} - s(k)\right) = a^{(k)} \tilde{e}^{(k)} = 0 \quad (4)$$

where $\tilde{e}^{(k)}$ is the aposteriori error. Thus from (4), it is implied that the NLMS algorithm forces the aposteriori error to zero (one step termination). Of course, in finite arithmetic environment, this will not be achieved, so the parameter $\mu$ will serve as an acceleration parameter to control speed of convergence (a regularization parameter $\varepsilon$ is also required for increased stability).

### *B. The Affine Projection Algorithm (APA) as a Multi-Dimensional Minimal Residual Method*

The Affine Projection algorithm (APA) in its standard form updates the coefficient vector based on $P$previous input vectors such that

$$\begin{aligned} x^{(k+1)} &= x^{(k)} + \alpha_1^{(k)} a^{(k)} + \alpha_2^{(k)} a^{(k-1)} + \cdots + \alpha_P^{(k)} a^{(k-P+1)} \\ &= x^{(k)} + V^{(k)} y^{(k)} \end{aligned} \quad (5)$$

where

$$V^{(k)} = \begin{pmatrix} | & | & & | \\ a^{(k)} & a^{(k-1)} & \cdots & a^{(k-P+1)} \\ | & | & & | \end{pmatrix}$$

$$y^{(k)} = \begin{pmatrix} \alpha_1^{(k)} \\ \alpha_2^{(k)} \\ \vdots \\ \alpha_P^{(k)} \end{pmatrix}$$

To minimize the residual 2-norm square

$$\left\|r^{(k)}\right\|_2^2 = \left\|p_P^{(k)} - \Phi_P^{(k)} x^{(k)}\right\|_2^2$$

where

$$p_P^{(k)} = V^{(k)} s^{(k)}, \quad \Phi_P^{(k)} = V^{(k)} V^{(k)T}$$

$$s^{(k)} = \left(s(k), s(k-1), \ldots, s(k-P+1)\right)^T,$$

the following (multidimensional) orthogonality condition is imposed,

$$\left(\tilde{r}^{(k)}\right)^T \Phi_P^{(k)} V^{(k)} = 0 \quad (6)$$

where $\tilde{r}^{(k)} = p_P^{(k)} - \Phi_P^{(k)} x^{(k+1)} = r^{(k)} - \Phi_P^{(k)} V^{(k)} y^{(k)}$. Condition (6) insists that $r^{(k+1)}$ is orthogonal to the span of $\left\{\Phi_P^{(k)} a^{(k)}, \ldots, \Phi_P^{(k)} a^{(k-P+1)}\right\}$.

By using the fact that $r^{(k)} = V^{(k)}\left(s^{(k)} - V^{(k)T} x^{(k)}\right) = V^{(k)} e^{(k)}$, followed by some simplification, Eqn. (6) leads to

$$\begin{aligned} V^{(k)} y^{(k)} &= \left(\Phi_P^{(k)2}\right)^{-1} \Phi_P^{(k)} r^{(k)} \\ &= \left(\Phi_P^{(k)2}\right)^{-1} \Phi_P^{(k)} V^{(k)} e^{(k)} \end{aligned} \quad (7)$$

Substituting for $V^{(k)} y^{(k)}$ in (5) gives rise to

$$\begin{aligned} x^{(k+1)} &= x^{(k)} + \left(\Phi_P^{(k)2}\right)^{-1} \Phi_P^{(k)} V^{(k)} e^{(k)} \\ &= x^{(k)} + \left(V^{(k)} V^{(k)T}\right)^{-1} V^{(k)} e^{(k)} \end{aligned}$$

The more common form of the APA coefficient vector update equation is

$$x^{(k+1)} = x^{(k)} + \mu \left(V^{(k)} V^{(k)T} + \varepsilon I\right)^{-1} V^{(k)} e^{(k)}$$

where $\mu$ and $\varepsilon$ are the acceleration and regularization parameters respectively, for controlling rate of convergence and stability in finite arithmetic environment.

