Hindawi Publishing Corporation
EURASIP Journal on Audio, Speech, and Music Processing
Volume 2007, Article ID 71495, 13 pages
doi:10.1155/2007/71495

Research Article

Detection-Guided Fast Affine Projection Channel Estimator for Speech Applications

Yan Wu Jennifer,[1] John Homer,[2] Geert Rombouts,[3] and Marc Moonen[3]

[1] *Canberra Research Laboratory, National ICT Australia and Research School of Information Science and Engineering, The Australian National University, Canberra ACT 2612, Australia*
[2] *School of Information Technology and Electrical Engineering, The University of Queensland, Brisbane QLD 4072, Australia*
[3] *Departement Elektrotechniek, Katholieke Universiteit Leuven, ESAT/SCD, Kasteelpark Arenberg 10, 30001 Heverlee, Belgium*

Received 9 July 2006; Revised 16 November 2006; Accepted 18 February 2007

Recommended by Kutluyil Dogancay

In various adaptive estimation applications, such as acoustic echo cancellation within teleconferencing systems, the input signal is a highly correlated speech. This, in general, leads to extremely slow convergence of the NLMS adaptive FIR estimator. As a result, for such applications, the affine projection algorithm (APA) or the low-complexity version, the fast affine projection (FAP) algorithm, is commonly employed instead of the NLMS algorithm. In such applications, the signal propagation channel may have a relatively low-dimensional impulse response structure, that is, the number m of active or significant taps within the (discrete-time modelled) channel impulse response is much less than the overall tap length n of the channel impulse response. For such cases, we investigate the inclusion of an active-parameter detection-guided concept within the fast affine projection FIR channel estimator. Simulation results indicate that the proposed detection-guided fast affine projection channel estimator has improved convergence speed and has lead to better steady-state performance than the standard fast affine projection channel estimator, especially in the important case of highly correlated speech input signals.

1. INTRODUCTION

For many adaptive estimation applications, such as acoustic echo cancellation within teleconferencing systems, the input signal is highly correlated speech. For such applications, the standard normalized least-mean square (NLMS) adaptive FIR estimator suffers from extremely slow convergence. The use of the affine projection algorithm (APA) [1] is considered as a modification to the standard NLMS estimators to greatly reduce this weakness. The built-in prewhitening properties of the APA greatly accelerate the convergence speed especially with highly correlated input signals. However, this comes with a significant increase in the computational cost. The lower complexity version of the APA, the fast affine projection (FAP) algorithm, which is functionally equivalent to APA, was introduced in [2].

The fast affine projection algorithm (FAP) is now, perhaps, the most commonly implemented adaptive algorithm for high correlation input signal applications.

For the above-mentioned applications, the signal propagation channels being estimated may have a "low dimensional" parametric representation [3–5]. For example, the impulse responses of many acoustic echo paths and communication channels have a "small" number m of "active" (nonzero response) "taps" in comparison with the overall tap length n of the adaptive FIR estimator. Conventionally, estimation of such low-dimensional channels is conducted using a standard FIR filter with the normalized least-mean square (NLMS) adaptive algorithm (or the unnormalized LMS equivalent). In these approaches, each and every FIR filter tap is NLMS-adapted during each time interval, which leads to relatively slow convergence rates and/or relatively poor steady-state performance. An alternative approach proposed by Homer et al. [6–8] is to detect and NLMS adapt only the active or significant filter taps. The hypothesis is that this can lead to improved convergence rates and/or steady-state performance.

Motivated by this, we propose the incorporation of an activity detection technique within the fast affine projection FIR channel estimator. Simulation results of the newly proposed detection-guided fast affine projection channel estimator demonstrate faster convergence and better steady-state error performance over the standard FAP FIR channel estimator, especially in the important case of highly correlated input signals such as speech. These features make this newly proposed detection-guided FAP channel estimator a good candidate for adaptive channel estimation applications such as acoustic echo cancellation, where the input signal is highly correlated speech and the channel impulse response is often "long" but "low dimensional."

The remainder of the paper is set out as follows. In Section 2 we provide a description of the adaptive system we consider throughout the paper as well as the affine projection algorithm (APA) [1] and the fast affine projection algorithm (FAP) [2]. Section 3 begins with a brief overview of the previous proposed detection-guided NLMS FIR estimators of [6–8]. We then propose our detection-guided fast affine projection FIR channel estimator. Simulation conditions are presented in Section 4, followed by the simulation results in Section 5. The simulation results include a comparison of our newly proposed estimator with the standard NLMS channel estimator, the earlier proposed detection-guided NLMS channel estimator [8], the standard APA channel estimator [1] as well as the standard FAP channel estimator [2] in 3 different input correlation level cases.

2. SYSTEM DESCRIPTION

2.1. Adaptive estimator

We consider the adaptive FIR channel estimation system of Figure 1. The following assumptions are made:

(1) all the signals are sampled: at sample instant k, $u(k)$ is the signal input to the unknown channel and the channel estimator; additive noise $v(k)$ occurs within the unknown channel;
(2) the unknown channel is linear and is adequately modelled by a discrete-time FIR filter $\Theta = [\theta_0, \theta_1, \dots, \theta_n]^T$ with a maximum delay of n sample intervals;
(3) the additive noise signal is zero mean and uncorrelated with the input signal;
(4) the FIR-modeled unknown channel, $\Theta[z^{-1}]$ is sparsely active:

$$\Theta[z^{-1}] = \theta_{t_1} z^{-t_1} + \theta_{t_2} z^{-t_2} + \cdots + \theta_{t_m} z^{-t_m}, \tag{1}$$

where $m \ll n$, and $0 \le t_1 < t_2 < \cdots t_m \le n$.

At sample instant k, an *active* tap is defined as a tap corresponding to one of the m indices $\{t_a\}_{a=1}^{m}$ of (1). Each of the remaining taps is defined as an *inactive* tap.

The observed output from the unknown channel is

$$y(k) = \Theta^T \underline{U}(k) + v(k), \tag{2}$$

where $\underline{U}(k) = [u(k), u(k-1), \dots, u(k-n)]^T$.

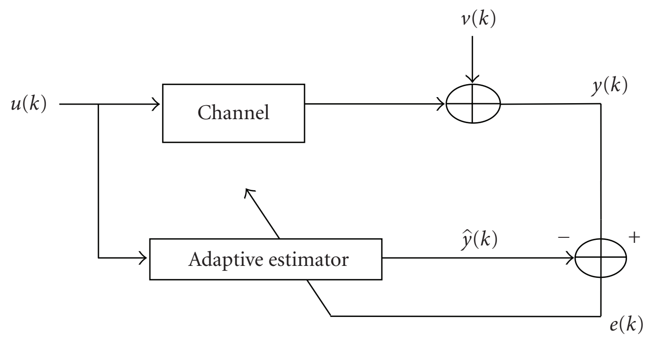

FIGURE 1: Adaptive channel estimator.

The standard adaptive NLMS estimator equation, as employed to provide an estimate $\hat{\theta}$ of the unknown channel impulse response vector Θ, is as follows [9]:

$$\hat{\theta}(k+1) = \hat{\theta}(k) + \frac{\mu}{\underline{U}^T(k)\underline{U}(k) + \delta}\underline{U}(k)[y(k) - \hat{y}(k)], \tag{3}$$

where $\hat{y}(k) = \hat{\theta}^T(k)\underline{U}(k)$ and where δ is a small positive regularization constant.

