HAL
open science

Adaptive and Non-adaptive ISI Sparse Channel Estimation Based on SL0 and Its Application in ML Sequence-by-Sequence Equalization

Rad Niazadeh, Sina Hamidi Ghalehjegh, Massoud Babaie-Zadeh, Christian Jutten

To cite this version:

Rad Niazadeh, Sina Hamidi Ghalehjegh, Massoud Babaie-Zadeh, Christian Jutten. Adaptive and Non-adaptive ISI Sparse Channel Estimation Based on SL0 and Its Application in ML Sequence-by-Sequence Equalization. LVA/ICA 2010 - 9th International Conference on Latent Variable Analysis and Signal Separation, Sep 2010, Saint Malo, France. pp.579-587. hal-00526065

HAL Id: hal-00526065

https://hal.science/hal-00526065

Submitted on 13 Oct 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Adaptive and Non-Adaptive ISI Sparse Channel Estimation Based on SL0 and Its Application in ML Sequence-by-Sequence Equalization

Rad Niazadeh[1], Sina Hamidi Ghalehjegh[1], Massoud Babaie-Zadeh[1] ⋆, and Christian Jutten[2]

[1] Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran
[2] GIPSA-Lab, Grenoble, and Institut Universitaire de France, France
radniazadeh@gmail.com, shghamidi@gmail.com, mbzadeh@yahoo.com, Christian.Jutten@inpg.fr

Abstract. In this paper, we firstly propose an adaptive method based on the idea of Least Mean Square (LMS) algorithm and the concept of smoothed l_0 (SL0) norm presented in [1] for estimation of sparse Inter Symbol Interface (ISI) channels which will appear in wireless and acoustic underwater transmissions. Afterwards, a new non-adaptive fast channel estimation method based on SL0 sparse signal representation is proposed. ISI channel estimation will have a direct effect on the performance of the ISI equalizer at the receiver. So, in this paper we investigate this effect in the case of optimal Maximum Likelihood Sequence-by-sequence Equalizer (MLSE) [2]. In order to implement this equalizer, we propose a new method called pre-filtered Parallel Viterbi Algorithm (or pre-filtered PVA) for *general ISI sparse channels* which has much less complexity than ordinary Viterbi Algorithm (VA) and also with no considerable loss of optimality, which we have examined by doing some experiments. Indeed, Simulation results clearly show that the proposed concatenated estimation-equalization methods have much better performance than the usual equalization methods such as Linear Mean Square Equalization (LMSE) for ISI sparse channels, while preserving simplicity at the receiver with the use of PVA.

1 Introduction

An sparse channel is a channel whose impulse response has only a few significant coefficients. More precisely, in the framework of digital communication there are special scenarios in which one can model the overall channel as a Finite Impulse Response (FIR) sparse filter which will produce interference with previous samples. Such channels may be encountered, for example, in wireless multipath fading channels, acoustic underwater channels, etc. In such scenarios, the operation of estimating the channel and equalizing the produced ISI is an important task of the receiver. Considering the problem of estimating the channel, many efforts have been done to design batch estimation algorithms [3–5]. Most of these algorithms try to exploit the sparsity of the channel or detecting the locations

⋆ This work has been partially funded by Iran NSF (INSF) and also by ISMO and French embassy in Tehran in the framework of a GundiShapour collaboration program.

of non-zero taps of the Channel Impulse Response(CIR). In addition to these, adaptive algorithms are also presented in the literature which are based on iterative estimating of the sparse filter [6]. At the other side, the problem of equalizing the output of sparse channel is in the point of interest. Specially optimal Viterbi equalization is desired [7].

In this work, we try to investigate these two problems (which are so related to each other) in the case of ISI sparse channels with a novel approach. According to the best knowledge of the authors, all of the works around the problems of estimating the channel and equalization have considered ISI channel as 'FIR causal filter' with no constraint on the taps. The first contribution of our work is to solve these two problems in the case of using a matched filter structure at the receiver. In this situation, the resulting sparse FIR channel can be assumed to be minimum phase [2] (the reason of this assumption is described in Sect. 2). This assumption can make our problem much easier to solve (as will be shown in Sect. 3 of this work). So, by the use of this assumption we develop our algorithms in order to find the solutions to both of the estimation and equalization problems. Followed by these, we experimentally examine the efficiency of the proposed algorithms in a concatenation of estimation and equalization.

