AN EXACT BLIND CHANNEL ESTIMATOR AND ITS EXTENSIONS TO ADDITIVE NOISE CHANNELS

Ahmet Kemal Özdemir and Orhan Arıkan

Department of Electrical and Electronics Engineering,
Bilkent University, Ankara, TR-06533 TURKEY.
Phone & Fax: 90-312-2664307,
e-mail: `kozdemir@ee.bilkent.edu.tr`
`oarikan@ee.bilkent.edu.tr`

ABSTRACT

Recent work on fractionally spaced blind equalizers have shown that it is possible to exactly identify the channel and its input sequence from the *noise–free* channel outputs [1–3]. However, the obtained results are based on a set of over–restrictive constraints on the channel. We have shown that the exact identification can be achieved in a much broader class of channels [4]. In this paper, we present performance results of the exact channel estimator in the presence of noise.

1. INTRODUCTION

Since the invention of digital communication, blind channel equalization has been an active area of research. Some of the proposed approaches have provided significant improvements in the removal of inter–symbol–interference (ISI). For a review of the past and present research on the blind channel equalization please refer to [5–7]. Although the recent research on this subject provides significant contributions, most of the proposed solutions are applicable only for a limited class of communication channels and input sequences even in the noise–free case [1–3, 8–11]. This observation prompted us to investigate the noise–free blind channel identification and input sequence estimation problem to fully characterize what can be done with the least set of assumptions on the channel model [4]. In our work, we have presented theoretical results on the exact identification of the channel response and input sequence based on the noise–free observation of the channel output sequence. In this paper, we extend the exact blind channel estimator proposed in [4] to the more realistic case of noisy channels and we investigate its performance with computer simulations.

Over–sampling the output of an FIR continuous–time channel at a rate M' times faster than the symbol rate $1/T$ provides channel diversity which can be equivalently represented as a single–input M'–output discrete–time multi–channel FIR filter [6]. Without loss of generality, assuming that first $M \leq M'$ of these sub–channels are to be identified, the corresponding multi–channel model is shown in Fig. 1 where the outputs of the multi–channel filter are the samples of the received signal $y(t)$:

$$y_i[n] \;=\; y(nT + (i \Leftrightarrow 1)\frac{T}{M'}) \quad , \quad 1 \leq i \leq M \;. \qquad (1)$$

In this model $\{a[n]\}_{n=0}^{\infty}$ is the input symbol sequence chosen from a finite alphabet, D represents the transmission delay and $v_i[n]$ is the additive channel noise. The FIR filter $h_c[n]$ in Fig. 1 corresponds to the common zeros of the sub–channels. If the sub–channels share L_1 common zeros, then the length of $h_c[n]$ is $L_1 + 1$. If the sub–channels do not share any common zeros then $h_c[n] = \delta[n]$. In this work we investigate the estimation of the uncommon parts of the channels $h_1[n], \ldots, h_M[n]$ based on the noisy channel output. We also discuss the estimation of the common input $x_c[n]$ once the uncommon parts of the channels are identified. The estimation of the common part of the channels $h_c[n]$ and the input sequence $a[n]$ will not be discussed here. These problems require further investigation.

The organization of the paper is as follows. In Section 2 we review the blind channel estimator proposed in [4] and summarize its key features for the noise–free problem. Then we discuss the extensions of these results to the noisy channels. After giving a simulation example in Section 3, we derive some conclusions in Section 4.

2. CONSTRAINED IDENTIFICATION OF CHANNELS $h_1[n], \ldots, h_m[n]$

For $1 \leq i \leq M$, let $g_i[n]$ be an estimate of $h_i[n]$. Here we assume that the estimated order $\hat{L}_2$ of the channel estimates is larger than or equal to L_2 which is the largest order of the channels $h_i[n]$. We will base the optimality of the channel estimates at time N to the following cost function:

$$J_{\hat{L}_2}(\boldsymbol{g}; N) \;=\; \frac{1}{M} \sum_{i=1}^{M} \sum_{j=i+1}^{M} J_{ij}(\boldsymbol{g}_i, \boldsymbol{g}_j; N) \;\; , \qquad (2)$$

