A NEW ALGORITHM FOR CONSTRAINED ADAPTIVE ARRAY PROCESSING

By

Y.M. Zhu

and

G. Yin

IMA Preprint Series # 443

August 1988

A NEW ALGORITHM FOR CONSTRAINED ADAPTIVE ARRAY PROCESSING

Y.M. Zhu* and G. Yin**

Abstract. A new approach for algorithms of linearly constrained adaptive array processing is developed in this paper. The new algorithm is the product of an effort to provide a more efficient procedure for real time implementation of the array processing. The essence of our approach is to improve the efficiency by utilizing a number of processors updating the same array element. A delay is introduced in the computation for each processor. The structure (delay and multi-processors) of our algorithm requires that all processors are operating asynchronously, and in a pipe line manner so that the recursively computed data flows in a rhythmic fashion, passing through each processor periodically. The almost sure convergence of the algorithm is proved under the assumption of non-stationary and correlated signals.

Keywords. adaptive array, multiple processors, pipelining, delay, non-stationary, correlated signals, strong convergence.

AMS (MOS) subject classification. 60G35, 93E10, 93E11.

1. Introduction

A new algorithm for linearly constrained adaptive array processing is developed in this paper. This algorithm is a distributed and asynchronous computation procedure, and it is an extension and generalization of the well-known adaptive filter algorithm with constraints, or adaptive beam-former algorithm. In lieu of using a single processor alone as in the conventional case, we propose to use a number of processors operating in pipe line to update the same array adaptively. The efficiency is improved by introducing a delay in the update for each individual processor. Suppose that r units of time is needed for each iteration if a single processor is used in the computation. By using the new multi-processor algorithm, we will use r processors to do the work. Consequently, only one unit of time is needed. Thus, the new approach provides significant savings and is preferable over the conventional approach in real time implementation.

Were the statistics of the signals known a priori, this array problem would be a classical constrained least squares problem. Therefore, to some extent, the algorithm can be called a constrained least squares procedure.

* Institute of Mathematical Sciences, Chengdu Branch of Academia Sinica, Sichuan, 610015, PRC.
Research of this author was supported in part by Science Foundation of Academia Sinica under Grant #85074.

** Department of Mathematics, Wayne State University, Detroit, MI 48202, U.S.A.
Research of this author was supported in part by Wayne State University under the Wayne State University Research Grant Award. The paper was completed when this author was visiting the Institute for Mathematics and Its Applications, at the University of Minnesota.

The adaptive filter type algorithms have a wide range of applications, these include signal estimation and detection, electromagnetic antenna arrays, under water signal processing, learning and pattern recognition, adaptive control, and in many other related fields.

We now describe the adaptive beam-former algorithm as follows. The problem is concerned with the approximation of target or reference signals. In what follows, $\{y_n\}$, $\{\psi_n\}$ will denote the output and reference signals respectively. We let y_n be weighted by a matrix x in order that the signals $x'y_n$ become the best approximation of ψ_n in the mean square sense. z' stands for the transpose of z. To avoid error accumulation while maintaining good approximation, we impose an additional constraint $x'C = \Phi$.

To be more precise, let $x \in \mathbb{R}^{r\times m}$, $y_n \in \mathbb{R}^{r\times l}$, $\psi_n \in \mathbb{R}^{m\times l}$. The aforementioned problem can be phrased as:

$$\text{minimize } E(x'y_n - \psi_n)(x'y_n - \psi_n)' \tag{1.1}$$

$$\text{subject to } x'C = \Phi. \tag{1.2}$$

In view of (1.2), $x'C(I - C^\dagger C) = 0$,

$$\Phi = \Phi C^\dagger C \tag{1.3}$$

where $z^\dagger$ denotes the pseudo-inverse of z. Conversely, if (1.3) holds, then (1.2) is true for $x' = \Phi C^\dagger$. As a consequence, (1.3) is a necessary and sufficient condition for (1.2).

To proceed, we need the following assumptions.