In the first section of this paper, motivated by Smoothed l_0-norm (SL0) algorithm [1] which is known as a fast sparse representation technique, we propose two new approaches for sparse channel estimation problem. Firstly, by modifying LMS algorithm introduced by Widrow and Hoff [8], we propose an adaptive algorithm similar to the algorithms proposed in [6] (which are named as Zero-Attracting LMS (ZA-LMS) and Re-weighted Zero-Attracting LMS (RZA-LMS) in [6]) called SL0-LMS. We show that by exploiting the sparsity information and using the concept of Smoothed l_0-norm we can improve the filtering performance. The second proposed algorithm is a non-adaptive algorithm which uses SL0 in a direct manner to estimate channel coefficients by finding the sparsest solution[3] of an under-determined system of linear equations [1]. Although this algorithm is non-adaptive, it uses the speed of SL0 while preserving enough accuracy for the act of equalization. It is important to mention that none of the proposed algorithms in this section depends on the assumption of modelling the channel as a minimum phase filter and can be used in the cases when we do not use the matched filter structure at the receiver.

In the next section, efficient equalization based on Maximum Likelihood Sequence-by-sequence Equalization (MLSE) will be described. In conventional ISI sparse channels, implementation of MLSE using ordinary VA is almost impossible (because of the fact that computational complexity of the receiver grows exponentially with the channel memory [2]). However, in some special cases known as zero-pad ISI channels, Parallel Viterbi Algorithm (PVA) is used instead of ordinary VA [9] which has much less complexity. In the case of *general* ISI sparse channels, at our best knowledge, no such solution is yet presented in the literature. So, we propose a new method for using PVA in *general* ISI sparse channels based on the modelling of the ISI channel with a minimum-phase FIR filter. Our idea is based on the usage of a pre-filter at the receiver. In fact, this filter will re-shape the channel structure to a zero-pad channel which is presented in [9] and so, applying the PVA method will be possible afterwards. Although

[3] Solution with minimum l_0-norm, i.e. minimum number of non-zero coefficients.

this method is not optimal, as we will see in experimental results, we do *not* have considerable amount of loss in optimality in most practical cases.

In the last section, simulation results will verify the efficiency of the presented algorithms in both fields of estimation of the channel and equalization of the produced ISI. It is important to note that the overall performance of the receiver system (as a concatenation of channel estimation and ISI equalization) depends on both of these parts which will be evaluated in these simulations and it is considered as the overall measure of efficiency of the proposed algorithms for both estimation and equalization tasks.

2 ISI Sparse Channel Estimation

In this section, we would like to estimate the ISI sparse channel coefficients while transmitting digital bits through this channel using Binary Phase Shift Keying (BPSK) modulation. Including the BPSK modulator and de-modulator with the channel and assuming that we are using matched filter structure at the receiver, the total ISI channel can be modelled as an FIR minimum-phase filter $\mathbf{f} = [f_0, \cdots f_{M-1}]^T$, which has only a few non-zero coefficients, followed by an additive white Gaussian noise ($v(n)$) which is independent of the input [2]. To show that this assumption is reasonable, we note that by sampling at the output of the matched filter with a suitable period, it is convenient to model the overall channel as a symmetric non-causal FIR filter with taps equal to $\{x(n) : -M \leq n \leq M\}$ and an additive coloured noise $z(n)$ with spectral density of $X(z) = \mathscr{Z}\{x(n)\}$ in which $\mathscr{Z}\{.\}$ denotes the Z-transform operator [2]. This model of describing an ISI channel is called the "X" model in the area of communication [2]. Additionally, we force the sparsity constraint on the taps of $x(n)$ which is a valid constraint in the scenarios that were mentioned in the previous section. Although this model is very common in the literature, we did not use this model in our work. Instead, we use a simpler model called "F" model [2] which can be made by decomposing the Z-transform of channel impulse response into a minimum phase filter and its conjugate reciprocal ($X(z) = F(z)F^*(\frac{1}{z^*})$)[4]. So by concatenating a whitening filter $G(z) = \frac{1}{F^*(\frac{1}{z^*})}$ to the end of sampler, the equivalent discrete channel would be a minimum phase FIR filter with taps equal to $\{f(n) : 0 \leq n \leq M\}$ and $F(z) = \mathscr{Z}\{f(n)\} = G(z)X(z)$.