Note: the standard initial channel estimate $\hat{\theta}(0)$ is the all-zero vector.

For stable 1st-order mean behavior, the step size μ should satisfy $0 < \mu \le 2$. In practice, however, to attain higher-order stable behavior, the step size is chosen to satisfy $0 < \mu \ll 2$.

For the standard discrete NLMS adaptive FIR estimator, every coefficient $\hat{\theta}_i(k)$ $[i = 0, 1, \dots, n]$ is adapted at each sample interval. However, this approach leads to slow convergence rates when the required FIR filter tap length n is "large" [6]. In [6–8], it is shown that if only the active or significant channel taps are NLMS estimated then the convergence rate of the NLMS estimator may be greatly enhanced, particularly when $m \ll n$.

2.2. Affine projection algorithm

The affine projection algorithm (APA) is considered as a generalisation of the normalized least-mean-square (NLMS) algorithm [2]. Alternatively, the APA can be viewed as an in-between solution to the NLMS and RLS algorithms in terms of computational complexity and convergence rate [10]. The NLMS algorithm updates the estimator taps/weights on the basis of a single-input vector, which can be viewed as a one-dimensional affine projection [11]. In APA, the projections are made in multiple dimensions. The convergence rate of the estimator's tap weight vector greatly increases with an increase in the projection dimension. This is due to the built-in decorrelation properties of the APA.

To describe the affine projection algorithm (APA) [1], the following notations are defined:

(a) N: affine projection order;

(b) $n+1$: length of the adaptive channel estimator excitation signal matrix of size $(n+1)\times N$;

(c) $U(k)$: $U(k) = [\underline{U}(k), \underline{U}(k-1), \ldots, \underline{U}(k-(N-1))]$, where $\underline{U}(k) = [u(k), u(k-1), \ldots, u(k-n)]^T$;

(d) $U^T(k)U(k)$: covariance matrix;

(e) Θ: the channel FIR tap weight vector, where $\Theta = [\theta_0, \theta_1, \ldots, \theta_n]^T$;

(f) $\hat{\theta}(k)$: the adaptive estimator FIR tap weight vector at sample instant k where $\hat{\theta}(k) = [\hat{\theta}_0(k), \hat{\theta}_1(k), \ldots, \hat{\theta}_n(k)]^T$;

(g) $\hat{\theta}(0)$: initial channel estimate with the all-zero vector;

(h) $\underline{e}(k)$: the channel estimation signal error vector of length N;

(i) $\underline{\varepsilon}(k)$: N-length normalized residual estimation error vector;

(j) $y(k)$: system output;

(k) $v(k)$: the additive system noise;

(l) δ: regularization parameter;

(m) μ: step size parameter.

The affine projection algorithm can be described by the following equations (see Figure 1).

The system output $y(k)$ involves the channel impulse response to the excitation/input and the additive system noise $v(k)$ and is given by (2).

The channel estimation signal error vector $\underline{e}(k)$ is calculated as

$$\underline{e}(k) = Y(k) - U(k)^T\hat{\theta}(k-1), \tag{4}$$

where $Y(k) = [y(k), y(k-1), \ldots, y(k-N+1)]^T$.

The normalized residual channel estimation error vector $\underline{\varepsilon}(k)$, is calculated in the following way:

$$\underline{\varepsilon}(k) = \left[U(k)^T - U(k) + \delta I\right]^{-1} \cdot \underline{e}(k), \tag{5}$$

where $I = N \times N$ identity matrix.

The APA channel estimation vector is updated in the following way:

$$\hat{\theta}(k+1) = \hat{\theta}(k) + \mu U(k)\underline{\varepsilon}(k). \tag{6}$$

A regularization term δ times the identity matrix is added to the covariance matrix within (5) to prevent the instability problem of creating a singular matrix inverse when $[U(k)^T \ U(k)]$ has eigenvalues close to zero. A well behaved inverse will be provided if δ is large enough.

From the above equations, it is obvious that the relations (4), (5), (6) reduce to the standard NLMS algorithm if $N = 1$. Hence, the affine projection algorithm (APA) is a generalization of the NLMS algorithm.

2.3. Fast affine projection algorithm

The complexity of the APA is about $2(n+1)\,N + 7N^2$, which is generally much larger than the complexity of the NLMS algorithm, $2(n+1)$. Motivated by this, a fast version of the APA was derived in [2]. Here, instead of calculating the error vector from the whole covariance matrix, the FAP only calculates the first element of the N-element error vector, where an approximation is made for the second to the last components of the error vector $\underline{e}(k)$ as $(1-\mu)$ times the previously computed error [12, 13]:

$$\underline{e}(k+1) = \begin{bmatrix} e(k+1) \\ (1-\mu)\overline{\underline{e}}(k) \end{bmatrix}, \tag{7}$$

where the $N-1$ length $\overline{\underline{e}}(k)$ consists of the $N-1$ upper elements of the vector $\underline{e}(k)$.

Note: (7) is an exact formula for the APA if and only if $\delta = 0$.

The second complexity reduction is achieved by only adding a weighted version of the last column of $U(k)$ to update the tap weight vector. Hence there are just $(n+1)$ multiplications as opposed to $N \times (n+1)$ multiplications for the APA update of (6). Here, an alternate tap weight vector $\hat{\theta}_1(k)$ is introduced.

Note: the subscript 1 denotes the new calculation method.

$$\underline{\hat{\theta}_1}(k+1) = \underline{\hat{\theta}_1}(k) - \mu\underline{U}(k-N+2)E_{N-1}(k+1), \tag{8}$$

where

$$\begin{aligned} E_{N-1}(k+1) &= \sum_{j=0}^{N-1} \varepsilon_j(k-N+2+j) \\ &= \varepsilon_{N-1}(k+1) + \varepsilon_{N-2}(k) + \cdots + \varepsilon_0(k-N+2) \end{aligned} \tag{9}$$

is the $(N-1)$th element in the vector

$$\underline{E}(k+1) = \begin{bmatrix} \varepsilon_0(k+1) \\ \varepsilon_1(k+1) + \varepsilon_0(k) \\ \vdots \\ \varepsilon_{N-1}(k+1) + \varepsilon_{N-2}(k) + \cdots + \varepsilon_0(k-N+2) \end{bmatrix}. \tag{10}$$

Alternatively, $\underline{E}(k+1)$ can be written as

$$\underline{E}(k+1) = \begin{bmatrix} 0 \\ \overline{\underline{E}}(k) \end{bmatrix} + \underline{\varepsilon}(k+1), \tag{11}$$

where $\overline{\underline{E}}(k)$ is an $N-1$ length vector consisting of the upper most $N-1$ elements of $\underline{E}(k)$ and $\underline{\varepsilon}(k+1) = [\varepsilon_{N-1}(k+1), \varepsilon_{N-2}(k+1) + \cdots + \varepsilon_0(k+1)]^T$ as calculated via (5).

Hence, it can be shown that the relationship between the new update method and the old update method of APA can be viewed as

$$\underline{\hat{\theta}}(k) = \underline{\hat{\theta}_1}(k) + \mu\overline{U}(k)\overline{\underline{E}}(k), \tag{12}$$

where $\overline{U}(k)$ consists of the $N-1$ leftmost columns of $U(k)$.

