A Fast and Low Complexity Adaptive Filters Implementation in Wavelet Packet domain with Partial updating and Selective Regressors

JAFAR RASI, BEHZAD MOZAFFARI TAZEHKAND

Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran

J_rasi@tabrizu.ac.ir, mozaffary@tabrizu.ac.ir

Abstract: Selective regressors and selective coefficient update (partial updating) of affine projection algorithm (APA) is two effective schemes which can reduce the computational load and power consumption in adaptive filter implementations. In addition to selective regressors and selective coefficient update, computational load of APA can be reduced by making use of sub-band processing techniques with applying the set-membership concepts in all sub-bands. In this paper, at first, the new implementation of adaptive algorithms technique will be considered in wavelet packet domain using multiple independent adaptive filters. In the second step, we apply the APA to the proposed structure and incorporate a data selection strategy based on the set-membership concepts with considering reduction of the computational complexity using selective coefficient update and selective regressors in each packets, simultaneously. Also selective coefficient update NLMS Algorithm is used in the proposed structure for further confirmation of proposed algorithm. Simulation experiments confirm the effectiveness of the proposed algorithm in terms of reduced computational complexity and rate of convergence.

Key words: *Selective Regressor, Selective Coefficient Update, Partial Updating, Data-Selective, Affine Projection, MSE Analysis, Set-membership Filtering, Wavelet Packet, Filter Banks.*

1. Introduction.

Adaptive filters have been found many applications in a wide range of diverse fields such as communications, control, radar, sonar, acoustic, and speech processing. The least mean-square (LMS) and the normalized LMS (NLMS) algorithms are the most widely used adaptive-filter algorithms in engineering applications due to their low computational complexity. Unfortunately, the convergence speed of the algorithms deteriorates significantly when the input signals to be color or highly correlated [1]. For highly correlated input signals the affine projection (AP) algorithm offers faster convergence than the LMS and NLMS algorithms. The computational complexity, however, has been a weak point in the implementation of APA. As an attempt to solve this important problem, sub-band processing schemes have been proposed in adaptive filter structures [2-3]. Multirate filtering techniques have been used over the past few years with attention to sub-band decomposition [4]. In sub-band processing, not only computational complexity can be reduced but also the rate of convergence can be improved [4-6].

In [7], we proposed an efficient structure for implementation of adaptive algorithm based on wavelet packet transform. This method increases the initial convergence speed without any more computations, approximately. The basic idea underlying in our proposed structure [7], is to decompose the spectral content of the signal into uniform frequency sub-bands, according to the desired frequency resolution. In the last step we applied several adaptive algorithms to corresponding packets, simultaneously.

A number of partial update algorithms have been proposed to reduce the computational complexity. The partial update algorithms deal with updating a selected subset of the filter coefficients at sequential iterations and hence they can reduce the computational complexity [8 -11]. Because of the great demand placed on expensive real-time resources such as power consumption and memory, computational complexity of adaptation of filter coefficients can become prohibitively expensive [8,11].

Another efficient approach to reduce the computational complexity is applying a set-membership filtering (SMF) approach to adaptive filtering [11]. In addition to these mentioned methods, in order to reduce complexity of APA, a subset of regression inputs based on optimal selection of regressions has been selected in sequential iterations [12].

Besides of wavelet packet implementation advantage, the proposed method in this paper enjoys from two useful methods. First, low computational complexity is obtained with considering the selective coefficient update, selective regressors for APA and second, the sparse updating related to the set-membership framework.

This paper is organized as follows. Section 2 briefly reviews the LMS, NLMS, APA, Set-Membership filtering technique [11], SCU-APA algorithms [13], SCU-NLMS [11], SM-APA method [1, 14], wavelet packet transform [15] and SR-APA [12]. The proposed algorithms named as WPTD-SM-SCU-NLMS, WPTD-SM-SCU-APA and WPTD-SM-SR-APA will be described in section 3. Complexity requirements are represented in section 4. Simulation experiments which illustrate the performance and computational complexity of the proposed method will be shown in section 5. Finally conclusions are given in section 6.

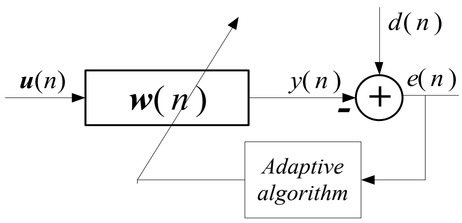

Fig 1. A typical block diagram of adaptive filter

2. Background materials

2.1. LMS Algorithm

Figure (1) shows a block diagram of typical adaptive filter where $\boldsymbol{u}(n)$, $d(n)$ and $e(n)$ are the input, desired and output error signals, respectively. Here, $\boldsymbol{w}(n)$ is the $M\times1$ column vector of the filter coefficients at iteration 'n'. The desired signal is assumed to have the following linear data model:

$$d(n) = \boldsymbol{u}^T(n)\boldsymbol{w}(n) + v(n) \quad (1)$$

Where $\boldsymbol{u}(n)=[u(n), u(n-1), ..., u(n-M+1)]^T$ is the input signal, $v(n)$ is the measurement noise which is assumed to be zero mean, white, Gaussian, and independent of $\boldsymbol{u}(n)$, and $\boldsymbol{w}(n)$ is the unknown filter coefficients.

The least-mean-square (*LMS*) algorithm is the most widely used among various adaptive algorithms because of its simplicity and robustness. The *LMS* algorithm is based on the steepest-descent technique which was proposed by Widrow and Stearns [16] to study the pattern-recognition machine. The *LMS* algorithm updates the weight vector of an adaptive filter as follows:

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) + \mu\boldsymbol{u}(n)e(n) \quad (2)$$

Where μ is the step size of the algorithm which can controls the stability and the convergence rate of the adaptive filter.

2.2. NLMS Algorithm

It is well known that the NLMS algorithm can be derived from the solution of the following optimization problem:

$$J(w) = \|\boldsymbol{w}(n+1) - \boldsymbol{w}(n)\|^2 \quad (3)$$

Subject to

$$d(n) = \boldsymbol{u}^T(n)\boldsymbol{w}(n+1) \quad (4)$$

Using the method of Lagrange multipliers, solving the Optimization problem leads to the following recursion equation:

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) + \frac{\mu}{\|\boldsymbol{u}(n)\|^2}\boldsymbol{u}(n)e(n) \quad (5)$$

Where $e(n)=d(n)-\boldsymbol{u}^T(n)\boldsymbol{w}(n)$ and μ is the step-size that determines the rate of convergence and excess MSE.