According to the relation between $x(n)$ and $f(n)$ in the frequency domain, we have $x(n) = f(n) * f^*(-n)$ in which $*$ denotes the convolution operation. So according to this relation, when $x(n)$ is very sparse it is very unlikely that $f(n)$ would not be sparse. Although we have currently no mathematical proof for this, it can heuristically seen from the convolution sum in which it is very unlikely that the convolution of two non sparse signal produces an sparse signal. It can also be experimentally verified. According to all of these, we use the above mentioned minimum phase sparse FIR model for the ISI channel which will make our problem much easier to solve.

Now, let $d(n)$ be the last observed sample of the noisy output signal of the channel and let $\mathbf{u}(n)$ be a vector that contains the last M samples of the input signal of the ISI sparse channel, that is:

$$\mathbf{u}(n) = [u(n), u(n-1), \cdots, u(n-M+1)]^T \quad . \tag{1}$$

[4] The possibility of this decomposition is due to the symmetry of $x(n)$.

Consequently, we have the following input-output relation for the channel:

$$d(n) = \mathbf{f}^T \mathbf{u}(n) + v(n) \ . \tag{2}$$

In order to estimate the channel, we use a semi-random training sequence (which is known to the receiver) that is generated by producing a random block of 0, 1 bits with length equal to M and transmitting it periodically (note that the BPSK modulator will map these bits to ± 1 symbols).

2.1 Adaptive Sparse Channel Estimation

Review of ZA-LMS and RZA-LMS. In the state of transmitting the training sequence, the standard LMS, which is based on iteratively minimizing the cost function $J(\mathbf{w}(n)) = E\{e^2(n)\}$, adaptively estimates $\mathbf{f}$ by the following recursion:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu e(n)\mathbf{u}(n) \ . \tag{3}$$

where $\mathbf{w}(n)$ is the estimated adaptive filter at the nth iteration, μ is the step size parameter and $e(n) = d(n) - \mathbf{w}(n)^T\mathbf{u}(n)$ [8]. Then, in ZA-LMS, the cost function is modified by adding a penalty term based on l_1 norm to enforce some sparsity on $\mathbf{w}(n)$ which is an estimation of the sparse vector $\mathbf{f}$:

$$J_1(\mathbf{w}(n)) = \frac{1}{2}E\{e^2(n)\} + \gamma\|\mathbf{w}(n)\|_1 \ . \tag{4}$$

Using steepest descent, the channel coefficients update equation will then be [6]:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu e(n)\mathbf{u}(n) - \mu\gamma \ \mathrm{sgn}(\mathbf{w}(n)) \ . \tag{5}$$

In RZA-LMS the new cost function is defined as [6]:

$$J_2(\mathbf{w}(n)) = \frac{1}{2}E\{e^2(n)\} + \gamma\sum_{i=1}^{M}\log(1+\varepsilon|w_i|) \ . \tag{6}$$

The log-sum term has been used because it behaves more similarly to the l_0 norm than $\|\mathbf{w}(n)\|_1$. So, the update equation will be

$$w_i(n+1) = w_i(n) + \mu e(n)u_i(n) - \mu\gamma\varepsilon\frac{\mathrm{sgn}(w_i(n))}{1+\varepsilon|w_i(n)|} \ . \tag{7}$$

The New Smoothed l_0-LMS Algorithm (SL0-LMS). Inspired from the idea of the SL0 algorithm [1], we propose replacing the above mentioned cost function by:

$$J_3(\mathbf{w}(n)) = \frac{1}{2}E\{e^2(n)\} + \gamma\|\mathbf{w}(n)\|_0 \ . \tag{8}$$

in which, $\|\mathbf{w}\|_0$ is replaced by its smooth approximation as [1] in order to exploit the sparsity nature of the estimated channel, i.e. we use this approximation:

$$\|\mathbf{w}\|_0 \approx M - \sum_{i=1}^{M} e^{-w_i^2/2\sigma^2} \ . \tag{9}$$

So, the update equation will be

$$w_i(n+1) = w_i(n) + \mu e(n) u_i(n) - \frac{\rho\gamma}{\sigma^2} w_i(n) e^{-w_i^2(n)/2\sigma^2} \quad . \tag{10}$$

As shown in [1], (9) tends to equality when $\sigma \to 0$. Consequently, we expect the SL0-LMS have a better performance than RZA-LMS and the experimental results prove this. Unfortunately, this algorithm requires adjusting the parameter σ, which involves nested loops and thus increases the computational complexity.