A new efficient method to calculate $e(k)$ using $\hat{\theta}_1(k)$ rather than $\hat{\theta}(k)$ is also derived:

$$\underline{\tilde{r}}_{xx}(k+1) = \underline{\tilde{r}}_{xx}(k) + u(k+1)\underline{\tilde{\alpha}}(k+1) - u(k-n)\underline{\tilde{\alpha}}(k-n), \tag{13}$$

where

$$\underline{\tilde{\alpha}}(k+1) = [u(k), u(k-1), \ldots, u(k-N+2)]^T \tag{14}$$

$$e_1(k+1) = y(k+1) - \underline{U}(k+1)^T\hat{\theta}_1(k) \tag{15}$$

$$e(k+1) = e_1(k+1) - \mu\underline{\tilde{r}}^t_{xx}(k+1)\underline{\bar{E}}(k). \tag{16}$$

(Further details can be found in [2].)

The following is a summary of the FAP algorithm:

(1) $\underline{\tilde{r}}_{xx}(k+1) = \underline{\tilde{r}}_{xx}(k) + u(k+1)\underline{\tilde{\alpha}}(k+1) - u(k-n)\underline{\tilde{\alpha}}(k-n)$,

(2) $e_1(k+1) = y(k+1) - \underline{U}(k+1)^T\hat{\theta}_1(k)$,

(3) $e(k+1) = e_1(k+1) - \mu\underline{\tilde{r}}^t_{xx}(k+1)\underline{\bar{E}}(k)$,

(4) $\underline{e}(k+1) = \begin{bmatrix} e(k+1) \\ (1-\mu)\underline{\bar{e}}(k) \end{bmatrix}$,

(5) $\underline{\varepsilon}(k+1) = [U(k+1)^T U(k+1) + \delta I]^{-1}\underline{e}(k+1)$,

(6) $\underline{E}(k+1) = \begin{bmatrix} 0 \\ \underline{\bar{E}}(k) \end{bmatrix} + \underline{\varepsilon}(k+1)$,

(7) $\underline{\hat{\theta}_1}(k+1) = \underline{\hat{\theta}_1}(k) - \mu\underline{U}(k-N+2)E_{N-1}(k+1)$.

The above formulae are in general only approximately equivalent to the APA; they are exactly equal to the APA if the regularization δ is zero. Steps (2) and (7) of the FAP algorithm are each of complexity $(n+1)$ MPSI (multiplications per symbol interval). Step (1) is of complexity $2N$ MPSI and steps (3), (4), (6) are each of complexity N MPSI. Step (5), when implemented in the Levinson-Dubin method, requires $7N^2$ MPSI [2]. Thus, the complexity of FAP is roughly $2(n+1) + 7N^2 + 5N$. For many applications like echo cancellation, the filter length $(n+1)$ is always much larger than the required affine projection order N, which makes FAP's complexity comparable to that of NLMS. Furthermore, the FAP only requires slightly more memory than the NLMS.

3. DETECTION-GUIDED ESTIMATION

3.1. Least-squares activity detection criteria review

The original least-squares-based detection criterion for identifying active FIR channel taps for white input signal conditions [6] is as follows.

The tap index j is defined to be detected as a member of the active tap set $\{t_a\}_{a=1}^m$ at sample instant k if

$$X_j(k) > T^{(k)}, \tag{17}$$

where

$$X_j(k) = \frac{\{\sum_{i=1}^k [y(i)u(i-j)]^2\}}{\sum_{i=1}^k u^2(i-j)},$$
$$T(k) = \frac{2\log(k)}{k}\sum_{i=1}^k y^2(i). \tag{18}$$

However, the original least-square-based detection criterion suffers from tap coupling problems when colored or correlated input signals are applied. In particular, the input correlation causes $X_j(k)$ to depend not only on θ_j but also the neighboring taps.

The following three modifications to the above activity detection criterion were proposed in [7, 8] for providing enhanced performance for applications involving nonwhite input signals.

Modification 1. Replace $X_j(k)$ by

$$\tilde{X}_j(k) = \frac{\{\sum_{i=1}^k [y(i) - \hat{y}(i) + \hat{\theta}_j(i)u(i-j)]u(i-j)\}^2}{\sum_{i=1}^k u^2(i-j)}. \tag{19}$$

The additional term $-\hat{y}(i) + \theta_j(i)u(i-j)$ in the numerator of $\tilde{X}_j(k)$ is used to reduce the coupling between the neighboring taps [7, 8].

Modification 2. Replace $T(k)$ by

$$\tilde{T}(k) = \frac{2\log(k)}{k}\sum_{i=1}^k [y(i) - \hat{y}(i)]^2. \tag{20}$$

This modification is based on the realization that for *inactive taps*, the numerator term of $\tilde{X}_j(k)$ is approximately

$$N_j(k) \approx \left\{\sum_{i=1}^k [y(i) - \hat{y}(i)]u(i-j)\right\}^2, \quad j = \text{inactive tap index}. \tag{21}$$

Combining this with the LS theory on which the original activity criterion (17) is based suggests the following modification [8].

Modification 3. Apply an exponential forgetting operator $W_k(i) = (1-\gamma)^{k-i}$, $0 < \gamma \ll 1$ within the summation terms of the activity criterion [8].

Modification 2 is theoretically correct only if $\Theta - \hat{\theta}(k)$ is not time varying. Clearly this is not the case. Modification 3 is included to reduce the effect of $\Theta - \hat{\theta}(k)$ being time varying. Importantly, the inclusion of Modification 3 also improves the applicability of the detection-guided estimator to time-varying systems. (Note that the result of Modification 3 is denoted with superscript W in the next section.)

3.2. Enhanced detection-guided NLMS FIR channel estimator

The enhanced time-varying detection-guided NLMS estimation proposed in [8] is as follows.

For each tap index j and at each sample interval:

(1) label the tap index j to be a member of the active parameter set $\{t_a\}_{a=1}^m$ at sample instant k if

$$\tilde{X}_j^w(k) > \tilde{T}^w(k), \tag{22}$$

where

$$\tilde{X}_j^w(k) = \frac{\left\{\sum_{i=1}^{k} W_k(i)\left[y(i) - \hat{y}(i) + \hat{\theta}_j(i)u(i-j)\right]u(i-j)\right\}^2}{\sum_{i=1}^{k} W_k(i)u^2(i-j)}, \quad (23)$$

$$\tilde{T}^w(k) = \frac{2\log\left(L^w(k)\right)}{L^w(k)} \sum_{i=1}^{k} W_k(i)\left[\left(y(i) - \hat{y}(i)\right)\right]^2, \quad (24)$$

$$L^w(k) = \sum_{i=1}^{k} W_k(i), \quad (25)$$

and where $W_k(i)$ is the exponentially decay operator:

$$W_k(i) = (1-\gamma)^{k-i} \quad 0 < \gamma \ll 1; \quad (26)$$

(2) update the NLMS weight for each detected active tap index t_a:

$$\hat{\theta}_{t_a}(k+1) = \hat{\theta}_{t_a}(k) + \frac{\mu}{\sum_{t_a} u(k-t_a)^2 + \varepsilon} u(k-t_a)e(k), \quad (27)$$

where $\sum_{t_a}$ = summation over all detected active-parameter indices;

(3) reset the NLMS weight to zero for each identified inactive tap index.