2.3. Affine Projection Algorithm (APA)

Suppose that desired signal $d(k)$ can be modeled as following:

$$d(k) = \mathbf{u}^T(k)\boldsymbol{w} + v(k) \quad (6)$$

Where $v(k)$ denotes the measurement noise, $\boldsymbol{w}$ is an unknown column vector and $\boldsymbol{u}(k)$ denotes $M\times1$ column input regression vectors as: $\boldsymbol{u}(k)=[u(k), u(k-1), \dots, u(k-M+1)]^T$

Using the LMS algorithm, anyone can obtain the following recursion formula about APA [12] as:

$$\boldsymbol{w}(k) = \boldsymbol{w}(k-1) + \mu\boldsymbol{U}^*(k)(\boldsymbol{U}(k)\boldsymbol{U}^*(k))^{-1}\boldsymbol{e}(k) \quad (7)$$

where $\boldsymbol{e}(k)=\boldsymbol{d}(k)-\boldsymbol{U}^T(k)\boldsymbol{w}(k-1)$ is error vector and $\boldsymbol{d}(k)=[d(k), d(k-1), \dots, d(k-K+1)]^T$ is the desired vector signal and $\boldsymbol{U}(k)=[\boldsymbol{u}(k), \boldsymbol{u}(k-1), \dots, \boldsymbol{u}(k-K+1)]$ which is a $M\times K$ data matrix.

2.4. Set – Membership Filtering (SM)

In set-membership filtering (SMF) [17], the filter coefficient vector $\boldsymbol{w}$ is designed to achieve a specified bound on the magnitude of the output error. Several valid estimates of $\boldsymbol{w}$ satisfy the chosen bound on the output error at instant k. Let H_k denote the set containing all vectors $\boldsymbol{w}$ for which the associated output error at time instant k is upper bounded in magnitude by γ, i.e.,

$$H_k = \left\{w \in R^N : \left|d_k - \boldsymbol{w}^T\boldsymbol{u}_k\right| \le \gamma\right\} \quad (8)$$

Where H_k is referred to as the constraint set, and its boundaries are hyperplanes. Finally, defining the exact feasibility set as ψ_k to be the intersection of the constraint sets over the time instants such: $i = 1,...,k$, i.e.,

$$\psi_k = \bigcap_{i=1}^{k} H_i \quad (9)$$

The idea of set-membership adaptive recursion techniques (SMART) [17] is to adapt the coefficient vector such that it will always remain within the predefined feasibility set.

2.5. Selective coefficient update APA (SCU-APA)

Let us partition the regression vector $\boldsymbol{u}(k)$ and the assosiated filter coefficients vector $\boldsymbol{w}(k)$ into P blocks whose their length equal $L = M/P$ as the follows:

$$\boldsymbol{w}(k) = \left[\boldsymbol{w}_1^T(k),\ \boldsymbol{w}_2^T(k),\ ...,\boldsymbol{w}_P^T(k)\right]^T$$

$$\mathbf{u}(k) = \left[\mathbf{u}_1^T(k),\ \mathbf{u}_2^T(k),\ ...,\mathbf{u}_P^T(k)\right]^T$$

Also with considering the optimisation problem with multiple constraints as:

$$\min_{1\le i\le P}\ \min_{\boldsymbol{w}_i(k+1)} \left\|\boldsymbol{w}_i(k+1)-\boldsymbol{w}_i(k)\right\|_2^2 \qquad (10)$$

Subject to $\boldsymbol{U}^T(k)\boldsymbol{w}(k+1) = \mathbf{d}(k)$

with the number of constraints $1\le k\le L$.

For each block $i\in\{1,2,...,P\}$, the cost function is

$$J_i(k) = \left\|\boldsymbol{w}_i(k+1)-\boldsymbol{w}_i(k)\right\|_2^2 + \boldsymbol{\lambda}^T(\mathbf{d}(k)-\boldsymbol{U}^T(k)\boldsymbol{w}(k+1))$$

Where $\boldsymbol{\lambda} = [\lambda_1\ \ \lambda_2\ \ ...\ \ \lambda_K]$ is the vector of Lagrange multipliers. Setting $\partial J_i(k)\big/\partial \boldsymbol{w}_i(k+1) = 0$ gives

$$2(\boldsymbol{w}_i(k+1)-\boldsymbol{w}_i(k)) = \boldsymbol{U}_i(k)\boldsymbol{\lambda} \qquad (11)$$

Where

$$\boldsymbol{U}_i(k) = \left[\mathbf{u}_i(k)\quad \mathbf{u}_i(k-1)\quad ...\quad \mathbf{u}_i(k-K+1)\right]_{L\times K}$$

Premultiplying both sides of (11) by $\boldsymbol{U}_i^T(k)$ and assuming that the $K\times K$ matrix $\boldsymbol{U}_i^T(k)\boldsymbol{U}_i(k)$ is full-rank, we can obtain equation (12) as:

$$\boldsymbol{\lambda} = 2(\boldsymbol{U}_i^T(k)\boldsymbol{U}_i(k))^{-1}\boldsymbol{U}_i^T(k)(\boldsymbol{w}_i(k+1)-\boldsymbol{w}_i(k)) \qquad (12)$$

Substituting (12) into (11) and using the constraint as (10), we can get the equation (13) as:

$$\boldsymbol{w}_i(k+1) = \boldsymbol{w}_i(k) + \boldsymbol{U}_i(k)(\boldsymbol{U}_i^T(k)\boldsymbol{U}_i(k))^{-1}\boldsymbol{e}(k) \qquad (13)$$

Where $\boldsymbol{e}(k) = \boldsymbol{U}^T(k)\boldsymbol{w}(k)$ is the $K\times 1$ error vector.