where $J_{ij}(\boldsymbol{g}_i, \boldsymbol{g}_j; N)$, the cost function associated with channels i and j, is defined as:

$$J_{ij}(\boldsymbol{g}_i, \boldsymbol{g}_j; N) \;=\; \frac{1}{C_{w,N}} \sum_{k=0}^{N} w_{N-k} \left| \boldsymbol{g}_i^T \boldsymbol{y}_j[k] \Leftrightarrow \boldsymbol{g}_j^T \boldsymbol{y}_i[k] \right|^2 \;\; , \qquad (3)$$

where $\boldsymbol{g}_i$ and $\boldsymbol{y}_i[k]$ are defined as:

$$\boldsymbol{g}_i \;=\; \left[\, g_i[0] \quad g_i[1] \quad \cdots \quad g_i[\hat{L}_2] \,\right]^T \qquad (4)$$

$$\boldsymbol{y}_i[k] \;=\; \left[\, y_i[k] \quad y_i[k \Leftrightarrow 1] \quad \cdots \quad y_i[k \Leftrightarrow \hat{L}_2] \,\right]^T \;. \qquad (5)$$

$C_{w,N}$ in (3) is a normalization constant defined as $C_{w,N} = \sum_{k=0}^{N} w_{N-k}$ and w_k is a weighting sequence that satisfies

$$0 \leq w_k \leq 1 \quad , \quad 0 \leq k \leq N \ . \tag{6}$$

By using (4) a more compact representation for the cost $J_{\hat{L}_2}(\boldsymbol{g}; N)$ can be given as:

$$J_{\hat{L}_2}(\boldsymbol{g}; N) \ = \ \boldsymbol{g}^H \boldsymbol{R_{yy}}[N] \ \boldsymbol{g} \ , \tag{7}$$

where $\boldsymbol{g} = [\boldsymbol{g}_1^T \ \boldsymbol{g}_2^T \ \cdots \ \boldsymbol{g}_M^T]^T$ is the concatenated channel vector estimates and $\boldsymbol{R_{yy}}[N]$ is the hermitian nonnegative definite matrix with i^{th} diagonal entry $\sum_{j \neq i} \boldsymbol{R}_{\boldsymbol{y}_j \boldsymbol{y}_j}[N]$ and $(i,j)^{\text{th}}$ off–diagonal entry $\Leftrightarrow \boldsymbol{R}_{\boldsymbol{y}_j \boldsymbol{y}_i}[N]$, where $\boldsymbol{R}_{\boldsymbol{y}_j \boldsymbol{y}_i}[N]$ is the weighted cross–correlation matrix of the multi–channel filter outputs $\boldsymbol{y}_j$ and $\boldsymbol{y}_i$:

$$\boldsymbol{R}_{\boldsymbol{y}_j \boldsymbol{y}_i}[N] \ = \ \frac{1}{C_{w,N}} \sum_{k=0}^{N} w_{N-k} \boldsymbol{y}_j^*[k] \boldsymbol{y}_i^T[k] \ . \tag{8}$$

In the noise–free case, the minimizers of the cost function given by (7) are fully characterized by the following theorems whose proofs are given in [4, 12].

Theorem 1.

$$J_{\hat{L}_2}(\boldsymbol{g}\,; N) \ = \ 0 \Leftrightarrow g_i[n] = h_i[n] * f[n] \tag{9}$$

provided that

$$1) \quad \hat{L}_2 \geq L_2 \tag{10}$$

$$2) \quad N \geq D + L_2 + \hat{L}_2 \tag{11}$$

where $f[n]$ is an arbitrary FIR filter of order at most $\hat{L}_2 \Leftrightarrow L_2$.

