Subspace Based Blind Sparse Channel Estimation

Hayashi, Kazunori; Matsushima, Hiroki; Sakai, Hideaki; De Carvalho, Elisabeth; Popovski, Petar

Published in:
Proceedings of Asia-Pacific Signal & Information Processing Association Annual Summit and Conference (APSIPA ASC) 2012

Publication date:
2012

Document Version
Early version, also known as pre-print

Subspace Based Blind Sparse Channel Estimation

Kazunori Hayashi*, Hiroki Matsushima*, Hideaki Sakai*, Elisabeth de Carvalho† and Petar Popovski†
* Graduate School of Informatics, Kyoto University, Kyoto, Japan
E-mail: kazunori@i.kyoto-u.ac.jp Tel/Fax: +81-75-753-5509
E-mail: matsushima@sys.i.kyoto-u.ac.jp Tel/Fax: +81-75-753-4755
E-mail: hsakai@i.kyoto-u.ac.jp Tel/Fax: +81-75-753-5492/4755
† Aalborg University, Aalborg, Denmark
E-mail: edc@es.aau.dk
E-mail: petarp@es.aau.dk

***Abstract*—The paper proposes a subspace based blind sparse channel estimation method using ℓ_1–ℓ_2 optimization by replacing the ℓ_2–norm minimization in the conventional subspace based method by the ℓ_1–norm minimization problem. Numerical results confirm that the proposed method can significantly improve the estimation accuracy for the sparse channel, while achieving the same performance as the conventional subspace method when the channel is dense. Moreover, the proposed method enables us to estimate the channel response with unknown channel order if the channel is sparse enough.**

I. Introduction

Blind channel estimation is a method to estimate the channel impulse response by using received signals corresponding to unknown transmitted signals. Compared to the method using known pilot signals, a blind estimation approach has an advantage that it could be possible to improve the frequency usage efficiency of wireless communications systems, which motivates studies on the blind methods [1]. Among them, the subspace based blind channel estimation method [2] is one of the most popular methods, since it can identify the channel response using second-order statistics of the received signals.

Recently, sparse reconstruction using ℓ_1 optimization has been receiving a lot of attention triggered by studies on compressed sensing [3] [4], where the problem is to reconstruct a sparse vector based on linear observations of dimension smaller than the size of the unknown sparse vector. The ℓ_1 optimization problem is a linear programming problem and hence can be solved by various algorithms. Moreover, because of the simplicity of the problem setting, the sparse reconstruction has a huge impact on very wide range of fields including signal processing, information theory and communication. For the problem with noisy observations, the linear equality constraint is replaced by an inequality, and the problem can be reduced to ℓ_1–ℓ_2 optimization problem, where the cost function is a weighted sum of ℓ_1 and ℓ_2 norms, and which is solved by using computationally efficient algorithms such as ISTA (iterative shrinkage-thresholding algorithm) [5].

In this paper, assuming the sparsity of the channel response, we propose a subspace based blind channel estimation method using ℓ_1–ℓ_2 optimization by replacing an ℓ_2–norm minimization in the original subspace method by an ℓ_1–norm minimization. Since linear constraints in the problem are obtained by the orthogonality condition coming from subspace fitting, they are homogeneous equations (i.e., $\mathbf{A}\mathbf{x} = \mathbf{0}$), while linear heterogeneous constraints (i.e., $\mathbf{A}\mathbf{x} = \mathbf{y} \neq \mathbf{0}$) typically appear in the context of compressive sensing, because $\mathbf{y}$ is obtained by observations. Thus, since the conventional optimization algorithms cannot be directly applied to our problem, we propose an algorithm by modifying ISTA. Moreover, we also propose a practical method to determine the weight in the cost function of the ℓ_1–ℓ_2 optimization. Numerical results show that the proposed method can significantly improve the estimation accuracy for the sparse channel, while achieving the same performance as the conventional subspace method when the channel is dense. Moreover, the proposed method can achieve blind estimation even when the channel order is unknown as far as the channel is sufficiently sparse.

II. Subspace-based Blind Channel Estimation Method

Here, we briefly review the conventional subspace-based blind channel estimation method [2].

