HAL
open science

# An Alternating Minimization Method for Sparse Channel Estimation

Rad Niazadeh, Massoud Babaie-Zadeh, Christian Jutten

## To cite this version:

Rad Niazadeh, Massoud Babaie-Zadeh, Christian Jutten. An Alternating Minimization Method for Sparse Channel Estimation. LVA/ICA 2010 - 9th International Conference on Latent Variable Analysis and Signal Separation, Sep 2010, Saint Malo, France. pp.319-327. hal-00525949

**HAL Id: hal-00525949**

**https://hal.science/hal-00525949v1**

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.

# An Alternating Minimization Method for Sparse Channel Estimation

Rad Niazadeh[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
mbzadeh@yahoo.com, Christian.Jutten@inpg.fr

**Abstract.** The problem of estimating a sparse channel, i.e. a channel with a few non-zero taps, appears in many fields of communication including acoustic underwater or wireless transmissions. In this paper, we have developed an algorithm based on Iterative Alternating Minimization technique which iteratively detects the location and the value of the channel taps. In fact, at each iteration we use an approximate Maximum A posteriori Probability (MAP) scheme for detection of the taps, while a least square method is used for estimating the values of the taps at each iteration. For approximate MAP detection, we have proposed three different methods leading to three variants for our algorithm. Finally, we experimentally compared the new algorithms to the Cramér-Rao lower bound of the estimation based on knowing the locations of the taps. We experimentally show that by selecting appropriate preliminaries for our algorithm, one of its variants almost reaches the Cramér-Rao bound for high SNR, while the others always achieve good performance.

## 1 Introduction

In this paper, we are going to investigate the problem of sparse channel estimation which appears in acoustic underwater or wireless transmissions [1]. In this problem, we want to estimate channel taps by observing the output of the channel while the channel is sparse, i.e. it has a few non-zero taps. Mathematically, we have the following model:

$$y(k) = u(k) * h(k) + n(k) \ . \tag{1}$$

in which $y(k)$ is the output signal of the channel, $h(k)$ is the $K$-sparse[1] channel with $N$ taps, $u(k)$ is $L$-length training sequence at the input of the channel and

⋆ 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.

[1] By $K$-sparse we mean that there are at most $K$ non-zero elements in $\mathbf{h} = [h(0), h(1), \ldots, h(N)]^T$.

$n(k)$ is a Gaussian noise. By observing all of the $M = N + L - 1$ output symbols of the channel, we have the following model in matrix form:

$$\mathbf{y} = \mathbf{U}\mathbf{h} + \mathbf{n} = \mathbf{U}_{\mathbf{h}}\mathbf{b} + \mathbf{n} \ . \tag{2}$$

in which $\mathbf{y}$ is the $M \times 1$ vector of output symbols of the channel, $\mathbf{n} \sim N(\sigma^2 \mathbf{I}_M, 0)$ is an $M \times 1$ Gaussian noise vector and $\mathbf{U}$ is the $M \times N$ full-rank training matrix (as in [2]). Additionally, $\mathbf{U}_{\mathbf{h}} \triangleq \mathbf{U}\text{diag}(\mathbf{h})$ and $\mathbf{b} \in \{0,1\}^N$ is a binary vector that indicates the locations of non-zero taps i.e. :

$$\forall i \in \{1, 2, \ldots N\} : \quad b_i = \begin{cases} 1 \text{ if } h_i \neq 0, \\ 0 \text{ if } h_i = 0. \end{cases} \tag{3}$$

The goal of our problem is to find an appropriate estimation for $\mathbf{h}$ based on the observation vector, $\mathbf{y}$. If the estimator has no prior information about $\mathbf{h}$ i.e. it is unstructured, then it is well known [2] that the Least Square (LS) estimator will attain the best estimate for $\mathbf{h}$ in the sense of Mean Square Error (MSE). This estimator finds the solution of the following problem:

$$\hat{\mathbf{h}} = \underset{\mathbf{h}}{\text{argmin}} \|\mathbf{y} - \mathbf{U}\mathbf{h}\|_2^2 \ . \tag{4}$$

