SEMI-BLIND SPARSE CHANNEL ESTIMATION WITH CONSTANT MODULUS SYMBOLS

Müjdat Çetin[†] *and Brian M. Sadler*[‡]

[†]Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
77 Massachusetts Ave., Cambridge, MA 02139, USA
[‡]Army Research Laboratory, AMSRL-CI-CN, Adelphi, MD 20783, USA

ABSTRACT

We propose two methods for estimation of sparse communication channels. In the first method, we consider the problem of channel estimation based on training symbols, and formulate it as an optimization problem. In this formulation, we combine the objective of fidelity to the received data, with a non-quadratic constraint reflecting the prior information about the sparsity of the channel. This approach leads to accurate channel estimates with much shorter training sequences than conventional methods. The second method we propose is aimed at taking advantage of any available training-based data, as well as any "blind" data based on unknown, constant modulus symbols. We propose a semi-blind optimization framework making use of these two types of data, and enforcing the sparsity of the channel as well as the constant modulus property of the symbols. This approach improves upon the channel estimates based only on training sequences, and also produces accurate estimates for the unknown symbols.

1. INTRODUCTION

In many wireless communication systems, the propagation channels involved exhibit a large delay spread, but a sparse impulse response consisting of a small number of dominant echoes. A primary example is terrestrial transmission of high definition television (HDTV) signals [1,2]. In Fig. 1 we show an example of such a sparse channel impulse response. Conventional least-squares channel estimation techniques do not exploit the sparse structure of such channels and require the transmission of many training symbols to generate an accurate estimate. Recently, a matching pursuit algorithm that exploits the sparse structure of such channels has been proposed [3]. This approach has also been extended to multiuser environments [4]. In the initial portion of our work presented in this paper, we propose a channel estimation technique that has a similar goal of exploiting sparsity. In contrast with the approach in [3], we formulate the channel estimation problem as a non-quadratic optimization problem involving a data fidelity term, and an ℓ_1-norm-based, sparsity-enforcing regularization term. Both matching pursuit and ℓ_1-norm regularization (also called basis pursuit) can be viewed as trying to solve a combinatorial sparse signal representation problem in a suboptimal fashion. Matching pursuit provides a greedy solution, while ℓ_1-norm-based methods replace the original problem with a relaxed version for tractability. Recent work has illuminated interesting theoretical properties of both approaches. The ℓ_1-norm-based approach we propose for channel estimation has regularization built into it to provide robustness against noise. Furthermore, the optimization-based nature of our framework makes it easy to extend these ideas to the semi-blind case, which we discuss next.

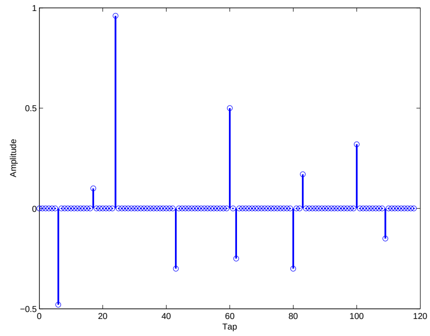

Fig. 1. A sample sparse channel impulse response (adapted from [3]).

The discussion above was implicitly focused on channel estimation in the presence of training symbols. Most current wireless communication systems depend on the transmission of such known symbols for channel estimation and equalization. However, use of training symbols limits the effective transmission bandwidth. Therefore, it is of interest to reduce the number of training symbols. On the other hand, blind equalization techniques do not require training. One of the most popular blind techniques is based on the so-called constant modulus algorithms [5]. While such blind algorithms have good performance with long data sequences, they may not achieve equalization in a short burst.[1] In order to alleviate this problem, a number of researchers have recently proposed methods which

This work was supported by the Army Research Office under Grant DAAD19-00-1-0466.

[1]Although most constant modulus algorithms are applied to long data sequences, there is also some recent work on finite-interval constant modulus algorithms [6].

couple training-based and blind techniques, leading to so-called *semi-blind* methods [7,8]. These methods provide an attractive tradeoff between training-based and blind techniques. However we are not aware of any semi-blind technique designed explicitly for, and exploit the characteristics of, sparse communication channels. We propose a semi-blind sparse channel estimation technique for the case of constant modulus symbols. In particular, we formulate an optimization problem which contains terms for fidelity to the training-based as well as blind data, an ℓ_1-norm term for enforcing channel sparsity, and a term enforcing the constant modulus property of the symbols. The solution of this optimization problem yields both an estimate for the channel impulse response, and estimates for the transmitted unknown symbols. Our experiments on simulated data demonstrate both the improvements of our sparse channel estimation framework over conventional techniques, and also how our semi-blind approach improves over our channel estimation technique based only on training data.