We consider a multi-channel model with L channels. Let d_n and $\mathbf{h}^{(i)} = [h_0^{(i)} \ \cdots \ h_M^{(i)}]^{\mathrm{T}}$ respectively denote the transmitted symbol at time n and the discrete-time impulse response of the i-th channel with order M, where $[\cdot]^{\mathrm{T}}$ is the transpose operation. The received signal $x_n^{(i)}$ sampled at time n on the i-th channel is given by

$$x_n^{(i)} = \sum_{m=0}^{M} d_{n-m} h_m^{(i)} + b_n^{(i)}, \tag{1}$$

where $b_n^{(i)}$ is a zero-mean white noise with variance σ^2. By stacking N ($\geq M$) successive received signals at the output of the i-th channel, we obtain the received signal vector

$$\begin{aligned} \mathbf{x}_n^{(i)} &= [x_n^{(i)} \ \cdots \ x_{n-N+1}^{(i)}]^{\mathrm{T}} \\ &= \mathcal{H}_N^{(i)} \mathbf{d}_n + \mathbf{b}_n^{(i)}, \end{aligned} \tag{2}$$

This work was supported in part by the Grants-in-Aid for Scientific Research no. 24560457 from the Ministry of Education, Science, Sports, and Culture of Japan.

where $\mathcal{H}_N^{(i)}$ is an $N \times (N+M)$ Toeplitz matrix defined as

$$\mathcal{H}_N^{(i)} = \begin{bmatrix} h_0^{(i)} & \cdots & h_M^{(i)} & & \mathbf{0} \\ & \ddots & & \ddots & \\ \mathbf{0} & & h_0^{(i)} & \cdots & h_M^{(i)} \end{bmatrix}, \tag{3}$$

$\mathbf{d}_n = [d_n \ \cdots \ d_{n-N-M+1}]^{\mathrm{T}}$ and $\mathbf{b}_n^{(i)} = [b_n^{(i)} \ \cdots \ b_{n-N+1}^{(i)}]^{\mathrm{T}}$. Moreover, by defining $\mathcal{H}_N = [\mathcal{H}_N^{(0)\mathrm{T}} \ \cdots \ \mathcal{H}_N^{(L-1)\mathrm{T}}]^{\mathrm{T}}$ and $\mathbf{b}_n = [\mathbf{b}_n^{(0)\mathrm{T}} \ \cdots \ \mathbf{b}_n^{(L-1)\mathrm{T}}]^{\mathrm{T}}$, we have

$$\begin{aligned} \mathbf{x}_n &= [\mathbf{x}_n^{(0)\mathrm{T}} \ \cdots \ \mathbf{x}_n^{(L-1)\mathrm{T}}]^{\mathrm{T}} \\ &= \mathcal{H}_N \mathbf{d}_n + \mathbf{b}_n. \end{aligned} \tag{4}$$

Hereafter, we assume $\mathcal{H}_N$ and $\mathcal{H}_{N-1}$ are of full rank unless noted otherwise.

The purpose of the blind channel identification is to obtain the estimate of $\mathbf{h} = [\mathbf{h}^{(0)\mathrm{T}} \ \cdots \ \mathbf{h}^{(L-1)\mathrm{T}}]^{\mathrm{T}}$ by using $\mathbf{x}_n$ corresponding to unknown d_n up to a complex multiplicative constant, which is inherent to the problem.

The correlation matrix $\mathbf{R}_{\mathrm{x}}$ of the received signal vector $\mathbf{x}_n$ is written as

$$\mathbf{R}_{\mathrm{x}} = \mathrm{E}[\mathbf{x}_n \mathbf{x}_n^{\mathrm{H}}] = \mathcal{H}_N \mathbf{R}_{\mathrm{d}} \mathcal{H}_N^{\mathrm{H}} + \mathbf{R}_{\mathrm{b}}, \tag{5}$$

where $\mathbf{R}_{\mathrm{d}} = \mathrm{E}[\mathbf{d}_n \mathbf{d}_n^{\mathrm{H}}]$ and $\mathbf{R}_{\mathrm{b}} = \mathrm{E}[\mathbf{b}_n \mathbf{b}_n^{\mathrm{H}}] = \sigma^2 \mathbf{I}$. Here, $[\cdot]^{\mathrm{H}}$ and $\mathrm{E}[\cdot]$ denote Hermitian transpose and expectation operations, respectively. Let $\lambda_0 \geq \cdots \geq \lambda_{LN-1}$ denote the eigenvalues of $\mathbf{R}_{\mathrm{x}}$. Since we have rank $(\mathcal{H}_N \mathbf{R}_{\mathrm{d}} \mathcal{H}_N^{\mathrm{H}}) = N+M$, the eigenvalues of $\mathbf{R}_{\mathrm{x}}$ can be separated into two groups as