Theorem 2. *The set of vectors $[\ \boldsymbol{g}_1^T \quad \boldsymbol{g}_2^T \quad \cdots \quad \boldsymbol{g}_M^T\ ]^T$ that satisfies* (9) *constitutes an $\hat{L}_c + 1$ dimensional vector space, where $\hat{L}_c = \hat{L}_2 \Leftrightarrow L_2$.*

An important implication of the second theorem is stated as:

Corollary 1. *The matrix $\boldsymbol{R_{yy}}[N]$ has an $\hat{L}_c + 1$ dimensional null-space.*

These theoretical results establish a basis for the exact blind channel estimator in the noise–free case: By starting with $\hat{L}_2$ which is larger than L_2, $\boldsymbol{R_{yy}}[N]$ is obtained. Then by using Corollary 1, the true channel order is identified as

$$L_2 = \hat{L}_2 \Leftrightarrow \eta + 1 \ , \tag{12}$$

where η is the dimension of the null–space of $\boldsymbol{R_{yy}}[N]$. Once the actual order L_2 is obtained, the minimization of (7) is solved with $\hat{L}_2 = L_2$. Since $\hat{L}_c = \hat{L}_2 \Leftrightarrow L_2 = 0$, Theorem 1 states that any minimizer of $J_{L_2}(\boldsymbol{g}; N)$ would be in the form

$$g_i[n] \ = \ f[0] h_i[n] \quad , \quad \text{for } i = 1, \ldots, M \tag{13}$$

where $f[0]$ is an arbitrary constant. To avoid the undesired trivial solution $f[0] = 0$, we have to introduce some constraints into the minimization problem. The constraints should be imposed in a way that a non–zero multiple of the actual channels $h_1[n], \ldots, h_M[n]$ should be in the feasible set. As it can be shown easily, the following set of constraints meets this requirement [12]:

(i) $||\boldsymbol{g}||^2 \ = \mu^2$

(ii) $||\boldsymbol{g}_i||^2 = \mu_i^2$ for some i

(iii) $\boldsymbol{C}^H \boldsymbol{g} \ = \alpha$ for some known matrix $\boldsymbol{C}$.

First two of them places energy constraints on the the sub–channels and the last one linearly constraints the true channel coefficients which can be of use in certain applications such as the training phase of a Binary Phase Shift Keying (BPSK) communication system.

Several complications arise, while we are extending the above algorithm to the more realist case of additive noise channels. For instance, in the noisy case the true channel order is not given by $L_2 = \hat{L}_2 \Leftrightarrow \eta + 1$, where η is the dimension of the null space of $\boldsymbol{R_{yy}}[N]$. However, if the channel noise is additive and white we can still make use of the same equation provided that we redefine η as the multiplicity of the smallest singular value of $\boldsymbol{R_{yy}}[N]$. Because in this case the smallest η singular values of $\boldsymbol{R_{yy}}[N]$ will be clustered together especially at high signal to noise ratios and/or large N. Another point is that, in [12] it has been shown that all of the above constraints are equivalent in the noise–free case in the sense that any of them can be used to obtain a non–zero multiple of the true channel. However, in the presence of noise, Theorem 1 does not hold and the minimization of (7) under different constraints in general produces different channel estimates. In the following, closed form solutions corresponding to these constraints are provided. As we show in the simulation section, the quality of the obtained estimates may vary considerably in the noisy case.

2.1. Energy Constraint 1

Let the eigen decomposition of $\boldsymbol{R_{yy}}$ be given as[1]:

$$\boldsymbol{R_{yy}} \quad = \quad \sum_{i=1}^{M(L_2+1)} \lambda_i \boldsymbol{q}_i \boldsymbol{q}_i^H, \quad \lambda_1 \geq \cdots \geq \lambda_{M(L_2+1)} \geq 0,$$
$$||\boldsymbol{q}_i||^2 = 1 \text{ for } 1 \leq i \leq M(L_2+1) \ . \tag{14}$$

It is well known that the minimizer of (7) under the total energy constraint is the eigenvector of $\boldsymbol{R_{yy}}$ corresponding to the smallest eigenvalue $\lambda_{M(L_2+1)}$:

$$\hat{\boldsymbol{g}} \ = \ \mu \, \boldsymbol{q}_{M(L_2+1)} \ . \tag{15}$$

For computational efficiency, it is better to compute the smallest eigenvalue and the corresponding eigenvector in (15) by using an algorithm specially designed for that purpose [13], rather than carrying out a full eigen decomposition as in (14).

