DECISION FEEDBACK BLIND SYMBOL ESTIMATION BY ADAPTIVE LEAST SQUARES SMOOTHING

Qing Zhao and Lang Tong

Department of Electrical Engineering
Cornell University Ithaca, NY 14853

Abstract

A decision feedback blind symbol estimation algorithm based on the least squares smoothing approach is proposed for single-input multiple-output finite impulse response systems. With the finite alphabet property, the input signal is estimated based on the past detected symbols and the least squares smoothing error of the observation. Implemented both time and order recursively, the proposed algorithm is adaptive to channel variation and has low complexity both in computation and in VLSI implementation. Based on a deterministic model, this algorithm has the finite sample convergence property, i.e., the input signal can be perfectly detected with a small set of data samples in the absence of noise

1. INTRODUCTION

One important requirement on blind channel equalization is the convergence rate. Methods based on the statistical modeling, such as the constant modulus algorithm [2], the subspace-based MMSE detector [1], rely on the convergence of statistics. Therefore, a large set of data samples are required in these methods. Deterministic algorithms, on the other hand, has the *finite sample convergence* property, *i.e.*, the channel or the input symbols can be perfectly estimated using a finite number of samples in the absence of noise. However, the focus of most deterministic algorithms, such as the subspace algorithm (SS) [5], the cross relation (CR) algorithm [8], the joint order detection and channel estimation algorithm (J-LSS) [6], and the adaptive least squares smoothing channel estimation algorithm (A-LSS) [9] has been on the channel identification. The input signal is estimated in the second step based on the identified channel. Perhaps the main drawback of this two-step approach to equalization is the performance degradation of the symbol detector when the estimated channel is not accurate.

Liu and Xu [4], Van der Veen *et al.* [7] proposed subspace-based direct symbol estimation algorithms. By exploiting the connection between the input and output data structures, input symbols can be perfectly estimated from a finite number of output samples. However, relying on the singular value decomposition of the data matrix, these algorithms have high computation cost and do not have an efficient adaptive implementation.

In this paper, we propose a decision feedback blind symbol estimation algorithm that is fast convergent, adaptive and computationally efficient. This algorithm has two components: a least squares smoother applied to the observation process and a decision feedback detector that takes advantage of the finite alphabet structure of the input signal. The smoothing error of the input signal is obtained from the observation by the least squares smoother based on the isomorphic relation between the input and output subspaces. The input signal is then detected from its smoothing error by the decision feedback detector. It turns out that the smoothing error of the input signal has no channel-induced intersymbol interference. Therefore, the knowledge of channel is no longer necessary for the decision feedback technique. This feature distinguishs the proposed algorithm from the non-blind decision feedback equalizer. Like all deterministic methods, the proposed algorithm has the finite sample convergence property. Furthermore, this least squares smoothing approach (LSS) naturely leads to an order and time adaptive implementation without matrix operations. The efficient adaptive implementation enables the proposed algorithm to track a wide range of channel variation, both in parameters and in length. Also, the low computation complexity and regular structure of this algorithm make it a good candidate for VLSI implementation.

Notations used in this paper are mostly standard. Upper- and lower-case bold letters denote matrices and vectors, respectively with $(\cdot)^t$ and $(\cdot)'$ denoting transpose and Hermitian operations. Given a matrix $\mathbf{A}$, $\mathcal{R}\{\mathbf{A}\}$ is the range space of $\mathbf{A}$. For a set of vectors $\mathbf{x}_1, \cdots, \mathbf{x}_n$, $sp\{\mathbf{x}_1, \cdots, \mathbf{x}_n\}$ denotes the linear subspace spanned by $\mathbf{x}_1, \cdots, \mathbf{x}_n$. For a vector $\mathbf{x}$ and a linear subspace $\mathcal{X}$, $\check{\mathbf{x}}_{|\mathcal{X}} \stackrel{\Delta}{=} \arg\min_{\mathbf{z}\in\mathcal{X}} ||\mathbf{x}-\mathbf{z}||^2$ is the orthogonal projection of $\mathbf{x}$ onto $\mathcal{X}$, and $\tilde{\mathbf{x}}_{|\mathcal{X}} \stackrel{\Delta}{=} \mathbf{x}-\check{\mathbf{x}}_{|\mathcal{X}}$ is its projection error. Finally, $||\cdot||$ denotes the 2-norm.

