Paper

Blind Estimation of Linear and Nonlinear Sparse Channels

Kristina Georgoulakis

Department of Science and Technology of Telecommunications, University of Peloponnese, Tripolis, Greece

Abstract—This paper presents a Clustering Based Blind Channel Estimator for a special case of sparse channels – the zero pad channels. The proposed algorithm uses an unsupervised clustering technique for the estimation of data clusters. Clusters labelling is performed by a Hidden Markov Model of the observation sequence appropriately modified to exploit channel sparsity. The algorithm achieves a substantial complexity reduction compared to the fully evaluated technique. The proposed algorithm is used in conjunction with a Parallel Trellis Viterbi Algorithm for data detection and simulation results show that the overall scheme exhibits the reduced complexity benefits without performance reduction.

Keywords—blind estimation and equalization, clustering techniques, sparse zero pad channels.

1. Introduction

The last years an intense research effort has risen in communications problems involving the estimation and equalization of sparse channels, i.e., channels with a large delay spread but with a small non zero support. Estimation and equalization of sparse channels is a challenging problem and sparsity aware estimators and equalizers should be used in order to improve system performance and reduce complexity. Sparse channels are encountered, among others, in High Definition Television (HDTV) [1], in broadband wireless communications [2] and in underwater acoustic channels [3].

Various training based estimators have been developed for the estimation of sparse channels [4]–[6]. Recently a blind algorithm based on the Expectation-Maximization (EM) algorithm for sparse channel estimation has been proposed [7]. The main drawback of the algorithm is its computational burden growing exponentially with the channel length. In this paper the Cluster Based Blind Channel Estimation algorithm (CBBCE) [8] is evaluated for a special class of sparse channels, the zero pad channels. The proposed algorithm, in order to exploit the structured sparsity of zero pad channels, uses a modification of the cluster based blind channel estimation procedure leading to a much lower complexity.

The CBBCE algorithm consists of two steps. Data clusters are first estimated via an unsupervised learning technique and next labelling of the estimated clusters is achieved by unravelling the information hidden in the sequence of received data. Labelling is performed using a Hidden Markov Model (HMM) of the estimation process and by relating data clusters to HMM states. The probability of each cluster to correspond to a specific label is treated as the unknown parameter of the HMM learning task implemented by the EM algorithm [8], [9].

Assuming an *M-ary* alphabet for the data symbols and a channel, with length $L+1$, the HMM is typically evaluated with M^L states [7]. However, in the sparse channel case where only a small fraction of the channel taps is active (i.e., $s+1 \ll L+1$), the full evaluation of the HMM is computationally inefficient [10]. In the zero pad channels all the non zero taps are placed on a regular grid [11], [12]. In this case, the memory of the channel concerns only the transmitted data that correspond to the non-zero taps. Thus, by involving in the HMM states only the $(s+1)$ data corresponding to the non zero taps, the number of states is reduced to M^s. Then, the reduced states HMM results in the appropriate labelling of the data clusters.

The outline of the paper is as follows. In Section 2 system description is given. The proposed CBBCE algorithm for sparse channels is presented in Section 3. In Section 4 the various channel cases where the algorithm can be applicable are referred (i.e., linear and non linear zero pad sparse channels, and a special case of group sparse channels). The tremendous complexity reduction achieved by the algorithm, compared to the full evaluated HMM, is also discussed in the same Section. In Section 5 the use of a Parallel Trellis Viterbi Algorithm (PTVA) [11] employing the channel estimates of the proposed algorithm is evaluated. The performance of the entire scheme in terms of the achieved Bit Error Rate (BER) is illustrated. Finally, conclusions are drawn in Section 6.

