ON THE PERFORMANCE COMPARISON OF GRADIENT-TYPE JOINT PROCESS ESTIMATORS IN ADAPTIVE SIGNAL PROCESSING

Deniz Gençağa, Ayşın Ertüzün

Electrical and Electronic Engineering Dept., Boğaziçi University, Bebek, 34342, İstanbul, Turkey
gencagao@boun.edu..tr, ertuz@boun.edu.tr

ABSTRACT

In adaptive signal processing, joint process estimation plays an important role in various estimation problems. It is well known that a joint process estimator consists of two structures, namely the orthogonalizer and the regression filter. In literature, orthogonalization step is performed either by orthogonal transformations or by linear predictors. While the orthogonal transformations do not preserve entropy; the predictors, such as the lattice, do preserve it. However, the steady-state performance of such linear predictors is not as good as those of the orthogonal transformations. Lattice filters do not perform perfect orthogonalization when they operate as gradient-based adaptive predictors. In this work, adaptive escalator predictor is proposed to be used as the orthogonalizer of the joint process estimator. The proposed method preserves the entropy and achieves perfect orthogonalization at all times. Moreover it has good steady-state performance compared to those structures utilizing gradient adaptive lattice filters.

1. INTRODUCTION

Adaptive filtering is widely used in many areas such as system identification, echo cancellation, channel equalization, linear prediction and spectral estimation. In these applications, generally a desired signal is estimated from its observed form. It is well known that orthogonalization of signals provides great number of advantages in the estimation of signals [1].

The Least Mean Square (LMS) algorithm [1] is widely used in adaptive signal processing applications. The LMS algorithm is preferred as a tool of adaptation because of its simplicity and low computational complexity. However it has slow convergence for highly correlated input signals, such as speech. Decorrelation of the input signal will reduce the eigenvalue spread and hence will speed up the convergence. If the processing is performed in a batch mode, Karhunen Loeve Transform (KLT) is the optimal way of orthogonalization [1]. However, when the processing needs to be done online, as in case of most adaptive processing methods, the orthogonalization also needs to be done sequentially, as observation data arrives. Joint Process Estimation (JPE) is developed for this purpose and performs jointly the online orthogonalization of the data and the estimation of the signal by using a multiple regression filter[1]. The scheme of JPE is illustrated in Fig. 1.

Two main approaches are utilized for the orthogonalization stage, which may be realized either by an orthogonal transformation or by linear prediction [2], which may be performed by a lattice filter. Within the first group, various transformations are used in literature, such as the Discrete Cosine Trasform (DCT), Discrete Wavelet Transform (DWT) and others [3,4]. It is well known that the DCT approximates the KLT in the best possible way, among these. However, these transformations do not preserve the entropy, thus information is lost after these operations. That is why, this approach is also known as a “prewhitening” method. Lattice filter, however, preserves the entropy after the orthogonalization and thus preferred in applications, where the loss of information needs to be avoided. The Gradient Adaptive Lattice (GAL) is proposed to be used as the orthogonalization stage of the online JPE’s [1]. Despite its preferred low computational complexity, GAL cannot perform perfect orthogonalization at all times [5].

In this work, we propose a new JPE structure, exploiting the use of an Adaptive Escalator Predictor (AEP) as the orthogonalization stage of the JPE structure. In literature, AEP is merely used as a linear predictor [6] and as a postprocessor after the DCT block of a transform-domain adaptive filter [7]. Since perfect orthogonalization can be accomplished by an AEP, it has superiority against the widely used GAL predictor. It is known that the AEP utilizes the use of Gram-Schmidt orthogonalization as in GAL. It is also known that as a result of using Gram-Schmidt orthogonalization, the signal is represented in terms of its backward prediction errors, which span the same information space [1, pp. 178-180]. So, by directly applying the correlated signal to the input of AEP, the possible information loss that can arise due to the use of an ortohogonal transform, such as DCT, is avoided. Thus, in the proposed JPE, the DCT block of [7] is discarded in order to preserve the information.