With introducing a small positive stepsize as μ, we obtain the AP algorithm for a fixed block update as:

$$\boldsymbol{w}_i(k+1) = \boldsymbol{w}_i(k) + \mu\boldsymbol{U}_i(k)(\boldsymbol{U}_i^T(k)\boldsymbol{U}_i(k))^{-1}\boldsymbol{e}(k) \qquad (14)$$

The updated block in sequential iterations can be selected by finding the blocks with the smallest squared-Euclidean-norm as similar equation (10), i.e.,

$$\begin{aligned} i &= \arg\min_{1\le j\le P} \left\|\boldsymbol{w}_j(k+1)-\boldsymbol{w}_j(k)\right\|_2^2 \\ &= \arg\min_{1\le j\le P} \boldsymbol{e}^T(k)(\boldsymbol{U}_j^T(k)\boldsymbol{U}_j(k))^{-1}\boldsymbol{e}(k) \end{aligned} \qquad (15)$$

where we used (13).

We can extend the constrained optimisation problem in (10) to multiple blocks, whose solution leads to the following selective-partial-update AP algorithm:

$$\boldsymbol{w}_{I_B}(k+1) = \boldsymbol{w}_{I_B}(k) + \mu\boldsymbol{U}_{I_B}(k)(\boldsymbol{U}_{I_B}^T(k)\boldsymbol{U}_{I_B}(k))^{-1}\boldsymbol{e}(k)$$

where

$$I_B = \arg\min_{I_B\in S} \boldsymbol{e}^T(k)(\sum_{I_B\in S}\boldsymbol{U}_{I_B}^T(k)\boldsymbol{U}_{I_B}(k))^{-1}\boldsymbol{e}(k) \qquad (16)$$

Here S contains all B-subsets of $\{1,2,...,P\}$. The full implementation of (16) can be computationally too expensive because of the high complexity associated with the subset selection. To reduce the complexity of (16), we need to simplify the calculation of $\sum_{I_B\in S}\boldsymbol{U}_{I_B}^T(k)\boldsymbol{U}_{I_B}(k)$. Assuming that the diagonal components of $\sum_{I_B\in S}\boldsymbol{U}_{I_B}^T(k)\boldsymbol{U}_{I_B}(k)$ is much larger than the off-diagonal components, we can focus only on the diagonal components of $\sum_{I_B\in S}\boldsymbol{U}_{I_B}^T(k)\boldsymbol{U}_{I_B}(k)$. Therefore the equation can be approximated as follows [10]:

$$\boldsymbol{e}^T(k)(\sum_{I_B\in S}\boldsymbol{U}_{I_B}^T(k)\boldsymbol{U}_{I_B}(k))^{-1}\boldsymbol{e}(k) \approx \sum_{I_B\in S}\frac{\left\|\boldsymbol{e}(k)\right\|^2}{\left\|\boldsymbol{U}_{I_B}(k)\right\|^2} \qquad (17)$$

2.6. SCU-NLMS

The collection of coefficients for SCU-NLMS [11] algorithm must be updated in sequential iteration as equation (18).

$$\begin{aligned} I_B &= \arg\min_{I_B\in S}\left\|\boldsymbol{w}_{I_B}(k+1)-\boldsymbol{w}_{I_B}(k)\right\|_2^2 \\ &= \arg\max_{I_B\in S,\, j\in I_B}\sum_{j\in I_B}\left\|\boldsymbol{u}_j(k)\right\|_2^2 \end{aligned} \qquad (18)$$

where S is the collection of all B-subsets, i.e., $I_B\in S$

The update equation of SCU-NLMS is as (19)

$$\boldsymbol{w}_{I_B}(k+1) = \boldsymbol{w}_{I_B}(k) + \frac{1}{\left\|\boldsymbol{u}_{I_B}(k)\right\|_2^2}\boldsymbol{u}_{I_B}(k)e(k) \qquad (19)$$

2.7. Set-membership Affine Projection Algorithm (SM-APA)

The update equation can be written for SM–APA [1], [14] algorithm as:

$$\boldsymbol{w}(k+1) = \begin{cases}\boldsymbol{w}(k) + \mu\boldsymbol{U}^*(k)(\boldsymbol{U}(k)\boldsymbol{U}^*(k))^{-1}\boldsymbol{e}(k)\mathbf{u}_1. & if\ \left|\boldsymbol{e}(k)\mathbf{u}_1\right| > \gamma \\ \boldsymbol{w}(k) & otherwise\end{cases} \qquad (20)$$

Where $\mathbf{u}_1 = [1,0,0,...,0]_{K\times 1}^T$ which is an indicator vector.

2.8. Selective Regressors APA (SR-APA)

Suppose that we select Q input regressors among given 'K' input regressors where $\boldsymbol{w}_{i-1} = \boldsymbol{w}_{i-1}^{AP}$. Also let $\tau_Q = \{t_1, t_2, ..., t_Q\}$ denote a Q-subset (subset with K members) of the set $\{0,1,...,K-1\}$, that is, t_k denotes the delay of the selected input regressor. Let S be the collection of all Q-subsets, i.e., $\tau_Q\in S$.

The update equation of the SR – APA is obtained as

$$w_i(k+1) = w_i(k) + U_{i,\tau_Q}(k)(U^T_{i,\tau_Q}(k)U_{i,\tau_Q}(k))^{-1}e_{i,,\tau_Q}(k) \quad (21)$$

where $e_{i,\tau_Q} = d_{i,\tau_Q} - U_{i,\tau_Q} w_{i-1}$.

$$d_{i,\tau_Q} = \begin{bmatrix} d(i-t_1) \\ d(i-t_2) \\ \vdots \\ d(i-t_Q) \end{bmatrix}, \quad U_{i,\tau_Q} = \begin{bmatrix} u_{i-t_1} \\ u_{i-t_2} \\ \vdots \\ u_{i-t_Q} \end{bmatrix} \quad (22)$$

Optimal collection of input regressors can be obtained as:

$$\tau_Q^{opt} = \arg\max_{\tau_Q \in S} e^*_{i,\tau_Q}\left(U_{i,\tau_Q}U^*_{i,\tau_Q}\right)^{-1} e_{i,\tau_Q} \quad (23)$$

Therefore, the SR – APA with best input regressors is given by:

$$w_i = w_{i-1} + \mu U^*_{i,\tau_Q^{opt}}\left(U_{i,\tau_Q^{opt}}U^*_{i,\tau_Q^{opt}}\right)^{-1} e_{i,\tau_Q^{opt}} \quad (24)$$

The term in equation (23) can be approximated [12] as:

$$e^*_{i,\tau_Q}\left(U_{i,\tau_Q}U^*_{i,\tau_Q}\right)^{-1} e_{i,\tau_Q} \approx \frac{e^2_{t_1}(i)}{\|u_{i-t_1}\|^2} + \frac{e^2_{t_2}(i)}{\|u_{i-t_2}\|^2} + \ldots + \frac{e^2_{t_K}(i)}{\|u_{i-t_Q}\|^2} \quad (25)$$

where $e_t(i) = d(i-t) - u_{i-t}w_{i-1}$.