(A1) The signals are non-stationary in that

$$Ey_ny_n' = m_nQ_1 \tag{1.4}$$

$$Ey_n\psi_n' = m_nQ_2 \tag{1.5}$$

$$E\psi_n\psi_n' = m_nQ_3. \tag{1.6}$$

In (1.4)-(1.6), Q_i, $i \leq 3$ are constant but unknown matrices, such that they are non-negative definite; $\{m_n\}$ is a sequence of positive real numbers satisfying equation (1.9).

(A2) The signals are correlated in the following way:

$$\sum_{n=1}^{\infty} a_n(y_ny_n' - m_nQ_1) \text{ converges a.s.} \tag{1.7}$$

$$\sum_{n=1}^{\infty} a_n(y_n\psi_n' - m_nQ_2) \text{ converges a.s.} \tag{1.8}$$

with

$$a_n > 0,\ a_nm_n \xrightarrow{n} 0,\ \sum_{n=1}^{\infty} a_nm_n = \infty. \tag{1.9}$$

Our assumptions (A1)(A2) cover a wide class of non-stationary and correlated signals. In fact, they are rather general. $\{y_n y_n' - m_n Q_1\}$ and $\{y_n \psi_n' - m_n Q_2\}$ can be ARMA type processes or moving average type of processes with infinitely many terms. The behavior of $\{m_n\}$ can be pretty wild. It can diminish to 0 or grow to infinity as long as (1.9) is satisfied.

By means of diagonalization, it can be shown that there exist constant matrices H, L, such that

$$Q_1 = HH', \; Q_2 = HL', \; Q_3 = LL' \tag{1.10}$$

(cf. [1] and the references therein).

It turns out that the optimization problem (1.1) (1.2) can be solved analytically. The minimizer x^* is given by

$$x^* = C^{\dagger,\prime}\Phi' + (PHH'P)^{\dagger}(HL' - HH'C^{\dagger,\prime}\Phi') \tag{1.11}$$

with $P = I - CC^{\dagger}$. For a derivation of the above formula, the readers are referred to [1], [5] and the references therein.

Although x^* is expressed in a closed form, (1.11) is not very valuable from an application point of view. Since H, L are unknown, there is no way for us to obtain x^* from (1.11) directly. The next thing that one wants to do is perhaps to approximate H, L sequentially. However, to compute pseudo-inverse sequentially for large dimensional systems is practically not permissible. We thus approximate x^* by a sequence of matrices $\{x_n\}$ which can be adjusted on the basis of the outputs (measurements) y_n, such that

$$x_n \xrightarrow{n} x^* \text{ a.s.} \tag{1.12}$$

Such a sequence can be constructed recursively as:

$$x_{n+1} = C^{\dagger,\prime}\Phi' + P\Big(x_n + a_n(y_n\psi_n - y_n y_n' x_n)\Big) \tag{1.13a}$$

$$x_0 = C^{\dagger,\prime}\Phi'. \tag{1.13b}$$

(1.13) is called the algorithm for linearly constrained adaptive array or adaptive beamformer algorithm. In case $\Phi = C = 0$, (1.13) reduces to the usual adaptive filter algorithm

$$x_{n+1} = x_n + a_n(y_n\psi_n - y_n y_n' x_n). \tag{1.14}$$

Due to their theoretical value and practical usefulness, adaptive filter and adaptive beam-former algorithms have been studied extensively by a host of authors, for example [1], [3-9], [11-12] among others.

In the past two decades, most contributions to the aforementioned topics have been concentrated on the convergence issue. Efforts have been made to relax the conditions on the signals. Independent signals were considered in [8]; correlated but bounded signals were dealt with in [6-7]; unbounded signals with finite memory and finite moments were treated in [4]. In a recent paper, [12], the with probability 1 convergence was proved for a wide class of correlated and non-stationary signals. To our knowledge, the conditions used in [12] are the weakest up to date.

Relatively, little attention has been given to the problem of improving the convergence properties and efficiency. As a consequence, few result available along this line. This issue is however important, especially when we are dealing with large dimensional arrays.

The objective of the present paper is to investigate this efficiency issue. In lieu of (1.13), we shall propose an alternative recursive procedure, namely constrained adaptive array with delay. The delay is introduced in order to shorten the computation time for each iteration.