$$\lambda_i = \begin{cases} \lambda_i^{\mathrm{s}} + \sigma^2, & \text{for } i = 0, \cdots, M+N-1 \\ \sigma^2, & \text{for } i = M+N, \cdots, LN-1 \end{cases}, \tag{6}$$

where $\lambda_0^{\mathrm{s}} \geq \cdots \geq \lambda_{M+N-1}^{\mathrm{s}}$ are non-zero eigenvalues of $\mathcal{H}_N \mathbf{R}_{\mathrm{d}} \mathcal{H}_N^{\mathrm{H}}$. Let $\mathbf{u}_0, \cdots, \mathbf{u}_{M+N-1}$ and $\mathbf{v}_0, \cdots, \mathbf{v}_{LN-M-N-1}$ respectively denote unit-norm eigenvectors corresponding to $\lambda_0, \cdots, \lambda_{M+N-1}$ and $\lambda_{M+N}, \cdots, \lambda_{LN-1}$. Since $\mathbf{R}_{\mathrm{x}}$ is a Hermitian matrix, a subspace spanned by $\{\mathbf{u}_0, \cdots, \mathbf{u}_{M+N-1}\}$ (signal subspace) is orthogonal to the subspace spanned by $\{\mathbf{v}_0, \cdots, \mathbf{v}_{LN-M-N-1}\}$ (noise subspace). The signal subspace is also spanned by columns of $\mathcal{H}_N$, thus we have

$$\mathbf{v}_i^{\mathrm{H}} \mathcal{H}_N = \mathbf{0}, \ i = 0, \cdots, LN-M-N-1. \tag{7}$$

If $\mathcal{H}_{N-1}$ is of full-rank, $\mathbf{h}$ can be uniquely determined from (7) except for a constant complex multiplication. In practice, $\mathbf{R}_{\mathrm{x}}$ is unknown and a sample correlation matrix $\hat{\mathbf{R}}_{\mathrm{x}}$ has to be used instead, however, eigenvectors $\hat{\mathbf{v}}_i$ of $\hat{\mathbf{R}}_{\mathrm{x}}$ are different from those of $\hat{\mathbf{R}}_{\mathrm{x}}$ in general. Therefore, the estimation approach using ℓ_2–norm minimization is proposed in [2]:

$$\begin{aligned} \hat{\mathbf{h}} = &\underset{\mathbf{h}}{\operatorname{argmin}} \sum_{i=0}^{LN-N-M-1} ||\hat{\mathbf{v}}_i^{\mathrm{H}} \mathcal{H}_N||_2^2 \\ &\text{subj. to} \quad ||\mathbf{h}||_2^2 = 1, \end{aligned} \tag{8}$$

where $||\mathbf{x}||_p$ denotes an ℓ_p–norm ($p > 0$) of $\mathbf{x} = [x_1 \ \cdots \ x_N]^{\mathrm{T}}$ defined as

$$||\mathbf{x}||_p = \left(\sum_{i=1}^{N} |x_i|^p \right)^{1/p}, \tag{9}$$

and $||\mathbf{h}||_2^2 = 1$ is a constraint to avoid the trivial solution of $\mathbf{h} = \mathbf{0}$.

For a vector $\hat{\mathbf{v}}_i = [\hat{v}_i^0 \ \cdots \hat{v}_i^{LN-1}]^{\mathrm{T}}$, by defining a matrix as

$$\mathbf{P}_i = \begin{bmatrix} \hat{v}_i^0 & & \mathbf{0} & & \hat{v}_i^{(L-1)N} & & \mathbf{0} \\ \vdots & \ddots & & & \vdots & \ddots & \\ \hat{v}_i^{N-1} & & \hat{v}_i^0 & \cdots & \hat{v}_i^{LN-1} & & \hat{v}_i^{(L-1)N} \\ & \ddots & \vdots & & & \ddots & \vdots \\ \mathbf{0} & & \hat{v}_i^{N-1} & & \mathbf{0} & & \hat{v}_i^{LN-1} \end{bmatrix}, \tag{10}$$

the cost function in the optimization problem can be rewritten as

$$\sum_{i=0}^{LN-N-M-1} ||\hat{\mathbf{v}}_i^{\mathrm{H}} \mathcal{H}_N||_2^2 = \mathbf{h}^{\mathrm{H}} \mathbf{Q} \mathbf{h}, \tag{11}$$

where

$$\mathbf{Q} = \sum_{i=0}^{LN-N-M-1} \mathbf{P}_i^{\mathrm{H}} \mathbf{P}_i. \tag{12}$$