Since the reduction of the computational complexity and the convergence rate in the SR-APA depend on the relative size of Q and K, therefore we introduce the selection ratio $r = Q/K$ For $r = 1$, the SR-APA becomes identical to the conventional APA as in (7). As r decreases, the computational complexity decreases rapidly while the convergence performance slowly gets worse. Thus, only a small amount in the convergence rate can be compromised while the computational complexity is greatly reduced.

2.9. Wavelet Packet Transform

Wavelets are transform methods that has received great deal of attention over the past decades. Wavelet transform is a time-scale representation that decomposes signals into basis functions of time and scale, which makes it useful in applications such as signal denoising, detection, data compression, feature extraction, etc.

There are many techniques based on wavelet theory, such as wavelet packets, wavelet approximation and decomposition, discrete and continuous wavelet transform, etc. The backbone of the wavelets theory is the following two equations

$$\varphi_{j,k}(t) = 2^{j/2}\varphi(2^j t - k) \quad (26)$$

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k) \quad (27)$$

Where $\phi(t)$ and $\psi(t)$ are the basic scaling function and mother wavelet function, respectively [15].

A wavelet system is a set of building blocks to construct or represent a signal. It is a two dimensional expansion set. A linear expansion would be as:

$$f(t) = \sum_{-\infty}^{+\infty} c_k \varphi(t-k) + \sum_{k=-\infty}^{+\infty}\sum_{j=0}^{+\infty} d_{j,k}\psi(2^j t - k) \quad (28)$$

Most of the results of wavelet theory are developed using filter banks. In applications one never has to deal directly with the scaling functions or wavelets, only the coefficients of the filters in the filter banks are needed. The full wavelet packet decomposition in two scales is shown in figure (2).

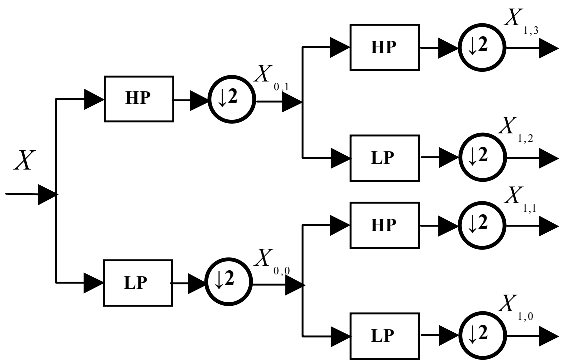

Fig 2. Wavelet packet decomposition

3. Proposed method.

In this paper, all the mentioned methods will be used to derive the block diagram of the proposed algorithm. In this stage, independent adaptive filters are applied to each pair of input signal and corresponding desired packets. Then, the selective coefficients update, selective regressors for APA and data selective updating methods are applied to each corresponding packets, independently.

Figure (3) indicates the proposed structure to implementation of adaptive algorithms with reducing the complexity. Since the reduction of the computational complexity and the convergence rate in the SCU-APA depends on the relative size of ‘B’ and ‘P’, therefore we introduce the selection ratio $r = B / P$.

For $r = 1$, the SCU-APA becomes identical to the conventional APA as in equation (7). As ‘r’ decreases, the computational complexity decreases rapidly while the convergence performance slowly gets worse. Thus, only a small amount in the convergence rate can be compromised while the computational complexity is

greatly reduced.
Also for SR-APA computational complexity and the convergence rate depends on the relative size of 'Q' and 'K', therefore we introduce the selection ratio $r = \frac{Q}{K}$. As 'r' decreases, the computational complexity of SR is decreased too.
For WPTD-SCU-APA with 'j' level decomposition, we want to update 'B' blocks out of 'P' blocks. here $I_B = \{j_1, j_2, ..., j_B\}$ determines the indices of the 'B' blocks out of 'P' blocks. Let us consider the optimization problem subject to the set of 2^j constraints imposed on the decimated filter output as:

$$\mathbf{d}_{l,D}(q) = \boldsymbol{U}_{l,I_B}^T(q)\boldsymbol{w}_{I_B}(q+1) \qquad l = 0,1,...,2^j - 1 \quad (29)$$

Again by applying the method of Lagrange multipliers [5] on the proposed criterion, we obtain the recursive relation for updating the tap weight vector of WPTD-SCU-APA in each corresponding packets as follows:

$$\boldsymbol{w}_{l,I_B}(q+1) = \boldsymbol{w}_{l,I_B}(q) + \mu \boldsymbol{U}_{l,I_B}(q)(\boldsymbol{U}_{l,I_B}^T(q)\boldsymbol{U}_{l,I_B}(q))^{-1}\boldsymbol{e}_{l,D,I_B}(q) \quad (30)$$

Where $l = 0,1,...,2^j - 1$ and

$$\boldsymbol{e}_{l,D}(q) = \mathbf{d}_{l,D}(q) - \boldsymbol{U}_{l,I_B}^T(q)\boldsymbol{w}_{l,I_B}(q)$$

and $\boldsymbol{U}_{l,I_B}(q) = \left[\boldsymbol{U}_{l,1}^T(q) \quad \boldsymbol{U}_{l,2}^T(q) \quad ... \quad \boldsymbol{U}_{l,B}^T(q)\right]^T$ (31)

Now, we consider the optimization problem for a 'B' multiple blocks and then derive a criterion for block selections. Then the selected blocks should be updated in each pair of packets at every iterations.
We now determine the blocks which should be updated in each subbands at every iterations:

$$l, I_B = \arg\min_{I_B} \left\| \boldsymbol{w}_{l,I_B}(q+1) - \boldsymbol{w}_{l,I_B}(q) \right\|_2^2$$

for $l = 0,1,...,2^j - 1$ (32)