2. OBSERVATION MODEL

Let the training symbols $s_T(n)$, $n = 0, ..., N_T - 1$ be transmitted through a channel with impulse response $c(m)$, $m = 0, ..., N_c - 1$. We can model the observed signal samples $y_T(n)$ as:

$$y_T(n) = \sum_{m=0}^{N_c-1} s_T(n-m)c(m) + v(n), \quad n = 0, ..., N_T - 1 \quad (1)$$

where $v(n)$ denotes the measurement noise. We can write this equation in matrix form as follows:

$$\mathbf{y}_T = \mathbf{A}_T\mathbf{c} + \mathbf{v} \quad (2)$$

where $\mathbf{A}_T$ is a Toeplitz matrix that depends on the transmitted symbols; and $\mathbf{c}$, $\mathbf{y}_T$, and $\mathbf{v}$ are the channel impulse response, observed data, and measurement noise, respectively, column-stacked as vectors. The conventional channel estimation method is based on the pseudo-inverse operation: $\hat{\mathbf{c}}_{LS} = \mathbf{A}_T^\dagger \mathbf{y}_T$, where "$\dagger$" denotes the pseudo-inverse. We will refer to this method as *least-squares*, as is customary, although we should note that when $N_T < N_c$, this is actually a least-squares, min-norm solution.

3. CHANNEL ESTIMATION BY SPARSITY-ENFORCING REGULARIZATION

We propose estimating the channel by minimizing the following cost function:

$$E_T(\mathbf{c}) = \|\mathbf{y}_T - \mathbf{A}_T\mathbf{c}\|_2^2 + \lambda\|\mathbf{c}\|_1 \quad (3)$$

where λ is a scalar parameter. The first term of $E_T(\mathbf{c})$ is a data fidelity term, and the second term is a sparsity-enforcing regularization term. It is well-known that minimizing the ℓ_1-norm leads to a preference for sparse structure [9]. By incorporating the prior information that the channel is sparse, we aim to achieve accurate channel estimation with a much smaller number of training symbols than conventional methods. We solve the optimization problem in Eqn. (3) by adapting and applying the numerical algorithm of [10].

4. SEMI-BLIND CHANNEL ESTIMATION WITH CONSTANT MODULUS SYMBOLS

The technique we proposed in the previous section relies entirely on known training symbols. In this section, we propose an extension, where we take advantage of some training symbols as well as a block of transmitted unknown symbols for channel estimation. The technique also produces estimates of the unknown symbols. We focus on the case where the transmitted symbols have constant modulus K, and build that information into our formulation as well. Let us assume that we have two data streams: $\mathbf{y}_T$ referring to received data associated with the training symbols, and $\mathbf{y}_B$ referring to received "blind" data associated with the unknown symbols $\mathbf{s}$. Then we construct the following "semi-blind" cost function:

$$\begin{aligned} E_{SB}(\mathbf{s}, \mathbf{c}) &= \|\mathbf{y}_T - \mathbf{A}_T\mathbf{c}\|_2^2 + \|\mathbf{y}_B - \mathbf{A}_B(\mathbf{s})\mathbf{c}\|_2^2 \\ &\quad +\lambda\|\mathbf{c}\|_1 + \gamma \sum_{i=1}^{N_s} \left[|\mathbf{s}_i|^2 - K^2\right]^2 \end{aligned} \quad (4)$$

where γ is a scalar parameter, N_s denotes the number of unknown symbols, and $\mathbf{s}_i$ denotes the i-th unknown symbol. The matrix $\mathbf{A}_B(\mathbf{s})$ is constructed in a similar fashion to $\mathbf{A}_T$, however it depends on the unknown symbols $\mathbf{s}$ rather than the training symbols, and we make that dependence explicit in our notation. The first two terms of $E_{SB}(\mathbf{s}, \mathbf{c})$ enforce fidelity to training-based and blind observations, the third term is the channel sparsity constraint, and the last term enforces the constant modulus property of the symbols. The major novelty of this formulation as compared to other semi-blind techniques is the channel sparsity constraint. Although we weight the training-based and blind data fidelity terms in $E_{SB}(\mathbf{s}, \mathbf{c})$ equally, one could generalize this cost function slightly to apply different weights, e.g. if the two data streams have different SNRs. We minimize $E_{SB}(\mathbf{s}, \mathbf{c})$ by applying a block coordinate descent algorithm on $\mathbf{s}$ and $\mathbf{c}$. In particular, we first find an initial channel estimate $\hat{\mathbf{c}}^{(0)}$ based only on the training symbols, using the technique described in Section 3. Then we run the following set of alternating minimizations at each iteration k, starting with $k = 0$:

$$\begin{aligned} \hat{\mathbf{s}}^{(k+1)} &= \arg\min_{\mathbf{s}} E_{SB}(\mathbf{s}, \hat{\mathbf{c}}^{(k)}) \\ \hat{\mathbf{c}}^{(k+1)} &= \arg\min_{\mathbf{c}} E_{SB}(\hat{\mathbf{s}}^{(k+1)}, \mathbf{c}) \end{aligned} \quad (5)$$

To solve the first minimization above, we use gradient descent; and to solve the second minimization, we use the same numerical algorithm utilized in Section 3. We run the coordinate descent iterations in (5) until $\|\hat{\mathbf{c}}^{(k+1)} - \hat{\mathbf{c}}^{(k)}\|_2 / \|\hat{\mathbf{c}}^{(k)}\|_2 < \delta$, where $\delta > 0$ is a small constant.

5. EXPERIMENTAL RESULTS

We present simulations where the channel to be estimated and equalized is given in Fig. 1. This impulse response has length $N_c = 119$, and 10 non-zero taps. For the transmitted symbols, we use BPSK sequences with equiprobable symbols.

5.1. Training-based Channel Estimation

We first present the results of our training-based sparse channel estimation technique of Section 3. We start with a remark about the observation model. Note that in Section 2, we assumed we had the training symbols $s_T(n)$ for $n = 0, ..., N_T - 1$. However, the observation model in Eqn. (1) also depends on $s_T(n)$ for $n < 0$. As a result, the matrix $\mathbf{A}_T$ in Eqn. (3) also depends on those symbols. These symbols may be obtained from previous decodings in a data stream or can be assumed zero if this is the first packet received [3]. Which assumption is made can have an impact on the results, especially when short data sequences are of interest. Therefore we present results based on both assumptions.

We compare our ℓ_1-norm-based approach with matching pursuit [3], as well as conventional least squares. We run these techniques on 100 different realizations of the training symbols and measurement noise. We repeat these experiments at a range of SNRs, where SNR is defined as the variance ratio of the signal and noise components in Eqn. (1). Given an estimated channel impulse response produced by each technique, we find the locations of the 10 largest magnitude taps for each technique. We then count how many of these locations match the actual 10 non-zero tap locations of the channel in Fig. 1.

In the first set of experiments, we assume that $s_T(n)$ for $n < 0$ are obtained from previous decodings. Fig. 2 shows the number of matches averaged over 100 realizations for each technique, as a function of SNR. Fig. 2(a),(b) and (c) correspond to different numbers of training symbols N_T. We observe that both our ℓ_1-norm-based method and matching pursuit provide much more accurate channel estimates than least squares. The ℓ_1 method provides slightly better performance than matching pursuit, which is more significant when the number of training symbols is relatively small. We should note that we have not optimized the choice of the free parameter λ in Eqn. (3), and the performance of the ℓ_1 technique might be improved further by better parameter choices. In Fig. 3, we present similar results for the second set of experiments where $s_T(n)$ for $n < 0$ are assumed to be zero.

5.2. Semi-blind Channel Estimation

We now present the results of the semi-blind algorithm proposed in Section 4. For the training portion of the experimental setup, we assume that $s_T(n)$ for $n < 0$ are available from previous decodings. We consider the transmission of two BPSK sequences: a training sequence of N_T symbols and an unknown sequence $\mathbf{s}$ of $N_s = 200$ symbols. Based on the training and the blind data, we minimize $E_{SB}(\mathbf{s}, \mathbf{c})$ to find estimates of both the channel and the unknown symbols. We consider a range of SNRs, as well as training sequence lengths N_T, and find the average number of matches in the channel impulse response over 100 realizations, as in the experiments of Section 5.1. We present these results in Fig. 4. The dashed curves correspond to the ℓ_1 results presented in Section 5.1, based only on training data. The solid curves are based on the semi-blind experiments. We observe that we are able to exploit the unknown symbols in the semi-blind framework to improve upon the results of our training-based method. As a result, our semi-blind

