SECOND ORDER STATISTICS BASED BLIND CHANNEL EQUALIZATION WITH CORRELATED SOURCES

Roberto López-Valcarce and Soura Dasgupta

Department of Electrical and Computer Engineering
University of Iowa, IA-52242, USA
e-mail: valcarce@icaen.uiowa.edu, dasgupta@eng.uiowa.edu

ABSTRACT

In this paper the blind equalization method of Tong, Xu and Kailath, which was derived under the assumption of a white transmitted sequence, is generalized in order to include correlated sources. The resulting algorithm can be applied to arbitrarily colored sources, assuming that the source second-order statistics are known. Because it makes explicit use of this information, the new algorithm is expected to outperform subspace based methods which neither require nor exploit such knowledge. This is verified via simulations.

1. INTRODUCTION

We consider the blind equalization / estimation of 1-input p-output channels, from the second order statistics (SOS) of the channel output. It is assumed that the channel input statistics *are colored but known.* The channel model is

$$x_n = \sum_{i=0}^{l} h_i a_{n-i} + w_n, \tag{1}$$

where $\{a_n\}$ is the zero mean, wide sense stationary sequence of transmitted symbols, and x_n, h_n, w_n are $p \times 1$ vectors of channel outputs, channel impulse response and white noise respectively. Such a multi-channel model may arise through the deployment of multiple sensors or through fractional sampling when the continuous-time channel has excess bandwidth [6]. A typical reformulation of this problem [6] involves the vector processes

$$\begin{aligned}
\mathbf{x}_n &\triangleq [\; x_n^T \quad x_{n-1}^T \quad \cdots \quad x_{n-m+1}^T \;]^T, \\
\mathbf{w}_n &\triangleq [\; w_n^T \quad w_{n-1}^T \quad \cdots \quad w_{n-m+1}^T \;]^T, \\
\mathbf{s}_n &\triangleq [\; a_n \quad a_{n-1} \quad \cdots \quad a_{n-d+1} \;]^T,
\end{aligned}$$

where $d = m + l$. These vectors are related via

$$\mathbf{x}_n = \mathcal{H}\mathbf{s}_n + \mathbf{w}_n, \tag{2}$$

Supported by NSF grants ECS-9970105 and CCR-9973133.

where $\mathcal{H}$ is an $mp \times (m + l)$ generalized Sylvester matrix:

$$\mathcal{H} \triangleq \begin{bmatrix} h_0 & \cdots & h_l & & 0 \\ & \ddots & & \ddots & \\ 0 & & h_0 & \cdots & h_l \end{bmatrix}. \tag{3}$$

Two types of techniques that have received significant attention are the so called subspace based algorithms, pioneered by [4], and the genre of algorithms initiated by [6]. Both assume that $\mathcal{H}$ has full column rank, and use the SOS of the channel output. Subspace methods have the advantage of not requiring knowledge of the channel input statistics. However, in many instances such knowledge is available, and which if exploited should result in equalizers with better performance. The method of [6] does exploit this knowledge, but assumes the sequence $\{a_n\}$ is white. It uses the lag zero and lag one autocorrelations of the channel output,

$$\mathcal{R}_x(0) \triangleq E[\mathbf{x}_n \mathbf{x}_n^H], \qquad \mathcal{R}_x(1) \triangleq E[\mathbf{x}_n \mathbf{x}_{n-1}^H], \tag{4}$$

to estimate $\mathcal{H}$ to within an unitary scaling constant. Here $(\cdot)^H$ indicates conjugate transpose.

The whiteness assumption on $\{a_n\}$ is crucial for the algorithm from [6]. This was noted in [3] where a modification was proposed in order to deal with weakly correlated sources with unknown correlation (which must then be estimated). Another approach is found in [1], which uses the information contained in $\mathcal{R}_x(0)$, $\mathcal{R}_x(1)$, $\ldots$, $\mathcal{R}_x(d)$ with the corresponding increase in computational complexity, By contrast, our algorithm is a natural extension of the original method from [6], making use of $\mathcal{R}_x(0)$ and $\mathcal{R}_x(1)$ only. It is shown via simulations that the new algorithm outperforms the subspace approach of [4].