2. System Description

Consider the discrete time system described by:

$$g(t) = c(t) + w(t)\,, \tag{1}$$

where

$$c(t) = F(I(t), I(t-1), \ldots, I(t-L)) \tag{2}$$

is the noiseless channel output sequence, $F(.)$ is the function representing the channel action, $I(t)$ is an equiprobable sequence of independent and identically distributed (i.i.d.) transmitted data taken from an *M-ary* alphabet, and $w(t)$ is Additive White Gaussian Noise (AWGN). The channel length is assumed to be $L+1$, however with only $s+1$ taps being non zero. The received data form $Q = M^{L+1}$

clusters in the one dimensional space [8]. Each cluster is represented by a suitably chosen representative which corresponds to the noiseless channel response, i.e.:

$$c(t) \in (c_k,\ k = 1, 2, \ldots, Q).$$

Here, due to channel sparsity, the actual number of clusters formed is:

$$Q = M^{s+1},$$

since the zero valued taps do not contribute to the formation of clusters.

Zero pad channels are sparse channels of a specific form, whose channel impulse response is described by [11]–[14]:

$$H = [h_1 \underbrace{0\ldots0}_{f\ zeros} h_2 \underbrace{0\ldots0}_{f\ zeros} \ldots h_s \underbrace{0\ldots0}_{f\ zeros} h_{s+1}]^T, \qquad (3)$$

where f is the number of zeros between the non zero valued taps and $L = s(f+1)$. In the case of zero pad channels, the received data depend on alternated data symbols [11], [13]. In this case, the noiseless channel output sequence (2) takes the form:

$$c(t) = F(I(t), I(t-(f+1)), \ldots, I(t-s(f+1))). \qquad (4)$$

3. Clustering Based Blind Channel Estimator for Zero Pad Sparse Channels

Channel estimation can be performed either using a known training sequence of data or identifying the channel based only on the received data (blind mode). Blind channel estimation based on data clustering techniques has been developed for general (non-sparse) channels [8], [9], [15]. The data clustering technique will be adopted in this paper to evaluate a blind estimator for structured sparse channels. Clustering based blind channel estimation is performed in two steps as it is detailed in [16] where initially the clusters representatives c_k are estimated via an unsupervised clustering technique following by clusters labelling, where each cluster is mapped to a specific sequence of transmitted data. When this technique is applied in the case of channels exhibiting a zero pad sparsity profile, the first step remains unaltered. In the second step, the structured sparsity of channels under investigation is taken into account resulting in a novel, reduced complexity labelling procedure. These two tasks are detailed in the sequel.

3.1. Unsupervised Clustering

An unsupervised learning technique is adopted for the estimation of the clusters representatives such as the Isodata algorithm, the Neural Gaz network, etc. [17], [18]. The clusters formed is the contribution of the non-zero taps of the channel only and the number of clusters estimated by the unsupervised clustering technique equals M^{s+1}.

3.2. Clusters Labelling through a Structured Sparsity Aware HMM

The transmitted data input vector:

$$\mathbf{I}(t) = [I(t)\ I(t-1) \ldots I(t-L)]^T,$$

can be described as a first order Markov chain having M^L states denoted by $S(t)$. Since the received data $g(t)$ are a probabilistic function of the state vector $\mathbf{I}(t)$, the channel estimation problem can be formulated as a HMM parameter estimation problem. Thus, a standard HMM parameter estimation algorithm, referred to hereafter as fully evaluated HMM algorithm (FE-HMM), considering M^L states can be applied, being however impractical from the computational point of view, apart from the case when the channel memory L is sufficiently small, which is not the typical case of sparse channels. Since the actual number of the clusters formed is $Q = M^{s+1}$ only, and for reasons of complexity reduction, the proposed algorithm considers M^s states in the HMM, resulting to a novel scheme referred to hereafter as the reduced evaluated HMM algorithm (RE-HMM). This task can be achieved considering instead the cluster model (4), where the actual channel memory pattern is taken into account. In this way, the labelling of the states of the HMM considers transmitted data that are $f+1$ time units apart. Based on the above remarks, the discrete observations RE-HMM for the sparse zero pad channel is characterized by the following elements:

- The states of the model, which according to Eq. (4) are formed as:

  $$S(t) \rightarrow (I(t-(f+1)), \ldots, I(t-s(f+1))). \qquad (5)$$

  The number of states in this case equals to:

  $$N = M^s,$$

  as opposed to the number of states $N' = M^L$ required by the FE-HMM approach.