Note that (23)–(25) can be implemented in the following recursive form:

$$\begin{aligned} N_j(k) &= (1-\gamma)N_j(k-1) \\ &\quad + \left[y(k) - \hat{y}(k) + \hat{\theta}_j(k)u(k-j)\right]u(k-j), \\ D_j(k) &= (1-\gamma)D_j(k-1) + u^2(k-j), \\ q(k) &= (1-\gamma)q(k-1) + \left[y(k) - \hat{y}(k)\right]^2, \quad (28) \\ L^w(k) &= (1-\gamma)L^w(k-1) + 1, \end{aligned}$$

$$\begin{aligned} \tilde{X}_j^w(k) &= \frac{N_j^2(k)}{D_j(k)}, \\ \tilde{T}^w(k) &= \frac{2q(k)\log\left[L^w(k)\right]}{L^w(k)}. \quad (29) \end{aligned}$$

Note, as suggested in [8], that a threshold scaling constant η may be introduced on the right-hand side of (24) or (29). If $\eta > 1$, the system may avoid the incorrect detection of "non-active" taps. This, however, may come with an initial delay in detecting the smallest of the active taps, leading to an initial additional error increase. If $\eta < 1$, it may improve the detectibility of "weak" active taps. However, it has the risk of incorrectly including inactive taps within the active tap set, resulting in reduced convergence rates.

3.3. Proposed detection-guided FAP FIR channel estimator

The enhanced detection-guided FAP estimation is derived as follows.

The tap index j is detected as being a member of the active parameter set $\{t_a\}_{a=1}^m$ at sample instant k if

$$\tilde{X}_j^W(k) > \tilde{T}^W(k), \quad (30)$$

where

$$\tilde{X}_j^W(k) = \frac{\left\{\sum_{i=1}^{k} W_k(i)\left[e_1(i) + \hat{\theta}_{1j}(i)u(i-j)\right]u(i-j)\right\}^2}{\sum_{i=1}^{k} W_k(i)u^2(i-j)}, \quad (31)$$

$$\tilde{T}^w(k) = \frac{2\log\left(L^w(k)\right)}{L^w(k)} \sum_{i=1}^{k} W_k(i)\left[\left(e_1(i)\right)\right]^2, \quad (32)$$

$$L^w(k) = \sum_{i=1}^{k} W_k(i), \quad (33)$$

and where $W_k(i)$ is the exponentially decay operator

$$W_k(i) = (1-\gamma)^{k-i} \quad 0 < \gamma \ll 1 \quad (34)$$

and $\hat{\theta}_{1j}(i)$ is the jth element of $\hat{\theta}_1(i)$ as defined in (8), (11), and $e_1(i)$ is as defined in (15).

We propose to apply this active detection criterion to the fast affine projection algorithm. This involves creating an $(n+1)\times(n+1)$ diagonal activity matrix $B(k)$, where the jth diagonal element $B_j(k) = 1$ if the jth tap index is detected as being active at sample instant k, otherwise $B_j(k) = 0$. This matrix is then applied within the FAP algorithm as follows.

Replace (5) with

$$\underline{\varepsilon}_d(k) = \left\{\left[B(k)U(k)\right]^T\left[B(k)U(K)\right] + \delta I\right\}^{-1}\underline{e}(k). \quad (35)$$

Replace (11) with

$$\underline{E}_d(k) = \begin{bmatrix} 0 \\ \overline{\underline{E}}_d(k-1) \end{bmatrix} + \underline{\varepsilon}_d(k). \quad (36)$$

Replace (8) with

$$\underline{\hat{\theta}}_d(k) = B(k)\underline{\hat{\theta}}_d(k-1) - \mu B(k)\underline{U}(k-N+1)E_{d,N-1}(k), \quad (37)$$

where

$$E_{d,N-1}(k) = \sum_{j=0}^{N-1} \varepsilon_{d},_j\,(k-N+1+j) \quad (38)$$

and $E_{d,j}(k)$ is the jth element of $\underline{\varepsilon}_d(k)$.

As with the detection-guided NLMS algorithm, a threshold scaling constant η may be introduced on the right-hand side of (32) based on different conditions. The effectiveness of this scaling constant is considered in the simulations.

3.4. Computational complexity

The proposed system requires $4(n+1)+4$ MPSI to perform the detection tasks required in the recursive equivalent of (30)–(33). By including the sparse diagonal matrix $B(k)$ in (37), the system only needs to include m multiplications rather than $(n+1)$ multiplications for (15) and (8). Thus, the proposed detection-guided FAP channel estimator requires $2m + 7N^2 + 5N + 4(n+1) + 4$ MPSI while the complexity of FAP is $2(n+1) + 7N^2 + 5N$ MPSI. Hence, for sufficiently long, low-dimensional active channels $n \gg m \geq 1$, $n \gg N$, the computational cost of the proposed detection-guided FAP channel estimator is essentially twice that of the FAP and of the standard NLMS estimators.

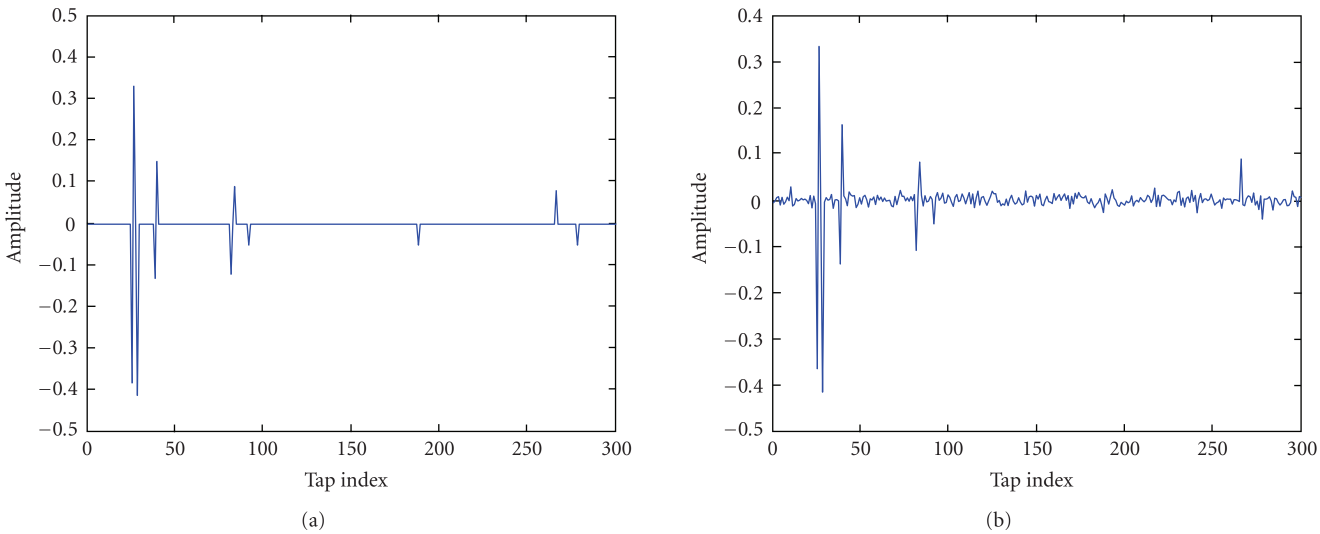

Figure 2: channel impulse response showing sparse structure: (a) is derived from the measured impulse response shown in (b) via the technique of the appendix.

4. SIMULATIONS

Simulations were carried out to investigate the performance of the following channel estimators when different input signals with different correlation levels are applied.