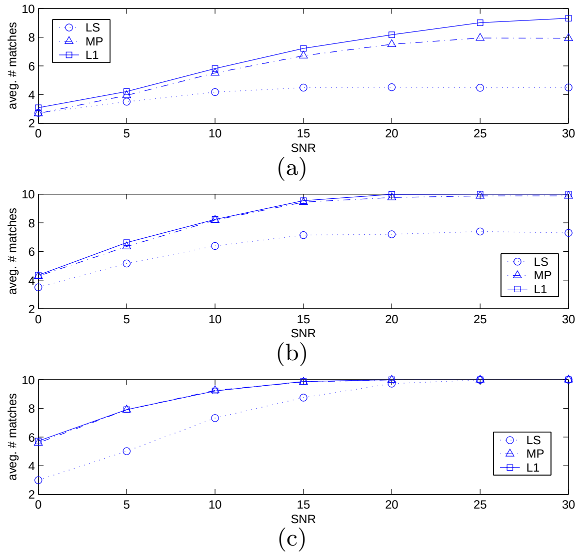

Fig. 2. Average number of matches between the locations of the 10 largest magnitude taps estimated by the three methods (ℓ_1, matching pursuit (MP), least squares (LS)) and the actual non-zero taps of the channel. It is assumed that $s_T(n)$ for $n < 0$ are obtained from previous decodings. Each plot is based on 100 trials. (a) $N_T = 40$. (b) $N_T = 80$. (c) $N_T = 130$.

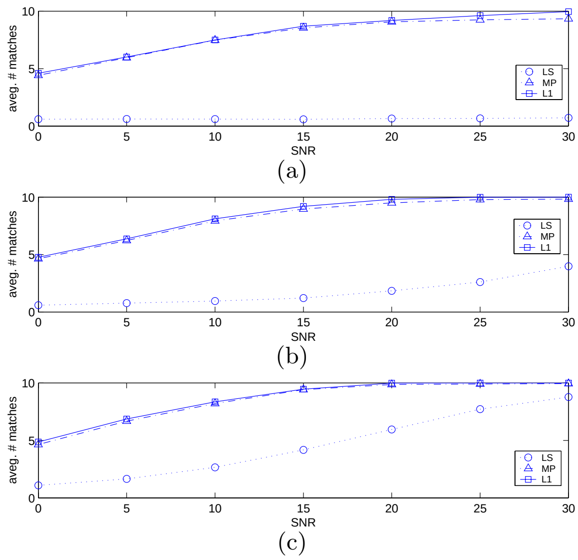

Fig. 3. Average number of matches as in Fig. 2, but for the case where $s_T(n) = 0$ for $n < 0$. (a) $N_T = 120$. (b) $N_T = 130$. (c) $N_T = 140$.

framework can achieve the accuracy of our training-based channel estimates with a much smaller number of training symbols, resulting in communication bandwidth savings.

Finally, we demonstrate the performance of our semi-blind technique in estimating the unknown symbol sequences $\mathbf{s}$ of length $N_s = 200$. After quantizing the continuous-valued symbol estimates produced by the minimization of

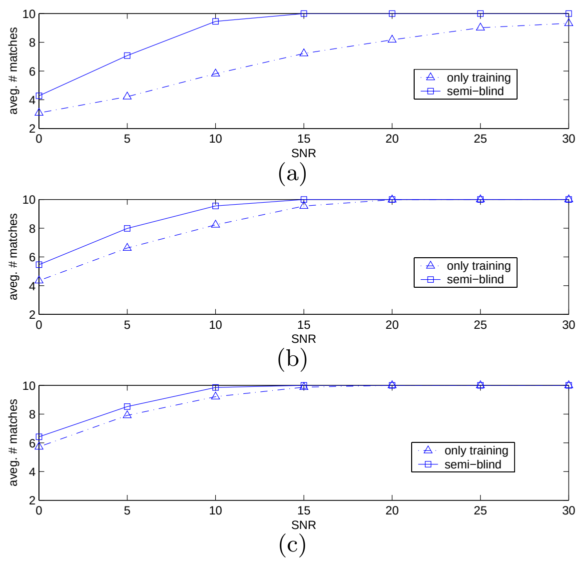

Fig. 4. Average number of matches obtained by minimizing the semi-blind cost function $E_{SB}(\mathbf{s}, \mathbf{c})$ compared with the results of the training-based cost function $E_T(\mathbf{c})$. Each plot is based on 100 trials. (a) $N_T = 40$. (b) $N_T = 80$. (c) $N_T = 130$.