It is possible to show [2] that the above LS estimator achieves the Cramér-Rao lower bound of the problem. For the case when the estimators have prior knowledge about the location or number of taps of $\mathbf{h}$, it is possible that we obtain better estimations compared to the least square solution. Indeed, if a genie aids us with the location of the non-zero taps of $\mathbf{h}$, then the Structural Least Square (SLS) estimator, which will find the solution of the following problem:

$$\hat{\mathbf{h}}_\tau = \underset{\mathbf{h}_\tau}{\text{argmin}} \|\mathbf{y} - \mathbf{U}_\tau \mathbf{h}_\tau\|_2^2 \ . \tag{5}$$

will be an efficient estimator [1]. In (5), $\tau \subset \{1, 2, \ldots N\}$ is the support of $\mathbf{h}$, $\mathbf{U}_\tau$ is a sub-matrix of $\mathbf{U}$ that includes columns corresponding to the indices in $\tau$ and $\mathbf{h}_\tau$ is a $K \times 1$ vector that contains the non-zero taps of $\mathbf{h}$. Here, by 'efficient' we mean that this estimator will reach CRB of the estimation problem in which, the estimator will estimate $\mathbf{h}$ based on $\mathbf{y}$ and $\tau$. In literature, this bound is known as CRB-S[2].

By comparing the Cramér-Rao bound of the unstructured estimation problem, known as CRB-US, with CRB-S, it is shown that CRB-S is much less than CRB-US [2] and their difference will increase as the $\mathbf{h}$ becomes sparser. Consequently, it is conceivable that we design an estimation technique based on $\mathbf{y}$ and some prior information about the sparsity degree of $\mathbf{h}$ so that its MSE will be close to CRB-S. Unfortunately, structured LS estimator which achieves CRB-S is not practical, because its estimation is based on knowing the exact location of the taps. The design of a practical estimator with MSE close to that of SLS estimator is of interest. So, many efforts have been done to find a practical solution for this estimation problem based on minimum knowledge about the sparsity of the original vector. Candès *et al.* [3] and Haupt *et al.* [4] proposed estimators that can achieve CRB-S to a factor of $\log M$. Similarly, Babadi *et al.* [5] showed that with the use of an estimator known as 'Typical' Estimator we can asymptotically achieve the Cramér-Rao bound of the genie aided estimation. Additionally, the

work done by Carbonelli *et al.* [2] proposed practical algorithms for this problem which can reach close to CRB-S in the MSE sense.

In this paper, we try to find practical solutions for the problem of estimating the channel taps from noisy outputs of the channel. Our proposed algorithm is based on Alternating Minimization [6] technique for joint estimation of $\mathbf{b}$ and $\mathbf{h}$ at each iteration step. More precisely, our algorithm detects non-zero taps locations i.e. $\mathbf{b}$, iteratively based on a MAP scheme. Simultaneously, we will use the detected locations for finding the original sparse vector using an structured least square estimation at each iteration. For approximate MAP detection, we propose three methods. We will compare the MSE curve of all of the variations of the proposed algorithm to that of CRB-S, CRB-US and also with MSE curve of ITD-SE algorithm introduced in [2] and we will discuss our results.

So, this paper is organized as follows. In the next section, we will investigate our proposed MAP estimation of the taps based on the Bernoulli model for the channel taps. We will introduce an alternating minimization technique that we used at each iteration for joint estimation of the location and the value of the taps. Then, we will develop our algorithms to find an approximate MAP solution for $\mathbf{b}$ in each iteration. We will discuss the theory behind each one and the steps of the proposed variants of the main algorithm. Finally, we have an experimental results section in which we compare our algorithm with CRB-S, CRB-US and ITD-SE and we will discuss the results. Note that experimental results are computed on simulated data.