The rest of the paper will be organized as follows. the new approach is described next, so does its modified version-delayed algorithm with randomly varying truncation bounds. The boundedness is proved in section 3, and the almost sure convergence is obtained in section 4. Finally, some concluding remarks are made in section 5.

2. Algorithm with delay

Let r be a fixed positive integer. In view of (1.13), we propose the following algorithm for adaptive linearly constrained adaptive array with delay in the following way.

$$x_{n+1} = C^{\dagger,\prime}\Phi' + P\Big(x_n + a_{n-r}(y_{n-r}\psi'_{n-r} - y_{n-r}y'_{n-r})\Big) \tag{2.1a}$$

$$x_j = C^{\dagger,\prime}\Phi',\ 0 \le j \le r. \tag{2.1b}$$

The motivation for proposing algorithm (2.1) is as follows. Suppose that to implement (1.13), r units of time is needed for completing each iteration. If instead, we use r processors operating in a pipe line way to carry out the recursion defined as (2.1). After each iteration, the computed data is fed forward to the next processor, and it is used for the next iteration. Thus, starting from an initial value, the recursively computed data flows in a rhythmic fashion, passing through each processoring element (processor) periodically (with period r). For each individual processing element, the computations are based on the iteration obtained r units ago, and the values obtained in the last iteration from previous processor. By this arrangement, the available information is used in a rather efficient way. In the same amount of time, the new algorithm iterates r times as many as in the conventional approach. As a consequence, only one unit of time is needed for each iteration of the newly developed procedure with delay. Thus, the efficiency has been improved r times. In a way, our algorithm is an asynchronous and distributed computation procedure.

Since our main concern is of asymptotic in nature, the first a few values are not very important. Consequently, the initial conditions are chosen according to (2.1b).

It is easily seen that (2.1) is an extension of (1.13). When $r = 0$, (2.1) reduces to (1.13).

Let $S_n = Px_n$. Since

$$S'_n = x'_n(I - CC^\dagger) = x'_n - \Phi C^\dagger$$

to study the convergence of x_n, we only need to consider the convergence of S_n.

It follows from (2.1),

$$S_{n+1} = S_n + a_{n-r}P\Big(y_{n-r}\psi'_{n-r} - y_{n-r}y'_{n-r}(S_{n-r} + C^{\dagger,\prime}\Phi')\Big) \tag{2.2a}$$

$$S_j = 0, \ 0 \le j \le r. \tag{2.2b}$$

The optimal weighting matrix S^* is given by

$$S^* = (PHH'P)^{\dagger}(HL' - HH'C^{\dagger,\prime}\Phi'). \tag{2.3}$$

Comparing (1.13) with (2.1) or (2.2), the only difference is the r-step delay. However, the convergence analysis for (2.1) is much harder than that of (1.13). In order to establish the almost sure convergence result, the first step is to prove the boundedness. Although the delay allows us to improve the efficiency of the algorithm, it creates much of the troubles for the convergence analysis. To circumvent the difficulties, we shall work with a modified version of the algorithm via the technique of randomly varying truncations. Such an idea was first appeared in [2] to deal with strong convergence of the RM and KW algorithms, and it was further generalized to treat more general stochastic approximation problems with non-additive noise in [10]. To obtain the convergence result for adaptive beam-forming by means of randomly varying truncations was first discussed in [11].

For future use, we define

$$O_n = P\Big(y_n\psi'_n - y_n y'_n(S_n + C^{\dagger,\prime}\Phi')\Big). \tag{2.4}$$

To fix the idea, let $\{B_n\}$ be a sequence of monotone increasing real numbers satisfying $B_n \xrightarrow{n} \infty$. Define a sequence of integer valued random variables $\sigma(n)$, and a sequence of matrices S^*_n as follows.

$$\sigma(0) = 0 \ \sigma(n) = \sum_{j=0}^{n-1} I_{\{|S^*_{j+1}| > B_{\sigma(j)}\}} \tag{2.5}$$

$$S_{n+1} = \begin{cases} S_n + a_{n-r}O_{n-r}, & \text{if } |S^*_{n+1}| < B_{\sigma(n)}; \\ S_r = 0, & \text{otherwise.} \end{cases} \tag{2.6}$$

Remark: the idea of randomly varying truncations is rather simple. At the nth iteration, if the approximation exceeds certain bound, we force the iterates returning to a fixed point; otherwise we keep iterating as in the usual algorithm.