We denote by $\mathbf{J}$ the square shift matrix with ones in the first subdiagonal and zeros elsewhere; $\mathbf{e}_k$ denotes the k-th unit vector. $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ denote respectively the conjugate, the transpose, and the conjugate transpose.

2. PRELIMINARIES

Define the $d \times d$ lag i autocorrelation matrix of $\mathbf{s}_n$ as

$$\mathcal{R}_s(i) \triangleq E[\mathbf{s}_n \mathbf{s}_{n-i}^H],$$

Let $\sigma_a^2 = E[|a_n|^2]$ and define also the autocorrelation vector

$$\mathbf{r}^H \triangleq \mathbf{e}_1^H \mathcal{R}_s(1) = E[a_n \mathbf{s}_{n-1}^H]. \quad (5)$$

We shall make the following standard assumptions.

Assumption 1 $\begin{bmatrix} \sigma_a^2 & \mathbf{r}^H \\ \mathbf{r} & \mathcal{R}_s(0) \end{bmatrix} > 0.$

Assumption 2 *$\mathcal{H}$ is tall and has full column rank.*

We shall assume $\mathbf{w}_n = 0$, as the white noise component can be subtracted from the output autocorrelation matrices using a standard device, [6]. In this case one has for all $i \geq 0$

$$\mathcal{R}_x(i) = \mathcal{H}\mathcal{R}_s(i)\mathcal{H}^H. \quad (6)$$

Our goal is to find an estimate of $\mathcal{H}$ from (6) for $i = 0, 1$ and from the knowledge of $\mathcal{R}_s(0)$ and $\mathcal{R}_s(1)$. To this end we first undertake a whitening step. Under Assumption 1, with $\mathcal{L}$ a lower triangular matrix with positive diagonal elements, one has the Cholesky decomposition

$$\mathcal{R}_s(0) = \mathcal{L}\mathcal{L}^H. \quad (7)$$

Introduce the normalized matrices

$$\mathbf{H} \triangleq \mathcal{H}\mathcal{L}, \qquad \mathbf{R}_s(1) \triangleq \mathcal{L}^{-1}\mathcal{R}_s(1)\mathcal{L}^{-H}. \quad (8)$$

Then from (6) one has

$$\mathcal{R}_x(0) = \mathbf{H}\mathbf{H}^H, \qquad \mathcal{R}_x(1) = \mathbf{H}\mathbf{R}_s(1)\mathbf{H}^H. \quad (9)$$

Since $\mathcal{L}$ is known, the problem amounts to identifying $\mathbf{H}$ from (9). Introduce the coefficient vector of the order d forward prediction error filter (FPEF)

$$\alpha = [\ \alpha_1 \ \ \cdots \ \ \alpha_d\ ]^T \triangleq -\mathcal{R}_s^{-1}(0)\mathbf{r}. \quad (10)$$

One can verify the following relation:

$$\mathcal{R}_s(1) = \mathbf{J}\mathcal{R}_s(0) + \mathbf{e}_1\mathbf{r}^H. \quad (11)$$

Consequently, because of (5), the matrix $\mathbf{R}_s(1)$ becomes

$$\mathbf{R}_s(1) = \mathcal{L}^{-1}(\mathbf{J} - \mathbf{e}_1\alpha^H)\mathcal{L}. \quad (12)$$

Note that $\mathbf{J} - \mathbf{e}_1\alpha^H$ is a companion matrix whose eigenvalues coincide with the zeros of the FPEF. The following fact about the order d FPEF is well known [2] and is very important to the subsequent development.

Theorem 1 *Under Assumption 1, all the zeros of $F_d(z) = 1 + \sum_{k=1}^{d} \alpha_k^* z^{-k}$, the order d FPEF for the process $\{a_n\}$, lie strictly inside the unit circle, whence*

$$|\alpha_d| < 1. \quad (13)$$

3. INTERMEDIATE RESULTS

As in [6], consider a singular value decomposition (SVD) of $\mathcal{R}_x(0)$:

$$\mathcal{R}_x(0) = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U}_1^H \\ \mathbf{U}_2^H \end{bmatrix}. \quad (14)$$