- The state transition probabilities a_{ij}, which are defined as

  $$a_{ij} = P[S(t+(f+1)) = j \mid S(t) = i], \qquad (6)$$
  $$1 \leq i, j \leq N.$$

  Notice that for each allowable transition ($a_{ij} = 1/M$) a specific noiseless channel output occurs. In other words, each state transition specifies uniquely a cluster label and a cluster transition arises every $f+1$ samples. The cluster labels are specified by:

  $$X(t) \rightarrow (I(t), I(t-(f+1)), \ldots, I(t-s(f+1))) \qquad (7)$$

  and each cluster label, $X(t)\ \varepsilon\ (n_k,\ k = 1, 2, \ldots, Q)$, corresponds to a specific cluster c_k.

- The distinct observation symbols per transition, which in this case are the clusters, c_k.

- The probabilities for each symbol to occur and for each state transition i to j, which denote the probability of a specific cluster to correspond to a specific label, i.e.:

$$\begin{aligned} b_{n_k}(c_k) &= P[c_k \mid S(t)=i, S(t+(f+1))=j] \\ &= P[c_k \mid X(t)=n_k], \end{aligned} \tag{8}$$

$$1 \le k \le Q, \quad 1 \le i,j \le N.$$

- The initial state distribution:

$$\pi_i = P[S(1)=i], \tag{9}$$

$$1 \le i \le N.$$

In the cluster based blind channel estimation procedure clusters' labelling is treated as a HMM learning problem. The EM algorithm is a commonly used numerical iterative scheme to obtain Maximum Likelihood (ML) estimates of a HMM. The resulting ML estimate is given by:

$$\hat{\theta} = argmax_{\theta} P(G \mid \theta), \tag{10}$$

where $P(G \mid \theta)$ denotes the probability of the observation sequence G of length T:

$$G = (g(1), g(2), \ldots g(T))^T,$$

given the model parameters (θ) with:

$$\theta = [b_{n_k}(c_k)],\ k = 1, \ldots, Q.$$

Thus, θ is the $Q \times Q$ probability matrix that maps labels to clusters and it is expected to converge to a matrix whose elements converge either to one, for the case when a specific symbol corresponds to a specific label, or to zero, otherwise. Convergence of the algorithm is achieved when:

$$P(G \mid \theta) > p, \tag{11}$$

with p a predetermined threshold [19].

Clusters' labelling using the RE-HMM approach described by Eqs. (5)–(9) requires the knowledge of the structure of the comb type channel response, which in turn requires the estimation of the distance or the number of unit time delays between all successive non zero elements of the model. In the case of zero pad sparse channels treated in this paper, and due to the specific form of the sparsity structure only a single time delay parameter has to be determined. The required time delay parameter d is estimated using an exhaustive search procedure, starting from $d = 1$, where for each candidate value d, a RE-HMM estimate $\hat{\theta}$ is obtained using Eqs. (5)–(9). When the algorithm converges (11), then the correct value of d is reached ($d = f+1$) and the correct channel structure is obtained. Then, θ provides the correct labelling.

The proposed CBBCE algorithm is summarized in Table 1. This procedure is further illustrated by a simple example using a channel with impulse response:

$$H = [h_1 \quad h_2 \quad h_3]^T = [1 \quad 0 \quad 0.5]^T,$$

Table 1
The proposed Clustering Based Blind Channel Estimation algorithm

CBBCE algorithm

1. Unsupervised clustering:

- Estimation of the clusters representatives by an unsupervised learning technique.
- The number of the estimated clusters reveals the number of non-zero taps $(s+1)$.

2. Labelling through a RE-HMM

Initialization

Set: Number of states, $N = M^s$

Time delay parameter, $d = 1$

Main

Repeat until convergence (11)

HMM formulation ((5)–(9)) with states:

$S(t) = (I(t-d)I(t-2d)\ldots I(t-sd))$.