2.2 Non-Adaptive SL0 Based Sparse Channel Estimation

In this section, we propose a non-adaptive channel estimation algorithm which will estimate the channel coefficients *after* observing $m \ll M$ successive samples of the output signal of the channel. The relation of theses observation to the input sequence can be described in the matrix form as follows :

$$\mathbf{d}_{m\times 1} = \mathbf{A}_{m\times M}.\mathbf{f}_{M\times 1} + \mathbf{v} \quad . \tag{11}$$

in which, $\mathbf{v}$ is a vector including samples of the Gaussian noise, $\mathbf{d}$ is the vector of observations, $\mathbf{f}$ is the vector of channel taps and $\mathbf{A}$ is a Toeplitz random matrix known to the receiver whose rows are circular shift of $\mathbf{q} = [q_0, \cdots, q_{M-1}]^T$ and $\mathbf{q}$ is a M-length random block of ± 1s (see Sect. 2). Consequently, if we start observing the channel output after l transmissions, then $\mathbf{A}$ can be expressed as the following matrix:

$$\mathbf{A} = \begin{bmatrix} q_l & q_{l-1} & q_{l-2} & \cdots & q_0 & q_{M-1} & \cdots & q_{l+1} \\ q_{l+1} & q_l & q_{l-1} & \cdots & q_1 & q_0 & \cdots & q_{l+2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ q_{l-1} & q_{l-2} & q_{l-3} & \cdots & q_{M-1} & q_{M-2} & \cdots & q_l \end{bmatrix} \quad . \tag{12}$$

Now, (11) is a noisy under-determined system of linear equations which has to be solved under the sparsity constraints. Now, we want to find the sparsest solution of (11) while having a constraint on square error, i.e. we are trying to solve the following optimization problem:

$$\text{argmin } \|\tilde{f}\|_0 \quad s.t \quad \|\mathbf{A}\tilde{\mathbf{f}} - \mathbf{d}\|_2^2 \leq \epsilon \quad . \tag{13}$$

in which we assume that in (11), $\|\mathbf{v}\|_2^2 \leq \epsilon$. There are many methods for finding the sparsest solution of a noisy under-determined system of linear equations such as Basis Pursuit De Noising (BPDN) [10], Least-Absolute Shrinkage and Selection Operator (LASSO) [11] and Robust SL0 [1, 12]. By the use of simulations, we have seen that there is not much difference in the accuracy of Robust SL0, LASSO and BPDN in our application and so, we use Robust-SL0 according to its speed in comparison to the other two algorithms. So, in our experimental results we just mentioned the results of Robust-SL0 algorithm.

3 Pre-Filtered PVA Equalization in General ISI Sparse Channels

After estimating channel coefficients using one of the methods of the previous section, in this part, efficient equalization in *general* ISI sparse channels will be

investigated (in which the channel has a large amount of memory while having a few number of significant coefficients). In these channels, implementation of the optimum ISI equalizer (MLSE) by the use of Viterbi algorithm is almost impossible (because computational complexity of the receiver grows exponentially with channel memory [2]). PVA which uses parallel trellises can be used in a special case of such channels known as *the zero-pad channel* [9] in which channels have equally spaced coefficients. Assume that $\mathbf{w}$ is a zero-pad channel with the length $M = K.L + 1$, then it has a structure as follows:

$$\mathbf{w} = [\underbrace{w_0\ 0\ 0\ \ldots}_{K \text{ coefficients}}\ \underbrace{w_K\ 0\ 0\ \ldots}_{K \text{ coefficients}}\ \underbrace{w_{2K}\ 0\ 0\ \ldots}_{K \text{ coefficients}} \ldots\ 0\ 0\ w_{L.K}]\ . \tag{14}$$