Here $\mathbf{\Sigma} > \mathbf{0}$ is $d \times d$ diagonal. By Asumption 2, it follows from (9) that, with $\mathbf{V}$ some unitary matrix,

$$\mathbf{H} = \mathbf{U}_1\mathbf{\Sigma}\mathbf{V}. \quad (15)$$

In order to estimate $\mathbf{V}$, consider as in [6] the matrix

$$\mathbf{R} \triangleq \mathbf{\Sigma}^{-1}\mathbf{U}_1^H\mathcal{R}_x(1)\mathbf{U}_1\mathbf{\Sigma}^{-1}. \quad (16)$$

By direct verification using (9), one finds that $\mathbf{R} = \mathbf{V}\mathbf{R}_s(1)\mathbf{V}^H$. Thus, in view of (12),

$$\mathbf{R} = \mathbf{V}\mathcal{L}^{-1}(\mathbf{J} - \mathbf{e}_1\alpha^H)\mathcal{L}\mathbf{V}^H = \tilde{\mathbf{V}}(\mathbf{J} - \mathbf{e}_1\alpha^H)\bar{\mathbf{V}}^H, \quad (17)$$

where we have introduced the matrices

$$\tilde{\mathbf{V}} \triangleq \mathbf{V}\mathcal{L}^{-1}, \qquad \bar{\mathbf{V}} \triangleq \mathbf{V}\mathcal{L}^H. \quad (18)$$

Note that $\bar{\mathbf{V}}^H\tilde{\mathbf{V}} = \mathbf{I}$. Thus, (17) implies

$$\mathbf{R}\tilde{\mathbf{V}} = \tilde{\mathbf{V}}(\mathbf{J} - \mathbf{e}_1\alpha^H), \quad \mathbf{R}^H\bar{\mathbf{V}} = \bar{\mathbf{V}}(\mathbf{J} - \mathbf{e}_1\alpha^H)^H. \quad (19)$$

Specifically, if we partition columnwise $\tilde{\mathbf{V}} = [\ \tilde{\mathbf{v}}_1 \ \ \cdots \ \ \tilde{\mathbf{v}}_d\ ]$, then (19) yields

$$\mathbf{R}\tilde{\mathbf{v}}_i = \tilde{\mathbf{v}}_{i+1} - \alpha_i^*\tilde{\mathbf{v}}_1, \quad i = 1, \ldots, d-1; \quad (20)$$

$$\mathbf{R}\tilde{\mathbf{v}}_d = -\alpha_d^*\tilde{\mathbf{v}}_1. \quad (21)$$

Since $\mathbf{R}$, α are available, if we had $\tilde{\mathbf{v}}_1$ the remaining columns of $\tilde{\mathbf{V}}$ could be obtained via (20). This we proceed to do.

First we expose the structure of the singular values of $\mathbf{R}$, which are the same as those of $\mathbf{R}_s(1)$. Therefore they are independent of the channel matrix, being determined by the autocorrelation of the symbols $\{a_n\}$ alone.

Lemma 1 *There exists a $d \times d$ unitary matrix $\mathbf{Q}$ such that $\mathbf{R}_s(1) = \mathbf{Q}\mathbf{D}$, where $\mathbf{D}$ is $d \times d$ diagonal given by*

$$\mathbf{D} = \mathrm{diag}(\ 1 \ \ \cdots \ \ 1 \ \ |\alpha_d|\ ). \quad (22)$$

Note that $\mathbf{R}_s(1) = \mathbf{Q}\mathbf{D}$ is a singular value decomposition of $\mathbf{R}_s(1)$. Therefore $\mathbf{R} = (\mathbf{V}\mathbf{Q})\mathbf{D}\mathbf{V}^H$ is a singular value decomposition of $\mathbf{R}$. The significance of lemma 1 resides in the fact that the smallest singular value of $\mathbf{R}$ is given by $|\alpha_d|$, which is unique because of Theorem 1. This uniqueness allows us to extract the matrix $\mathbf{V}$ from $\mathbf{R}$, up to a unitary scaling constant. The key result is as follows:

Lemma 2 *Let $\beta_0 \triangleq \mathbf{e}_d^H\mathcal{L}^{-1}\mathbf{e}_d$. The vector $\beta_0^{-1}\tilde{\mathbf{v}}_1$ is a unit-norm left singular vector of the matrix $\mathbf{R}$ associated with its unique smallest singular value (under Assumption 1) $|\alpha_d|$.*

4. THE BLIND ALGORITHM

In view of the previous results, it is possible to estimate the columns of $\tilde{\mathbf{V}}$ as follows: first extract $\tilde{\mathbf{v}}_1$ as β_0 times the left singular vector of $\mathbf{R}$ associated with the smallest singular value; then use the recurrence (20) to estimate the remaining columns. For convenience, the algorithm is detailed next:

1. Perform an SVD of $\mathcal{R}_x(0)$ as in (14) and form the matrix $\mathbf{R} = \mathbf{\Sigma}^{-1}\mathbf{U}_1^H\mathcal{R}_x(1)\mathbf{U}_1\mathbf{\Sigma}^{-1}$.
2. Let $\hat{\mathbf{v}}_1$ be a unit-norm left singular vector of $\mathbf{R}$ associated to the smallest singular value.
3. For $i = 1, 2, \ldots, d-1$, let $\hat{\mathbf{v}}_{i+1} = \mathbf{R}\hat{\mathbf{v}}_i + \alpha_i^*\hat{\mathbf{v}}_1$.
4. The channel matrix estimate is given by
$$\hat{\mathcal{H}} = \beta_0\mathbf{U}_1\mathbf{\Sigma}[\ \hat{\mathbf{v}}_1 \quad \cdots \quad \hat{\mathbf{v}}_d\ ].$$
5. The zero-forcing equalizers are the columns of the matrix
$$\mathcal{G}_{\mathrm{ZF}} \triangleq \beta_0\mathbf{U}_1\mathbf{\Sigma}^{-1}[\ \hat{\mathbf{v}}_1 \quad \cdots \quad \hat{\mathbf{v}}_d\ ]\mathcal{R}_s(0).$$

Note that if $\{a_n\}$ is white, one has $\mathcal{R}_s(0) = \sigma_a^2\mathbf{I}$, $\mathcal{L} = \sigma_a\mathbf{I}$, $\alpha = 0$ and $\beta_0 = 1/\sigma_a$, and the algorithm above recovers as a special case the one in [6]. We can state:

Theorem 2 *Consider the algorithm above. Under Assumptions 1 and 2, there exists a real θ such that $\hat{\mathcal{H}}$ obtained in step 4 obeys $\hat{\mathcal{H}} = e^{j\theta}\mathcal{H}$. Further with $\mathbf{w}_n = 0$ in (2), the matrix $\mathcal{G}_{\mathrm{ZF}}$ obtained in step 5 obeys $\mathcal{G}_{\mathrm{ZF}}^H\mathbf{x}_n = e^{-j\theta}\mathbf{s}_n$.*

5. SIMULATION RESULTS

Several simulation experiments were performed to test the new algorithm. For comparison, the subspace algorithm of [4] was also implemented in the same environment.

The channel impulse response corresponds to a two-ray multipath environment and was taken from [6]. The number of subchannels is $p = 4$ and the corresponding channel order is $l = 5$. The input symbols $\{a_n\}$ are drawn from a QPSK constellation as follows. Let $\{b_n\}$ be the input stream of i.i.d. bits, i.e. $b_n \in \{0, 1\}$. Then

$$a_n = \begin{cases} -1+j & \text{if} \quad (b_n\, b_{n-2}) = (0\,0) \\ +1+j & \text{if} \quad (b_n\, b_{n-2}) = (0\,1) \\ -1-j & \text{if} \quad (b_n\, b_{n-2}) = (1\,0) \\ +1-j & \text{if} \quad (b_n\, b_{n-2}) = (1\,1) \end{cases}$$

This scheme generates a colored symbol sequence:

$$E[a_n a_{n-k}^*] = \begin{cases} 2, & k = 0, \\ \pm j, & k = \pm 2, \\ 0, & \text{else.} \end{cases}$$

We consider an equalizer of order $m = 4$, which yields $d = m + l = 9$. The vector α is then given by

$$\alpha = [\ 0 \quad 0.8j \quad 0 \quad -0.6 \quad 0 \quad -0.4j \quad 0 \quad 0.2 \quad 0\ ]^T.$$