$d = d+1$

End

- The correct value of d is reached $(f+1)$,
- The ML estimate, (θ), reveals the labels – clusters correspondence.

where the input data are assumed to be bipolar (i.e., $I(t) = \pm 1$) and the Signal to Noise Ratio (SNR) is set equal to 17 dB. In this particular case, the number of non-zero taps is $s+1 = 2$, while the channel length is $L+1 = 3$. Since the data alphabet consists of two symbols, the number of clusters assumed is $Q = 4$. Following the first step of the proposed algorithm, the clusters representatives are estimated using an unsupervised clustering technique. Specifically, Isodata is used for clusters estimation, using $T = 30$ received data, obtaining the estimates:

$$\hat{c}_1 = 1.512,\ \hat{c}_2 = 0.49,\ \hat{c}_3 = -0.507,\ \hat{c}_4 = -1.49.$$

Following the second step of the proposed algorithm, the RE-HMM is formed, with $N = 2$. Initially, d is set equal to 1, thus, a cluster transition is assumed to arise every single sample. Since the assumed channel model is not the correct $((d=1) \neq (f+1=2))$, the algorithm does not converges according to Eq. (11). The probability matrix θ, after 15 iterations, is shown in Table 2. Obviously, the identification procedure does not converge and no labels – clusters correspondence can be derived. Then, the time

Table 2
Probabilities matrix for channel H = $[1\ 0\ 0.5]^T$ after 15 iterations and $d = 1$. Clusters labels cannot be unravelled

Label		Cluster representative			
I(t)	I(t–1)	c_1	c_2	c_3	c_4
–1	–1	0.1235	0.2393	0.4301	0.2071
–1	1	0.2703	0.0632	0.2423	0.4242
1	–1	0.1841	0.5817	0.1136	0.1206
1	1	0.3477	0.1588	0.1175	0.3759

delay parameter, d, is set equal to 2 and a new RE-HMM is formed. This time the algorithm converges. The probability matrix θ, after 10 iterations, converges as it appears in Table 3. In this case the labels – clusters mapping is easily achieved.

Table 3
Probabilities matrix after convergence, for the channel with impulse response H = $[1\ 0\ 0.5]^T$ and time delay parameter $d = 2$

Label		Cluster representative			
I(t)	I(t–1)	c_1	c_2	c_3	c_4
–1	–1	0	0	0	1
–1	1	0	0	1	0
1	–1	0	1	0	0
1	1	1	0	0	0

Once the channel estimation process is accomplished, signal detection can be performed by employing a PTVA with reduced complexity [10]. The PTVA is a computationally improved reformulation of the Viterbi Algorithm (VA) which operates into a set of independent trellises for the zero pad sparse channels. The PTVA is optimum for zero pad sparse channels and results in complexity reduction compared to the ordinary VA which uses a single trellis. The evaluation of a channel equalizer employing the CBBCE algorithm followed by a PTVA is described in Section 5.

4. Case Studies

We proceed further our developments on case studies, where a variety of channels amenable to the application of the proposed channel identification method is considered. Complexity issues are also discussed. Notice that, for the sake of simplicity, the symbol values are assumed to be drawn from a binary alphabet set (i.e., $M = 2$).

4.1. Linear Zero Pad Sparse Channels

Linear zero pad sparse channels are successfully identified using the proposed method. Note that, the presence of an arbitrary time delay (number of zeros) at the edges of the non zero taps of the channel does not affect the algorithm, and channels with impulse response of the form:

$$H = [\underbrace{0\ldots0}_{x\ zeros} h_1 \underbrace{0\ldots0}_{f\ zeros} h_2 \underbrace{0\ldots0}_{f\ zeros} \ldots h_{s+1} \underbrace{0\ldots0}_{y\ zeros}]^T\,, \tag{12}$$

can be tackled by the method, including, for $f = 0$, a special case of group sparse channels where the non zero taps are located in a single cluster [6]. Consider for example a channel with impulse response given by:

$$H = [0.2\ 0\ 0\ 0\ 0\ 0.5\ 0\ 0\ 0\ 0\ 0.9]^T. \tag{13}$$

In this case we get $L+1 = 11$, $s+1 = 3$ and $f = 4$. Here, the clusters representatives $\hat{c}_k$, $k = 1,..,Q$ are estimated using the Isodata algorithm [18] from a received data sequence of length $T = 300$ [20]. Once clusters identification is completed, the task of clusters labelling is subsequently addressed. A RE-HMM with $N = 2^2$ states is formulated. Application of the proposed algorithm, as it is summarized in Table 1, results in $d = f+1 = 5$ and the states of the RE-HMM are formed by the (non-successive) data $S(t) = (I(t-5), I(t-10))$. The probabilities matrix θ resulting from the proposed algorithm, after convergence, is tabulated in Table 4.

