Qiu et al. EURASIP Journal on Advances in Signal Processing 2012, 2012:104
http://asp.eurasipjournals.com/content/2012/1/104

EURASIP Journal on Advances in Signal Processing
a SpringerOpen Journal

RESEARCH Open Access

Identification of MIMO systems with sparse transfer function coefficients

Wanzhi Qiu*, Syed Khusro Saleem and Efstratios Skafidas

Abstract

We study the problem of estimating transfer functions of multivariable (multiple-input multiple-output–MIMO) systems with sparse coefficients. We note that subspace identification methods are powerful and convenient tools in dealing with MIMO systems since they neither require nonlinear optimization nor impose any canonical form on the systems. However, subspace-based methods are inefficient for systems with sparse transfer function coefficients since they work on state space models. We propose a two-step algorithm where the first step identifies the system order using the subspace principle in a state space format, while the second step estimates coefficients of the transfer functions via L1-norm convex optimization. The proposed algorithm retains good features of subspace methods with improved noise-robustness for sparse systems.

Keywords: system identification, MIMO system, sparse representation, L1-norm optimization

1. Introduction

The problem of identifying multiple-input multiple-output (MIMO) systems arises naturally in spatial division multiple access architectures for wireless communications. Subspace system identification methods refer to the category of methods which obtain state space models from subspaces of certain matrices constructed from the input-output data [1]. Being based on reliable numerical algorithms such as the singular value decomposition (SVD), subspace methods do not require nonlinear optimization and, thereby, are computationally efficient and stable without suffering from convergence problems. They are particularly suitable for identifying MIMO systems since there is no need to impose on the system a canonical form and, therefore, they are free from the various inconveniences encountered in classical parametric methods.

Another good feature of subspace methods is they incorporate a reliable order estimation process. Although this is largely ignored by other identification methods, system order estimation should be an integral part of a system identification algorithm. In fact, correctly identifying the system order is essential to ensure that subsequent parameter estimation will yield a well-defined set of model coefficient estimates. While it is obvious that an underestimated order will result in large modeling errors, it is equally dangerous to have an over-parameterized model as a result of selecting an improperly large order. Over-parameterization not only creates a larger set of parameters to be estimated, but also leads to poorly defined (high variance) coefficient estimates and surplus unvalidated content in the resulting model [2]. For parameterized linear models, a usual way of estimating the order is to conduct a series of tests on different orders, and select the best one based on the goodness-of-fit using certain criterion such as the Akaike's information criterion [3]. Subspace-based order identification determines the order as the number of principle eigenvalues (or singular values) of a certain input-output data matrix. This mechanism has been proven to be a simple and reliable way of order estimation and the same principle has been applied to detect the number of emitter sources in array signal processing [4], to determine the number of principal components in signal and image analysis [5,6] and to estimate the system order in blind system identification [7,8].

It has become well known that many systems in real applications actually have sparse representations, i.e., with a large portion of their transfer function coefficients equal to zero [9]. For example, communication channels exhibit great sparseness. In particular, in high-

* Correspondence: wanzhiqiu128@yahoo.com
National ICT Australia, Department of Electrical and Electronic Engineering, The University of Melbourne, Parkville, VIC 3010, Australia

definition television, there are few echoes but the channel response spans many hundreds of data symbols [10-12]. In broadband wireless communications, a hilly terrain delay profile consists of a sparsely distributed multipath [13]. Underwater acoustic channels also exhibit sparseness [14].

Despite of their good features, subspace methods are not suited for systems with sparse transfer function coefficients since they deal with state space models. One fact is that for a given transfer function matrix, there exist infinite number of state space representations related by a similarity transform. As a result, when the system to be identified has sparse transfer function coefficients, the state space model produced by a subspace method is almost certain to be non-sparse due to the inherent arbitrary similarity transform associated with the model.

Work on sparse system identification has been reported in [15-17], however, they consider only single-input single-out (SISO) systems. In addition, these methods assume the system order is known *a priori*, which is not necessarily true in practice. This article studies the problem of the identification of MIMO systems with sparse transfer functions coefficients. In particular, we leverage on reliable system order identification of subspace methods to build input-output relationship in terms of transfer function coefficients, and exploit L1-norm optimization proven to be able to produce robust sparse solutions [9] to estimate these coefficients. The resulting method consists of a systematic way of order identification and efficient coefficient estimation for sparse systems.

The rest of this article is organized as follows. Section 2 gives an overview of subspace identification methods. Section 3 proposes the LRL1 algorithm. Section 4 presents simulation results and Section 5 draws conclusions.