Therefore, the solution of (8) is obtained as a unit-norm eigenvector corresponding to the minimum eigenvalue of $\mathbf{Q}$.

III. Proposed Blind Sparse Channel Estimation Method

A. Problem Setting

Assuming that the channel coefficients are sparse, i.e., the number of non-zero elements of $\mathbf{h}$ is much smaller than $L(M+1)$, and that only $\hat{\mathbf{R}}_{\mathrm{x}}$ instead of $\mathbf{R}_{\mathrm{x}}$ is available, we consider ℓ_1–norm minimization problem under the relaxed orthogonality condition of (7) with the inequality. Specifically, the minimization problem can be formulated as follows:

$$\begin{aligned} \hat{\mathbf{h}} = &\arg\min_{\mathbf{h}} ||\mathbf{h}||_1 \\ &\text{subj. to} \quad ||\mathbf{P}\mathbf{h}||_2^2 \leq \epsilon, \ ||\mathbf{h}||_2^2 = 1, \end{aligned} \tag{13}$$

where ϵ is a non-negative constant and $\mathbf{P}$ is obtained by Cholesky decomposition as $\mathbf{Q} = \mathbf{P}^{\mathrm{H}} \mathbf{P}$. Note that we assume the channel coefficients are real for simplicity hereafter, while the proposed approach can be easily extended to the case with complex coefficients [6]. By replacing the inequality constraint to the penalty, we obtain

$$\begin{aligned} \hat{\mathbf{h}} = &\arg\min_{\mathbf{h}} \lambda ||\mathbf{h}||_1 + ||\mathbf{P}\mathbf{h}||_2^2 \\ &\text{subj. to} \quad ||\mathbf{h}||_2^2 = 1, \end{aligned} \tag{14}$$

where λ is a non-negative constant, which determines the trade-off between the sparsity and the error of the estimated channel response.

B. Modified ISTA

The optimization problem with the form

$$\hat{\mathbf{z}} = \arg\min_{\mathbf{z}} \lambda ||\mathbf{z}||_1 + ||\mathbf{y} - \mathbf{A}\mathbf{z}||_2^2 \tag{15}$$

is called ℓ_1–ℓ_2 optimization problem, since the cost function consists of terms with ℓ_1– and ℓ_2–norms. In general, the ℓ_1–ℓ_2 optimization problem can be solved efficiently with various algorithm. In this paper, we consider to apply ISTA to our problem. The steps in ISTA are summarized as follows:
(i) Initialization: Set $\mathbf{z}[0] \in \mathbb{R}^N$ and a constant $c > \sigma_{\max}(\mathbf{A})$, where $\sigma_{\max}(\mathbf{A})$ denotes the maximum singular value of $\mathbf{A}$.
(ii) Repeat (a) and (b) below:
(a) Back projection of residual:

$$\mathbf{z}'[k] = \frac{1}{c}\mathbf{A}^{\mathrm{T}}(\mathbf{y} - \mathbf{A}\mathbf{z}[k-1]) + \mathbf{z}[k-1]$$

(b) Shrinkage：

$$\mathbf{z}''[k] = \mathrm{F}(\mathbf{z}'[k]),$$

where

$$F(\mathbf{z}'[k]) = \begin{bmatrix} \mathrm{sgn}(z_1'[k])(|z_1'[k]| - \frac{2\lambda}{c})_+ \\ \vdots \\ \mathrm{sgn}(z_N'[k])(|z_N'[k]| - \frac{2\lambda}{c})_+ \end{bmatrix},$$

$\mathbf{z}'[k] = [z_1'[k] \cdots z_N'[k]]^{\mathrm{T}}$, and $(b)_+ = \max\{b, 0\}$.
In the algorithm, F is a vector valued nonlinear function, whose input-output relation of the element is shown in Fig. 1, and is employed to obtain sparse solution by forcing elements in the vector to be zero if their absolute values are less than a certain value.