## 2 Iterative Approximate MAP-SLS Algorithms

### 2.1 MAP Detection and Iterative Alternating Minimization

In this section we introduce our main strategy for jointly estimating of the channel taps and the location of the taps i.e. $\mathbf{h}$ and $\mathbf{b}$ in the model described in (2). We use an iterative procedure based on alternating minimization for jointly finding both of the $\mathbf{h}$ and $\mathbf{b}$. To develop our algorithm, first assume that we have an appropriate estimate for $\mathbf{b}$ i.e. $\hat{\mathbf{b}}$ at a given iteration step. So, we estimate $\mathbf{h}$ at this iteration by finding the solution of the following problem:

$$\hat{\mathbf{h}} = \underset{\mathbf{h}}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{U}\operatorname{diag}(\hat{\mathbf{b}})\mathbf{h}\|_2 = \underset{\mathbf{h}}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{U}_{\hat{\mathbf{b}}}\mathbf{h}\|_2 = ({\mathbf{U}_{\hat{\mathbf{b}}}}^T\mathbf{U}_{\hat{\mathbf{b}}})^\dagger {\mathbf{U}_{\hat{\mathbf{b}}}}^T\mathbf{y} \ . \quad (6)$$

in which $(.)^\dagger$ denotes the pseudo-inverse operator. Note that this estimate is equal to the structured least square solution based on the location vector $\mathbf{b}$. On the other hand, if we have an appropriate estimation for $\mathbf{h}$ at a given iteration step and we want to obtain a more accurate estimate for $\mathbf{b}$, then we can make a MAP estimation for $\mathbf{b}$ based on the following observation:

$$\mathbf{y} \approx \mathbf{U}_{\hat{\mathbf{h}}}\mathbf{b} + \mathbf{n} \ . \quad (7)$$

for the MAP estimation, we need prior knowledge on the probability distribution of $\mathbf{b}$. In this paper, we assume an i.i.d Bernoulli distribution for the channel location vector, $\mathbf{b}$, based on the sparsity degree of the channel. Mathematically, if we define $P_a = \frac{K}{N} < \frac{1}{2}$ then we assume that $\mathbb{P}\{b_i = 1\} = P_a$ and so:

$$\mathbb{P}\{\mathbf{b}\} = \prod_{i=1}^{N} \mathbb{P}\{b_i\} = (1 - P_a)^{(N - \|\mathbf{b}\|_0)} P_a^{\|\mathbf{b}\|_0} \ . \quad (8)$$

the probability of vector $\mathbf{b}$ is computed for a given vector with exactly $K$ components and not any vector with exactly $K$ components. Now, we can find the MAP solution as:

$$\begin{aligned}\hat{\mathbf{b}}_{\text{MAP}} &= \underset{\mathbf{b}\in\{0,1\}^N}{\text{argmax}}\ \mathbb{P}\{\mathbf{y}|\hat{\mathbf{h}},\mathbf{b}\}\mathbb{P}\{\mathbf{b}\} = \underset{\mathbf{b}\in\{0,1\}^N}{\text{argmax}}\ \exp\left(-\frac{1}{2\sigma^2}\|\mathbf{y}-\mathbf{U}_{\hat{\mathbf{h}}}\mathbf{b}\|_2^2\right)\left(\frac{P_a}{1-P_a}\right)^{\|\mathbf{b}\|_0} \\ &= \underset{\mathbf{b}\in\{0,1\}^N}{\text{argmin}}\ \|\mathbf{y}-\mathbf{U}_{\hat{\mathbf{h}}}\mathbf{b}\|_2^2 + \lambda\|\mathbf{b}\|_0\ . \end{aligned} \tag{9}$$

in which $\lambda = 2\sigma^2\ln(\frac{1-P_a}{P_a}) > 0$, according to the fact that $P_a < \frac{1}{2}$. Note that in (9) we minimize a function which is a combination of sparsity degree of $\mathbf{b}$ and $l_2$-norm of error $\mathbf{y}-\mathbf{U}_{\hat{\mathbf{h}}}\mathbf{b}$. Indeed, $\lambda$ is an increasing function of $\sigma$ which intuitively seems correct: as the noise increases, the importance of sparsity becomes greater in the estimation according to the fact that $\mathbf{y}$ is more noisy.