The formulation is completed. We are now ready to present the consistency result.

3. Boundedness

In what follows, we shall always assume that $PH \neq 0$. If $PH = 0$, then

$$EPy_ny'_nP = m_nPHH'P = 0.$$

This implies that $Py_n = 0$ a.s., therefore $S_n = 0$ a.s.

The boundedness is proved by establishing a series of lemmas. Henceforth, K_i, $i \le 9$ will stand for various positive constants. trA will stand for the trace of a matrix A.

Lemma 3.1. *Consider the sequence $\{S_n\}$ defined by (2.5) (2.6). If conditions (A1) (A2) are satisfied, and $\{S_{n_k}\}$ is a convergent subsequence of $\{S_n\}$; and if $\{S_{n_k-j}; j \le r\}$ is bounded uniformly in k, then there exist $\delta > 0$, such that for any $0 \le \eta < \delta$, and some K_η, whenever $k \ge K_\eta$, the following equations (3.2a) (3.2b) hold for any $m \in [n_k, m(n_k, \eta)]$, with*

$$m(n_k, \eta) = \max\{m; \sum_{j=n}^{m} a_{j-r} m_{j-r}\}. \tag{3.1}$$

$$|\sum_{j=n_k}^{m+1} a_{j-r} O_{j-r}| \le K_1 \tag{3.2a}$$

$$|S_{m+1} - S_{n_k}| \le K_2 \eta. \tag{3.2b}$$

Proof: Observe that $\sigma(n)$ is strictly increasing, and thus $\sigma(n) \xrightarrow{n} \sigma$. Either $\sigma < \infty$ or $\sigma = \infty$. If $\sigma < \infty$, then for sufficiently large k,

$$|\sum_{j=n_k}^{m+1} a_{j-r} O_{j-r}| = |S_{m+2} - S_{n_k}| \le 2B_\sigma. \tag{3.3}$$

Hence, (3.2a) holds with $K_1 = 2B_\sigma$.

Now, consider the case of $\sigma = \infty$. If (3.2a) were not true, there would be a contradiction. To see this, denote $\bar{S} = \lim_k S_{n_k}$. Choose $K_3 > |\bar{S}|$. For $\eta = 2^{-l}$, there exist k_l and m_l, such that $m_l \in [n_{k_l}, m(n_{k_l}, 2^{-l})]$, and for any $k \ge k_l$,

$$|S_{n_k}| \le \frac{(K_3 + |\bar{S}|)}{2} \tag{3.4}$$

$$|\sum_{j=n_{k_l}}^{m_l+1} a_{j-r} O_{j-r}| > \frac{K_3 - |\bar{S}|}{2}. \tag{3.5}$$

Without loss of generality, we may assume that

$$m_l = \inf\{m; |\sum_{j=n_{k_l}}^{m} a_{j-r} O_{j-r}| > \frac{K_3 - |\bar{S}|}{2}\}. \tag{3.6}$$

For any $m \in [n_{k_l}, m_l]$,

$$|S_{n_{k_l}} + \sum_{j=n_k}^{m+1} a_{j-r} O_{j-r}| \le K_3. \tag{3.7}$$

The fact $\sigma(n_k) \to \infty$ implies that for l large enough, $B_{\sigma(n_{k_l})} \ge K_3$, and hence the right hand side of (3.7) is bounded above by $B_{\sigma(n_{k_l})}$ for any $m \in [n_{k_l}, m_l]$. It follows that

$$|S_{m+1}| \le B_{\sigma(n_{k_l})}. \tag{3.9}$$

In view of the boundedness of $\{S_{n_k-i}; i \le r\}$ in the assumption, there exists a positive constant which we may assume to be equal to K_3 without loss of generality. Now, for any $m \in [n_{k-i}, m_{l+1}]$, for $i \le r$,

$$|S_m| \le K_3. \tag{3.10}$$

By virtue of (1.7) and (3.6),

$$\begin{aligned} &tr \sum_{j=n_{k_l}}^{m_l} a_{j-r}(y_{j-r}y'_{j-r} - m_{j-r}HH') \\ &\le \sum_{j=n_{k_l}}^{m_l} a_{j-r}(|y_{j-r}|^2 - m_{j-r}|H|^2) \xrightarrow{l} 0 \text{ a.s.} \end{aligned} \tag{3.11}$$