The reason of the performance difference between the two JPE’s that utilize the use of GAL or AEP as their first stage, can be explained by their structures. Lattice filters consist of cascaded stages and the backward prediction errors are obtained sequentially, the convergence rate of a particular lattice stage depends on the convergence rates of the preceding stages thus perfect orthogonalization of the backward errors cannot be satisfied simultaneously at all times [5]. AEP, however, has a parallel structure and thus, produces backward prediction errors of different orders in parallel. That is, the convergence of each stage is independent from the preceding ones.

After the orthogonalization stage, power normalization should be performed in order to be able to increase the convergence speed of the LMS algorithm used to adapt the coefficients of the regression filter at the second stage of JPE. As a result of this power normalization, the eigenvalue spread of the signal at the input of the regression filter is reduced. In other words, the signal that is composed of the backward prediction errors are whitened. Although the whitening procedure may seem to be contradictory with the argument of "not losing information", the best estimate can be reconstructed by applying the inverse operation of Gram-Schmidt orthogonalization.

The rest of the paper is as follows: In Section 2, the proposed JPE structure, utilizing the use of an AEP as its first stage, is explained, which will be referred to as the Adaptive Escalator Joint Process Estimator (AEJPE). In Section 3, computer simulations are given and finally the conclusions are drawn in Section 4.

2. Adaptive Escalator Joint Process Estimator (AEJPE):

The proposed AEJPE is a structure to solve the well-known joint-process estimation problem. It consists of an AEP and a JPE. AEP is a subsystem used to achieve orthogonalization (decorrelation) of the input and to generate orthogonal backward prediction errors, and JPE is a regression filter which simultaneously does estimation/filtering using the orthogonal backward prediction errors which span the same space as the correlated input (cf. Fig. 2). In AEJPE, the escalator filter rather than the lattice filter accomplishes the Gram-Schmidt orthogonalization. It exploits the use of parallel adaptation resulting in smaller misadjustment values. Moreover this method is not affected by the characteristics of the system to be identified and provides a very good steady-state performance with some increase in computational complexity.

The proposed structure can be summarized as follows:

The backward prediction errors obtained at the output of each stage of AEP, can be order updated as follows:

$$b_{k,i}(n) = b_{k-1,i}(n) - b_{k,k-1}(n)\alpha_{k,i-k+1} \qquad (1)$$

where $i\text{-}k+1 \geq 1$; $k = 1,2,........,M$; $i = 0,1,2,..........,M$ and $b_{k,i}(n)$ is the error at the k-th stage of an i-th order filter.

$\alpha_{k,i-k+1}$ is the $(i\text{-}k+1)$-st coefficient of the k-th stage. Backward prediction errors can be initialized as $b_{0,i} = u(n\text{-}i)$, where $u(n\text{-}i)$ is the input signal. The update equations for the coefficients of the AEP, utilizing an LMS type of adaptation are given as:

$$\alpha_{k,i-k+1}(n+1) = \alpha_{k,i-k+1}(n) + \mu(n) b_{k,k-1}(n) b_{k,k}(n) \qquad (2)$$

where the step-size parameter $\mu(n)$ is selected as a time-varying parameter in terms of the power estimate, $\xi_{k,k-1}(n)$, for the input of coefficient $\alpha_{k,i-k+1}$ as follows

$$\mu(n) = \frac{\hat{\mu}}{\xi_{k,k-1}(n)}, \qquad k = 1,2,...,M \qquad (3)$$

where $\hat{\mu} \leq 0.1$. This is where the power normalization takes place in order to achieve whiteness and time update of the power can be recursively estimated as follows:

$$\xi_{k,k-1}(n) = \beta \xi_{k,k-1}(n-1) + (1-\beta)\left|b_{k,k-1}(n-1)\right|^2 \qquad (4)$$
$$k = 1,2,...,M$$

where the smoothing factor β is a constant between zero and one. It should be noted that the AEP converges to a better result than ALP [6] since each stage is updated independently from the other stages. On the other hand, error propagation is involved in ALP and thus causes performance deterioration [5]. In AEP, the backward prediction errors, $b_{M,i}(n)$, are generated in parallel thus they are simultaneously orthogonal at all times, which is not the property of lattice filters due to having sequential time and order updates.