Table 4
Probabilities matrix after convergence, for the channel with impulse response H = $[0.2\ 0\ 0\ 0\ 0\ 0.5\ 0\ 0\ 0\ 0\ 0.9]^T$, $d = 5$

Label			Cluster representative							
I(t)	I(t–5)	I(t–10)	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8
–1	–1	–1	0	0	0	1	0	0	0	0
–1	–1	1	0	0	1	0	0	0	0	0
–1	1	–1	0	0	0	0	0	0	1	0
–1	1	1	0	1	0	0	0	0	0	0
1	–1	–1	0	0	0	0	1	0	0	0
1	–1	1	0	0	0	0	0	0	0	1
1	1	–1	0	0	0	0	0	1	0	0
1	1	1	1	0	0	0	0	0	0	0

The Average Squared Error (ASE) is adopted as a metric of the accuracy of the estimated clusters representatives:

$$m = \frac{1}{Q}\sum_{k=1}^{Q}(c_k - \hat{c}_k)^2\,, \tag{14}$$

where $\hat{c}_k$ are the estimated clusters representatives and c_k are the noiseless clusters representatives that correspond to Eq. (13).

The ASE versus SNR is illustrated in Fig. 1. Figure 2 shows the impact of the received data sequence length, T, to the accuracy of the estimated clusters values, for SNR = 20 dB.

For the sake of comparison, a supervised Least - Absolute Shrinkage and Selection Operator (LASSO) [21] estimator is used as benchmark. Since LASSO is not capable

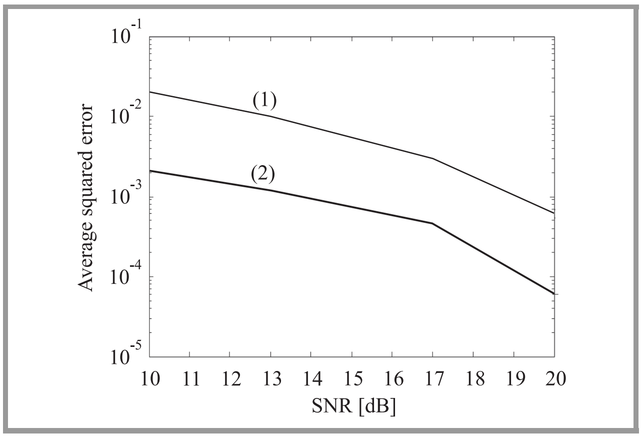

Fig. 1. Average squared error for a channel with impulse response H = [0.2 0 0 0 0 0.5 0 0 0 0 0.9]^T, using a sequence of $T = 300$ data. (1) – proposed CBBCE algorithm, (2) – supervised LASSO channel estimation.

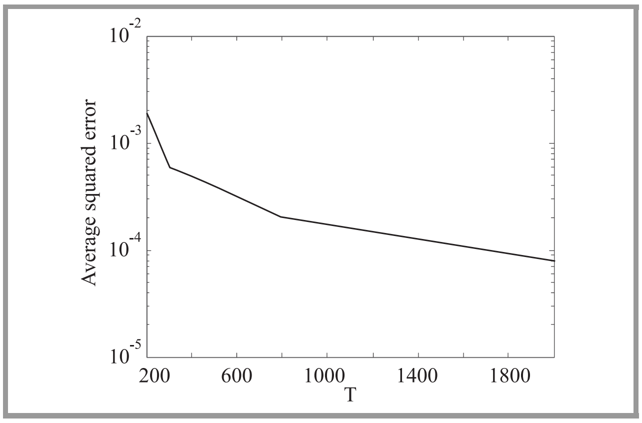

Fig. 2. Average squared error versus T, (number of received data used by the proposed CBBCE), for a channel with impulse response H = [0.2 0 0 0 0 0.5 0 0 0 0 0.9]^T, SNR = 20 dB.