$E_{SB}(\mathbf{s}, \mathbf{c})$, we compare them to the true symbols, and count the number of bit errors. Dividing the number of bit errors by N_s, and averaging over 100 realizations, we obtain an estimate of the bit error rate (BER). Fig. 5 shows plots of BER for three different sizes of training sequences, and a range of SNRs. This result demonstrates that our semi-blind approach can provide reasonable estimates of the unknown symbols. We have not optimized the choice of the free parameters λ and γ in Eqn. (4), and the performance of our technique might be improved further by better parameter choices.

6. CONCLUSION

We have developed new techniques for the estimation of sparse communication channels for training-based and for semi-blind scenarios. We have defined appropriate optimization problems taking advantage of both training-based and blind data streams, and incorporating prior information we have about the structure of the propagation channels and about the unknown symbols. This work provides a principled framework synthesizing recent ideas on semi-blind equalization with ideas on sparse channel estimation. Our preliminary experiments have demonstrated the promise of this framework in generating accurate channel estimates, as well as symbol estimates with short data streams. However, a much more detailed analysis of the proposed framework is needed, and is a subject of our current research. We are interested in comparing our approach to other semi-blind techniques, illuminating the role of sparsity-enforcing regularization. We are also interested in characterizing potential performance improvements provided by this framework over the fully-blind case. Finally, we are interested in exploring further performance improvements of our approach by utilizing effective automatic techniques for choosing the free parameters involved.

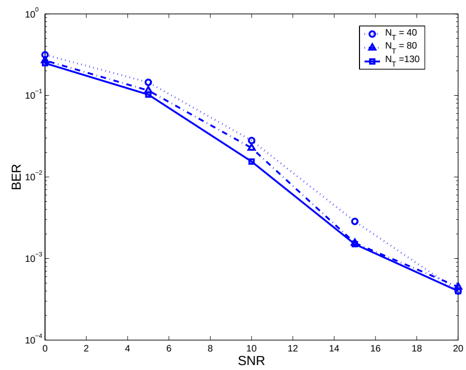

Fig. 5. Performance of the semi-blind technique in estimating the unknown symbols, in terms of BER.

7. REFERENCES

[1] A. Rontogiannis and K. Berberidis, "Efficient decision feedback equalization for sparse wireless channels," *IEEE Trans. Wireless Comunications*, vol. 2, no. 3, pp. 570–581, May 2003.

[2] M. Ghosh, "Blind decision feedback equalization for terrestrial television receivers," *Proc. IEEE*, vol. 86, pp. 2070–2081, Oct. 1998.

[3] S. F. Cotter and B. D. Rao, "Sparse channel estimation via matching pursuit with application to equalization," *IEEE Trans. Communications*, vol. 50, no. 3, pp. 374–377, Mar. 2002.

[4] S. Kim and R. A. Iltis, "A matching pursuit/GSIC-based algorithm for DS-CDMA sparse-channel estimation," *IEEE Signal Processing Letters*, vol. 11, no. 1, pp. 12–15, Jan. 2004.

[5] C. R. Johnson, Jr., P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown, and R. A. Casas, "Blind equalization using the constant modulus criterion: a review," *Proc. IEEE*, vol. 86, no. 10, pp. 1927–1950, 1998.

[6] P. A. Regalia, "A finite-interval constant modulus algorithm," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 2002, vol. 3, pp. 2285–2288.

[7] A. Kuzminskiy, L. Féty, P. Forster, and S. Mayrargue, "Regularized semi-blind estimation of spatio-temporal filter coefficients for mobile radio communications," in *Proc. GRETSI*, 1997, pp. 127–130.

[8] V. Buchoux, O. Cappé, Éric Moulines, and A. Gorokhov, "On the performance of semi-blind subspace-based channel estimation," *IEEE Trans. Signal Processing*, vol. 48, no. 6, pp. 1750–1759, June 2000.

[9] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ^1 minimization," *Proc. Nat. Acad. Sci.*, vol. 100, no. 5, pp. 2197–2202, Mar. 2003.

[10] C. R. Vogel and M. E. Oman, "Fast, robust total variation-based reconstruction of noisy, blurred images," *IEEE Trans. Image Processing*, vol. 7, no. 6, pp. 813–824, June 1998.