From (30) and (31) we can obtain the equation as:

$$l, I_B = \arg\min_{I_B} \boldsymbol{e}_{l,D,I_B}^T \left(\sum_{l,I_B \in S} \boldsymbol{U}_{l,I_B}^T(q)\boldsymbol{U}_{l,I_B}(q) \right)^{-1} \boldsymbol{e}_{l,D,I_B}(q)$$

$l = 0,1,...,2^j - 1$ (33)

The term to be minimized in (33) can be approximated as equation (34).

$$l, I_B = \boldsymbol{e}_{l,D,I_B}^T \left(\sum_{l,I_B \in S} \boldsymbol{U}_{l,I_B}^T(q)\boldsymbol{U}_{l,I_B}(q) \right)^{-1} \boldsymbol{e}_{l,D,I_B}(q) \approx \sum_{l,I_B \in S} \frac{\left\| \boldsymbol{e}_{l,D,I_B}(q) \right\|^2}{\left\| \boldsymbol{U}_{l,I_B} \right\|^2} \quad for \quad l = 0,1,...,2^j - 1 \quad (34)$$

With applying the set–membership filtering in each subband we can obtain the following equations for WPTD-SM-SCU-APA.

$$\boldsymbol{w}_{l,I_B}(q+1) = \boldsymbol{w}_{l,I_B}(q) + \alpha_{l,q} \boldsymbol{U}_{l,I_B}(q)(\boldsymbol{U}_{l,I_B}^T(q)\boldsymbol{U}_{l,I_B}(q))^{-1}\boldsymbol{e}_{l,D,I_B}(q).\mathbf{u}_1 \quad (35)$$

where

$$\alpha_{l,q} = \begin{cases} 1 - \dfrac{\gamma}{\left| \boldsymbol{e}_{l,D,I_B}(q)\boldsymbol{u}_1 \right|} & if \ \left| \boldsymbol{e}_{l,D,I_B}(q)\boldsymbol{u}_1 \right| > \gamma \\ 0 & otherwise \end{cases} \quad (36)$$

for $l = 0,1,...,2^j - 1$.

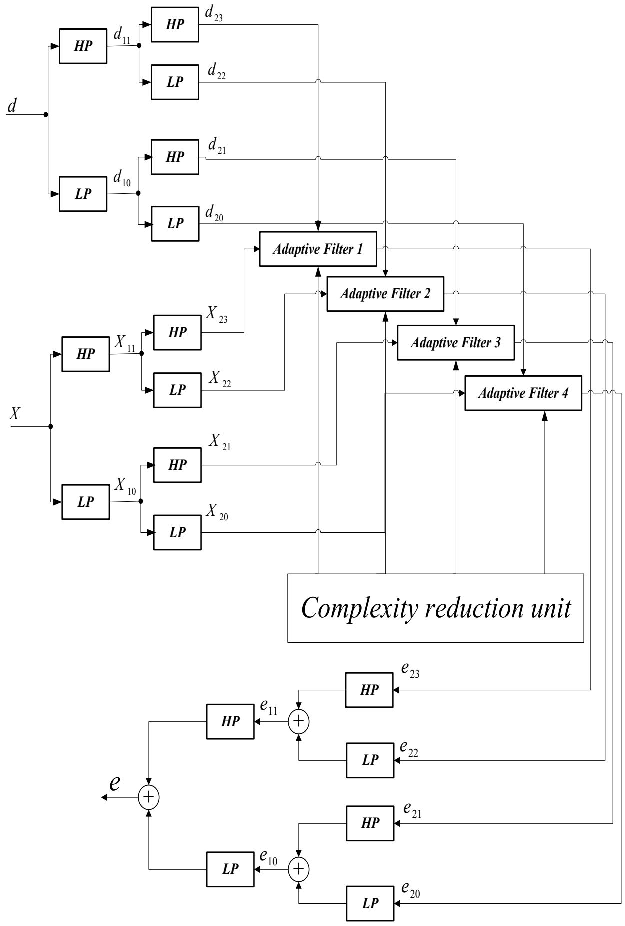

Fig 3. Proposed method in wavelet packet domain

Also, we obtain the equations of WPTD-SM-SR-APA algorithm similar to SR-APA as follows:

$$\boldsymbol{w}_l(q+1) = \boldsymbol{w}_l(q) + \alpha_{l,q} \boldsymbol{U}_{q+1,l,\tau_Q^{opt}} (\boldsymbol{U}_{q+1,l,\tau_Q^{opt}} \boldsymbol{U}_{q+1,l,\tau_Q^{opt}}^*)^{-1} \boldsymbol{e}_{q,l,\tau_Q^{opt}}(q).\mathbf{u}_1 \quad (37)$$

where

$$\alpha_{q,l} = \begin{cases} 1 - \dfrac{\gamma}{\left| \boldsymbol{e}_{l,\tau_Q^{opt}}(q)\boldsymbol{u}_1 \right|} & if \ \left| \boldsymbol{e}_{l,D,\tau_Q^{opt}}(q)\boldsymbol{u}_1 \right| > \gamma \\ 0 & otherwise \end{cases}$$

$$\tau_Q^{opt} = \arg\max_{\tau_Q \in S} \boldsymbol{e}_{l,i,\tau_Q}^* \left(\boldsymbol{U}_{l,i,\tau_Q} \boldsymbol{U}_{l,i,\tau_Q}^* \right)^{-1} \boldsymbol{e}_{l,i,\tau_Q}. \quad (38)$$

for $l = 0,1,...,2^j - 1$.

Where τ_Q^{opt} is optimum selected regressors which can be used in equation (37).

Briefly speaking, we use set–membership filtering concepts in each pair of corresponding packets, in wavelet domain independently. Then we will determine whose coefficients should be selected to update in each independent applied adaptive filter at sequential iteration. But in WPTD-SM-SR-APA all of the coefficients are updated by selected regressors along with set-membership filtering framework.

4. Computational complexity

Computational complexity is an important issue in adaptive algorithms therefore in this section we consider it in all of the mentioned algorithms. The number of calculations in SCU-APA method is calculated based on dimensions of block matrixes. Using the mentioned statements for NSAF-NLMS in [6], we note that, the NSAF-APA algorithm requires additional 2N×k×K multiplications in the analysis filter banks related to input and desired signal and N×k multiplications for synthesis filter banks in the first row of error vector. Hence, compared to the full-band APA algorithm, the NSAF algorithm requires a slight number of extra multiplications (i.e., an additional N×k×(2×K+1) multiplications) for the filter banks implementation.