The conventional ISTA cannot be directly applied to our problem because the constraints obtained from the orthogonality condition are linear homogeneous equations, while heterogeneous equations appear in the original problem. Therefore, we propose a modified ISTA to cope with the problem (14) as follows:
(i) Initialization: Set $\mathbf{h}[0] = \{\mathbf{h} | \mathbf{h} \in \mathbb{R}^{L(M+1)},\ ||\mathbf{h}||_2^2 = 1\}$ and a constant $c > \sigma_{\max}(\mathbf{P})$.
(ii) Repeat (a)–(c) below:
(a) Back projection of residual:

$$\mathbf{h}'[k] = \left(\mathbf{I}_{L(M+1)} - \frac{1}{c}\mathbf{P}^{\mathrm{H}}\mathbf{P}\right)\mathbf{h}[k-1]$$

(b) Shrinkage：

$$\mathbf{h}''[k] = \mathrm{G}(\mathbf{h}'[k]),$$

where

$$G(\mathbf{h}'[k]) = \begin{bmatrix} \mathrm{sgn}(h_1'[k])(|h_1'[k]| - \frac{2\lambda}{c}||\mathbf{h}'[k]||_2)_+ \\ \vdots \\ \mathrm{sgn}(h_N'[k])(|h_N'[k]| - \frac{2\lambda}{c}||\mathbf{h}'[k]||_2)_+ \end{bmatrix},$$

$\mathbf{h}'[k] = [h_1'[k] \cdots h_N'[k]]^{\mathrm{T}}$.
(c) Normalization:

$$\mathbf{h}[k] = \frac{\mathbf{h}''[k]}{||\mathbf{h}''[k]||_2}$$

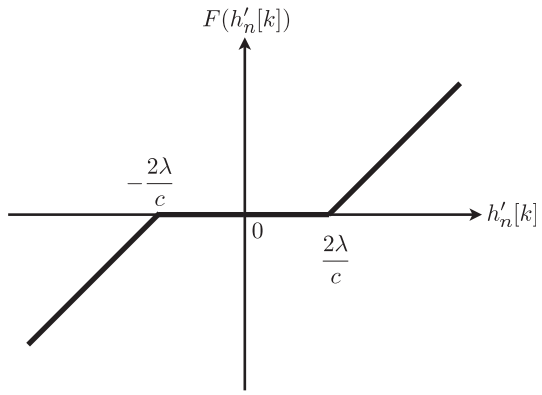

Fig. 1. Shrinkage in ISTA

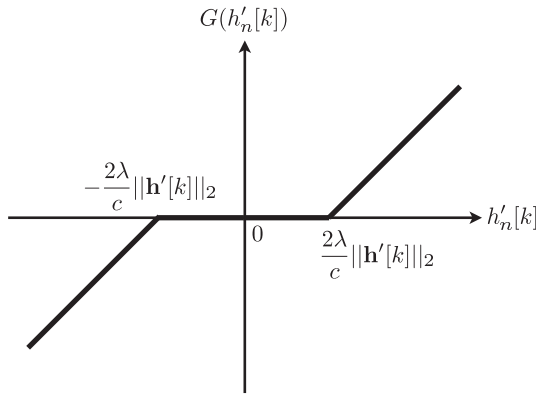

Fig. 2. Shrinkage in modified ISTA

In the algorithm, G is a vector valued nonlinear function, whose input-output relation of the element is shown in Fig. 2, where we have replaced $2\lambda/c$ in the conventional algorithm with $2\lambda||\mathbf{h}'[k]||_2/c$, because the ℓ_2–norm of the estimated vector varies by the back projection operation in our problem. Also, we have introduced the normalization operation in the end of each iteration to meet the constraint of $||\mathbf{h}||_2^2 = 1$.

In the ℓ_1–ℓ_2 optimization problem, how to select λ is one of the crucial issues. We discuss the selection of λ based on numerical observations in the next section.