This work was supported in part by the National Science Foundation under Contract NCR-9321813 and by the Office of Naval Research under Contract N00014-96-1-0895.

2. PROBLEM STATEMENT

Considered in this paper is the estimation of the input signal $\{s(t)\}$ to a single-input P-output linear system given by

$$\mathbf{x}(t) = \sum_{i=0}^{L} \mathbf{h}_i s_{t-i}, \quad \mathbf{y}(t) = \mathbf{x}(t) + \mathbf{n}(t), \ \ t = 0, \cdots, \quad (1)$$

where $\mathbf{x}(t) = [x^{(1)}(t), \cdots, x^{(P)}(t)]^t$ is the output of P subchannels, $\mathbf{n}(t)$ is the additive noise, $\mathbf{y}(t)$ is the received signal, and $\{\mathbf{h}_t\}$ is the channel impulse response. Given N samples of the system input and output, define $\mathbf{s}_t \triangleq [s(t), \cdots, s(t-N+1)]^t$, $\mathbf{x}_t \triangleq [\mathbf{x}(t), \cdots, \mathbf{x}(t-N+1)]^t$. The input and output subspaces spanned by p consecutive vectors are then defined as

$$\mathcal{S}_{t,p} \triangleq sp\{\mathbf{s}_t, \cdots, \mathbf{s}_{t-p+1}\}, \quad \mathcal{X}_{t,p} \triangleq sp\{\mathbf{x}_t, \cdots, \mathbf{x}_{t-p+1}\}. \quad (2)$$

Note that for $p < 0$, the subspaces are spanned by p consecutive future data vectors. Our goal is to estimate $\mathbf{s}_t$ from $\mathbf{y}_t$.

Two assumptions are made in this paper.

- A1: <u>Channel Diversity:</u> The subchannel transfer functions do not share common zeros, *i.e.*, , $\{\mathbf{h}_i(z)\}$ are co-prime.

The following property results directly from A1.

Property 1 *Under A1, there exists a (smallest) $w_0 \geq \frac{L}{P-1}$ such that for any $w \geq w_0$, we have the isomorphic relation between the input and output subspaces: $\mathcal{X}_{t,w} = \mathcal{S}_{t,w+L}$.*

The second assumption is about the linear complexity of the input sequence.

- A2: <u>Linear Complexity:</u> The input sequence $\{s(t)\}$ has linear complexity greater than $2w_0 + 2L$.

.

3. THE LEAST SQUARES SMOOTHING

The isomorphism between the input and output subspaces leads to a specific relation between the smoothing error of the input and the smoothing error of the output, as given in the following theorem.

Theorem 1 *Consider $L+1$ consecutive output data vectors $\mathbf{x}_{t+L}, \cdots, \mathbf{x}_t$ and the input data vector $\mathbf{s}_t$. For $w \geq w_0$, define the input and the output subspace as $\mathcal{Z}_s(t) \triangleq \mathcal{S}_{t-1,w+L} \cup \mathcal{S}_{t+1,-(w+L)}$ and $\mathcal{Z}_x(t) \triangleq \mathcal{X}_{t-1,w} \cup \mathcal{X}_{t+L+1,-w}$, respectively. We have the following relation between the smoothing error of $\mathbf{x}_{t+L}, \cdots, \mathbf{x}_t$ and the smoothing error of $\mathbf{s}_t$:*

$$\begin{aligned} \mathbf{E} &\triangleq [\tilde{\mathbf{x}}_{t+L|\mathcal{Z}_x(t)}, \cdots, \tilde{\mathbf{x}}_{t|\mathcal{Z}_x(t)}] & (3) \\ &= \tilde{\mathbf{s}}_{t|\mathcal{Z}_s(t)} [\mathbf{h}_L^t, \cdots, \mathbf{h}_0^t]. & (4) \end{aligned}$$