Additive white Gaussian noise $\mathbf{w}_n$ was added to the channel output, so that the model becomes $\mathbf{x}_n = \mathcal{H}\mathbf{s}_n + \mathbf{w}_n$. The SNR is defined as

$$\mathrm{SNR} = 10\log\frac{\mathrm{trace}\, E[(\mathcal{H}\mathbf{s}_n)(\mathcal{H}\mathbf{s}_n)^H]}{\mathrm{trace}\, E[\mathbf{w}_n\mathbf{w}_n^H]}.$$

The noise variance estimate $\hat{\sigma}_w^2$ was taken as the smallest eigenvalue of $\mathcal{R}_x(0)$ and then substracted to provide the algorithms with denoised autocorrelation estimates. For simplicity, knowledge of the channel length l was assumed.

In figure 1 the normalized root-mean-square error (NRMSE) of the channel estimate, defined in eq. (86) of [6], is depicted as a function of the number of samples K used for the estimation of $\mathcal{R}_x(0)$, $\mathcal{R}_x(1)$. The SNR was fixed at 25 dB. The new algorithm clearly outperforms the subspace approach; this is as expected since the latter does not exploit knowledge about the symbol autocorrelation.

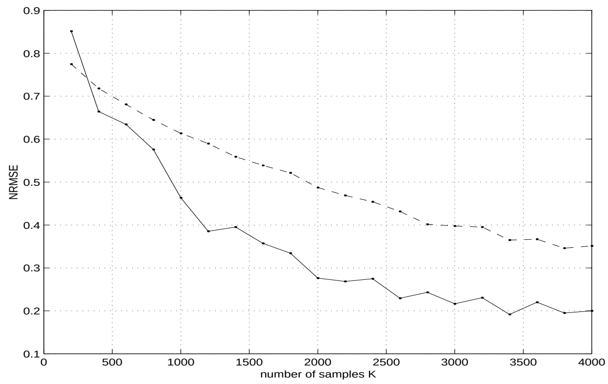

Figure 1: NRMSE versus number of samples K for the new algorithm (solid) and the subspace method (dashed). SNR = 25 dB; results averaged over 200 runs.

Once the channel matrix has been estimated by the subspace algorithm, the zero-forcing equalizers are obtained as the rows of the pseudoinverse: $\hat{\mathcal{H}}^{\#} = \mathcal{G}_{\mathrm{ZF}}^H$. For both algorithms the minimum mean-squared error (MMSE) equalizers are computed as [5]

$$\mathcal{G}_{\mathrm{MMSE}} = \left(\mathbf{I} - \hat{\sigma}_w^2\mathcal{R}_x^{-1}(0)\right)\mathcal{G}_{\mathrm{ZF}}, \qquad (23)$$

where $\mathcal{R}_x(0)$ represents the *undenoised* autocorrelation matrix of the channel output. The different rows of $\mathcal{G}_{\mathrm{MMSE}}$ correspond to different equalization delays and therefore they present different MSE values at their outputs. Figure 2 shows the symbol error rate (SER) obtained with the equalizers of delays 0, 3, 5 and 8, as a function of SNR and using $K = 2000$ samples for autocorrelation estimation. It is