(A) Standard NLMS channel estimator.
(B) Active-parameter detection-guided NLMS channel estimator (as presented in Section 3.2).
(C) APA channel estimator with $N = 10$.
(D) FAP channel estimator with $N = 10$.
(E) Active-parameter detection-guided FAP channel estimator with $N = 10$ (without threshold scaling).
(F) Active-parameter detection-guided FAP channel estimator with $N = 10$, with threshold scaling constant.
(G) FAP channel estimator with $N = 14$. In this case, it has almost the same computational complexity[1] as that of the active-parameter detection-guided FAP channel estimator with $N = 10$.

Simulation conditions are the following.

(a) The channel impulse response considered, as given in Figure 2(a), was based on a real acoustic echo channel measurement made by CSIRO Radiophysics, Sydney, Australia. The impulse response of Figure 2(a) was derived from a measured acoustic echo path impulse response, Figure 2(b), by applying the technique based on the Dohono thresholding principle [14], as presented in the appendix. This technique essentially removes the effects of estimation/measurement noise. The measured impulse response of Figure 2(b) was obtained from a room approximately 5 m × 10 m × 3 m. The noise thresholded impulse response of Figure 2(a) consists of $m = 11$ active taps and a total tap length of $n = 300$.
The channel response used in the simulations is an example of a room acoustic impulse response which displays a sparse-like structure. Note, whether or not a room acoustic impulse response is sparse-like depends on the room configuration (size, placement of furniture, wall/floor coverings, microphone and speaker positioning). Nevertheless, a significant proportion of room acoustic impulse responses are, to varying degrees, sparse-like.
(b) Adaptive step size $\mu = 0.005$.
(c) Regularization parameter $\delta = 0.1$
(d) Initial channel estimate $\hat{\theta}(0)$ is the all-zero vector.
(e) Noise signal $v(k)$ = zero mean Gaussian process with variance of either 0.01 (Simulations 1 to 3) or 0.05 (Simulation 4).
(f) The squared channel estimator error $\|\theta - \hat{\theta}\|^2$ is plotted to compare the convergence rate. All plots are the average of 10 similar simulations.
(g) For the simulations of the detection-guided NLMS channel estimator and the detection-guided FAP channel estimator, the forgetting parameter $\gamma = 0.001$.

Simulation 1. Lowly correlated coloured input signal $u(k)$ described by the model $u(k) = w(k)/[1-0.1z^{-1}]$, where $w(k)$ is a discrete white Gaussian process with zero mean and unit variance.

Simulation 2. Highly correlated input signal $u(k)$ described by the model $u(k) = w(k)/[1 - 0.9z^{-1}]$, where $w(k)$ is a discrete white Gaussian process with zero mean and unit variance.

[1] The complexity is calculated based on the discussion in Section 3.4. The computational complexity of the active-parameter detection-guided FAP channel estimator with $N = 10$ is 1980 MPSI, which is slightly lower than the complexity of standard FAP with $N = 14$ of 2044 MPSI.

Simulation 3. Tenth-order AR-modelled speech input signal.

Simulation 4. Tenth-order AR-modelled speech input signal under noisy conditions. That is, with higher noise variance = 0.05.

In all four simulations, two detection-guided scaling constants were employed: $\eta = 1$ (i.e., no scaling) and $\eta = 4$.

5. RESULT AND ANALYSIS

Simulation 1 (lowly correlated input signal case). The results of the simulations for channel estimators (a) to (g) with $\mu = 0.005$ are shown in Figure 3.

(a) Channel estimators (b) to (f) show faster convergence than the standard NLMS channel estimator (a).
(b) The detection-guided NLMS estimator (b) provides faster convergence rate than the APA channel estimator (c) with $N = 10$ and the FAP channel estimator (d) with $N = 10$. It is clear that the APA channel estimator (c) with $N = 10$ and FAP channel estimator (d) with $N = 10$ still have not reached steady state at the 20000 sample mark.
(c) The detection-guided FAP channel estimators with $N = 10$ (e), (f) show a better convergence rate than channel estimators (b), (c), and (d).
(d) Detection-guided FAP estimator (e) and detection-guided FAP estimator with threshold scaling constant $\eta = 4$ (f) both can detect all the active taps and almost have the same performance.
(e) With almost the same computational cost, detection-guided FAP estimator (e) significantly outperforms standard FAP estimator with $N = 14$ in terms of convergence rate.

Simulation 2 (highly correlated input signal case). The results of the simulations for channel estimators (a) to (g) with $\mu = 0.005$ are shown in Figure 4.

(a) The active-parameter detection-guided NLMS channel estimator (b) does not provide suitably enhanced improved convergence speed over the standard NLMS channel estimator (a). This is due to the incorrect detection of many of the inactive taps with the highly correlated input signals.
(b) The APA channel estimator with $N = 10$ (c) and the FAP channel estimator with $N = 10$ (d) show significantly improved convergence over (a) and (b). This is due to the autocorrelation matrix inverse $[U(k)^T U(k) + \delta I]^{-1}$ in (5) essentially prewhitening the highly colored input signal.
(c) The detection-guided FAP channel estimators with $N = 10$ (e), (f) show better convergence rates than the standard APA channel estimator with $N = 10$ (c) and the standard FAP channel estimator with $N = 10$ (d). In addition, the detection-guided FAP estimators (e), (f) appear to provide better steady-state error performance.
(d) The detection-guided FAP channel estimator (e) without threshold scaling detects extra "nonactive" taps. In the simulation, it detects 32 active taps, which are 21 in excess of the true number. This leads to slower convergence rate. In comparison, the detection-guided FAP channel estimator (f) with threshold scaling $\eta = 4$, it shows the ability to detect the correct number of active taps, however, this comes with a relative initial error increase.
(e) The detection-guided FAP channel estimator (e) with $N = 10$ provides noticeably better convergence rate performance than the standard FAP channel estimator (d) with $N = 14$ in terms of the convergence rate and the steady-state error.

Simulation 3 (highly correlated speech input signal case). The results of the simulations for channel estimators (a) to (g) with $\mu = 0.005$ are shown in Figure 5. The trends shown here are similar to those of Simulations 1 and 2, although here the convergence rate and steady-state benefits provided by detection guiding are further accentuated.

(a) When the speech input signal is applied, the active parameter detection-guided NLMS channel estimator (b) suffers from very slow convergence, similar to that of the standard NLMS channel estimator (a). This is due to the incorrect detection of many of the inactive taps.
(b) The detection-guided FAP channel estimators (e) and (f) significantly outperform channel estimators (c) and (d) in terms of convergence speed. The results also indicate that the newly proposed detection-guided FAP estimators may have better steady state error performance than the standard APA and FAP estimators.
(c) For detection FAP estimator (e) and detection FAP estimator with threshold scaling constant $\eta = 4$ (f), the trends are similar to those observed for Simulation 2: detection FAP estimator (e) detects extra 23 active taps, resulting in reduced convergence rate and there is an initial error increase occurring in detection FAP estimator with threshold scaling constant $\eta = 4$ (f).
(d) Again, with the same computational cost, the detection-guided FAP channel estimator (e) with $N = 10$ shows a faster convergence rate and reduced steady state error relative to standard FAP channel estimator (d) with $N = 14$.

Simulation 4 (highly correlated speech input signal case with higher noise variance). The results of the simulations for channel estimators (a) to (g) with $\mu = 0.005$ are shown in Figure 6, which confirm the similar good performance of our newly proposed channel estimator under noisy conditions. The detection FAP estimator with threshold scaling constant $\eta = 4$ (f) performs noticeably better than the detection estimator FAP without threshold scaling (e) due to the ability to detect the correct number of active taps.