The above theorem results from the isomorphism between the input and output subspaces: $\mathcal{Z}_s(t) = \mathcal{Z}_x(t)$. Detailed proof was given in [9] [6]. This theorem implies that the smoothing error $\tilde{\mathbf{s}}_{t|\mathcal{Z}_s(t)}$ of the input can be obtained (up to a scalaring factor) from the smoothing error $\mathbf{E}$ of the output without knowing the input sequence. It is also important to note that $\tilde{\mathbf{s}}_{t|\mathcal{Z}_s(t)}$ is independent of the channel, *i.e.*, the least squears smoother removes the channel effect. The following question arises: *can we obtain the input signal from its smoothing error?* An answer to this question is presented in Section 4.

4. SYMBOL DETECTION FROM THE SMOOTHING ERROR

4.1. The Basic Idea

From (4) we can see each column in $\mathbf{E}$ is an estimate of $\tilde{\mathbf{s}}_{t|\mathcal{Z}_s(t)}$. For an uncorrelated input sequence, $\tilde{\mathbf{s}}_{t|\mathcal{Z}_s(t)}$ converges to $\mathbf{s}_t$ when N goes to infinity. However, in this case $\mathbf{s}_t$ can only be perfectly estimated with infinite number of samples even when there is no noise. Can we obtain $\mathbf{s}_t$ from $\tilde{\mathbf{s}}_{t|\mathcal{Z}_s(t)}$ while preserving the finite sample convergence property? The finite alphabet structure of the input signal can help us to answer this question.

Suppose that the input symbols belong to a finite alphabet $\mathcal{A} = \{a^1, \cdots, a^M\}$. One way to obtain $\mathbf{s}_t$ is to search for a vector whose elements belong to $\mathcal{A}$ to fit $\tilde{\mathbf{s}}_{t|\mathcal{Z}_s(t)}$. This approach involves the computation of M^N projections, which is obviously impractical even for a small set of data samples. However, if we have detected all the input symbols up to time $t-1$, then only one symbol in $\mathbf{s}_t$ needs to be determined. Therefore, to detect the current symbol, only M projection errors needs to be calculated and compared with $\tilde{\mathbf{s}}_{t|\mathcal{Z}_s(t)}$. Furthermore, these projection errors can be calculated both time and order recursively, as presented in the next section.

4.2. Adaptive Decision Feedback Symbol Estimation

For simplicity and without loss of generality, here we only consider obtaining $\tilde{\mathbf{s}}_{t|\mathcal{Z}_s(t)}$ from one observation vector, for example, $\mathbf{y}_t^{(1)}$.

To obtain $\tilde{\mathbf{s}}_{t|\mathcal{Z}_s(t)}$, the output subspace $\mathcal{Z}_x(t)$ is estimated from the observation: $\hat{\mathcal{Z}}_x(t) \triangleq \mathcal{R}\{\mathbf{Z}(t)\}$ with $\mathbf{Z}(t)$

defined as

$$\begin{pmatrix} \mathbf{y}(0)^t & \cdots & \mathbf{y}(w-1)^t & \mathbf{y}(w+L+1)^t & \cdots & \mathbf{y}(2w+L)^t \\ \mathbf{y}(1)^t & \cdots & \mathbf{y}(w)^t & \mathbf{y}(w+L+2)^t & \cdots & \mathbf{y}(2w+L+1)^t \\ \vdots & & \vdots & \vdots & & \vdots \\ \mathbf{y}(t-w)^t & \cdots & \mathbf{y}(t-1)^t & \mathbf{y}(t+L+1)^t & \cdots & \mathbf{y}(t+L+w)^t \end{pmatrix}, \quad (5)$$

where w is the smoothing order satisfying Property 1. Note that the knowledge of the channel order is required in constructing $\mathbf{Z}(t)$. The extension to the unknown order case is discussed in [10]. Define

$$\mathbf{y}_t^{(1)} \triangleq [y^{(1)}(w), y^{(1)}(w+1), \cdots, y^{(1)}(t)]^t \quad (6)$$