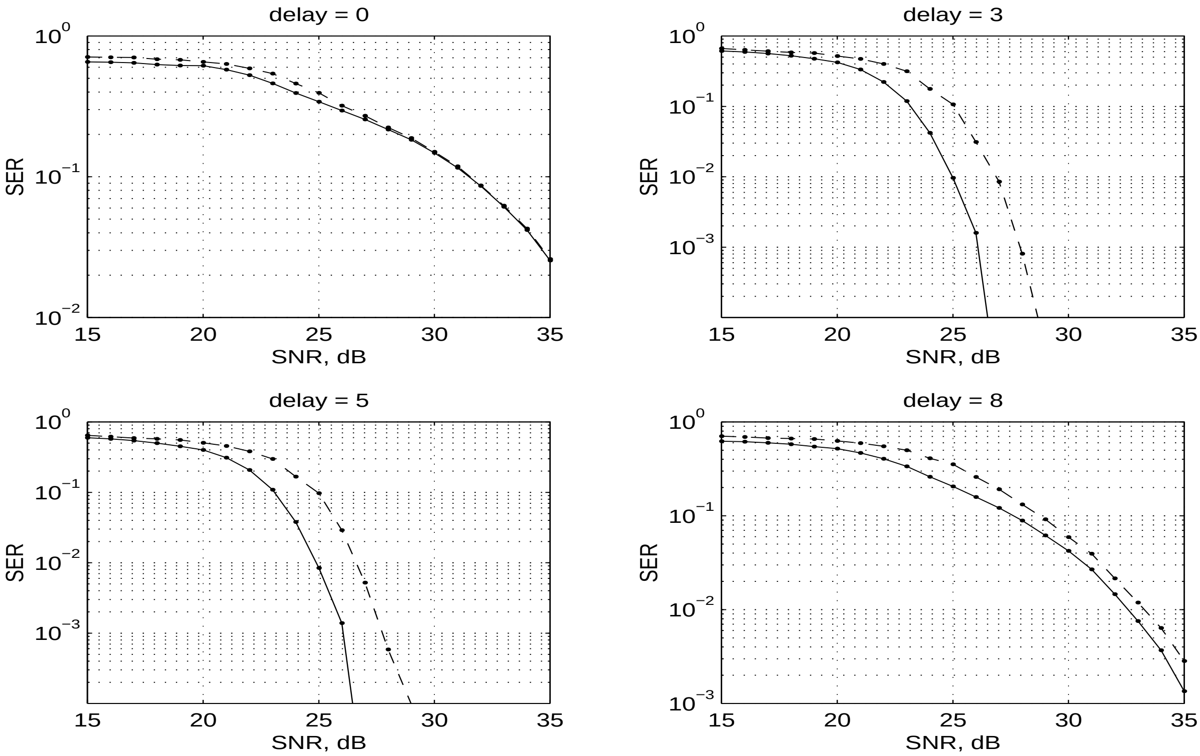

Figure 2: SER obtained by the MMSE equalizers from the new algorithm (solid) and from the subspace algorithm (dashed) for different equalization delays. The number of samples for autocorrelation estimation was $K = 2000$. Results averaged over 200 independent trials.

seen that for intermediate delays, where the equalizer performance is best for either method, the new algorithm provides an advantage of about 2 dB compared to the subspace algorithm.

6. CONCLUSION

An extension of the algorithm of [6] for blind channel identification has been developed in order to account for source correlation. The new algorithm estimates the channel from second-order statistics of the observed signal for arbitrary but known transmitted symbol coloring. The computational complexity is comparable to that of the original algorithm for white sources. The feasibility of the algorithm was tested via simulations, which show that by using the knowledge of input statistics, the new algorithm provides better channel estimates and equalizers than those obtained by subspace based methods.

7. REFERENCES

[1] K. H Afkhamie and Z.-Q. Luo, "Blind identification of FIR systems driven by Markov-like input signals", *IEEE Trans. Signal Processing,* vol. 48 no. 6, pp. 1726-1736, June 2000.

[2] S. Haykin, *Adaptive Filter Theory, 3rd ed.,* Prentice Hall, Upper Saddle River, NJ, 1996.

[3] Y. Hua, H. Yang and W. Qiu, "Source correlation compensation for blind channel identification based on second-order statistics", *IEEE Signal Processing Letters,* vol. 1 no. 8, pp. 119-120, August 1994.

[4] E. Moulines, P. Duhamel, J.-F. Cardoso and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters", *IEEE Trans. Signal Processing,* vol. 43 no. 2, pp. 516-525, Feb. 1995.

[5] C. B. Papadias and D. T. M. Slock, "Fractionally spaced equalization of linear polyphase channels and related blind techniques based on multichannel linear prediction", *IEEE Trans. Signal Processing,* vol. 47 no. 3, pp. 641-654, March 1999.

[6] L. Tong, G. Xu and T. Kailath, "Blind identification and equalization based on second-order statistics: A time-domain approach", *IEEE Trans. Information Theory,* vol. 40 no. 2, pp. 340-350, March 1994.