Thus

$$\sum_{j=n_{k_l}}^{m_l} a_{j-r}|y_{j-r}|^2 \le K_4 2^{-l}|H|^2 \xrightarrow{l} 0. \tag{3.12}$$

In view of (3.10),

$$\begin{aligned} &|\sum_{j=n_{k_l}}^{m_l} a_{j-r}Py_{j-r}y'_{j-r}S_{j-r}| \\ &\le \sum_{j=n_{k_l}}^{m_l} a_{j-r}|P||y_{j-r}|^2|S_{j-r}| \xrightarrow{l} 0. \end{aligned} \tag{3.13}$$

This together with (1.7) (1.8) imply

$$|S_{m_l+1} - S_{n_{k_l}}| = |\sum_{j=n_{k_l}}^{m_l} a_{j-r}O_{j-r}| \xrightarrow{l} 0 \tag{3.14}$$

and

$$a_{m_l+1-r}O_{m_l+1-r} \xrightarrow{l} 0 \tag{3.15}$$

Thus,

$$|S_{m_l+1} - S_{n_{k_l}} + a_{m_l+1-r}O_{m_l+1-r}| \xrightarrow{l} 0. \tag{3.16}$$

But, by virtue of (3.6),

$$\begin{aligned} &|S_{m_l+1} - S_{n_{k_l}} + a_{m_l+1-r}O_{m_l+1-r}| \\ &= |\sum_{j=n_{k_l}}^{m_l+1} a_{j-r}O_{j-r}| > \frac{K_3 - |\bar{S}|}{2}. \end{aligned} \tag{3.17}$$

This contradicts (3.16). Hence (3.2a) follows. By the similar kind of argument, (3.2b) can be verified as well.

Lemma 3.2. *Under the conditions of Lemma 3.1, for any convergent subsequence* $\{S_{n_k}\}$ *with limit* S*, the following equation holds.*

$$\sum_{j=n}^{m} a_{j-r}\Big(P(y_{j-r}\psi'_{j-r} - y_{j-r}y'_{j-r}(S + C^{\dagger,\prime}\Phi') - m_{j-r}PHH'P(S^* - S)\Big) \xrightarrow{n,m} 0 \text{ a.s.} \quad (3.18)$$

Proof: Since (cf. [1])

$$PHH'P(PHH'P)^{\dagger}H = PH, \; PS = S \quad (3.19)$$

$$PHH'P(PHH'P)^{\dagger}HH'PS = PHH'PS. \quad (3.20)$$

It then follows that

$$\begin{aligned}
&\sum_{j=n}^{m} a_{j-r}\left(P\big(y_{j-r}y'_{j-r} - y_{j-r}y'_r r(S + C^{\dagger,\prime}\Phi')\big) - m_{j-r}PHH'P(S^* - S)\right)\\
&= \sum_{j=n}^{m} a_{j-r}\left(P\big(y_{j-r}y'_{j-r} - y_{j-r}y'_{j-r}(S + C^{\dagger,\prime}\Phi')\big) - m_{j-r}P\big(HL' - HH'(S + C^{\dagger,\prime}\Phi')\big)\right.\\
&\qquad \left. + m_{j-r}P\big(HL'HH'(S + C^{\dagger,\prime}\Phi')\big) - m_{j-r}PHH'P(S^* - S)\right)\\
&= \sum_{j=n}^{m} a_{j-r}\left(P\big(y_{j-r}y'_{j-r} - y_{j-r}y'_{j-r}(S + C^{\dagger,\prime}\Phi')\big) - m_{j-r}P\big(HL' - HH'(S + C^{\dagger,\prime}\Phi')\big)\right)\\
&\qquad \xrightarrow{n,m} 0 \text{ a.s.}
\end{aligned} \quad (3.21)$$

This lemma is thus proved.

Theorem 3.3. *If (A1) (A2) are satisfied, then the sequence* $\{S_n\}$ *defined by (2.5) (2.6) is bounded uniformly in* n.

