

A THEORETICAL INVESTIGATION ON EXACT BLIND CHANNEL AND INPUT SEQUENCE ESTIMATION

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: 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. However, the obtained results are based on a set of over-restrictive constraints on the channel. In this paper it is shown that the exact identification can be achieved in a broader class of channels.

1 Introduction

Since the invention of digital communication, blind channel equalization has been an active area of research. Here, we present theoretical results on the exact identification of the channel response and input sequence based on the *noise-free* observation of the channel output sequence. Our purpose is to fully characterize what can be done with the least set of assumptions on the channel model.

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 [1]. 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 $y(t)$:

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

In this model $\{a[n]\}_{n=0}^{\infty}$ is the input symbol sequence chosen from a finite alphabet and D represents the transmission delay. The FIR filter $h_c[n]$ in Fig. 1 corresponds to the common zeros of the sub-channels. In order to clearly differentiate the present work from the previous ones, we state below some of the assumptions and/or constraints that are avoided here:

1. The knowledge of the exact channel order as in [2], [3], [4], [5], [6].
2. The constraints on the length of the channels and the number of the channels as in [7].
3. The assumption that the sub-channels do not share any common zeros (i.e., $h_c[n] = \delta[n]$) as in almost every second-order statistics based algorithm [2], [6], [7], [8]. It has been shown that these algorithms lose their robustness when this assumption is not true [9]. This is a severe limitation because there exists classes of multi-path

communication systems [10] for which this assumption is not valid for any over-sampling factor M' .

4. The assumption that the number of sub-channels is the same as the over-sampling factor.

If the sub-channels do not share any common zeros then $h_c[n] = \delta[n]$. It will be shown that in this latter case, an efficient tree-structured algorithm can be used to identify the exact channel and the input sequence. Also, the least number of channel output samples required in the identification is found. In the case of common zeros, it is shown that the tree-structured algorithm can still be used to identify the channels $h_1[n], \dots, h_M[n]$ and their common input $x_c[n]$. Then, the blind identification of $h_c[n]$ and the input sequence can be carried out by using a pruning algorithm which converges almost surely under very mild set of assumptions on the probability distribution of the channel coefficients.

2 Exact Identification of $h_1[n], \dots, h_M[n]$

As shown in Fig. 1, the output signals $y_1[n], \dots, y_M[n]$ are the responses of the channels $h_1[n], \dots, h_M[n]$ to the same input $x_c[n]$. Thus it can be conjectured that the input-output relationship between pairs of channels might produce sufficient information [6], [11] to estimate the channels $h_1[n], \dots, h_M[n]$ without any prior knowledge about their input $x_c[n]$. In this and the next section we will show that indeed this is the case.

For $1 \leq i \leq M$, let $g_i[n]$ be an estimate of $h_i[n]$. Here we assume that the assumed 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 a set of channel estimates at the sampling index N of the received data $y_i[n]$, $1 \leq i \leq M$, to the following cost function:

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

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

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

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

$$\mathbf{g}_i = \begin{bmatrix} g_i[0] & g_i[1] & \cdots & g_i[\hat{L}_2] \end{bmatrix}^T \quad (4)$$

$$\mathbf{y}_i[k] = \begin{bmatrix} y_i[k] & y_i[k \Leftrightarrow 1] & \cdots & y_i[k \Leftrightarrow \hat{L}_2] \end{bmatrix}^T \quad (5)$$

$C_{w,N}$ 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

$$0 \leq w_k \leq 1 \quad , \quad 0 \leq k \leq N \quad . \quad (6)$$

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

$$J_{\hat{L}_2}(\mathbf{g}; N) = \mathbf{g}^H \mathbf{R}_{\mathbf{y}\mathbf{y}}[N] \mathbf{g} \quad , \quad (7)$$

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

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

The minimizers of the cost function given by (7) are fully characterized by the following Theorem:

Theorem 1.

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

provided that $\hat{L}_2 \geq L_2$ and $N \geq D + L_2 + \hat{L}_2$, where $f[n]$ is an arbitrary FIR filter of order at most $\hat{L}_2 \Leftrightarrow L_2$.