It is important to mention that the problem in (9), is exponentially complex with respect to $N$ according to the search over $\{0,1\}^N$ for finding $\hat{\mathbf{b}}_{\text{MAP}}$. Although this problem seems to be impossible to solve, one may use an approximation method to find the MAP solution, as we will do in this paper.

To find the solutions of (9) and (6) simultaneously, we use an iterative alternating minimization approach. In fact, at each iteration step, by the use of the approximation for $\mathbf{b}$ in the previous iteration, we will find the solution to (6), then by using this solution we will find an approximate solution for (9). We have proposed 3 methods for finding an approximate solution to (9) i.e. MAP detection of the location vector at each iteration step of the proposed alternating minimization method. These methods are based on the algorithms introduced in [7] for finding the the user activity vector in CDMA[2], although they have been changed for being applied to our problem.

Finally, for the initialization of the algorithm, we use $\hat{\mathbf{b}}_0 = [1,\ldots,1]^T$. This algorithm, which we call Approximate MAP-SLS, is summarized in Algorithm 1 with all its variants. Each of its variants, uses one of the the proposed methods for approximate MAP detection, which are described in the following sections.

### 2.2 Approximate $l_2$-MAP With Threshold

Based on the idea of Ridge Detector introduced in [7], we replace $\|\mathbf{b}\|_0$ by $\|\mathbf{b}\|_2^2$ and then, we will find the solution of this new problem in the real field. In fact, we find the solution of the following problem:

$$\hat{\mathbf{b}}_{\mathbf{r}} = \underset{\mathbf{b}\in\mathbb{R}^N}{\text{argmin}}\ \|\mathbf{y}-\mathbf{U}_{\hat{\mathbf{h}}}\mathbf{b}\|_2^2 + \lambda\|\mathbf{b}\|_2^2\ . \tag{10}$$

By taking the gradient of the target function for optimization with respect to $\mathbf{b}$, we can find a closed form solution for (10). So, we have:

$$\hat{\mathbf{b}}_{\mathbf{r}} = ({\mathbf{U}_{\hat{\mathbf{h}}}}^T\mathbf{U}_{\hat{\mathbf{h}}} + \lambda I)^{-1}{\mathbf{U}_{\hat{\mathbf{h}}}}^T\mathbf{y}\ .^3 \tag{11}$$

After that, we will obtain $\hat{\mathbf{b}}$ by quantizing the solution of (11). More accurately:

$$\forall i \in \{1,2,\ldots N\}: \quad \hat{b}_i = \begin{cases} 1 & \text{if} \quad \hat{b}_r(i) \geq \gamma(\sigma), \\ 0 & \text{if} \quad \hat{b}_r(i) < \gamma(\sigma). \end{cases} \tag{12}$$

[2] Code Division Multiple Access.

[3] Note that ${\mathbf{U}_{\hat{\mathbf{h}}}}^T\mathbf{U}_{\hat{\mathbf{h}}} + \lambda I$ is non-singular, because all of its eigenvalues are greater or equal to $\lambda > 0$ and so, we can use inverse operation in (11).

in which, $\gamma(\sigma)$ is a threshold. By the use of simulation, we see that by choosing a suitable function for $\gamma(\sigma)$, we can find an approximately accurate estimate for $\mathbf{b}$, although there is no mathematical convenient way to show the accuracy of our proposed algorithm. This algorithm is so simple, although it needs a suitable function for $\gamma(\sigma)$. It is almost obvious from (11) that $\gamma(\sigma)$ needs to be increasing with respect to $\sigma$. So, in our algorithm we choose a simple increasing function for $\gamma(\sigma)$ i.e. $\gamma(\sigma) = \alpha\sigma^2$. Note that the parameter $\sigma$ is assumed to be known to the detector.