As in [9], it is evident that received symbols $\{u_0, u_K, u_{2K} \ldots\}$ will have interference between themselves while having no interference on any other symbols and they will generate output symbols $\{d_0, d_K, d_{2K} \ldots\}$. Similarly, symbols $\{u_1, u_{K+1}, u_{2K+1} \ldots\}$ have ISI between themselves and have no interference with any other symbols and so on. So, we can equalize this channel using K parallel trellises each using Viterbi algorithm for a channel with coefficients as:

$$\mathbf{w}' = [w_0\ \ w_K\ \ w_{2K}\ \ \ldots\ \ w_{L.K}]^T\ . \tag{15}$$

and the input of the jth trellis are symbols $\{d_{i.k+j} : \forall i \in \mathrm{N}\}$. So, this PVA structure will have less overall complexity than a single trellis in these channels. In fact, the complexity of a single trellis is of $\mathrm{O}(2^M)$ and the complexity of PVA is of $\mathrm{O}(K.2^L)$ which is much less than the single trellis case (for example, in our experiment $K = L = 64$ and hence, implementation of normal Viterbi algorithm which requires 2^{64} states is impossible. But having 8 trellises each with $2^8 = 64$ states is practical and possible). In the case of general ISI sparse channels, reducing the complexity of Viterbi algorithm is much more sophisticated and according to the authors' best knowledge, no exact solution has been presented in the literature [9]. To find a solution for the case of general ISI sparse channel, firstly we use the matched filter structure at the receiver and so, as was mentioned before, we can use the "F" model of the channel in which the channel impulse response can be assumed to be an FIR minimum-phase sparse filter. In this way, we can change the channel to be similar to the *comb* shape channel described above. Precisely, our idea is based on a pre-filtered equalization. This pre-filter will *re-shape* the channel impulse response to a zero-pad channel which has coefficients similar to the original channel, but they are moved so that the coefficients of the resulting channel will be equally spaced. In order to build such a filter, we use an IIR filter structure which has the estimated channel as its denominator and re-shaped channel as its numerator. In other words we use the following filter:

$$W_{\text{pre}}(z) = \frac{\tilde{F}(z)}{F(z)}\ . \tag{16}$$

in which $F(z)$ and $\tilde{F}(z)$ are the Z-transforms of the channel impulse response and its reshaped version, respectively. According to the fact that the channel impulse response is minimum-phase, this IIR filter will be stable and causal. So, PVA can be implemented afterwards using a few number of trellises which will cause a great amount of computational complexity reduction, but the equalization structure will be the dependent to the CIR which is somehow impractical.

4 Experimental Results

In this section, firstly we compare the performance of channel estimators mentioned in Sect. 2.1 which are SL0-LMS (Eq. 10), ZA-LMS (Eq. 5), RZA-LMS (Eq. 7) and standard LMS. In this experiment, the input signal is an equiprobable random vector of ± 1 with length 5000 and additive noise is white Gaussian random sequence of length 5000 and variance 10^{-3}. The channel has 256 taps where only 28 of them are non zero (which are selected random). The result of various estimation methods with equal step-sizes is shown in Fig. 1. It is

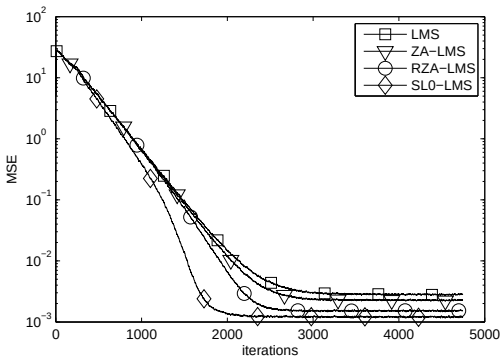
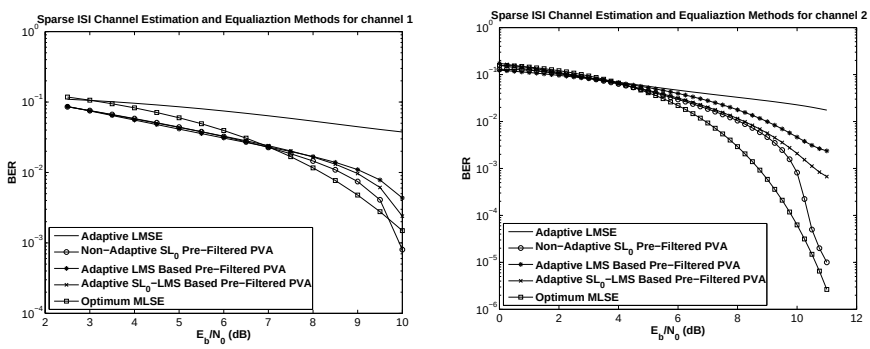


Fig. 1. Tracking and steady-state behaviors of estimation methods.