IV. Numerical Results

A. System Parameters

We have conducted computer simulations to demonstrate the performance of the proposed method. System parameters are shown in Table IV. We have set the number of samples of the received signal vector for each channel to be $N = 35$ and the number of multichannel to be $L = 2$. Also, the order of each channel is assumed to be $M = 30$, and nonzero elements of the channel vector are generated from $\mathcal{N}(0, 1)$. As the performance measure, we use normalized mean-square-error (NMSE) defined as

$$\mathrm{NMSE} = E\left[\min_{\alpha}\left(\frac{||\alpha\hat{\mathbf{h}} - \mathbf{h}^*||_2^2}{||\mathbf{h}^*||_2^2}\right)\right],$$

TABLE I
SYSTEM PARAMETER

# of samples for each channel	N =35
Channel order	M =30
# of multichannel	L =2
Channel model	Rayleigh fading
Delay power profile	uniform

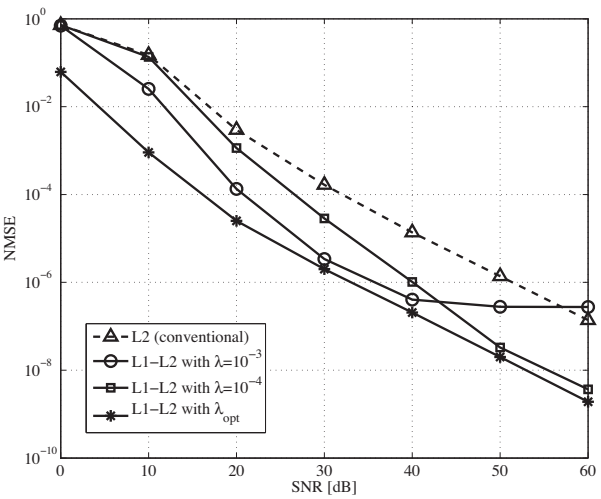

Fig. 3. NMSE performance for various values of λ

where $\hat{\mathbf{h}}$ and $\mathbf{h}^*$ respectively denote the estimated and the true channel vectors, and α is to identify channel vectors, which differ only by a scalar multiplication. We have evaluated the NMSE by averaging normalized estimation errors for 1,000 independent realizations of channel vectors.

B. Selection of λ

In order to see the impact of the selection of λ, we firstly show the NMSE performance of the $\ell_1 - \ell_2$ approach using various values of λ in Fig. 3. Here, we have assumed the channel order L to be known, and show the performance of the conventional subspace method with ℓ_2 optimization in the same figure for comparison purpose. In the figure, "$\lambda = 10^{-3}$" and "$\lambda = 10^{-4}$" show the performance by using fixed values of λ of 10^{-3} and 10^{-4}, respectively. On the other hand, "$\lambda_{\rm opt}$" is the performance obtained by exhaustive numerical search of λ, which achieves the minimum NMSE, for each realization of the sample correlation matrix. From the results, we can see that, although $\lambda_{\rm opt}$ is not feasible in practical applications, the proposed $\ell_1 - \ell_2$ approach has a potential to largely outperform the conventional subspace method with ℓ_2-norm by adequately choosing λ. Moreover, from the observations of the performance with $\lambda = 10^{-3}$ and $\lambda = 10^{-4}$, which achieves close performance as $\lambda_{\rm opt}$ for the SNR regions of 30-40 dB and 50-60 dB, respectively, we can expect that the optimum λ will depend on the received SNR.

In Fig. 4, we show the optimum values of λ versus SNR, which are obtained numerically, for different values of the

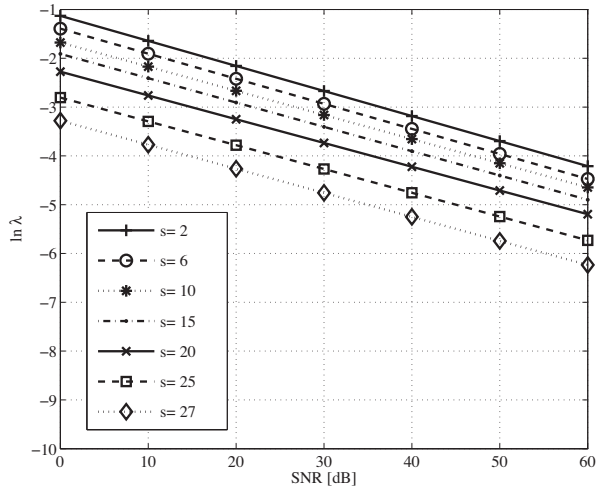

Fig. 4. Optimum values of λ versus SNR for different sparsities

number of nonzero elements s in the channel response. Note that, here we have assumed to use the same value of λ for each SNR, thus these are different from $\lambda_{\rm opt}$ in Fig. 3, which can select different value for each realization. From the figure, we can observe that there is a linear relation between the decibel value of SNR and the natural logarithm of λ. The relation also depends on the number of nonzero elements s, however, the impact is not on the slope but only on the bias.