Based on this method, we propose a variant for the main algorithm discussed in Sect. 2.1. The results of this variant of the main algorithm is examined for two values of $\alpha$ and it is experimentally seen that by choosing an appropriate $\alpha$, we can have an acceptable amount of accuracy. The exact steps of the algorithm is summarized in Algorithm 1.

### 2.3 Approximate LASSO-MAP With Threshold

Inspired by LASSO Detector introduced in [7] and similar to the algorithm introduced in Sect. 2.2, we replace $\|\mathbf{b}\|_0$ by $\|\mathbf{b}\|_1$ in (9) and find the solution in the real field. In fact we have:

$$\hat{\mathbf{b}}_{\mathbf{r}} = \underset{\mathbf{b}\in\mathbb{R}^N}{\operatorname{argmin}} \ \|\mathbf{y} - \mathbf{U}_{\hat{\mathbf{h}}}\mathbf{b}\|_2^2 + \lambda\|\mathbf{b}\|_1 \ . \tag{13}$$

The problem in (13) can be considered as LASSO problem introduced in [8]. LASSO is a shrinkage and selection method for linear regression. It minimizes the usual sum of squared errors, with a bound on the sum of the absolute values of the coefficients. It has connections to soft-thresholding of wavelet coefficients, forward stage wise regression, and boosting methods [8]. We use block coordinate [9] algorithm for LASSO to find the solution of (13). Afterwards, we use the quantization method of (12) to obtain $\hat{\mathbf{b}}$. This algorithm is summarized in Algorithm 1, and the result of this algorithm is examined experimentally in Sect. 3.

### 2.4 Approximate Backward-Detection MAP

In this section, we try to find an approximate search method for finding the solution of (9). Similar to [7], assume that we have a QR decomposition of $\mathbf{U}_{\hat{\mathbf{h}}}$ i.e. $\mathbf{U}_{\hat{\mathbf{h}}} = \mathbf{Q}\mathbf{R}_{\hat{\mathbf{h}}}$ in which $\mathbf{Q}$ is a unitary matrix and $\mathbf{R}_{\hat{\mathbf{h}}}$ is an upper triangular matrix. Now, we have:

$$\begin{aligned}
\hat{\mathbf{b}} &= \underset{\mathbf{b}\in\{0,1\}^N}{\operatorname{argmin}} \ \|\mathbf{y} - \mathbf{U}_{\hat{\mathbf{h}}}\mathbf{b}\|_2^2 + \lambda\|\mathbf{b}\|_0 = \underset{\mathbf{b}\in\{0,1\}^N}{\operatorname{argmin}} \ \|\mathbf{Q}^T\mathbf{y} - \mathbf{Q}^T\mathbf{Q}\mathbf{R}_{\hat{\mathbf{h}}}\mathbf{b}\|_2^2 + \lambda\|\mathbf{b}\|_0 \\
&= \underset{\mathbf{b}\in\{0,1\}^N}{\operatorname{argmin}} \ \sum_{i=1}^{M} \big(y_i' - \sum_{j=i}^{N} R_{\hat{h}}(i,j)b_j\big)^2 + \lambda\sum_{1}^{N} b_i \\
&= \underset{\mathbf{b}\in\{0,1\}^N}{\operatorname{argmin}} \ \sum_{i=1}^{N} \Big[\big(y_i' - \sum_{j=i}^{N} R_{\hat{h}}(i,j)b_j\big)^2 + \lambda b_i\Big] + \sum_{i=N+1}^{M} {y_i'}^2 \\
&= \underset{\mathbf{b}\in\{0,1\}^N}{\operatorname{argmin}} \ \sum_{i=1}^{N} \Big[\big(y_i' - \sum_{j=i}^{N} R_{\hat{h}}(i,j)b_j\big)^2 + \lambda b_i\Big] \ .
\end{aligned} \tag{14}$$