Then these errors are applied as input to the JPE where an estimate of the desired signal, $y_i(n)$, is obtained from the available data as follows:

$$y_i(n) = b_{M,i}(n) w_i(n), \quad i=0,1,....M \qquad (5)$$

Here w_i's are the coefficients of the JPE. In this work, the local adaptation of each regression coefficient is utilized, where each coefficient is updated by the corresponding error utilizing an LMS type of adaptation as:

$$w_i(n+1) = w_i(n) + \mu(n) b_{M,i}(n) e_i(n) \qquad (6)$$

for $i=0,1,...,M$. Here e_i's are the local estimation error for the i-th regression coefficient. These errors can be defined for the first coefficient and the other coefficients as follows, respectively.

$$e_0(n) = d(n) - w_0(n) b_{M,0}(n)$$
$$e_i(n) = e_{i-1}(n) - w_i(n) b_{M,i}(n) \qquad (7)$$

for $i=1,2,....,M$. $\mu(n)$ in Eq. (6), is the normalized step-size parameter as defined by Eq. (3) for $k=M$ and $k\text{-}1=i$. In this case, $\hat{\mu}$ is a constant usually selected to be less than 0.1 for a well-behaved convergence and $\xi_{M,,i}(n)$ is the power of the i-th input $b_{M,i}(n)$ of the JPE part. This power can be estimated using Eq. (4) where $k= M$ and $k\text{-}1=i$. The power normalization is done in order to increase the convergence speed of the LMS algorithm as in the transformed domain LMS (TRLMS) [3,4].

$7M(M+1)/2$ multiplications and $2M(M+1)$ additions are required for each update of AEJPE where M is the order of the structure.

It should be noted that in this work the adaptation of both the AEP and the JPE coefficients are performed simultaneously.

3. SIMULATIONS AND DISCUSSION

The proposed algorithm is used to identify an unknown plant, which is taken as a Finite Impulse Response (FIR) filter composed of 8 arbitrary tap weights. Both the adaptive filter and the unknown plant are driven by four different signals, that are generated by passing a white noise through four different filters, which are also used in the simulations of [4]. The magnitude response of each filter is shown in Figs. 3a-3d. The proposed AEJPE algorithm is compared with LMS, Normalized LMS, Transform Domain Least Mean Square (TRLMS) [3,4] where DCT is used as the transform (DCTLMS), and ALJPE [1] algorithms; the convergence of the mean squared error (MSE) as a function of time is presented in Figs.4a-4d where coloring filters shown in Figs. 3a-3d, are used, respectively. No parameter change (such as $\mu(n), \hat{\mu}, \beta$) has been performed in order to test the behaviour of various algorithms to the signal change. It is observed that AEJPE is robust to the different signals that are used, whereas all other methods can behave differently. As can be seen from these figures, the proposed AEJPE algorithm has a very good steady state MSE performance. The convergence of AEJPE is fast and the steady-state MSE value is very low. It is also seen that, only for the low-pass signals, the convergence rate of DCTLMS is around that of AEJPE, since DCT approximates KLT very well for such signals. In all other cases, AEJPE outperforms when compared with others. This is due to the fact that escalator has a parallel structure and the backward prediction errors are orthogonal at all times. The updates of the regression and the escalator coefficients can be done simultaneously, thus these two sets of coefficients are optimum at the same time. However this is not the case in ALJPE where both regression and the lattice coefficients can not be optimum at the same time since lattice is a serial structure. Thus, perfect orthogonalization cannot be obtained with ALJPE. On the other hand, perfect orthogonalization can be obtained by AEP. In addition to these, since the proposed AEJPE utilizes joint process estimation with simultaneous Gram-Schmidt type of orthogonalization, it does not destroy the information content of the signal contrary to the prewhitening methods. The entropy of the input signal in AEJPE is preserved during the process of the orthogonalization. So, the best estimate can be reconstructed without losing any information.

4. CONCLUSIONS

A new joint process estimator that utilizes the AEP to perform Gram Schmidt orthogonalization process is proposed in order to use the computationally simple LMS algorithm in the regression part of JPE's. The developed algorithm presents a compromise between the information preserving methods with poor steady-state performances, such as the ALJPE, and those with good convergence results, but entropy losing prewhitening methods, such as TRLMS techniques. Finally, the proposed AEJPE structure can be utilized where an online JPE is needed and provides the best performance among all algorithms based on gradient type of adaptation.