Proof of if part. $g_i[n] = h_i[n] * f[n] \Rightarrow$

$$\begin{aligned} J_{\hat{L}_2}(\mathbf{g}; N) &= \sum_{i=1}^M \sum_{j=i+1}^M \|h_i * y_j * f \Leftrightarrow h_j * y_i * f\|_{W,N}^2 \\ &= \sum_{i=1}^M \sum_{j=i+1}^M \|(h_i * y_j \Leftrightarrow h_j * y_i) * f\|_{W,N}^2 = 0 \quad . \end{aligned}$$

where we defined the norm $\|\cdot\|_{W,N}$ as $\|u\|_{W,N}^2 = 1/C_{w,N} \sum_{k=0}^N w_{N-k} |u[k]|^2$. \square

Proof of only if part. $J_{\hat{L}_2}(\mathbf{g}; N) = 0$ for $N \geq D + L_2 + \hat{L}_2$

$$\begin{aligned} \Rightarrow & \sum_{i=1}^M \sum_{j=1}^{i-1} \|g_i * y_j \Leftrightarrow g_j * y_i\|_{W,N}^2 = 0 \\ \Rightarrow & \|g_i * y_j \Leftrightarrow g_j * y_i\|_{W,N}^2 = 0 \quad , \quad \forall i, j \\ \Rightarrow & \|(g_i * h_j \Leftrightarrow g_j * h_i) * x_c\|_{W,N}^2 = 0 \quad , \quad \forall i, j \quad . \quad (10) \end{aligned}$$

To complete the proof of the theorem we will make use of several lemmas whose proofs are given in [12].

Lemma 1. *If $\|(g_i * h_j \Leftrightarrow g_j * h_i) * x_c\|_{W,N}^2 = 0$ for $N \geq D + L_2 + \hat{L}_2$ then $g_i[n] * h_j[n] \Leftrightarrow g_j[n] * h_i[n] = 0 \quad \forall n$.*

This lemma can be used in (10) to conclude the equality of $g_i[n] * h_j[n] = g_j[n] * h_i[n]$ for all n . Hence, in the z -transform domain:

$$G_i(z)H_j(z) = G_j(z)H_i(z) \quad . \quad (11)$$

Since for $i \neq j$, $H_i(z)$ and $H_j(z)$ may have common zeros, let monic polynomial $H_{ij}(z)$ be the greatest common divisor of $H_i(z)$ and $H_j(z)$. Hence, $H_i(z)$ can be decomposed as:

$$H_i(z) = H_{ij}(z)Q_{i/j}(z) \quad (12)$$

where the quotient polynomial $Q_{i/j}(z)$ is equal to $H_i(z)/H_{ij}(z)$. Then, (11) can be written as:

$$G_i(z)H_{ij}(z)Q_{j/i}(z) = G_j(z)H_{ij}(z)Q_{i/j}(z) \quad . \quad (13)$$

After cancellation of the $H_{ij}(z)$ we get;

$$G_i(z)Q_{j/i}(z) = G_j(z)Q_{i/j}(z) \quad . \quad (14)$$

Since, $Q_{j/i}(z)$ and $Q_{i/j}(z)$ have no common zeros;

$$Q_{i/j}(z)|G_i(z) \quad , \quad \forall i, j \quad . \quad (15)$$

At this stage we use the following lemma:

Lemma 2. *If $Q_{i/j}(z)|G_i(z)$ for all i, j then $H_i(z)|G_i(z)$ for all i .*

Lemma 2 implies that $G_i(z)$ can be factored as:

$$G_i(z) = F_i(z)H_i(z) \quad . \quad (16)$$

We complete the proof of Theorem 1 using the following lemma:

Lemma 3. $F_i(z) = F_j(z)$ for $1 \leq i, j \leq M$.