of estimating the clusters representatives directly, the estimates of channel taps are used to calculate the respective clusters representatives. The ASE versus SNR in the case where LASSO is used for channel identification, is also shown in Fig. 1. From a first glance it is evidence that the proposed estimator lacks behind its supervised counterpart, a result which is somehow expected since LASSO is a supervised learning algorithm, while the proposed scheme is a blind identification algorithm. However, as it is shown in the following Section the BER performance of an equalizer using the estimates of the proposed algorithm is very close to that using the supervised LASSO as a channel estimator.

4.2. Nonlinear Zero Pad Sparse Channels

Clustering based estimation algorithms do not adopt any assumption for the impulse response of the channel under consideration, thus, they can efficiently be employed in the case of nonlinear channels [8]. For example a sparse linear channel with impulse response given by H = [0 0 1 0 0 0 0 0.5 0 0 0]^T followed by the nonlinear action described by $g(t)+0.1g(t)^2+0.3g(t)^3$ is considered. In this case we get, $L = 10$, $s = 1$ and $f = 4$. The proposed CBBCE algorithm uses $T = 80$ received data and performs clusters estimation with only $N = 2^1$ states. The resulting ASE versus SNR is shown in Fig. 3.

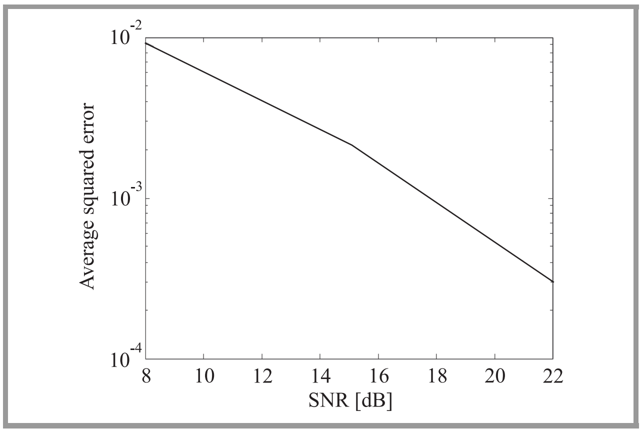

Fig. 3. Average squared error achieved by the proposed blind estimator, for a channel with impulse response H = [0 0 1 0 0 0 0 0.5 0 0 0]^T followed by the nonlinear action described by $g(t)+0.1g(t)^2+0.3g(t)^3$.

4.3. Complexity Assessment

The proposed algorithm reduces the complexity of the HMM scheme from $O(M^L \times T)$ required by the FE-HMM to $O(M^s \times (f+1) \times T)$ operations, required by the proposed RE-HMM, which is a tremendous reduction in the sparse channels case where $s \ll L$. For example, in the experiment described in Section 4.1, the proposed RE-HMM algorithm evaluates the HMM scheme using only $N = 2^2$ states. In this particular case, the HMM procedure is repeated $f+1 = 5$ times resulting in a complexity of $O(2^2 \times 5 \times T)$ operations which is a major improvement over the FE-HMM algorithm [7], [8] which requires $N' = 2^{10}$ states leading to a complexity of $O(2^{10} \times T)$ operations. In the experiment described in Section 4.2 the complexity of the RE-HMM reaches the $O(2 \times 5 \times T)$ operations while the FE-HMM involves $O(2^{10} \times T)$ operations.

5. Blind Clustering Based Equalizer for Zero Pad Channels

The proposed method can also be applied in the case when instead of channel estimation, channel equalization is under consideration. The proposed CBBCE algorithm combined with a PTVA performs blind clustering based sequence equalization, for the special case of zero pad channels, in an efficient way. We refer to the entire proposed scheme as Reduced Evaluated Blind Equalizer (REBE). Consider for example a channel with impulse response given by $H = [0.2\ 0\ 0\ 0\ 0\ 0.5\ 0\ 0\ 0\ 0\ 0.9]^T$ (13). The proposed CBBCE algorithm has already been described, for this channel, in Section 4.1. Since channel estimation is completed the PTVA algorithm is used for signal detection. The PTVA algorithm uses $f+1=5$ parallel trellises with $M^s=4$ states each [11]. The decision delay for the PTVA is 15. The resulting BER for the proposed REBE appears in Fig. 4.