### *C. Recursive Least Squares (RLS) Method with Refinement as a N-dimensional Minimal Residual Method with Euclidean Unit Vectors as Direction of Search*

Consider the $N$ dimensional minimal residual method for minimizing $\left\|p - \Phi x^{(k+1)}\right\|_2^2$, where the directions of search are the Euclidean unit vectors

$$\hat{e}_i = [0, \ldots, 1, \ldots, 0]^T, i = 1, \ldots, N$$

with 1 appearing in the $i$th place. Thus

$$x^{(k+1)} = x^{(k)} + \sum_{i=1}^{N} \alpha_i^{(k)} \hat{e}_i = x^{(k)} + y^{(k)}$$

The stepsize vector $y^{(k)}$ is calculated by insisting that $\left(p - \Phi x^{(k+1)}\right)^T \Phi y^{(k)} = \left(r^{(k)} - \Phi y^{(k)}\right)^T \Phi y^{(k)} = 0$, (the orthogonality condition). For a symmetric $n \times n$ matrix $\Phi$, the procedure above gives rise to an auxiliary system of linear equations in $y^{(k)}$ of the form $\Phi^2 y^{(k)} = \Phi r^{(k)}$.

We shall now apply the procedure above to the exponentially weighted least squares problem which is the objective function used for the Recursive Least Squares (RLS) method, i.e.,

$$\min_{x \in R^N} J_n(x) = \sum_{i=1}^{n} \lambda^{n-i} \left(a^{(i)T} x - s(i)\right)^2 \tag{8}$$

where the constant $\lambda \in [0,1]$ is the forgetting factor. Note that the solution to the minimization problem in (8) is also a solution to the following normal system,

$$\Phi^{(n)} x = p^{(n)} \tag{9}$$

where

$$\Phi^{(n)} = \sum_{i=1}^{n} \lambda^{n-i} a^{(i)} a^{(i)T}$$

and

$$p^{(n)} = \sum_{i=1}^{n} \lambda^{n-i} a^{(i)} s(i)$$

Thus, applying the $N$ dimensional minimal residual method with Euclidean unit vectors as the search directions on the normal equation (9) at the $k$th instant, leads to the auxiliary problem

$$\Phi^{(k)2} y = \Phi^{(k)} r^{(k)} \tag{10}$$

where $r^{(k)} = p^{(k)} - \Phi^{(k)} x^{(k)}$. Eqn. (10) has the exact solution $y^{(k)} = \Phi^{-1} r^{(k)}$ ( which is unique when $\Phi^{(k)2}$ is nonsingular). In other words,

$$\tilde{r}^{(k)} = p^{(k)} - \Phi\left(x^{(k)} + \Phi^{-1} r^{(k)}\right) = r^{(k)} - r^{(k)} = 0 \tag{11}$$

Now, it is possible to update $\Phi^{(n)}$ and $p^{(n)}$ recursively [7] through the formulas,

$$\Phi^{(n)} = \lambda \Phi^{(n-1)} + a^{(n)} a^{(n)T} \tag{12}$$

$$p^{(n)} = \lambda p^{(n-1)} + a^{(n)} s(n) \tag{13}$$

By applying these formulas, a recursive update for the residual vector $r^{(k)}$ is also possible which is given by

$$\begin{aligned} r^{(k)} &= p^{(k)} - \Phi^{(k)} x^{(k)} \\ &= \lambda p^{(k-1)} + a^{(k)} s(k) - \left(\lambda \Phi^{(k-1)} + a^{(k)} a^{(k)T}\right) x^{(k)} \\ &= \lambda\left(p^{(k-1)} - \Phi^{(k-1)} x^{(k)}\right) + a^{(k)}\left(s(k) - a^{(k)T} x^{(k)}\right) \\ &= \lambda \tilde{r}^{(k-1)} + a^{(k)} e^{(k)} \end{aligned} \tag{14}$$

where $e^{(k)} = s(k) - a^{(k)T} x^{(k)}$, is the instantaneous apriori error between the desired signal and the adaptive filter output. From (11),

$\tilde{r}^{(k-1)} = p^{(k-1)} - \Phi^{(k-1)} x^{(k)} = 0$, thus Eqn. (14) is reduced to

$$r^{(k)} = a^{(k)} e^{(k)}$$

and,

$$y^{(k)} = \Phi^{(k)-1} a^{(k)} e^{(k)} \tag{15}$$

This results in an update equation for the coefficient vector of the form

$$x^{(k+1)} = x^{(k)} + \Phi^{(k)-1} a^{(k)} e^{(k)} \tag{16}$$

where $\Phi^{(k)-1}$ is updated using the matrix inversion lemma. Equation (16) is the refined RLS method which is commonly used as an alternative update equation for RLS [11], [12].