$$\mathbf{s}_t^i \triangleq [\hat{s}(w), \hat{s}(w+1), \cdots, \hat{s}(t-1), a^i]^t, i = 1, \cdots, M, \quad (7)$$

where $\mathbf{s}_t^i$ contains the past detected symbols and the tentative guess a^i of the current symbol. In order to obtain $\hat{s}(t)$, the projection error of $\mathbf{y}_t^{(1)}$ and $\mathbf{s}_t^i (i = 1, \cdots, M)$ onto $\hat{\mathcal{Z}}_x(t)$ needs to be calculated. Suppose that at time t, we have computed the projection errors $\tilde{\mathbf{y}}^{(1)}_{t|\hat{\mathcal{Z}}_x(t)}$ and $\tilde{\mathbf{s}}^i_{t|\hat{\mathcal{Z}}_x(t)} (i = 1, \cdots, M)$. The current symbol is then detected by choosing the vector $\tilde{\mathbf{s}}^k_{t|\hat{\mathcal{Z}}_x(t)} (1 \le k \le M)$ that contains the smallest angle with $\tilde{\mathbf{y}}^{(1)}_{t|\hat{\mathcal{Z}}_x(t)}$, *i.e.*, $\hat{s}(t) = a^k$. With the definition

$$\hat{\mathbf{s}}_t \triangleq [\hat{s}(w), \hat{s}(w+1), \cdots, \hat{s}(t)]^t, \quad (8)$$

it is straightforward that $\tilde{\hat{\mathbf{s}}}_{t|\hat{\mathcal{Z}}_x(t)} = \tilde{\mathbf{s}}^k_{t|\hat{\mathcal{Z}}_x(t)}$ which has been calculated. Now consider at time $t+1$ when a new observation $\mathbf{y}(t+L+w+1)$ becomes available. Then $\mathbf{Z}(t+1)$ is constructed by enhancing $\mathbf{Z}(t)$ with one row. Similarly, $\mathbf{y}^{(1)}_{t+1} = [\mathbf{y}_t^{(1)t}, y^{(1)}(t+1)]^t$, $\mathbf{s}^i_{t+1} = [\hat{\mathbf{s}}^t_t, a^i]^t (i = 1, \cdots, M)$. Because only the last elements in the projection error $\tilde{\mathbf{y}}^{(1)}_{t+1|\hat{\mathcal{Z}}_x(t+1)}$ and $\tilde{\mathbf{s}}^i_{t+1|\hat{\mathcal{Z}}_x(t+1)} (i = 1, \cdots, M)$ are the necessary quantities to detect $s(t+1)$ and we have already obtained $\tilde{\mathbf{y}}^{(1)}_{t|\hat{\mathcal{Z}}_x(t)}$ and $\tilde{\hat{\mathbf{s}}}_{t|\hat{\mathcal{Z}}_x(t)}$ at time t, we can take advantage of the previous results without repeat the entire computation. This is a classical least squares problem which has been well-investigated. We can use recursive least squares algorithms, for example, the Recursive Modified Gram-Schmidt (RMGS) algorithm [3] to implement the proposed decision feedback symbol estimator both time and order recursively. One implementation is summarized below, where $\tilde{y}^{(1)}_{|\hat{\mathcal{Z}}_x(t)}(t)$, $\tilde{\hat{s}}_{|\hat{\mathcal{Z}}_x(t)}(t)$, $\tilde{a}^i_{|\hat{\mathcal{Z}}_x(t)}$ denote the last element in $\tilde{\mathbf{y}}^{(1)}_{t|\hat{\mathcal{Z}}_x(t)}$, $\tilde{\hat{\mathbf{s}}}_{t|\hat{\mathcal{Z}}_x(t)}(t)$ $\tilde{\mathbf{s}}^i_{t|\hat{\mathcal{Z}}_x(t)} (i = 1, \cdots, M)$, respectively. We have also defined $\mathcal{N}_y(t) \triangleq \sum_{n=w}^t |\tilde{y}^{(1)}_{|\hat{\mathcal{Z}}_x(n)}(n)|^2, \mathcal{N}_{\hat{s}}(t) \triangleq \sum_{n=w}^t |\tilde{\hat{s}}_{|\hat{\mathcal{Z}}_x(n)}(n)|^2$, and $\mathcal{N}_{y,\hat{s}}(t) \triangleq \sum_{n=w}^t \tilde{y}^{(1)}_{|\hat{\mathcal{Z}}_x(n)}(n) \cdot \tilde{\hat{s}}'_{|\hat{\mathcal{Z}}_x(n)}(n)$.