This completes the proof of $G_i(z) = F(z)H_i(z)$ for all i . \square

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

Proof. Theorem can be proved, by expressing the minimizers of (7) as $\mathbf{g}_i = \mathbf{H}_i \mathbf{f}$, $1 \leq i \leq M$, where \mathbf{H}_i is the non-singular convolution matrix corresponding to the i^{th} channel. Since these relations can be compactly written as $\mathbf{g} = \mathbf{H} \mathbf{f}$, where $\mathbf{H} = [\mathbf{H}_1^T \ \cdots \ \mathbf{H}_M^T]^T$ is a non-singular channel matrix it follows that the dimension of the solution space is equal to the dimension of the arbitrary vector \mathbf{f} . \square

An important implication of this theorem is stated as:

Corollary 1. *The matrix $\mathbf{R}_{\mathbf{y}\mathbf{y}}[N]$ has an $\hat{L}_c + 1$ dimensional null-space.*

By starting with \hat{L}_2 which is larger than L_2 , one can form $\mathbf{R}_{\mathbf{y}\mathbf{y}}[N]$ in (7). Then by using Corollary 1, L_2 can be obtained as $L_2 = \hat{L}_2 \Leftrightarrow \eta + 1$, where η is the dimension of the null-space of $\mathbf{R}_{\mathbf{y}\mathbf{y}}[N]$. 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 $J_{L_2}(\mathbf{g}; N)$ would be in the form

$$g_i[n] = f[0]h_i[n] \quad , \quad \text{for } i = 1, \dots, M \quad (17)$$

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 $h_1[n], \dots, h_M[n]$ should be in the feasible set. As it can be shown easily, the constraint $\|\mathbf{g}\|^2 = \mu^2$ meets this requirement [12].

3 A Tree Algorithm

As pointed in the previous section, true channel coefficients can be obtained as the constrained minimizer of (7). For $\hat{L}_2 = L_2$, the direct implementation of the corresponding closed form solution requires $O(M^3 L_2^3 + M^2 L_2^2 N)$ operations for estimation of all the sub-channels $h_1[n], \dots, h_M[n]$ using N consecutive channel outputs [12]. In this section we propose to use a binary-tree algorithm to reduce the computational load to $O(ML_2^3 + ML_2^2 N)$.

The basic idea of the binary-tree algorithm will be illustrated for $M = 4$ using the re-arranged multi-channel filter model in Fig. 2. In this figure D represents the overall delay of the communication system, $H_c(z)$ denotes the common finite zeros to the 4 channels, $H_{ij}(z)$, $i \neq j$, denotes the zeros common to the i^{th} and j^{th} channels apart from those in $z^{-D}H_c(z)$ and finally $Q_{i/j}(z) \triangleq H_i(z)/H_{ij}(z)$, $i \neq j$, denotes the zeros of the i^{th} channel which are not also zeros of the j^{th} channel. With this arrangement, we note that, all the filters in the dashed Boxes 1, 2 and 3 are coprime.

Below we give a two stage algorithm which is used to compute the FIR filters in Box 1 and their common input $u_{12}[n]$:

- 1) Since $Q_{1/2}(z)$ and $Q_{2/1}(z)$ are coprime, identify them using the procedure outlined in Section .
- 2) Identify their common input $u_{12}[n]$ using the Bezout identity [12], [13].

The same algorithm is also used to compute filters in Box 2 and their common input $u_{34}[n]$. Since the outputs of the filters in Box 3 have been identified, these filters and their common input $x_c[n]$ can be identified again using the same algorithm. This completes the identification of the sub-channels $h_1[n], \dots, h_4[n]$.

4 Identification of the Common Zeros

In the previous section we have presented a method for blind identification of the sub-channels $h_1[n], \dots, h_M[n]$ in Fig. 1 and their common input $x_c[n]$. Thus the overall problem is reduced to the blind identification of the filter $h_c[n]$ in Fig. 1. In this section we give a solution to this problem.

We write the input-output relation of the filter $h_c[n]$ using the vector/matrix notation¹:

$$\underbrace{\begin{bmatrix} x_c[D] \\ x_c[D+1] \\ \vdots \\ x_c[D+L_1] \end{bmatrix}}_{\mathbf{x}_{c,L_1}} = \underbrace{\begin{bmatrix} a[0] & 0 & \cdots & 0 \\ a[1] & a[0] & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a[L_1] & a[L_1 \Leftrightarrow 1] & \cdots & a[0] \end{bmatrix}}_{\mathbf{A}_{L_1}} \underbrace{\begin{bmatrix} h_c[0] \\ h_c[1] \\ \vdots \\ h_c[L_1] \end{bmatrix}}_{\mathbf{h}_c} \quad (18)$$