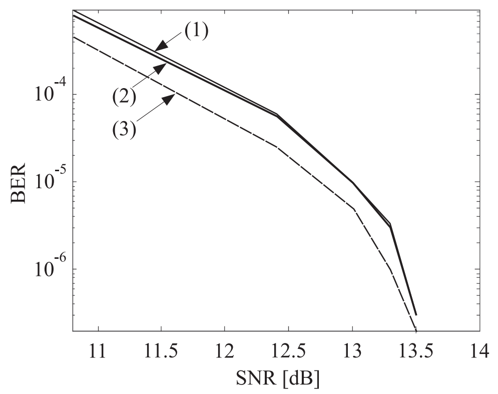

Fig. 4. BER versus SNR for a channel with impulse response H = $[0.2\ 0\ 0\ 0\ 0\ 0.5\ 0\ 0\ 0\ 0\ 0.9]^T$. (1) – proposed REBE (proposed reduced blind estimation algorithm and PTVA algorithm), (2) – FEBE (blind full evaluated clusters estimation algorithm and full evaluated VA), (3) – supervised LASSO channel estimation and full evaluated VA.

The performance of the conventional Full Evaluated Blind clustering based Equalization scheme [8] (FEBE) is also investigated. The FEBE performs clusters identification by an unsupervised clustering technique and clusters labelling using the FE-HMM and data detection through a conventional VA using $M^L = 2^{10}$ states. The decision delay for the VA is 30.

As seen from the Fig. 4 the performance of the FEBE is the same with the REBE.

Moreover, for the sake of comparison, an equalization scheme formed by a supervised estimator and a full – evaluated VA is realized and used as a benchmark. Channel estimation is achieved by the supervised LASSO algorithm [21]. Then the channel estimator is followed by a conventional VA, with 2^{10} states. The number of training data used for the estimator is 300. The decision delay for the VA is 30. As seen from Fig. 4 the resulting BERs of the three schemes are very close, however, the proposed REBE works at a substantially reduced complexity.

6. Conclusions

In this paper a novel reduced complexity blind estimator for zero pad channels is presented. The proposed scheme uses a Clustering Based Blind Channel Estimation algorithm extended to account for the structured sparsity of zero pad channels and exhibits a tremendous complexity reduction compared to the full evaluated counterpart. The algorithm is suitable both for linear and nonlinear channels. The proposed algorithm combined with a Parallel Trellis Viterbi Algorithm is used for signal detection and the proposed sequence equalization scheme exhibits, at a reduced complexity, a performance similar to that compared to other competitive schemes such as a full evaluated blind clustering based sequence equalizer and a supervised LASSO estimator accompanied by a Viterbi Algorithm. Modification of the algorithm for the expansion of its use to tackle general sparse channels is under investigation.

References

[1] W. F. Schreiber, "Advanced television systems for terrestrial broadcasting: Some problems and some proposed solutions", *Proc. IEEE*, vol. 83, pp. 958–981, 1995.

[2] S. Ariyavisitakul, N. R. Sollenberger, and L. J. Greenstein, "Tap selectable decision feedback equalization", *IEEE Trans. Commun.*, vol. 45, pp. 1497–1500, 1997.

[3] M. Kocic, D. Brady, and M. Stojanovic, "Sparse equalization for real time underwater acoustic communications", in *Proc. MTS/IEEE Conf. "Challenges of our changing global environment" OCEANS'95*, San Diego, CA, USA, 1995, pp. 1417–1422.

[4] J. A. Tropp and S. J. Wright, "Computational methods for sparse solution of linear inverse problems", *IEEE Proc.*, vol. 98, pp. 948–958, 2010.