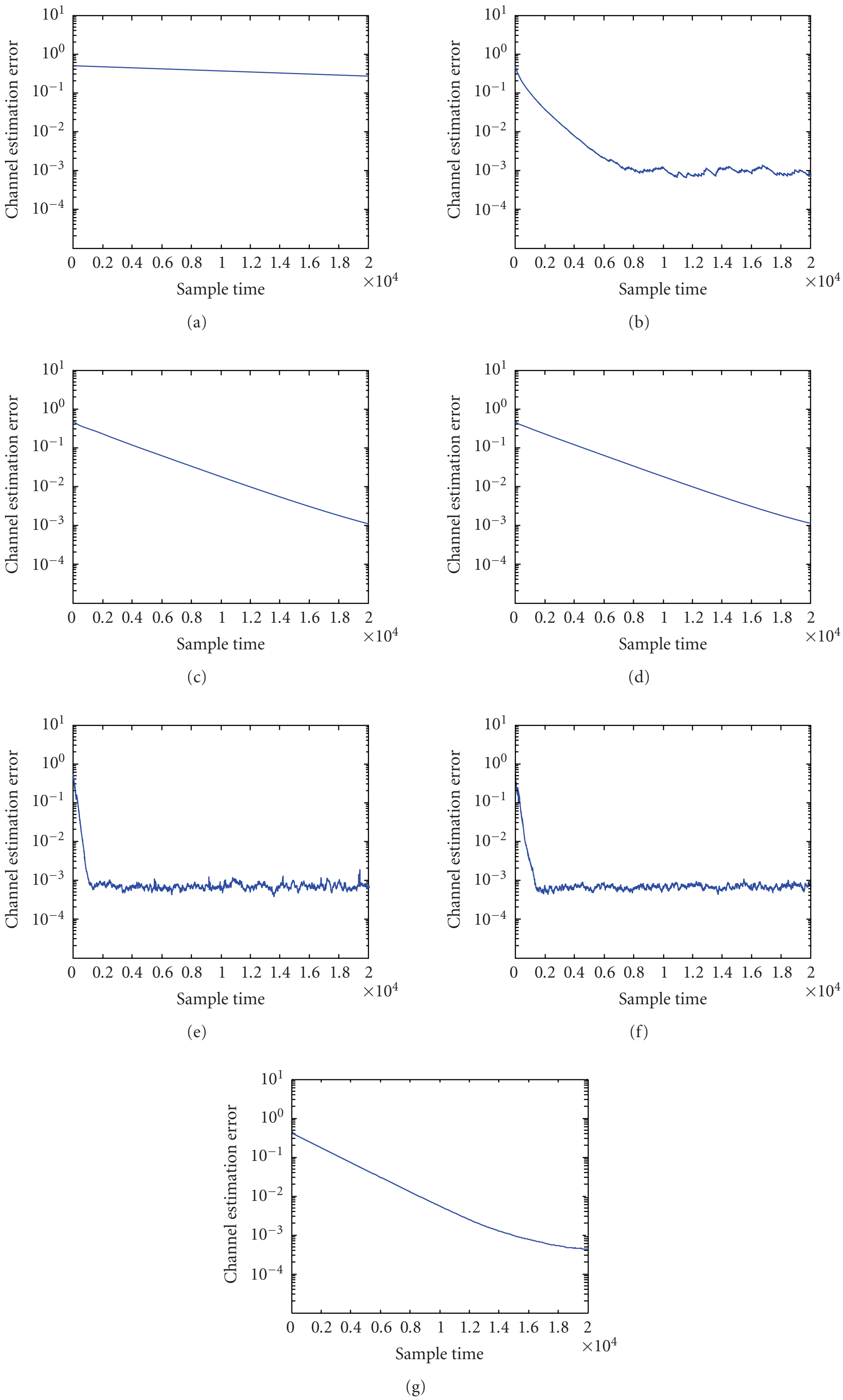

Figure 3: Comparison of convergence rates for lowly correlated input signal.

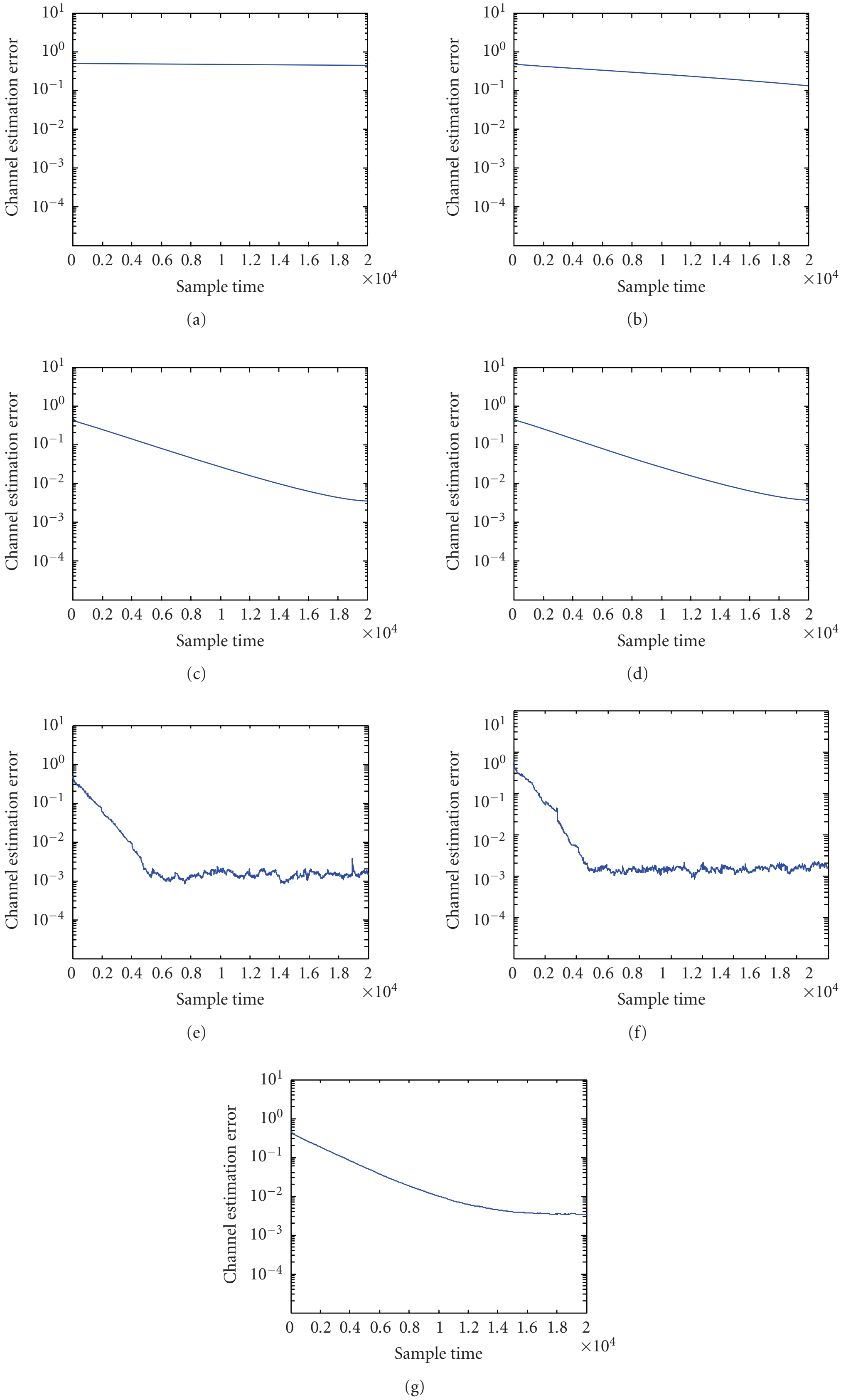

FIGURE 4: Comparison of convergence rates for highly correlated input signal.