For the case of WPTD with 'j' level decomposition, $N=2^j$ and using Shannon wavelet filter coefficients the 'k' (filters length) will be set to 2.
Table (1) shows the computational complexity of APA and SCU-APA and also Table (2) shows the computational complexity of proposed WPTD-SM-SCU-APA with 'j' level decomposition.

Table 1. Computational complexity of APA and SCU-APA

	Conventional APA	SCU-APA	
		Computations about weight updating	Additional computations
Multiplications	$(K^2+2K)M+K^3+K^2$	$(K^2+K)BL+KM+K^3+K^2$	$(1-r)K^2L$
Divisions	-	-	P
Comparisons	-	-	$Plog_2B+O(P)$

Table 2. Computational complexity of WPTD-SM-SCU-APA for the case that updates run for each sub-band, simultaneously

	WPTD-SM-SCU-APA, (j is level decomposition)	
	Computations about weight updating	Additional computations
Multiplications	$(K^2+K)BL+KM+K^3+K^2$	$(1-r)K^2L+2^{(j+1)}kK+2^jk$
Divisions	-	$P+2^j$
Comparisons	-	$Plog_2B+O(P)+2^j$

Based on above statements and table (3) [12], the computational complexity of WPTD-SM-SR-APA with 'j' level decomposition is shown in table (4).

Table 3. Computational complexity of APA and SR-APA [12]

	Conventional APA	SR-APA	
		Computations about weight updating	Additional computations
Multiplications	$(K^2+2K)M+K^3+K^2$	$r(rK^2+2K)M+r^2(rK^3+K^2)$	$(1-r)KM+K+1$
Divisions	-	-	K
Comparisons	-	-	$Klog_2rK+O(K)$

Table 4. Computational complexity of WPTD-SM-SR-APA for the case that updates run for each sub-band, simultaneously

	WPTD-SM-SR-APA (j level decomposition)	
	Computations about weight updating	Additional computations
Multiplications	$r(rK^2+2K)M+r^2(rK^3+K^2)$	$(1-r)KM+K+1+2^{(j+1)}kK+2^jk$
Divisions	-	$K+2^j$
Comparisons	-	$Klog_2rK+O(K)+2^j$

5. Simulation results

In this section, all of the mentioned adaptive algorithms in previous sections as: WPTD-LMS, WPTD-NLMS, SCU-NLMS, SCU-APA, SR-APA, WPTD-SM-SCU-APA and WPTD-SM-SR-APA algorithms will be applied to a system identification problem. The system is an unknown system which can be generated, randomly. The adaptive filter and the unknown system are assumed to have the same number of taps with M=32. We note that in WPTD structure, coefficients will be decimated by 2^j, so 'K' must be smaller than $\frac{M}{2^j}$ in WPTD structure. The input signal $u(k)$ is obtained by filtering a white, zero-mean, Gaussian random sequence through a first order autoregressive (AR) system with forget factor λ=0.8 as:

$u(k)=0.8u(k-1)+v(k)$

To make the comparison fair, the step-sizes were chosen to get approximately the same steady-state MSE. In all the experimental results, the bound on the output error in each sub-band were set to $\gamma = \sqrt{5\sigma_v^2}$.

In all the experimental results, the simulated learning curves were obtained by ensemble averaging over 500 independents trials and also the additive noise variance is set to $\sigma_v^2 = 0.001$.

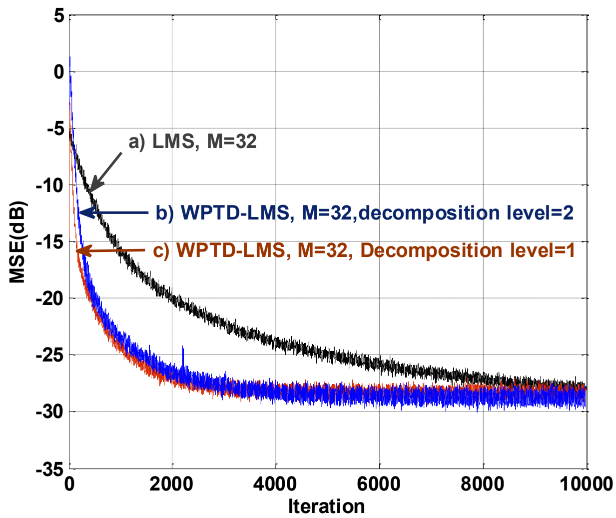

Fig 4. Learning curves, LMS and proposed method

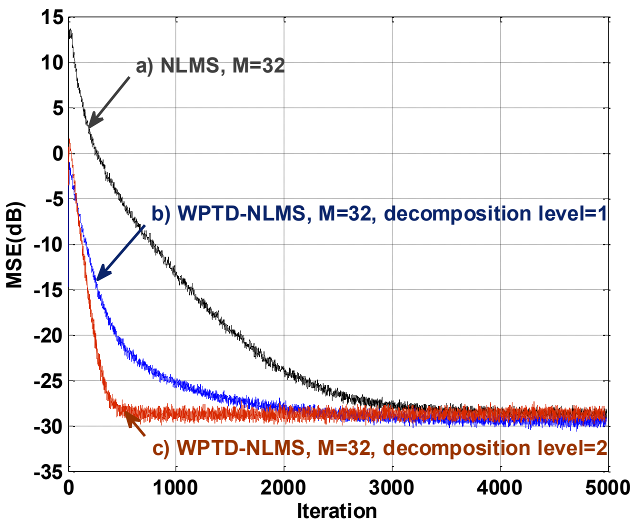

Fig 5. Learning curves, NLMS and WPTD-NLMS algorithms

Figure (4) shows a plot of MSE learning curves versus the number of iterations for the proposed method and LMS algorithm, respectively. Also in figure (5) a plot of MSE learning curves versus the number of iterations in the proposed WPTD-NLMS and NLMS is shown.

The obtained results in our simulations indicate that the rate of convergence in proposed method is better than the common earlier algorithms.