In the above equation the vector \mathbf{x}_{c,L_1} has been computed as explained in the previous section. Since the information symbols $\{a[0], \dots, a[L_1]\}$ are chosen from a finite alphabet with size \mathcal{N} , there are only \mathcal{N}^{L_1+1} distinct and possible \mathbf{A}_{L_1} matrices. The algorithm that we propose in this section first computes all possible channel estimates that can lead to the output sequence $x_c[D], \dots, x_c[D+L_1]$. For example, if $\mathbf{A}_{L_1}^{(k)}$

¹Using an over-estimate of L_1 does not change the results in this section. It only increases the computational load.

is one of the possible \mathbf{A}_{L_1} matrices then the corresponding channel estimate $\hat{h}_c^{(k)}$ is obtained by solving (18) using the forward-substitution. Since the sequence $x_c[n]$ can be computed exactly in the noise-free case one of the computed channel estimates, $\hat{h}_c^{(k)}$, should be the actual channel. The *pruning algorithm* given in Page 4 can be used to identify the actual channel among all those possible channel estimates. Basically at each sample index n , the algorithm discards the channel estimates which cannot produce the most recent output sample $x_c[n]$. The algorithm terminates when there remains only one channel estimate. Also it should be noted that when the correct channel is identified using the *pruning algorithm*, the configuration in Fig. 3 provides a solution for the input sequence $a[n]$.

There exist some pathological cases [12], where two or more more input sequence-channel pairs produce the same output $x_c[n]$. Now we show that for a given output the set of these pathological cases is discrete. Thus, under the assumption that the channel coefficients are realization of continuous random variables the total probability of this set of pathological cases is zero. Therefore in practice the *pruning algorithm* almost surely identifies $h_c[n]$.

Theorem 3. For a given output sequence $x_c[n]$ which is not identically zero, the set of input and channel pairs $(a[n], h_c[n])$ which accepts $x_c[n]$ as a valid output is a discrete set.

Proof. Since the allowed input sequences are all discrete, we have to show that the corresponding channel pairs are isolated points in the set of all possible channels. This can be shown by contradiction. Let $\epsilon > 0$ be given. Then there exists pairs $(a_1[n], h_{c_1}[n])$ and $(a_2[n], h_{c_2}[n])$ such that

$$0 < \sum_{n=0}^{L_1} |h_{c_1}[n] \Leftrightarrow h_{c_2}[n]|^2 < \epsilon \quad (19)$$

which satisfy the following transform domain identity:

$$X_c(z) = z^{-D} A_1(z) H_{c_1}(z) = z^{-D} A_2(z) H_{c_2}(z) \quad (20)$$

Let $H_{c_2}(z) = H_{c_1}(z) + \delta H_c(z)$, $A_2(z) = A_1(z) + \delta A(z)$. Substituting these into (20) yields:

$$H_{c_1}(z) \delta A(z) + \delta H_c(z) A_1(z) + \delta H_c(z) \delta A(z) = 0 \quad (21)$$

In this equation, $\delta A(z)$ is the z -transform of a sequence whose samples can take values only in a discrete set, and as implied by (19) the total energy in the coefficients of $\delta H_c(z)$ is less than ϵ . Hence, when $\epsilon \rightarrow 0$, (21) implies that $H_{c_1}(z) \delta A(z) = 0$. Since $\delta A(z) \neq 0$ (because $H_{c_1}(z) \neq H_{c_2}(z)$ in (20)), the above condition can be satisfied only if $H_{c_1}(z) = 0$ which implies $X_c(z) = 0$, hence a contradiction. \square

5 Conclusions

A theoretical investigation on the blind identification of the channel and input sequence from the noise free observation of the fractionally spaced channel outputs is presented. In the case of no common zeros between the channels, the channel identification problem is posed as a constrained minimization problem involving channel outputs and the channel estimates such that only at a constant multiple of the actual channel the global minima is reached. A novel *binary-tree algorithm* is

proposed for the computationally efficient identification of the channel and the input sequence. Also, the minimum number of channel output samples required by the algorithm is found. In the case of common zeros between the parallel branches of the fractionally spaced channel model, a second stage of processing is proposed for the almost sure identification of the channel and the input sequence.