Remark: Since the algorithm is truncated at randomly varying bounds, to show the boundedness is equivalent to show that the truncations cease for sufficiently large n. In order to have this assertion be valid, σ the limit of $\sigma(n)$ must be finite. As a consequence, there exists a N, such that for any $n \geq N$,

$$S_{n+1} = S_n + a_{n-r}P\Big(y_{n-r}\psi'_{n-r} - y_{n-r}y'_{n-r}(S_{n-r} + C^{\dagger,\prime}\Phi')\Big)$$

with $|S_n| \leq B_\sigma$.

Proof: For suppose not, we would arrive a contradiction. In view of (2.4) (2.5), $|S_n - S^*|^2$ would be across some interval $[\Delta_1, \Delta_2]$ infinitely often from the left, where

$\Delta_1 > |S^*|^2$. We observe that when $|S_n - S^*|^2 < \Delta_1$, S_n is bounded. Similarly to Lemma 3.1, for those n that $|S_n - S^*| < \Delta_1$,

$$|a_{n-r}O_{n-r}| \xrightarrow{n} 0. \tag{3.22}$$

Next, we extract $\{S_{n_k}\}$, $\{S_{m_k}\}$ subsequences of $\{S_n\}$ satisfying

$$n_k < m_k,\ |S_{n_{k-i}}| \leq K_5,\ i \leq r \tag{3.23}$$

$$|S_{m_k-1} - S^*|^2 < \Delta_1,\ \Delta_1 \leq |S_n - S^*|^2 \leq \Delta_2,\ n \in [n_{k-r}, m_k - 1] \tag{3.24}$$

$$|S_{m_k} - S^*|^2 > \Delta_2. \tag{3.25}$$

It is clear that

$$S_{n_k} = S_{n_k-1} + a_{n_k-r}O_{n_k-r}. \tag{3.26}$$

(3.22) (3.23) and (3.26) yield that

$$\begin{aligned}
&|S_{n_k} - S^*|^2 - |S_{n_k-1} - S^*|^2 \\
=&tr\Big((S_{n_k} + S_{n_k-1} - 2S^*)'(S_{n_k} - S_{n_k-1})\Big) \\
=&tr\Big((S_{n_k} + S_{n_k-1} - 2S^*)'a_{n_k-r}O_{n_k-r}\Big) \xrightarrow{k} 0.
\end{aligned} \tag{3.27}$$

Now, (3.24) implies that

$$|S_{n_k} - S^*|^2 \xrightarrow{k} \Delta_1. \tag{3.28}$$

The set

$$F = \{S; \Delta_1 \leq |S - S^*|^2 \leq \Delta_2\}$$

is a bounded and closed set.

Select a convergent subsequence of $\{S_{n_k}\}$. With no loss in generality, we may still denote the subsequence by $\{S_{n_k}\}$. Suppose that $\bar{S} = \lim_k S_{n_k}$, then $\bar{S} \in F$, and $|\bar{S} - S^*|^2 = \Delta_1$. By virtue of Lemma 3.1, for any $k \geq k_\eta$,

$$\begin{aligned}
&|S_{m(n_k,\eta)+1} - S^*|^2 - |S_{n_k} - S^*|^2 \\
=&tr\Big((S_{m(n_k,\eta)+1} + S_{n_k} - 2S^*)'(S_{m(n_k,\eta)+1} - S_{n_k})\Big) \\
=&tr\Big((S_{m(n_k,\eta)+1} + S_{n_k} - 2S^*)' \sum_{j=n_k}^{m(n_k,\eta)} a_r r O_{j-r}\Big) + tr\Big((2\bar{S} - 2S^*)' \sum_{j=n_k}^{m(n_k,\eta)} a_{j-r}O_{j-r}\Big) \\
\leq& K_6\eta^2 + 2tr\Big((\bar{S} - S^*) \sum_{j=n_k}^{m(n_k,\eta)} a_{j-r}P\big(y_{j-r}\psi'_{j-r} - y_{j-r}y'_{j-r}(\bar{S} + C^{\dagger,\prime}\Phi')\big)\Big) \\
&+ 2tr\Big((\bar{S} - S^*)' \sum_{j=n_k}^{m(n_k,\eta)} a_{j-r}Py_{j-r}y'_{j-r}(S_{j-r} - \bar{S})\Big).
\end{aligned} \tag{3.29}$$