It is seen from (14) that once the estimates $\{\hat{b}_j\}_{j=i+1}^{N}$ are available, the optimal solution for $\hat{b}_i$ can be obtained by the following equation, regardless of the values for $\{\hat{b}_j\}_{j=1}^{i-1}$:

$$\hat{b_i}: \quad (y_i' - R_{\hat{h}}(i,i) - \sum_{j=i+1}^{N} R_{\hat{h}}(i,j)\hat{b}_j)^2 + \lambda \underset{1}{\overset{0}{\gtrless}} (y_i' - \sum_{j=i+1}^{N} R_{\hat{h}}(i,j)\hat{b}_j)^2 \ . \tag{15}$$

by doing some manipulations, (15) can be simplified as :

$$\hat{b_i}: \quad y_i' \underset{0}{\overset{1}{\gtrless}} \frac{{R_{\hat{h}}}^2(i,i) + 2R_{\hat{h}}(i,i)\sum_{j=i+1}^{N} R_{\hat{h}}(i,j)\hat{b}_j + \lambda}{2R_{\hat{h}}(i,i)} \ . \tag{16}$$

Accordingly, by backward detection of $b_i$ for $i = N, N-1, \ldots, 1$ using (16), we can find a near exact solution for (9). Based on this method, we develop a variation for our main algorithm which is summarized in Algorithm 1. Additionally, this version of our main algorithm is examined experimentally in the Sect. 3.

**Algorithm 1** Main algorithm

**init**: $k \leftarrow$ max-iterations
  $\hat{\mathbf{b}}_0 = [1, 1, \ldots, 1]^T$
**for** $i \leq k$ **do**
  $\hat{\mathbf{h}}_i = (\mathbf{U}_{\hat{\mathbf{b}}_{i-1}}^T \mathbf{U}_{\hat{\mathbf{b}}_{i-1}})^\dagger \mathbf{U}_{\hat{\mathbf{b}}_{i-1}}^T \mathbf{y}$ ▷ Structural least square estimator

  **if** MAP-algorithm=$l_2$-MAP with Thresholding **then**
    $\hat{\mathbf{b}}_\mathbf{r} = (\mathbf{U}_{\hat{\mathbf{h}}}{}^T \mathbf{U}_{\hat{\mathbf{h}}} + \lambda I)^\dagger \mathbf{U}_{\hat{\mathbf{h}}}{}^T \mathbf{y}$

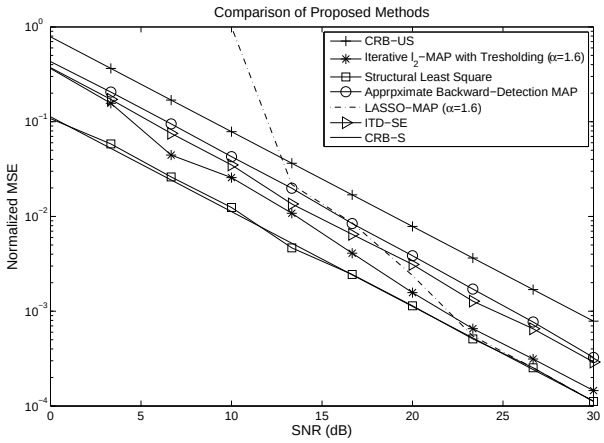
  **else if** MAP-algorithm=LASSO-MAP with Thresholding **then**
    $\hat{\mathbf{b}}_\mathbf{r} = \underset{\mathbf{b}\in\mathbb{R}^N}{\text{argmin}} \ \|\mathbf{y} - \mathbf{U}_{\hat{\mathbf{h}}}\mathbf{b}\|_2^2 + \lambda\|\mathbf{b}\|_1$
      ▷ Solution is found using block-coordinate algorithm for LASSO
  **end if**