The Decision Feedback Symbol Estimator by LSS

At time t:

1. Compute $\tilde{y}^{(1)}_{|\hat{\mathcal{Z}}_x(t)}(t)$, $\tilde{a}^i_{|\hat{\mathcal{Z}}_x(t)} (i = 1, \cdots, M)$ based on the results obtained at $t-1$.
2. Compute
$$\begin{aligned} \mathcal{N}_y(t) &= \mathcal{N}_y(t-1) + |\tilde{y}^{(1)}_{|\hat{\mathcal{Z}}_x(t)}(t)|^2, \\ \mathcal{N}_{s^i}(t) &= \mathcal{N}_{\hat{s}}(t-1) + |\tilde{a}^i_{|\hat{\mathcal{Z}}_x(t)}|^2, \\ \mathcal{N}_{y,s^i}(t) &= \mathcal{N}_{y,\hat{s}}(t-1) + \tilde{y}^{(1)}_{|\hat{\mathcal{Z}}_x(t)}(t)\tilde{a}^i_{|\hat{\mathcal{Z}}_x(t)}{}'. \end{aligned}$$
3. Compute
$$\cos\theta^i \triangleq \frac{\mathcal{N}_{y,s^i}(t)}{\sqrt{\mathcal{N}_y(t) \cdot \mathcal{N}_{s^i}(t)}} (i = 1, \cdots, M). \quad (9)$$
4. $k = \arg\max_{i=1,\cdots,M} \cos\theta^i$.
5. $\hat{s}(t) = a^k, \mathcal{N}_{\hat{s}}(t) = \mathcal{N}_{s^k}(t), \mathcal{N}_{y,\hat{s}}(t) = \mathcal{N}_{y,s^k}(t)$.

To gain a better understanding of this adaptive decision feedback symbol estimation algorithm, we make the following remarks.

1. It can be shown [9][6] that for $\mathcal{Z}_x(t) = \mathcal{S}_{t-1,L} \cup \mathcal{X}_{t+L+1,-w}$, Theorem 1 still holds. Therefore, the past detected symbols can be used to construct the projection space $\mathcal{Z}_x(t)$. With part of the projection space constructed from the input symbols directly, the dimension of the projection space is reduced, which leads to a lower computation cost.

2. In addition to the benefit of low complexity, the time- and order-recursive property also enables the proposed symbol estimator to track the channel variation. Furthermore, the order-recursive property is very useful in joint channel order and symbol detection [10].

3. This algorithm can be implemented with the same basic modules as in A-LSS [9]. The regular structure makes the proposed algorithm suitable for VISI implementation.

5. SIMULATION

Presented in this section are simulation studies of the proposed decision feedback symbol estimator by adaptive least squares smoothing (DF-LSS) as it was compared with the subspace-based symbol estimator (SS-s)[4]. We also compared the proposed algorithm with the MMSE FIR decision feedback equalizer based on the channel estimated by the subspace channel estimation algorithm [5] (SS-DFE). The channel was estimated with 100 data samples. The forward and backward filter of the decision feedback equalizer have order $2L$ and L, respectively with L denoting the channel order.

The channel used in this simulation was a three-ray $T/2$-spaced channel with gains $[1, 1, 1]$ and delays $[0.5T, T, 1.5T]$.

The pulse-shaping filter was a square-root raised-cosine function with 25% roll-off and being truncated at $\pm 3T$. Noise samples were generated from i.i.d. zero mean Gaussian random sequences and the signal-to-noise ratio (SNR) was defined as $SNR \triangleq \frac{E\{||\mathbf{x}_t||^2\}}{E\{||\mathbf{n}_t||^2\}}$. The input was generated from an i.i.d. binary phase shift keying (BPSK) sequence.