Based on the observations above, we propose to model λ as the function of the noise variance σ^2 and the number of nonzero elements of each channel s as

$$\ln\lambda(\sigma^2, s) = -a_1(s) + a_2 \ln\sigma^2, \tag{16}$$

where the bias $a_1(s)$ and the inclination a_2 are determined by linear regression from the numerical results. It should be noted that we can assume the signal power to be one without loss of generality, because channels, which differ only by a constant multiplication, are identified in the blind identification problem. Hereafter, we denote λ determined by (16) for given σ^2 and s as $\lambda_{\rm prop}$. In the blind channel identification problem, both σ^2 and s are unknown, and hence they have to be estimated. In the proposed method, σ^2 is estimated as the minimum eigenvalue of the sample correlation matrix $\hat{\mathbf{R}}_{\rm x}$, while s is estimated from the result obtained by using a fixed value of λ $(= 10^{-4})$.

C. NMSE Performance: Known Channel Order

Fig. 5 shows the NMSE performance of the proposed method using $\lambda_{\rm prop}$, assuming that channel order M is known and the number of nonzero elements s is 3. For comparison purpose, performance of the conventional subspace method with ℓ_2–norm, that of the proposed method using optimum lambda $\lambda_{\rm opt}$, and that of the proposed method using $\lambda_{\rm prop}$ but with known s and σ^2 in (16) are also plotted in the same figure. From the figure, we can see that the proposed approach

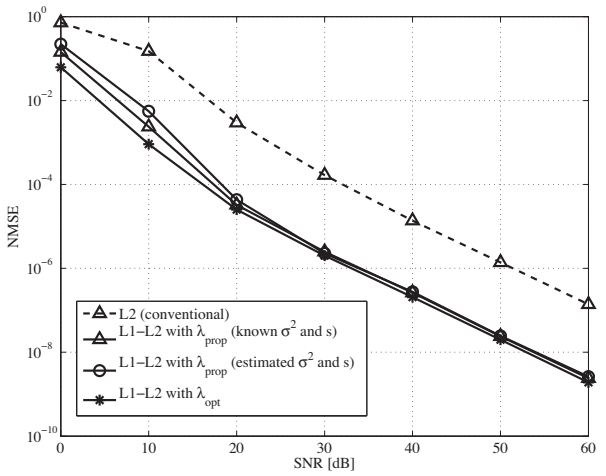

Fig. 5. NMSE versus SNR (channel order: known)

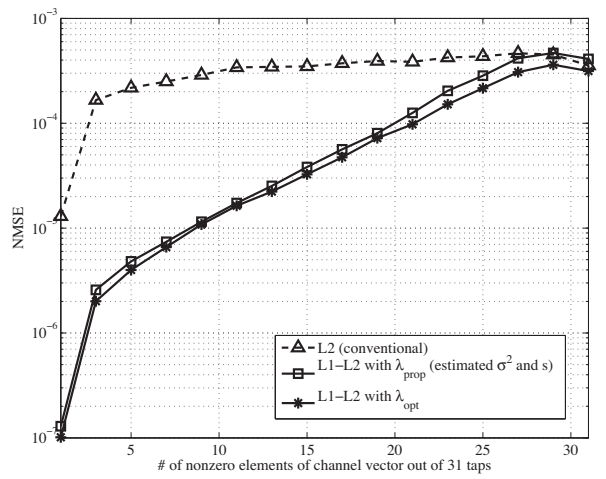

Fig. 6. NMSE versus sparsity (channel order: known)