  **if** MAP-algorithm=Backward-Detection MAP **then**
    $\mathbf{U}_{\hat{\mathbf{h}}} = \mathbf{Q}\mathbf{R}_{\hat{\mathbf{h}}}$
    $\mathbf{y}' = \mathbf{Q}^T\mathbf{y}$
    **for** $i = N, N-1, \ldots 1$ **do**
      Given $\{\hat{b}_j\}_{i+1}^{N}$
      $\hat{b_i}: \quad y_i' \underset{0}{\overset{1}{\gtrless}} \frac{{R_{\hat{h}}}^2(i,i)+2R_{\hat{h}}(i,i)\sum_{j=i+1}^{N} R_{\hat{h}}(i,j)\hat{b}_j+\lambda}{2R_{\hat{h}}(i,i)}$
    **end for**
  **else**
    $\forall i \in \{1, 2, \ldots N\}: \quad \hat{b}_i = \begin{cases} 1 & \text{if } \hat{b}_r(i) \geq \gamma(\sigma), \\ 0 & \text{if } \hat{b}_r(i) < \gamma(\sigma). \end{cases} \qquad \gamma(\sigma) = \alpha\sigma^2$
  **end if**
**end for**
$\hat{\mathbf{h}}_{\text{final}} = (\mathbf{U}_{\hat{\mathbf{b}}_k}^T \mathbf{U}_{\hat{\mathbf{b}}_k})^\dagger \mathbf{U}_{\hat{\mathbf{b}}_k}^T \mathbf{y}$

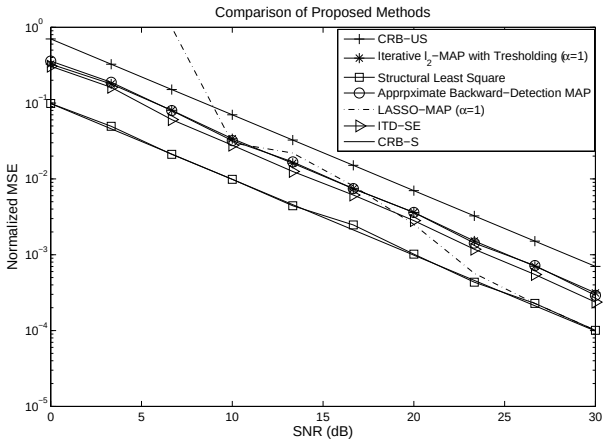
## 3 Experimental Results

In this part, we examine the efficiency of the proposed Iterative Approximate MAP-SLS algorithm and its variants, summarized in Algorithm 1 and we will compare it to ITD-SE [2]. For the sake of showing the effect of $\alpha$, we develop two

experiments. In both of the experiments, we choose $M = 50$, $N = 20$ and $K = 5$. We generate a separate $K$-sparse random signal as the sparse channel for each experiment. Indeed, elements of $\mathbf{u}$ are generated independently random according to $N(0, 1)$ as the training data[4] and $\mathbf{u}$ is generated from $\mathbf{U}$, as in [2]. For finding MSE, each experiment is repeated 100 times and the averaging over all these runs is taken. We choose $\alpha = 1.6$ for experiment 1 and $\alpha = 1$ for experiment 2. We run the simulations on an 2.8GHZ Intel Core2Duo CPU. The MSE vs SNR curves for both of the experiments are shown in Fig. 3. CRB-S and CRB-US for MSE are equal to $\sigma^2 \mathrm{Tr}\{(\mathbf{U_h}^T \mathbf{U_h})^{-1}\}$ and $\sigma^2 \mathrm{Tr}\{(\mathbf{U}^T \mathbf{U})^{-1}\}$ respectively as in [1]. Computational complexity of different algorithms are compared by the use of CPU time, shown in Tab. 3. We use CPU time as a metric for roughly comparison of computational complexity. Note that for the validity of our comparison, we use 10 iterations for all of the algorithms in the main loop. As it is seen from