Figure 1 shows the bit error rate (BER) of DF-LSS as it was compared with SS-s and SS-DFE. From Figure 1 we can see that SS-s performed rather poorly for this ill-conditioned channel. Based on an inaccurately estimated channel, the MMSE FIR decision feedback equalizer suffered from a considerable performance degradation. The proposed direct symbol estimator had a clear advantage in performance for this ill-conditioned channel.

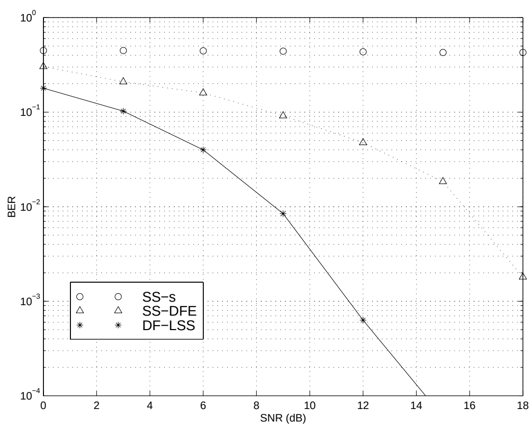

Figure 1: Performance Comparison

6. CONCLUSION

A decision feedback blind symbol estimation algorithm is proposed. Based on the least squares smoothing approach and taking advantage of the finite alphabet structure of the input signal, the proposed algorithm has several desirable properties including adaptivity to both channel order and channel parameter variation, fast convergence, good computational efficiency, and modular structure suitable for VLSI implementation. This algorithm can also be extended to the case when the channel order is unknown [10]. In this case, the channel order and the input signal are jointly detected to fit the smoothing error of the input signal.

7. REFERENCES

[1] D. Gesbert, J. Sorelius, and A. Paulraj. "Blind Multi-user MMSE Detection of CDMA Signals". In *Proc. ICASSP-98*, volume 6, pages 3161–3164, May 1998.

[2] D.N. Godard. "Self-recovering equalization and carrier tracking in two-dimensional data communication systems". *IEEE Trans. Commun.*, COM-28: 1867-1875, November 1980.

[3] F. Ling and J.G. Proakis. "A Generalized Multichannel Least Squares Lattice Algorithm based on Sequential Processing Stages". *IEEE Trans. ASSP*, ASSP-32(2):381–390, April 1984.

[4] H. Liu and G. Xu. "Closed-form blind symbol estimation in digital communications". *IEEE Trans. on Signal Processing*, 43(11):2714–2723, Nov. 1995.

[5] E. Moulines, P. Duhamel, J.F. Cardoso, and S. Mayrargue. "Subspace-Methods for the Blind Identification of Multichannel FIR Filters". *IEEE Trans. SP*, SP-43(2):516–525, Feb. 1995.

[6] L. Tong and Q. Zhao. "Joint Order Detection and Channel Estimation by Least Squares Smoothing". In *Proc. 30th Conf. Information Sciences and Systems*, Princeton, NJ., Mar. 1998.

[7] A. van der Veen, S. Talwar, and A. Paulraj. "A subspace approach to blind space-time signal processing for wireless communication systems". *IEEE Trans. Signal Processing*, SP-45(1):173–190, Jan 1997.

[8] G. Xu, H. Liu, L. Tong, and T. Kailath. "A Least-Squares Approach to Blind Channel Identification". *IEEE Trans. Signal Processing*, SP-43(12):2982–2993, December 1995.

[9] Q. Zhao and L. Tong. "Fast Convergent Adaptive Blind Channel Estimation and Deconvolution". In *Proc. 30th Conf. Information Sciences and Systems*, Princeton, NJ., Mar. 1998.

[10] Q. Zhao and L. Tong. "Joint Channel Order and Symbol Detection with Decision Feedback". *Signal Processing Workshop on Signal Processing Advances in Wireless Communications*, Annapolis, Maryland, May 9-12, 1999.