In view of Lemma 3.2, for large enough k,

$$2tr\Big((\bar{S}-S^*)'\sum_{j=n_k}^{m(n_k,\eta)} a_{j-r}P\big(y_{j-r}\psi_{j-r}-y_{j-r}y'_{j-r}(\bar{S}+C^{\dagger,\prime}\Phi')\big)\Big)$$

$$=2tr\big((\bar{S}-S^*)'\sum_{j=n_k}^{m(n_k,\eta)} a_{j-r}\big(P(y_{j-r}\psi'_{j-r}-y_{j-r}y'_{j-r}(\bar{S}+C^{\dagger,\prime}\Phi')-m_{j-r}PHH'P(S^*-\bar{S})\big)\Big)$$

$$+\,2tr\Big((\bar{S}-S^*)'\sum_{j=n_k}^{m(n_k,\eta)} a_{j-r}m_{j-r}PHH'P(S^*-\bar{S})\Big)\le -K_7\eta. \tag{3.30}$$

As for the third term on the right hand side of (3.29),

$$2tr\Big((\bar{S}-S^*)'\sum_{j=n_k}^{m(n_k,\eta)} a_{j-r}Py_{j-r}y'_{j-r}(S_{j-r}-\bar{S})\Big)$$
$$\le K_8\eta\max_{j\in[n_k,m(n_k,\eta)]}|S_{j-r}-\bar{S}|. \tag{3.31}$$

By virtue of (3.30) (3.31), taking $\limsup_k$ in (3.29),

$$\limsup_{k\to\infty}|S_{m(n_k,\eta)+1}-S^*|^2\le\Delta_1-K_7\eta<\Delta_1. \tag{3.32}$$

On the other hand, since $S_{n_k}\xrightarrow{k}\bar{S}$, (3.2b) and (3.29) yield

$$\max_{m\in[n_k,m(n_k,\eta)+1]}|S_m-S^*|^2-|S_{n_k}-S^*|^2\xrightarrow{\eta}0. \tag{3.33}$$

Noticing $|S_{n_k}-S^*|^2\xrightarrow{k}\Delta_1<\Delta_2$, for sufficiently small η, and $m\in[n_k,m(n_k,\eta)+1]$, we have $|S_m-S^*|^2<\Delta_2$. consequently,

$$\Delta_1\le|S_{m(n_k,\eta)+1}-S^*|^2\le\Delta_2. \tag{3.14}$$

This contradicts (3.32). The proof of Theorem 3.3 is thus completed.

4. Convergence

Theorem 4.1. *If (A1) (A2) are satisfied, then* $S_n\xrightarrow{n}S^*$ a.s.

Proof: The proof is divided into two steps. In step 1, we show that the limit of S_n exists. In step 2, we prove $S_n\xrightarrow{n}S^*$. Since the argument is very similar to the previous section, we shall keep our discussion very brief.

(1) Again, we prove the existence by contradiction. For suppose not, then Theorem 3.3 would imply that there exist two convergent subsequences $\{S_{n_k^1}\}$ and $\{S_{n_k^2}\}$, with limits S_1 and S_2 respectively, such that $S_1 \neq S_2$. Define $s_1 = |S_1 - S^*|^2$ and $s_2 = |S_2 - S^*|^2$, without loss of generality, we may assume that $s_2 > s_1$. It follows that

$$\lim_k |S_{n_k^2} - S^*|^2 - \lim_k |S_{n_k^1} - S^*|^2 = s_2 - s_1 > 0. \tag{4.1}$$

Since

$$\begin{aligned} &\Big| |S_{n+1} - S^*|^2 - |S_n - S^*|^2 \Big| \\ =&a_{n-r} K_9 tr\Big((S_{n+1} + S_n - 2S^*)'(S_{n+1} - S_n) \Big) \xrightarrow{n} 0. \end{aligned} \tag{4.2}$$

$|S_n - S^*|^2$ would be across some interval $[\Delta_1, \Delta_2]$ infinitely often, with $s_1 \leq \Delta_1 \leq \Delta_2 \leq s_2$. We can then extract two subsequences $\{S_{n_k}\}$ and $\{S_{m_k}\}$ of $\{S_n\}$, such that (3.23)-(3.25) hold, and

$$|S_{n_k} - S^*|^2 \xrightarrow{k} \Delta_1. \tag{4.3}$$

Without loss of generality, we may assume that $S_{n_k} \xrightarrow{k} S$ for some S. By Lemma 3.1, (3.2b) holds.