2.2. Energy Constraint 2

Without loss of generality we will assume that the energy constraint is imposed only on to the first channel. Then the constrained optimization problem can be restated as

$$\min_{\boldsymbol{g}} \quad \boldsymbol{g}^H \boldsymbol{R_{yy}} \boldsymbol{g} \tag{16}$$

$$\text{s.t.} \qquad ||\boldsymbol{g}_1||^2 = \mu_1^2 \ . \tag{17}$$

[1] For notational convenience, the dependence on N is suppressed.

If we partition the matrix $\boldsymbol{R}_{\boldsymbol{yy}}$ and the vector $\boldsymbol{g}$ as follows:

$$\boldsymbol{R}_{\boldsymbol{yy}} = \begin{bmatrix} \boldsymbol{R}_{11} & \boldsymbol{R}_{12} \\ \boldsymbol{R}_{21} & \boldsymbol{R}_{22} \end{bmatrix} \tag{18}$$

$$\boldsymbol{g} = \begin{bmatrix} \boldsymbol{g}_1^T & \boldsymbol{s}^T \end{bmatrix}^T \tag{19}$$

where $\boldsymbol{R}_{11}$ is the upper left L_2+1 by L_2+1 block of the correlation matrix $\boldsymbol{R}_{\boldsymbol{yy}}$, we can use the following procedure to obtain the optimal solution:

1. Find the smallest eigenvalue and the corresponding eigenvector of the Schur complement

$$\left(\boldsymbol{R}_{11} \Leftrightarrow \boldsymbol{R}_{12}\boldsymbol{R}_{22}^{-1}\boldsymbol{R}_{21}\right)\boldsymbol{q}_s = \lambda_s \boldsymbol{q}_s \ . \tag{20}$$

2. Compute $\hat{\boldsymbol{g}}_1$ and $\hat{\boldsymbol{s}}$ as

$$\left[\hat{\boldsymbol{g}}_1 \quad \hat{\boldsymbol{s}}\right] = \mu_1 \left[\boldsymbol{I} \quad \Leftrightarrow \boldsymbol{R}_{22}^{-1}\boldsymbol{R}_{21}\right]\boldsymbol{q}_s \ , \tag{21}$$

   where $\boldsymbol{I}$ is the L_2+1 dimensional identity matrix.

Since the dimension of the eigenvalue problem in (20) is M times smaller than that of in Section 2.1, computationally it is easier to obtain the solution under the second constraint than the first constraint.

2.3. Projection Matrix Constraint

The last constraint in Section 2 is a projection matrix constraint, which constrains the component of the channel vector $\boldsymbol{g}$ in the range–space of a known matrix $\boldsymbol{C}$. The solution of (7) under this constraint is given as

$$\hat{\boldsymbol{g}} = \alpha(\boldsymbol{R}_{\boldsymbol{yy}}^{-1}\boldsymbol{C})(\boldsymbol{C}^H\boldsymbol{R}_{\boldsymbol{yy}}^{-1}\boldsymbol{C})^{-1} \ . \tag{22}$$

If the constraint matrix is chosen as a vector $\boldsymbol{C} \equiv \boldsymbol{c}$, this expression simplifies to

$$\hat{\boldsymbol{g}} = \alpha \frac{\boldsymbol{R}_{\boldsymbol{yy}}^{-1}\boldsymbol{c}}{\boldsymbol{c}^H\boldsymbol{R}_{\boldsymbol{yy}}^{-1}\boldsymbol{c}} \ . \tag{23}$$

In the next section we conduct computer simulations to investigate the performance of these channel estimates obtained under different constraints.

3. SIMULATIONS

Consider the 4–channel FIR ($M = M' = 4$) filter with pole–zero locations shown in Fig. 2. In the simulation, samples of the input symbol sequence $a[n]$ with variance $\sigma_a^2 = 2$ are drawn from a QAM alphabet and the channel noise sequences $v_i[n]$ are chosen as the realizations of a white and Gaussian noise sequence with variance σ_v^2. In Fig. 3, 251 samples of the noisy channel outputs at a signal to noise ratio of 15 dB are shown, where the signal to noise ratio is defined as

$$SNR = 10\log_{10}\frac{\sigma_a^2\sum_{i=1}^4 ||h_i||^2}{\sigma_v^2} \ . \tag{24}$$