(a) experiment 1. (b) experiment 2.

Fig. 2. Comparison of the BER-SNR curves for our estimation/equalization methods.

clear that for this long sparse channel, the SL0-LMS is drastically better than the other algorithms, both in steady-state and convergence rate behaviour while needs a pre-adjustment of parameter σ.

After that, concatenation of the proposed channel estimation methods with PVA will be applied in two experiments in order to test the efficiency of the proposed channel estimation algorithms and pre-filtered PVA. In these experiments, our equalization method will be compared to adaptive LMS Equalizer (LMSE) and approximate Bit Error Rate (BER) bounds for MLSE introduced

in [2]. Channel taps for these two experiments are chosen at random, but with the constraint that the channel is minimum-phase according to our model. So, we generate random channels and select two of those which are minimum phase. In the case of LMSE, we have used an adaptive filter that it has the same number of taps comparing to the actual channel. Additionally the one for experiment 1 is so close to the comb shape, while the one for experiment 2 is not. The resulting BER vs Signal to Noise Ratio (SNR) per Bit is shown in Fig. 2(a) and Fig. 2(b). It is important to note that the derivation of the optimum MLSE's BER curves is impossible during the simulation (according to its complexity) and the "Optimum MLSE" curves in these figures are indeed approximate tight bounds for the exact curves, which result in preceding our algorithms from these approximate bounds at some SNRs (specially at low SNR). Advantages of our estimation methods and the proposed pre-filtered PVA equalizer could be seen in these results.

5 Conclusion

In this paper, we showed that with the use of proposed estimation/equalization methods, we benefit the speed of adaptive filtering, the optimality of ML equalizer and the complexity reduction of PVA. According to that, we have no significant loss of performance at the receiver while having much reduction in complexity. In fact, pre-filtering may increase the noise power, but as we have shown experimentally, the performance of our method is not much less than MLSE bound and so is appropriate in the sense of error performance.

References

1. Mohimani, H., Babaie-Zadeh, M., Jutten, C.: A fast approach for overcomplete sparse decomposition based on smoothed l_0-norm. IEEE Trans. Signal Processing **57**(1) (January 2009) 289–301
2. Proakis, J.: Digital Communication, 4th edition. Mc-Graw Hill, New york (2001)
3. C. Carbonelli, S.V., Mitra, U.: Sparse channel estimation with zero tap detection. IEEE Trans. Communication **6**(5) (May 2007) 1743–11754
4. Cotter, S., Rao, B.: Sparse channel estimation via matching pursuit with application to equalization. IEEE Trans. Communication **50**(3) (March 2002) 374–378
5. W. Li, S.V., Preisig, J.: Estimation of rapidly time-varying sparse channels. IEEE Journal of Oceanic Engineering **32**(4) (October 2007) 927–940
6. Chen, Y., Gu, Y., III, A.H.: Sparse LMS for system identification, ICASSP (2009) 3125–3128
7. Mcginty, N.C., Kennedy, R., Hoeher, P.: Parallel trellis viterbi algorithm for sparse channels. IEEE Communication Letters **2**(5) (May 1998) 143–145
8. Widrow, B., Streams, S.: Adaptive Signal Processing. Prentice Hall, New Jersey (1985)
9. Mietzner, J., Hoeher, S., Land, I., Hoeher, P.: Trellis-based equalization for sparse ISI channels revisited, ISIT (2005) 229–233
10. Chen, S., Donoho, D., Saunders, M.: Atomic decomposition by basis pursuit. SIAM review **43**(1) (2001) 129–159
11. Tibshirani, R.: Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. Series B (Methodological) **58**(1) (1996) 267–288
12. Eftekhari, A., Babaie-Zadeh, M., Jutten, C., Moghaddam, H.: Robust-SL0 for stable sparse representation in noisy settings, ICASSP (2009) 3433–3436