[5] B. Babadi, N. Kalouptsidis, and V. Trokh, "SPARLS: The sparse RLS algorithm", *IEEE Trans. Sig. Proces.*, vol. 58, pp. 4013–4025, 2012.

[6] Y. Chen and A. O. Hero, "Recursive $\ell_{1,\infty}$ Group lasso", *IEEE Trans. Sig. Proces.,* vol. 60, pp. 3978–3987, 2012.

[7] G. Mileounis, N. Kalouptsidis, B. Babadi, and V. Tarokh, "Blind identification of sparse channels and symbol detection via the EM algorithm", in *Proc. IEEE DSP Conf.*, Corfu, Greece, 2011, pp. 1–5.

[8] K. Georgoulakis and S. Theodoridis, "Blind and semi-blind equalization using hidden Markov models and clustering techniques", *Sig. Proces.*, vol. 80, pp. 1795–1805, 2000.

[9] P. Wang and W. Ser, "On cluster-based channel identification", *Procedia Engin.*, vol. 29, pp. 2699–2704, 2012.

[10] R. Niazadeh, S. H. Ghalehjegh, M. B. Zadeh, and C. Jutten, "ISI sparse channel estimation based on SLO and its application in ML sequence-by-sequence equalization", *Sig. Proces.*, vol. 92, pp. 1875–1885, 2012.

[11] N. C. McCinty, R. Kennedy, and P. Hoeher, "Parallel trellis Viterbi algorithm for sparse channels", *IEEE Communi. Lett.*, vol. 2, pp. 143–145, 1998.

[12] B. Cristea, D. Roviras, and B. Escrig, "Maximum likelihood sequence estimation based on periodic time-varying trellis for LPTVMA systems", in *Proc. EUSIPCO 2006*, Florence, Italy, 2006.

[13] J. Mietzner, S. Hoeher, S. Land, and P. Hoeher, "Trellis equalization for sparse ISI channels revisited", in *Proc. IEEE Int. Symp. Inform. Theory ISIT 2005*, Adelaide, Australia, 2005, pp. 229–233.

[14] F. K. H. Lee and P.J. McLane, "Iterative parallel-trellis MAP equalizers with nonuniformly-spaced prefilters for sparse multipath channels", in *Proc. 56th VTC 2002*, Vancouver, Canada, 2002, vol. 4, pp. 2201–2205.

[15] Y. L. Jeng and C. C. Yeh, "Cluster based blind nonlinear channel estimation", *IEEE Trans. Sig. Proces.*, vol. 45, pp. 1161–1172, 1997.

[16] K. Georgoulakis and S. Theodoridis, "Efficient clustering techniques for channel equalization in hostile environments", *Sig. Proces.*, vol. 58, pp. 153–164, 1997.

[17] T. M. Martinetz, S. G. Bekovich, and K. J. Schulten, "Neural-gas network for vector quantization and its application in time-series prediction", *IEEE Trans. Neural Netw.*, vol. 4, pp. 558–569, 1993.

[18] S. Theodoridis and K. Koutroubas, *Pattern Recognition*. New York: Academic Press, 1998.

[19] G. Kawas and R. Vallet, "Joint parameter estimation and symbol detection for linear or nonlinear unknown channels", *IEEE Trans. Commun.*, vol. 42, pp. 2406–2413, 1994.

[20] F. Balazs and S. Imre, "Modified radial basis network based blind channel estimation", in *Proc. SoftCOM 2004*, Dubrovnik, Croatia, 2004, pp. 449–452.

[21] R. Tibshirani, "Regression shrinkage and selection via the lasso", *J. Royal Stat. Soc.*, vol. 58, pp. 267-288, 1996.

Kristina Georgoulakis received the B.S. degree in Informatics from the Department of Informatics and Telecommunications, University of Athens, Greece, in 1990 and her M.Sc. and Ph.D. degree in 1993 and 1999, respectively, both from the same department. Her research interests are in signal processing for communications systems.

E-mail: kristina@uop.gr
Department of Science and Technology
of Telecommunications
University of Peloponnese
Tripolis, 22100 Greece