## IV. Refined RLS Method with Triangularization

In this section, we will show that the equivalence of the auxiliary equation (10) to a triangular system. First, observe that $y^{(k)}$ is chosen so that $\tilde{r}^{(k)T}\left(\Phi^{(k)} y^{(k)}\right) = 0$ which is equivalent to

$$\begin{aligned} r^{(k)T}(\Phi^{(k)} y^{(k)}) - \left(\Phi^{(k)} y^{(k)}\right)^T \left(\Phi^{(k)} y^{(k)}\right) \\ = y^{(k)T}(\Phi^{(k)} r^{(k)} - \Phi^{(k)2} y^{(k)}) = 0 \end{aligned} \tag{17}$$

Consider splitting the matrix $Q^k = \Phi^{(k)2}$ as,

$$Q^{(k)} = \Phi^{(k)2} = L^{(k)} + D^{(k)} + U^{(k)} \tag{18}$$

where $D^{(k)}$ is a diagonal matrix consisting of the main diagonal of $Q^{(k)}$, and, matrices $L^{(k)}$ and $U^{(k)}$ are the strict lower and upper triangular parts of $Q^{(k)}$ respectively. Now, substituting (18) into (17) leads to

$$y^{(k)T}\left(\Phi_k r^{(k)}\right) = y^{(k)T}\left(L^{(k)} y^{(k)} + D^{(k)} y^{(k)} + U^{(k)} y^{(k)}\right) \tag{19}$$

It is straightforward to see that the entries of $L^{(k)} y^{(k)}$ are given by

$$\left[L^{(k)} y^{(k)}\right]_1 = 0$$

$$\left[L^{(k)} y^{(k)}\right]_i = \sum_{j=2}^{i-1} y_j^{(k)}\left(\Phi_i^{(k)T} \Phi_j^{(k)}\right) = \sum_{j=2}^{i-1} y_j^{(k)} Q_{i,j}^{(k)};$$

$$i = 2, \ldots, N$$

where $Q_{i,j}^{(k)}$ is the entry of $Q^{(k)}$ in the $i$th row and $j$th column, and, $\Phi_i^{(k)}$ is the $i$th column of $\Phi^{(k)}$. Thus

$$y^{(k)T}L^{(k)}y^{(k)} = \sum_{i=2}^{N} y_i^{(k)} \sum_{j=1}^{i-1} y_j^{(k)} Q_{i,j}^{(k)}$$

Switching the order of summation leads to,

$$\begin{aligned} y^{(k)T}L^{(k)}y^{(k)} &= \sum_{i=2}^{N} y_i^{(k)} \sum_{j=1}^{i-1} y_j^{(k)} Q_{i,j}^{(k)} \\ &= \sum_{j=1}^{N-1} y_j^{(k)} \sum_{i=j}^{N} y_i^{(k)} Q_{i,j}^{(k)} \\ &= \sum_{j=1}^{N-1} y_j^{(k)} \sum_{i=j}^{N} y_i^{(k)} Q_{j,i}^{(k)} \\ &= y^{(k)T}U^{(k)}y^{(k)} \end{aligned} \tag{20}$$

The last line of (20) is obtained by using the fact that $Q^{(k)} = \Phi^{(k)2}$ is symmetric. Using (20), we are able to write (19) as either,

$$y^T \left(D^{(k)} + 2L^{(k)}\right) y = y^T \Phi^{(k)} r^{(k)} \tag{21}$$

or,

$$y^T \left(D^{(k)} + 2U^{(k)}\right) y = y^T \Phi^{(k)} r^{(k)} \tag{22}$$

It is straightforward to show that solutions to the following triangular systems,

$$\left(D^{(k)} + 2L^{(k)}\right) y = \Phi^{(k)} r^{(k)} \tag{23}$$

and,

$$\left(D^{(k)} + 2U^{(k)}\right) y = \Phi^{(k)} r^{(k)} \tag{24}$$

are also solutions to (21) and (22) which in turn solves (10). Moreover, (23) and (24) can be solved by forward and backward substitution respectively. For example, forward substitution procedure on (23) will provide the exact solution (in exact arithmetic)

$$y^{(k)} = \left(D^{(k)} + 2L^{(k)}\right)^{-1} \Phi^{(k)} r^{(k)}$$

and the coefficient vector update equation can be written in the exact form which is

$$x^{(k+1)} = x^{(k)} + \left(D^{(k)} + 2L^{(k)}\right)^{-1} \Phi^{(k)} r^{(k)} \tag{25}$$

thus providing a one-step solution to the normal equation $\Phi^{(k)}x = p^{(k)}$.