In Fig. 4, the singular values of the correlation matrix $\boldsymbol{R}_{\boldsymbol{yy}}[250]$ are plotted when the channel order estimate is $\hat{L}_2 = 5$. From this figure we estimate the number of noise singular values as $\eta = 3$ and identify the true channel order as $L_2 = \hat{L}_2 \Leftrightarrow \eta + 1 = 3$. After recomputing the correlation matrix $\boldsymbol{R}_{\boldsymbol{yy}}[250]$ for $\hat{L}_2 = 3$, we obtained estimates of the true channel under the constraints in Section 2. When obtaining the solution under the projection matrix constraint, $\boldsymbol{C}$ is chosen as a unit vector with a 1 at its first entry. Once an estimate for the true channel is obtained, several of the existing algorithms in the literature can be used to estimate the input sequence (for instance see [14]). In this simulation, we use the linear minimum mean square estimator [15] which is computationally intensive, but produces accurate results. The input symbol constellations estimated by this way are shown in Fig. 5. By comparing the obtained results, we conclude that the performance of the energy constraints are comparable whereas the performance of the projection matrix constraint is poorer.

4. CONCLUSIONS

An extension of the exact blind channel estimator [4] is given to the additive noise channels. The identification of the uncommon parts of the channels is posed as a constrained minimization problem and closed form solutions to this problem are obtained under different constraints. The quality of these estimates is investigated by using computer simulations. In the future, we plan to study the extension of the pruning algorithm in [4] to the noisy channels, which will be used for the identification of the common parts of the channels.

5. REFERENCES

[1] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second–order statistics: A time domain approah," *IEEE Trans. Information Theory*, vol. 40, pp. 340–349, Mar. 1994.

[2] G. Xu, H. Liu, L. Tong, and T. Kailath, "A least–squares approach to blind channel identification," *IEEE Trans. Signal Process.*, vol. 43, pp. 2982–2993, Dec. 1995.

[3] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Process.*, vol. 43, pp. 516–525, Feb. 1995.

[4] A. K. Özdemir and O. Arıkan, "A theoretical investigation on exact blind channel and input sequence estimation," in *Proc. IEEE Int. Symp. Circuits and Systems*, 1999.

[5] S. Haykin, ed., *Blind Deconvolution*. Englewood Cliffs, NJ: Prentice–Hall, 1994.

[6] S. Haykin, *Adaptive Filter Theory*. New Jersey: Prentice–Hall, third ed., 1996.

[7] H. Liu, G. Xu, L. Tong, and T. Kailath, "Recent developments in blind channel equalization: From cyclostationarity to subspaces," *Signal Process.*, vol. 50, pp. 83–99, Apr. 1996.

[8] D. T. M. Slock, "Blind fractionally–spaced equalization, perfect reconstruction filter banks and multi–channel linear prediction," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, vol. 4, pp. 585–588, 1994.

[9] H. Liu and G. Xu, "Closed–form blind symbol estimation in digital communications," *IEEE Trans. Signal Process.*, vol. 43, pp. 2714–2723, Nov. 1995.

[10] G. B. Giannakis and S. D. Halford, "Blind fractionally spaced equalization of noisy FIR channels: Direct and adaptive solutions," *IEEE Trans. Signal Process.*, vol. 45, pp. 2277–2292, Sept. 1997.

[11] D. Gesbert, P. Duhamel, and S. Mayrargue, "On–line blind multichannel equalization based on mutually referenced filters," *IEEE Trans. Signal Process.*, vol. 45, pp. 2307–2317, Sept. 1997.

[12] A. K. Özdemir, "Exact blind channel estimator," Master's thesis, Bilkent University, Dept. of Electrical and Electronics Engineering, June 1998.

[13] B. N. Datta, *Numerical linear algebra and applications*. Brooks/Cole Publishing Company, 1995.

[14] B. Farhang-Boroujney, "Channel equalization via channel identification: Algorithms and simulation results for rapidly fading HF channels," *IEEE Trans. on Communications*, vol. 44, pp. 1409–1412, Nov. 1996.

[15] L. L. Scharf, *Statistical signal processing : detection, estimation, and time series analysis*. Addison–Wesley, 1991.