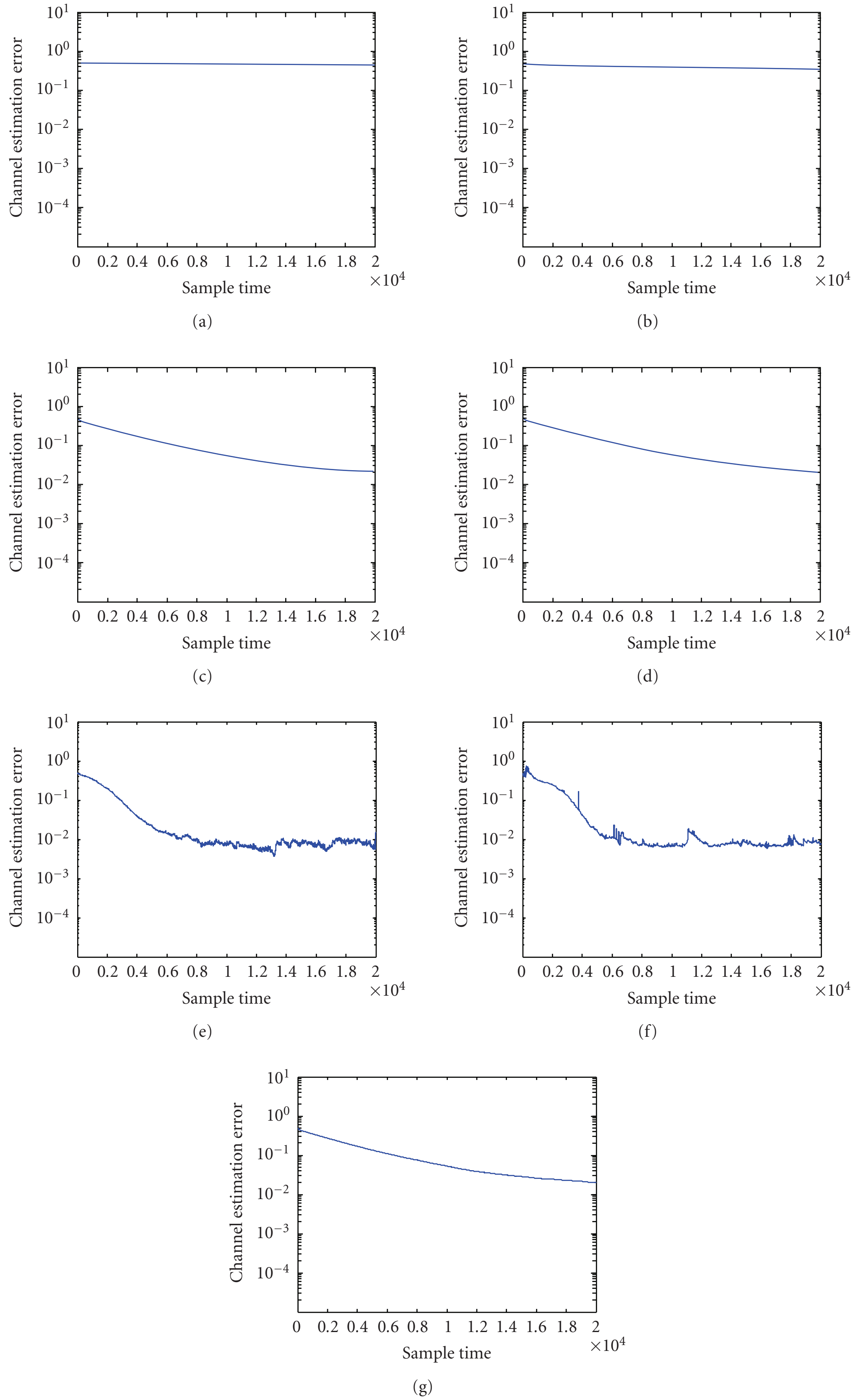

Figure 5: Comparison of convergence rates for speech input signal.

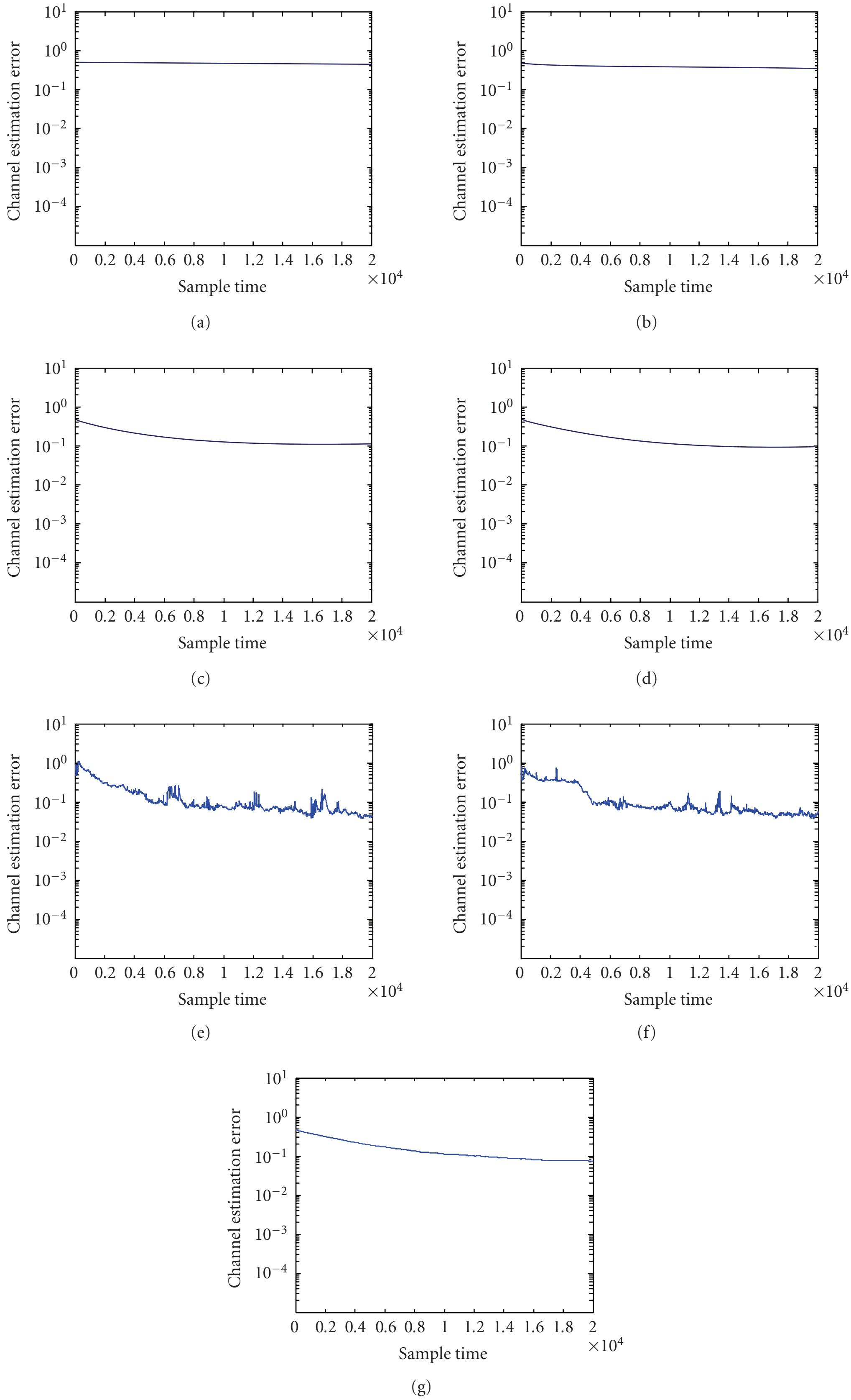

Figure 6: Comparison of convergence rates for speech input signal under noisy conditions.