Exactly the same argument as (3.29)-(3.32) yield that

$$\limsup_{k \to \infty} |S_{m(n_k, \eta)+1} - S^*|^2 < \Delta_1. \tag{4.4}$$

On the other hand, (3.33) (3.34) yield

$$\Delta_1 \leq |S_{m(n_k, \eta)+1} - S^*|^2 \leq \Delta_2. \tag{4.5}$$

This is a contradiction, and hence the limit of S_n exists.

(2) Now, we need to characterize the limit, i.e., show $|S_n - S^*| \to 0$. The argument is in the same spirit as above. We prove it by contradiction again. If there exists a $\Delta > 0$, such that $|S_n - S^*|^2 \to \Delta$, then similar argument as above yields that

$$\lim_n |S_{m(n_k, \eta)+1} - S^*|^2 < \Delta. \tag{4.6}$$

This is a contradiction. Thus the theorem follows.

5. Conclusion

Having developed the algorithm for adaptive array processing with delay and proved its consistency, we wish to make the following remarks.

(1) The algorithm uses extensive concurrency. Rather than sequentially using a single processor alone as in the conventional case, concurrent use of a number of processors

is suggested. The concurrency is obtained by pipelining the result for each individual processor.

(2) The computation of each processor has a simple form similar to that of the usual adaptive filter and adaptive beam-former algorithms. The distinction comes from the delay.

(3) The multiple use of data flow, the pipe line implementation and the regular structure of the algorithm make it be quite efficient. It seems that this algorithm is particularly useful for very large dimensional systems (like VLSI), which arise naturally from signal and imaging processing problems.

REFERENCES

[1] H.F. Chen, *Recursive Estimation and Control for Stochastic Systems*, John Wiley & Sons, 1985

[2] H.F. Chen and Y.M. Zhu, Stochastic approximation procedures with randomly varying truncations, Scientia Sinica (series A), vol. XXIX (1986), 914-926

[3] A.Z. Di, A new result on the iterative algorithm for adaptive array processing with correlated and nonstationary noise, J. Sys. Sci. and Math. Scis., 5(2) (1985), 151-160

[4] E. Eweda and O. Macchi, Quadratic mean and almost sure convergence of unbounded stochastic approximation algorithm with correlated observations, Ann. Inst. Henri Poincare, Section B, vol. 14, No 3 (1983), 235-255

[5] O.I. Frost, An algorithm for linear constrained adaptive array processing, Proc. IEEE, 60(1972), 8: 926-935

[6] H.J. Kushner and D.S. Clark, *Stochastic Approximation Method for Constrained and Unconstrained Systems*, Springer-Verlag, 1978

[7] L.Ljung, Analysis of stochastic approximation algorithm, IEEE Tran. on Automatic Control, AC-22 (1977), 551-575

[8] B. Widrow, P.E. Mantey, L.T. Griffiths and B.B. Goode, Adaptive antenna systems, Proc. IEEE, 55(1967), 2143-2158

[9] B. Widrow, J.M. McCool, M.G. Larimore and C.R. Johnson, Stationary learning characteristics of the LMS adaptive filter, Proc. IEEE, 64(1975), 8: 1692-1716

[10] G. Yin and Y.M. Zhu, On almost sure convergence of stochastic approximation algorithms with non-additive noise, to appear in Int. J. Control.

[11] Y.M. Zhu, An algorithm with randomly varying truncations for adaptive beam formers, Appl. Math. Mechanics 7(1986), 435-442.

[12] Y.M. Zhu and G. Yin, Adaptive filter with constraint and correlated nonstationary signals, Systems & Control Letters, 10(1988), 271-279.