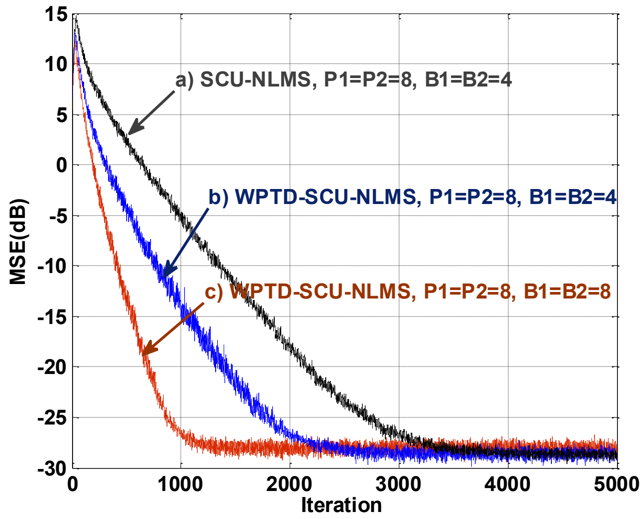

Fig 6. Learning curves, SCU-NLMS and proposed method

Figure (6) shows a plot of MSE learning curve versus the number of iterations for the proposed method 'WPTD-SCU-NLMS' and NLMS, respectively. We update all of the coefficients for case (a) and 50 percent of coefficient for case (b) and (c). Again it is shown that in proposed method the rate of convergence is better than the other methods therefore it can be obtain good results to system identification.

Again, figure (7) shows a plot of MSE learning curve versus the number of iterations using the proposed algorithm 'WPTD-SM-APA' and APA, respectively. In this simulation the order of APA is set to K=4.

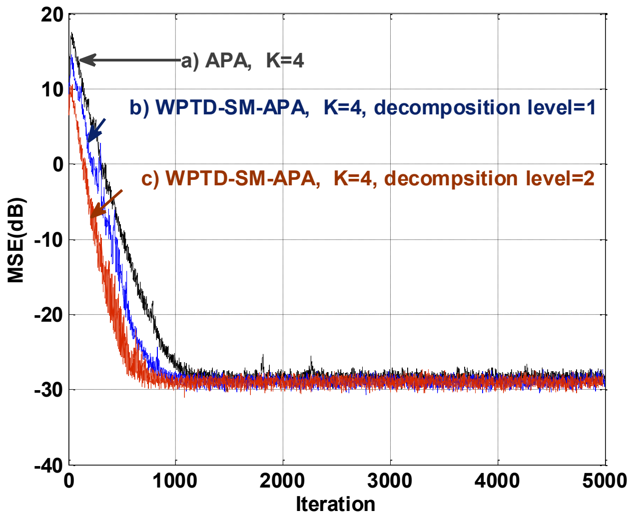

Fig 7. Learning curves, APA and proposed method

In this figure, the average number of updates for first and second adaptive filters in proposed method with first level decomposition is 736 and 723, respectively. In the second case, we apply the proposed algorithm in the second level of decomposition. Therefore we have

4 adaptive filters and the average numbers of updates in all used filters are 423, 367, 394 and 281, respectively.

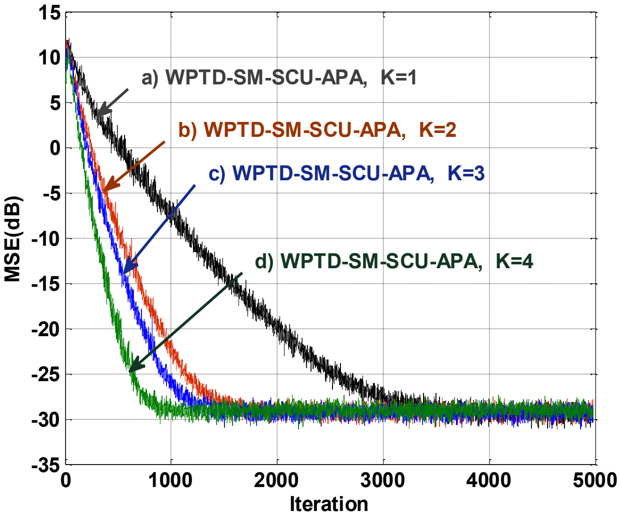

Fig 8. Learning curves, proposed method with various order of APA with 2 level decomposition and P=8, B1=8, B2=B3=B4=2

In the next simulation, we evaluate the proposed algorithm 'WPTD-SM-SCU-APA' in second decomposition level with 4 adaptive filters. In the first obtained sub-band all of the coefficients will be updated where high energy of the signal is focused and also other sub-bands have a few signal energies. Therefore in these sub-bands we update a few percent of coefficients i.e. 25 percent. In addition to selective coefficient method, data selective method is used too in this simulation. The results of the simulation are shown in the figure (8). The average number of updates in the proposed algorithm 'WPTD-SM-SCU-APA' with considering K=1 to 4 and j=2, are (250, 541, 422, 654), (370, 461, 582, 630), (153, 340, 464, 596) and (215, 295, 251, 257). In all of these sets the first number indicates the number of updates in the first adaptive filter and the second, third and fourth numbers show second, third and fourth adaptive filters updates, respectively.

In figure (9) the selection ratio is set to r=0.25 for a), b) and c) cases and r=0.5 for d) simulated result. As shown in the figure again, there is obtained good results in the rate of convergence.

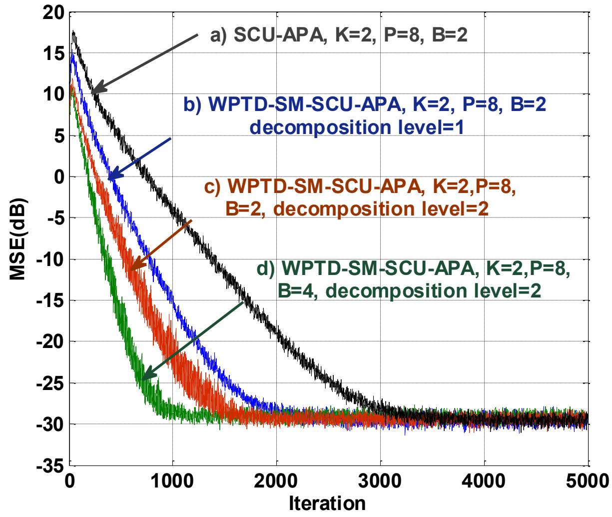

Fig 9. Learning curves, SCU-APA and WPTD-SM-SCU-APA (proposed algorithm) with j=1,2 and percent of selected coefficients for updates