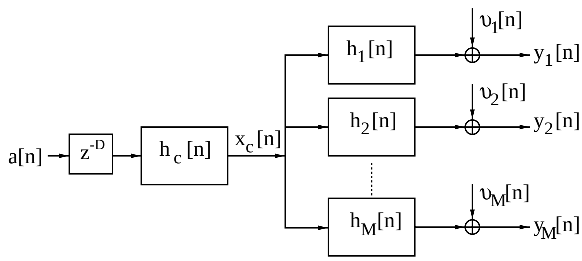

Figure 1: The multi–channel filter model.

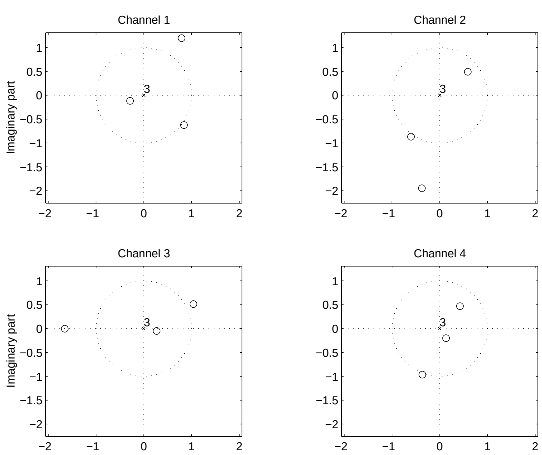

Figure 2: The pole–zero locations of the channels.

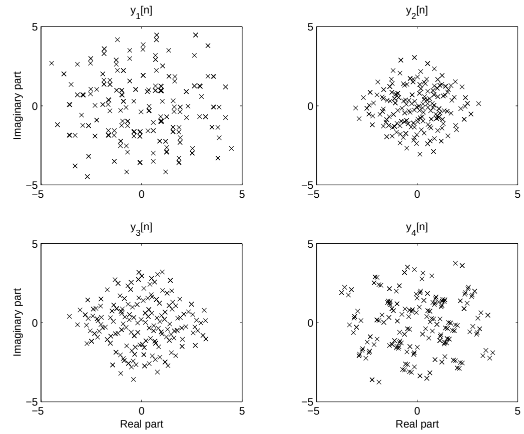

Figure 3: The scatter plot of the channel outputs.

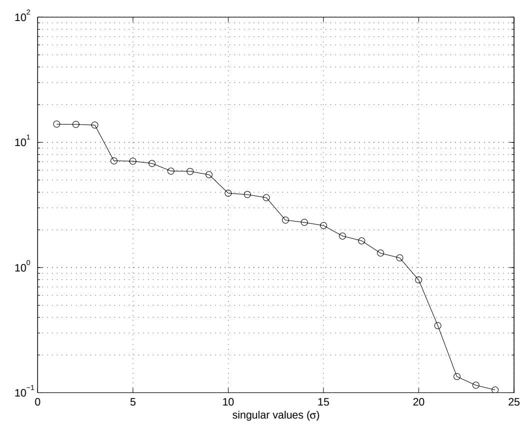

Figure 4: The singular values of the correlation matrix $\boldsymbol{R_{yy}}[250]$.

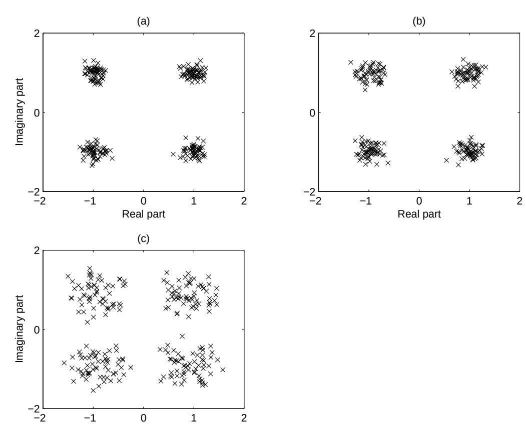

Figure 5: The scatter plot of the input sequence estimates for different constraint channel estimators: energy constraint 1 (a), energy constraint 2 (b) and projection matrix constraint (c).