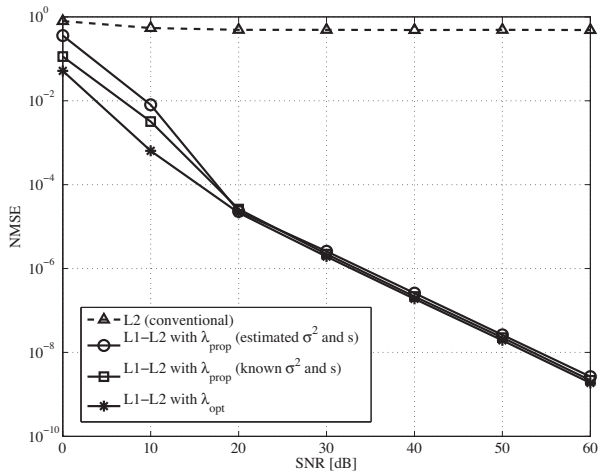

Fig. 7. NMSE versus SNR (channel order: unknown)

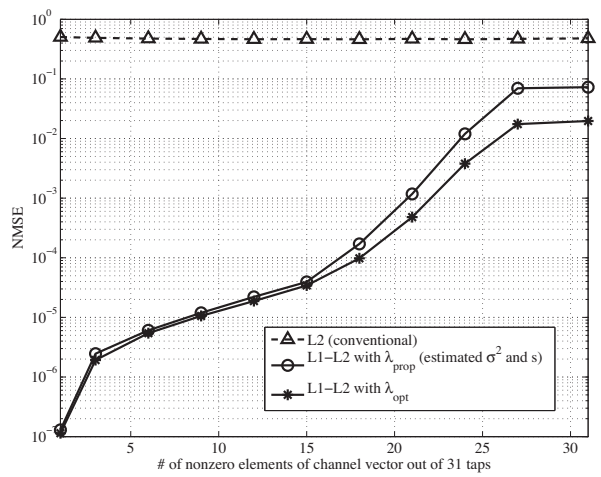

Fig. 8. NMSE versus sparsity (channel order: unknown)

using ℓ_1–ℓ_2 optimization largely outperform the conventional method using ℓ_2–norm. Also, the proposed method using $\lambda_{\rm prop}$ can achieve almost the same performance as that of the proposed method using $\lambda_{\rm opt}$ for high SNR.

Fig. 6 shows NMSE versus the number of nonzero elements s for a fixed value of SNR=30 dB. From the result, we see that the proposed method can significantly improve the estimation accuracy for the sparse channel, while achieving the same performance as the conventional method when the channel is dense. This means that the proposed method can be applied to channels with arbitrary sparsity by adequately adjusting λ. The proposed $\lambda_{\rm prop}$ can achieve the NMSE performance close to that of $\lambda_{\rm opt}$ for any sparsity.

D. NMSE Performance: Unknown Channel Order

Figs. 7 and 8 show NMSE versus SNR and NMSE versus sparsity s, assuming that the channel order is unknown and is overestimated by one. In this case, the orthogonality condition (7) becomes underdetermined, and thus the conventional method based on ℓ_2–norm fails to estimate the channel response. On the other hand, the proposed method can obtain an accurate estimate of the channel response if the channel is sparse enough (if around half of channel taps are zero). Therefore, with the proposed method, we can estimated the channel whose order is unknown, if the channel response is sparse to some extent and the order can be overestimated.

V. Conclusion

We have proposed a subspace blind channel estimation method using ℓ_1–ℓ_2 optimization assuming the channel impulse response to be sparse. We have modified the conventional ISTA to solve our problem, where the linear constraints are homogeneous, and provided a practical method

to determine λ, which plays an important role in the ℓ_1–ℓ_2 problem. From numerical results, we have confirmed that the proposed method can significantly improve the NMSE performance when the channel response is sparse. Moreover, with the proposed λ_{prop}, the proposed method can achieve good NMSE performance in channels with arbitrary sparsity. Furthermore, the proposed method can estimate the channel response even when the order is unknown, as far as the channel is sparse enough and the order can be overestimated.

REFERENCES

[1] L. Tong and S. Perreau, "Multichannel blind identification: From subspace to maximum likelihood methods", *Proc. IEEE*, vol. 86, pp. 1951–1968, Oct. 1998.

[2] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Process.*, vol. 43, no. 2, Feb. 1995.

[3] D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[4] E.J. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[5] M. Zibulevsky and M. Elad, "L1-L2 optimization in signal and image processing," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 76–88, May 2010.

[6] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Commun. Pure Appl. Math.*, vol. 57, no. 11, pp. 1413–1457, Nov. 2004.