REFERENCES

[1] Haykin, S., *Adaptive Filter Theory*, Prentice Hall International Editions, (1/e), 2002.

[2] Makhoul, J., " Linear Prediction: A Tutorial Review", *Proc. IEEE*, Vol. 63, pp. 561-580, April 1975.

[3] Narayan, S. S., A. M. Peterson and M. J. Narasimha, " Transform Domain LMS Algorithm", *IEEE Trans. Acoust., Speech and Signal Process.*, Vol. 31, pp. 609-615, June 1983.

[4] Marshall, D. F., W. K. Jenkins and J. J. Murphy, " The Use of Orthogonal Transforms for Improving Performance of Adaptive Filters", *IEEE Trans. Circuits and Systems*, Vol. 36, pp. 474-484, April 1989.

[5] Beaufays, F. and E. Wan, "An Efficient First-Order Stochastic Algorithm for Lattice Filters", *Proc. ICANN'94*, vol. 2, pp. 1021-1024, Sorrento, Italy, 26-29 May 1994.

[6] Ahmed, N. and D. H. Youn, "On a realization and Related Algorithm for Adaptive Prediction", IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. ASSP-28, No:5, pp. 493-497, October 1980.

[7] Parikh, V. N. and A. Z. Baraniecki, "The Use of the Modified Escalator Algorithm to Improve the Performance of Transform-Domain LMS Adaptive Filters", *IEEE Trans. Signal Process.*, Vol. 46, pp. 625-635, March 1998.

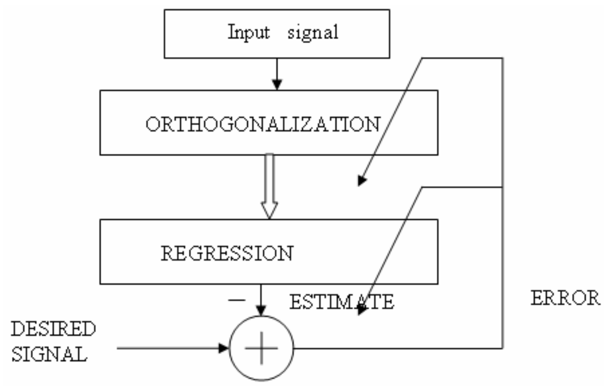
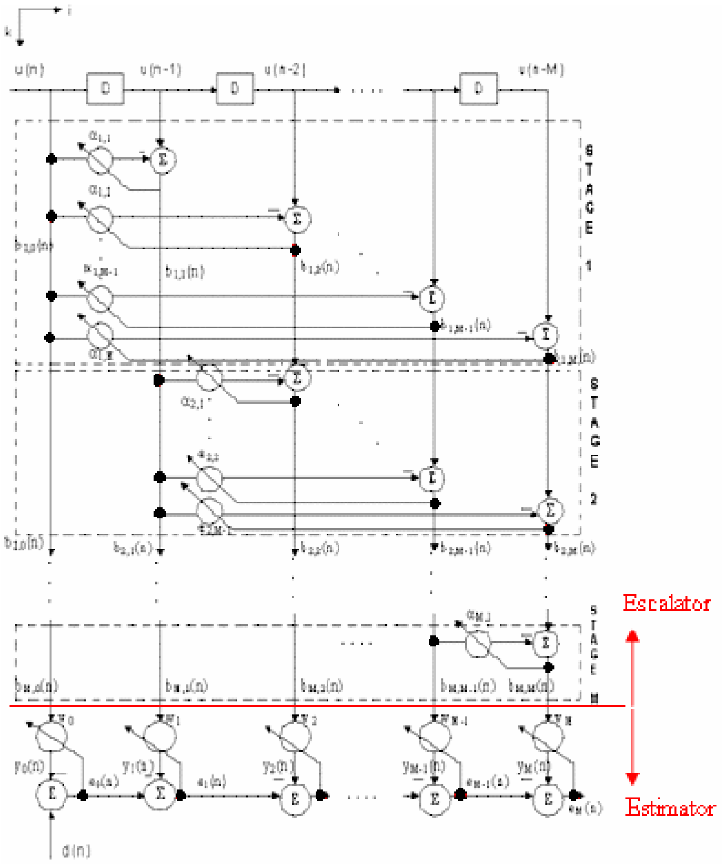