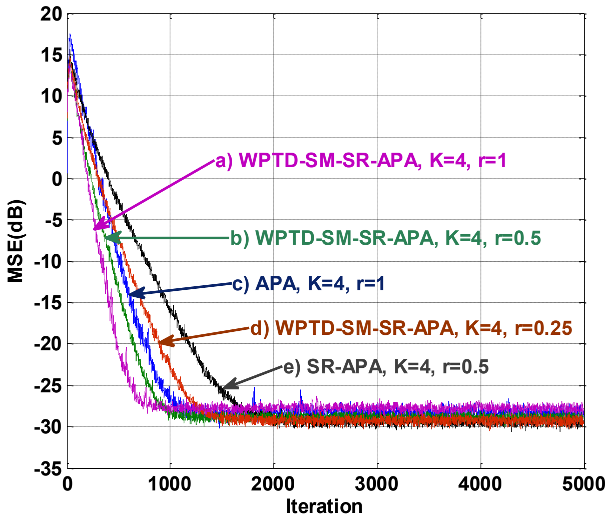

Fig 10. Learning curves, SR-APA and proposed algorithm 'WPTD-SM-SR-APA' with a percent of selected regressors

Figure (10) shows a plot of MSE learning curve respected to the number of iterations in proposed algorithm as: WPTD-SM-SR-APA and SR-APA method. The order of APA is set to K=4.

The averaged number of updates in our algorithm as WPTD-SM-SR-APA with $r=1$, $r=0.5$ and $r=0.25$ and j=1, are respectively (611, 567), (721, 703) and (814, 834). As shown in the figure when the 'r' parameter is increased the rate of convergence also is increased and therefore better results are obtained.

6. Conclusions

In this paper, we have developed a novel data selective, selective regressors and selective coefficient updates using wavelet packet decomposition for affine projection algorithm based on the concept of set-membership filtering. Also the SM-SCU-NLMS scheme that combines the data selective updating from set-membership filtering with the reduced computational complexity from partial updating is applied in proposed structure. It is verified not only the proposed algorithms can further reduce the averaged computational complexity as compared with the SCU-NLMS, SCU-APA and SR-APA, but it also retains fast convergence.

References

1. M. Z. A. Bhotto, A. Antoniou, "*A set-membership affine* projection algorithm with adaptive error bound," IEEE Conference on Electrical and Computer Engineering, pp. 894 - 897, May 2009.

2. A. Gilloire, M. Vetterli, "*Adaptive filtering in subbands with critical sampling: analysis, experiments, and application to acoustic echo cancellation,"* IEEE Trans. Signal Processing, vol. 40, pp. 1862–1875, Aug. 1992.

3. M. R. Petraglia, R. G. Alves, and P. S. R. Diniz, "*New Structures for Adaptive Filtering in Subbands with Critical Sampling,"* IEEE Trans. Signal Processing vol. 48, pp. 3316–3327, Dec. 2000.

4. A. O. Abid Noor, S. A. Samad and A. Hussain, "*Improved, Low Complexity Noise Cancellation Technique for Speech Signals,*'' World Applied Sciences Journal, pp. 272-278, 2009.

5. K. A. Lee, W. S. Gan, "*Improving Convergence of the NLMS Algorithm Using Constrained Subband Updates,*'' IEEE Signal Process. Lett., vol. 11, no. 9, Sep. 2004.

6. K. A. Lee, W. S. Gan and Y. Wen, "*Subband adaptive filtering using a multiple-constraint optimization criterion*," in Proc. EUSIPCO, pp. 1825–1828, 2004.

7. J. Rasi and B. M. Tazehkand, "*A New Efficient and Fast Adaptive Filtering Structure Based on Wavelet Packet Transform*," in Proc. 6th International Symposium on Telecommunications (IST), Iran, pp. 334-337, 2012.

8. K. Dogançay, O. Tanrikulu, "*Adaptive filtering algorithms with selective partial updates,"* IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process., vol. 48, no. 8, pp. 762–769, Aug. 2001.

9. S. Attallah, "*The wavelet transform-domain lms adaptive filter with partial subband-coefficient updating*," IEEE Trans. Circuits Syst. II, Exp. Briefs, vol. 53, no. 1, pp. 8–12, Jan. 2006.

10. P. A. Naylor and A. W. H. Khong, "*Affine Projection and Recursive Least Squares Adaptive Filters Employing Partial Updates,*'' IEEE Conference on Signals, Systems and Computers, vol. 1, pp. 950-954, 2004.

11. S. Werner, M. L. R. de Campos and P. S. R. Diniz, "*Partial-Update NLMS Algorithms with Data-Selective Updating,*'' IEEE Transactions On Signal Processing, vol. 52, no. 4, pp. 938-949, APRIL 2004.

12. K. Y. Hwang, W. J. Song, "*An Affine Projection Adaptive Filtering Algorithm With Selective Regressors*,'' IEEE Trans. Circuits And Systems-II: Express Briefs, Vol. 54, No. 1, pp. 43-46, Jan. 2007.

13. K. Dogancay, O. Tanrikulu, "*Selective-partial-update NLMS and Affine Projection algorithms for ecoustic echo cancellation,*'' IEEE International Conference on coustics, Speech, and Signal Processing, vol.1, pp. 448-451, 2000.

14. S. Werner, P. S. R. Diniz, "*Set-Membership Affine Projection Algorithm,*'' IEEE Signal Processing Letters, vol. 8, no. 8, Aug. 2001.

15. C. S. Burrus, R. A. Gopinath and H. Guo, "Introduction *to Wavelets and Wavelet Transforms, a Primer*," Prentice Hall, New Jersey, 1988.

16. Widrow and S. D. Stearns, *Adaptive Signal Processing, Upper Saddle River*, New Jersey:Prentice Hall, 1985.

17. S. Gollamudi, S. Nagaraj, S. Kapoor, and Y. F. Huang, "*Set-membership filtering and a set-membership normalized LMS algorithm with an adaptive step size*," IEEE Signal Processing Letter, vol. 5, pp. 111-114, May 1998.