(a) experiment 1., $\alpha = 1.6$




(b) experiment 2., $\alpha = 1$

**Fig. 1.** Normalized MSE vs SNR curves for proposed algorithms and ITD-SE

[4] It is better to use actual symbols as the training sequence, but, as we have seen in simulations not presented in this article, there is no big difference in the results

**Table 1.** Comparison of CPU time for various algorithms

| CPU time in seconds | |
|---|---|
| Iterative $l_2$-MAP with Thresholding | 0.0055 |
| Iterative LASSO-MAP with Thresholding | 0.1109 |
| Approximate Backward-Detection MAP | 0.0204 |
| ITD-SE | 0.0890 |

Fig. 1, by choosing $\alpha = 1.6$ the accuracy of Iterative $l_2$-MAP with Thresholding is better than that of ITD-SE and Approximate Backward-Detection MAP, although Approximate Backward-Detection MAP has the advantage that it does not need any thresholding and any pre-settings for $\gamma(\sigma)$. But by choosing $\alpha = 1$, all of these three algorithms will have the same accuracy. In both of the experiments, Iterative LASSO-MAP with Thresholding almost reaches to the CRB-S at high SNR, while it suffers from poor behaviour at low SNR. In fact, this algorithm is near-optimal at high SNR, but its complexity is much more than the others by Tab. 1.

## 4 Conclusion

In this paper, we have proposed a new method for solving the problem of channel estimation. Our method is based on an alternating minimization approach. In fact, at each iteration, MAP detection of the location of the taps and structured least square estimation are applied simultaneously. For the MAP detection part, we proposed three methods and so we proposed three variants for our algorithm. We compared our proposed variants with ITD-SE introduced in [2] which in known as an efficient method for our problem. All the proposed methods have better or at worst case equal accuracy (in the sense of MSE) compared to ITD-SE, while having less complexity except for LASSO-MAP with Thresholding. However, we experimentally see that LASSO-MAP with Thresholding almost reaches the CRB-S at high SNR and so it is near-optimal. For further work around this subject, one can perform further experiments on actual signals, coming form communication or seismic and examine the performance of our proposed methods.

## References

1. Sharp, M., Scaglione, A.: Estimation of sparse multipath channels. In: IEEE Military Communications Conference, 2008. MILCOM 2008. (2008) 1–7
2. Carbonelli, C., Vedantam, S., Mitra, U.: Sparse channel estimation with zero tap detection. IEEE Transactions on Wireless Communications **6**(5) (2007) 1743–1763
3. Candès, E., Tao, T.: The Dantzig selector: statistical estimation when p is much larger than n. Annals of Statistics **35**(6) (2007) 2313–2351
4. Haupt, J., Nowak, R.: Signal reconstruction from noisy random projections. IEEE Transactions on Information Theory **52**(9) (2006) 4036–4048
5. Babadi, B., Kalouptsidis, N., Tarokh, V.: Asymptotic achievability of the Cramér–Rao bound for noisy compressive sampling. IEEE Transactions on Signal Processing **57**(3) (2009) 1233–1236
6. Nocedal, J., Wright, S.: Numerical optimization. Springer Verlag (1999)
7. Zhu, H., Giannakis, G.: Sparsity-embracing multiuser detection for CDMA systems with low activity factor. In: Proceedings of ISIT 2009-Volume 1. (2009) 164–168
8. Tibshirani, R.: Regression shrinkage and selection via the LASSO. Journal of the Royal Statistical Society. Series B (Methodological) **58**(1) (1996) 267–288
9. Meier, L., Van de Geer, S., Buhlmann, P.: The group LASSO for logistic regression. Journal-Royal Statistical Society. Series B Statistical Methodology **70**(1) (2008) 53