Similarly, backward substitution procedure on (24) will result in the update equation of the form

$$x^{(k+1)} = x^{(k)} + \left(D^{(k)} + 2U^{(k)}\right)^{-1} \Phi^{(k)} r^{(k)} \tag{26}$$

It is worth noting that update equations (25) and (26) will track the unique solution of (9) as long as $\Phi^{(k)2}$ is nonsingular for all $k = 1, 2, \ldots$ .

## V. Conclusion

Our analysis in this paper has highlighted the relationship between NLMS, APA and RLS adaptive filtering algorithms with the minimal residual method. The NLMS is shown to be equivalent to one dimensional minimal residual direction search where the residual is the instantaneous residual vector associated with the instantaneous squared error norm. Extending the search into multiple directions leads to multidimensional minimal residual method. The APA is a result of setting the search directions to $P$ previous input vectors while the RLS algorithm is a result of setting the search directions along the $N$ Euclidean unit vectors, where $N$ is the order of the adaptive filter and $1 \leq P \leq N$. The minimal residual approach has also highlighted the equivalence of the refined RLS method with lower/upper triangular system which can be solved using forward/backward substitution respectively. The conventional RLS method updates the inverse autocorrelation matrix in (16) using the matrix inversion lemma. By representing the RLS normal equation using the triangular systems (23) or (24), direct solution of the normal equation may be obtained without directly computing the inverse of the autocorrelation nor will it require the use of the matrix inversion lemma.

## Acknowledgment

The authors would like to acknowledge the support of Universiti Sains Malaysia for the financial assistance.

## References

[1] B. Farhang-Boroujeny, *Adaptive filters: theory and applications*, John Wiley & Sons, Inc.,1998.
[2] G. K. Boray and M. D. Srinath, *Conjugate gradient techniques for adaptive filtering*, IEEE Trans. Circuits Syst., 1992.
[3] P. S. Chang and A. N. Wilson, Jr., *Analysis of conjugate gradient method for adaptive filtering*, IEEE Trans. Signal Processing, 2000.
[4] M. Q. Chen, *A direction set based algorithm for least squares problems in adaptive signal processing*, Linear Algebra and its Application, 1998.
[5] G. F. Xu, T. Bose and J. Schroeder, *The Euclidean direction search algorithm for adaptive filtering*, Proceedings of the 1999 IEEE Int. Symp. on Circuits and Systems, 1999.
[6] N. A. Ahmad, *Accelerated euclidean direction search algorithm and related relaxation schemes for solving adaptive filtering problem*, Proceedings of IEEE Int. Conf. on Information, Communication and Signal Processing (ICICS 2007), 2007.
[7] N. A. Ahmad, *A globally convergent stochastic pairwise conjugate gradient based adaptive filtering algorithm*, IEEE Signal Process. Letters, 2008.
[8] P. S. R. Diniz, M. L. R. de Campos and A. Antoniou, *Analysis of LMS-Newton Adaptive Filtering Algorithms with Variable Convergence Factor*, IEEE Trans. Signal Processing, 1995.
[9] C. W. Lee, Y. K. Lee and S. W. Kim, *An Affine Projection Algorithm with a Data-Selective Method of using the Condition Number*, Signal Processing 88, 2008.
[10] Y. Saad, *Iterative Methods for Sparse Linear Systems*, Englewood Cliffs, NJ: Prentice-Hall, 1991.
[11] S. Haykin, *Adaptive Filter Theory*, $2^{nd}$ edition, Prentice Hall,1991.
[12] P. S. R. Diniz,*Adaptive Filtering: Algorithms and Practical Implementation*, $3^{rd}$ Edition, Springer, 2008.