Fig.1. Joint Process Estimator



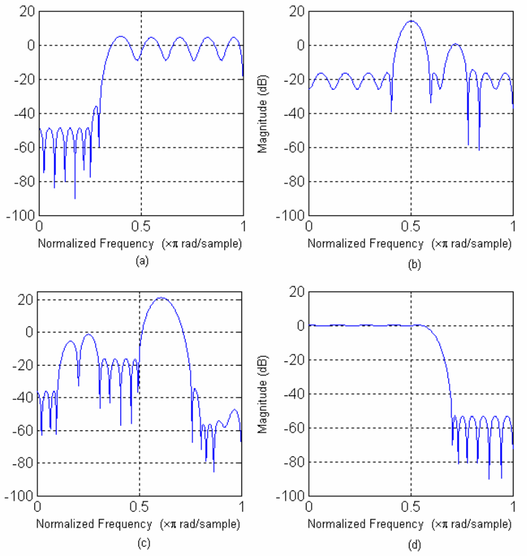


Fig. 3. Magnitude Responses of four different Coloring Filters, (a) High Pass, (b) Band Pass, (c) Band Stop, (d) Low Pass

Fig. 2. Adaptive Escalator Joint Process Estimator Architecture of order M

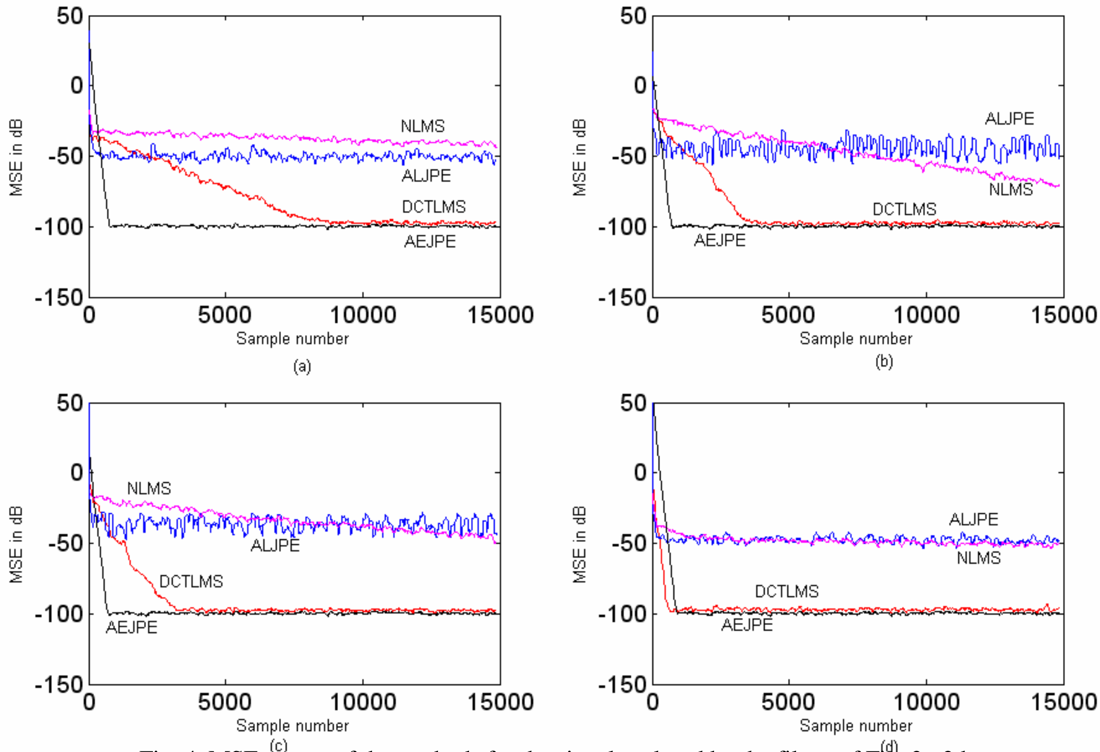


Fig. 4. MSE curves of the methods for the signals colored by the filters of Fig. 3a-3d