References

- [1] S. Haykin, *Adaptive Filter Theory*. New Jersey: Prentice-Hall, third ed., 1996.
- [2] 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.
- [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] 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.
- [5] 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, pp. 340-349, Mar. 1994.
- [6] 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.
- [7] H. Liu and G. Xu, "Closed-form blind symbol estimation in digital communications," *IEEE Trans. Signal Process.*, vol. 43, pp. 2714-2723, Nov. 1995.
- [8] 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.
- [9] A. Touzni and I. Fijalkow, "Robustness of blind fractionally-spaced identification/equalization to loss of channel disparity," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, pp. 3937-3940, 1997.
- [10] Z. Ding, "Characteristics of band-limited channels unidentifiable from second-order cyclostationary statistics," *IEEE Trans. Signal Processing Letters*, vol. 3, pp. 150-152, May 1996.
- [11] M. I. Gürelli and C. L. Nikias, "Evam: An eigenvector-based algorithm for multichannel blind deconvolution of input colored signals," *IEEE Trans. Signal Process.*, vol. 43, pp. 134-149, Jan. 1995.
- [12] A. K. Özdemir, "Exact blind channel estimator," Master's thesis, Bilkent University, Dept. of Electrical and Electronics Engineering, June 1998.
- [13] T. Kailath, *Linear Systems*. New Jersey: Prentice-Hall, 1980.

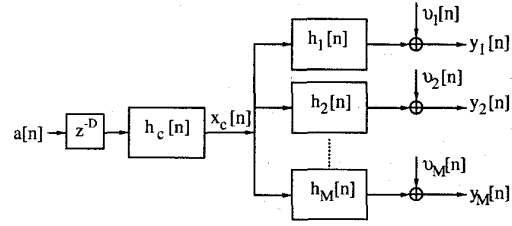


Figure 1: Multi-channel filter model.

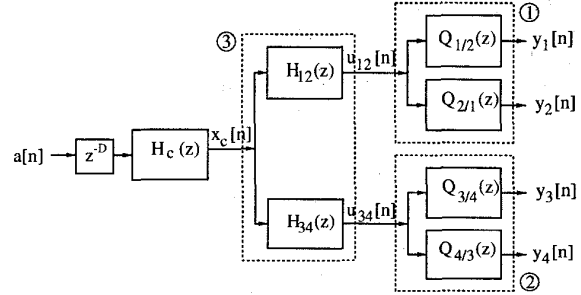


Figure 2: The binary-tree representation of a 4-channel FIR filter.

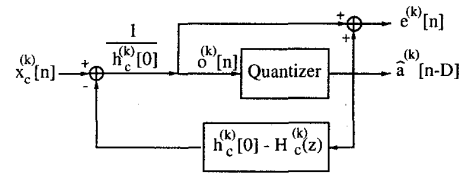


Figure 3: Identification of the common zeros.

Algorithm 1 The pruning algorithm

Initialization:

Define the current sample index n and the set of remaining channel estimates \mathcal{S}^{n-1} at the previous sample index:

$$n = D + L_1 + 1$$

$$\mathcal{S}^{n-1} = \{h_c^{(0)}, h_c^{(1)}, \dots, h_c^{N^{L_1+1}}\}$$

Pruning loop:

while $\text{size}(\mathcal{S}^{(n-1)}) > 1$ do

Set $\mathcal{S}^{(n)} := \mathcal{S}^{(n-1)}$.

Compute $x_c[n]$ using the results in Section .

for each channel estimate $\{h_c^{(k)}\} \in \mathcal{S}^{(n-1)}$ do

Compute $o^{(k)}[n]$ and $\hat{a}^{(k)}[n \leftrightarrow D]$ as in Fig. 3.

if the residual error

$$|e^{(k)}[n]| \triangleq |o^{(k)}[n] \leftrightarrow \hat{a}^{(k)}[n \leftrightarrow D]| \quad (22)$$

exceeds a threshold then

Set $\mathcal{S}^n := \mathcal{S}^n \leftrightarrow \{h_c^{(k)}\}$.

end if

end for

Set $n := n + 1$.

end while