6. CONCLUSION

For many adaptive estimation applications, such as acoustic echo cancellation within teleconferencing systems, the input signal is speech or highly correlated. In such applications, the standard NLMS channel estimator suffers from extremely slow convergence. To remove this weakness, the affine projection algorithm (APA) or the related computationally efficient fast affine projection (FAP) algorithm is commonly employed instead of the NLMS algorithm. Due to the signal propagation channels in such applications, sometimes having low dimensional or sparsely active impulse responses, we considered the incorporation of active-parameter detection with the FAP channel estimator. This newly proposed detection-guided FAP channel estimator is characterized with improved convergence speed and perhaps also better steady-state error performance as compared to the standard FAP estimator. The similar good performance is also achieved under noisy conditions. Additionally, simulations confirm these advantages of the proposed channel estimator under essentially the same computational cost. These features make this newly proposed channel estimator a good candidate for the adaptive estimation speech applications such as the acoustic echo cancellation problem.

APPENDICES

A. SPARSE CHANNEL IMPULSE RESPONSE ESTIMATION: REMOVING MEASUREMENT NOISE EFFECTS

In this appendix, a procedure for removing the measurements noise effect from the estimated time domain channel impulse response is presented. This procedure may be viewed as an offline scheme for active-tap detection of sparse channels and assumes that the true impulse response has a sufficiently large number of zero taps. Its applicability is restricted to channels which have a sparse structure.

In general, the presence of measurement noise or disturbance causes the tap coefficient estimate of each of the zero taps of the sparse channel to be nonzero. If we assume the estimate was obtained with a white input, then the discussion of Section 3 (more details can be found in [15]) suggests that asymptotically (at least for LS, LMS estimates) the zero-tap estimates have a zero mean i.i.d Gaussian distribution:

$$\{\hat{\theta}_i\} \sim N(0, \sigma^2), \quad \text{i.i.d, where } \theta_i = 0. \tag{A.1}$$

Under the validity of (A.1), we use the following results from the work of Donoho cited in [15], to develop a procedure for removing the effects of the noise, or, equivalently, for determining which taps are zero.

B. RESULT

Let $\{\hat{\theta}_i\} \sim N(0, \sigma^2)$, *i.i.d. Define the event* $A_M = \{\sup_{i \le M} |z_i| \le \sigma\sqrt{2\log M}\}$, *Then* , $\text{Prob}(A_M) \to 1$ as $M \to \infty$.

A priori knowledge of the indices i of the zero taps is required in order to use the threshold $\sigma\sqrt{2\log M}$ to determine which taps are zero. By applying the following iterative procedure, this requirement is avoided for sparse channels.

Algorithm 1. (1) Initially, include the indices of all n tap estimates $\{\hat{\theta}_i\}$ in the set S of zero taps and set $M = n$.

(2) Determine rms value σ_S of the estimates of the taps in Set S.

(3) Determine the indices i of those taps for which the estimates coefficients satisfy

$$|\hat{\theta}_i| \le \sigma_S\sqrt{2\log M}. \tag{B.1}$$

(4) Repeat steps (2) and (3) a given number of times or, alternatively, until the difference in σ_S from one iteration to the next has decreased to a given value.

ACKNOWLEDGMENT

The authors would like to acknowledge CSIRO Rdiophysics, Sydney for providing the measurement data of the simulation channel.

REFERENCES

[1] K. Ozeki and T. Umeda, "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties," *Electronics & Communications in Japan*, vol. 67, no. 5, pp. 19–27, 1984.

[2] S. L. Gay and S. Tavathia, "The fast affine projection algorithm," in *Proceedings of the 20th International Conference on Acoustics, Speech, and Signal Processing (ICASSP '95)*, vol. 5, pp. 3023–3026, Detroit, Mich, USA, May 1995.

[3] J. R. Casar-Corredera and J. Alcazar-Fernandez, "An acoustic echo canceller for teleconference systems," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '86)*, vol. 11, pp. 1317–1320, Tokyo, Japan, April 1986.

[4] A. Gilloire and J. Zurcher, "Achieving the control of the acoustic echo in audio terminals," in *Proceedings of European Signal Processing Conference (EUSIPCO '88)*, pp. 491–494, Grenoble, France, September 1988.

[5] S. Makino and S. Shimada, "Echo control in telecommunicaitons," *Journal of the Acoustic Society of Japan*, vol. 11, no. 6, pp. 309–316, 1990.

[6] J. Homer, I. Mareels, R. R. Bitmead, B. Wahlberg, and A. Gustafsson, "LMS estimation via structural detection," *IEEE Transactions on Signal Processing*, vol. 46, no. 10, pp. 2651–2663, 1998.

[7] J. Homer, "Detection guided NLMS estimation of sparsely parametrized channels," *IEEE Transactions on Circuits and Systems II*, vol. 47, no. 12, pp. 1437–1442, 2000.

[8] J. Homer, I. Mareels, and C. Hoang, "Enhanced detection-guided NLMS estimation of sparse FIR-modeled signal channels," *IEEE Transactions on Circuits and Systems I*, vol. 53, no. 8, pp. 1783–1791, 2006.

[9] S. Haykin, *Adaptive Filter Theory*, Prentice Hall Information and System Science Series, Prentice-Hall, Upper Saddle River, NJ, USA, 3rd edition, 1996.

[10] M. Bouchard, "Multichannel affine and fast affine projection algorithms for active noise control and acoustic equalization systems," *IEEE Transactions on Speech and Audio Processing*, vol. 11, no. 1, pp. 54–60, 2003.

[11] S. G. Sankaran and A. A. Beex, "Convergence behavior of affine projection algorithms," *IEEE Transactions on Signal Processing*, vol. 48, no. 4, pp. 1086–1096, 2000.

[12] G. Rombouts and M. Moonen, "A sparse block exact affine projection algorithm," *IEEE Transactions on Speech and Audio Processing*, vol. 10, no. 2, pp. 100–108, 2002.

[13] G. Rombouts and M. Moonen, "A fast exact frequency domain implementation of the exponentially windowed affine projection algorithm," in *Proceedings of IEEE Adaptive Systems for Signal Processing, Communications, and Control Symposium (AS-SPCC '00)*, pp. 342–346, Lake Louise, Alta., Canada, October 2000.

[14] M. R. Leadbetter, G. Lindgren, and H. Rootzen, *Extremes and Related Properties of Random Sequences and Processes*, Springer, New York, NY, USA, 1982.

[15] H. Cramer and M. R. Leadbetter, *Stationary and Related Stochastic Srocesses: Sample Function Properties and Their Applications*, John Wiley & Sons, New York, NY, USA, 1967.