2. Subspace identification

For an M-input $\{u_m(k), m = 1,...,M\}$ L-output $\{y_l(k), l = 1,...,L\}$ system described in the state space form:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k \tag{1a}$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k \tag{1b}$$

where $\mathbf{u}_k \stackrel{\text{def}}{=} [u_1(k), \ldots, u_M(k)]^T$, $\mathbf{y}_k \stackrel{\text{def}}{=} [\gamma_1(k), \ldots, \gamma_L(k)]^T$, and $\mathbf{x}_k \stackrel{\text{def}}{=} [x_1(k), \ldots, x_N(k)]^T$ are the input, output, and state vectors, respectively, with T denoting matrix transpose. $\mathbf{A} \in R^{N\times N}$, $\mathbf{B} \in R^{N\times M}$, $\mathbf{C} \in R^{L\times N}$, and $\mathbf{D} \in R^{L\times M}$ are the system, input, output, and direct feed-through matrices, respectively. Noise in (1) has been omitted for brevity purposes and will be added in the simulation in Section 4. Given S measurements (i.e., $k = 0,..., S - 1$) of the input and output, subspace identification methods estimate the system order N and matrices (**A, B, C, D**), and derive system transfer functions via $\mathbf{H}(z) = \mathbf{D} + \mathbf{C}(z\mathbf{I}_{NN} - \mathbf{A})^{-1}\mathbf{B}$, with $\mathbf{I}_{NN}$ denoting the identity matrix of size $N \times N$. The procedure is as follows.

Choose a positive integer i and define the input and output block Hankel matrices as:

$$\mathbf{U}_{0/i-1} = \begin{pmatrix} \mathbf{u}_0 & \mathbf{u}_1 & \cdots & \mathbf{u}_{j-1} \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_j \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{u}_{i-1} & \mathbf{u}_i & \cdots & \mathbf{u}_{i+j-2} \end{pmatrix} \tag{2}$$

$$\mathbf{Y}_{0/i-1} = \begin{pmatrix} \mathbf{y}_0 & \mathbf{y}_1 & \cdots & \mathbf{y}_{j-1} \\ \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_j \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{y}_{i-1} & \mathbf{y}_i & \cdots & \mathbf{y}_{i+j-2} \end{pmatrix}$$

Note that $\mathbf{U}_{0/i-1}$ has j $(= S - i + 2)$ columns with the first one formed by $\mathbf{u}_k$ $(k = 0,..., i - 1)$. Similarly, we can define the j-column block Hankel matrix $\mathbf{U}_{i/2i-1}$ using $\mathbf{u}_k$ $(k = i,..., 2i - 1)$ to form the first column. For convenience, the following short-hand notations are adopted:

$$\mathbf{U}_p \stackrel{\text{def}}{=} \mathbf{U}_{0/i-1}, \quad \mathbf{Y}_p \stackrel{\text{def}}{=} \mathbf{Y}_{0/i-1} \tag{3}$$

$$\mathbf{U}_f \stackrel{\text{def}}{=} \mathbf{U}_{i/2i-1}, \quad \mathbf{Y}_f \stackrel{\text{def}}{=} \mathbf{Y}_{i/2i-1} \tag{4}$$

For each of the four matrices in (3) and (4), by deleting or adding one block row at the bottom and keeping the number of columns unchanged, we have two additional notations. For example,

$$\mathbf{Y}_f^- \stackrel{\text{def}}{=} \mathbf{Y}_{i/2i-2}, \quad \mathbf{Y}_f^+ \stackrel{\text{def}}{=} \mathbf{Y}_{i/2i}$$

Note that i is a user-specified parameter and is required to be larger than the system order, i.e., $i > N$. In general, a subspace identification algorithm consists of three steps. The first step performs projections of the row spaces of the data matrices and estimates the order of the system. In particular, the following oblique projections are first calculated:

$$\mathbf{O}_i = \mathbf{Y}_f / \mathbf{u}_f \mathbf{W}_p \tag{5}$$

$$\mathbf{O}_{i+1} = \mathbf{Y}_f^- / \mathbf{u}_f^- \mathbf{W}_p^+$$

where $\mathbf{W}_p = \begin{pmatrix} \mathbf{U}_p \\ \mathbf{Y}_p \end{pmatrix}$, $\mathbf{W}_p^+ = \begin{pmatrix} \mathbf{U}_p^+ \\ \mathbf{Y}_p^+ \end{pmatrix}$, and $\mathbf{Y}_f/\mathbf{u}_f\mathbf{W}_p$ are the oblique projection of the row space of $\mathbf{Y}_f$ along the row space of $\mathbf{U}_f$ on the row space of $\mathbf{W}_p$ and can be calculated according to

$$\mathbf{Y}_f/\mathbf{u}_f\mathbf{W}_p = [\mathbf{Y}_f/\mathbf{U}_f^T] \cdot [\mathbf{W}_p/\mathbf{U}_f^T] \cdot \mathbf{W}_p$$

with $[\mathbf{Y}_f/\mathbf{U}_f^T]$ denoting the projection of the row space of $\mathbf{Y}_f$ on the row space of the orthogonal compliment of the row space of $\mathbf{U}_f$ [1].

The SVD of the weighted oblique projection is calculated and the order of the system (N) is determined as the number of principal singular vales in $\mathbf{\Sigma}$, leading to the identification of the principal subspaces $\mathbf{U}_1$ and $\mathbf{V}_1$,

$$\mathbf{W}_1\mathbf{O}_i\mathbf{W}_2 = \mathbf{U}\mathbf{\Sigma}\mathbf{V} = (\mathbf{U}_1\ \mathbf{U}_2)\begin{pmatrix} \mathbf{\Sigma}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_2 \end{pmatrix}\begin{pmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{pmatrix} = \mathbf{U}_1\mathbf{\Sigma}_1\mathbf{V}_1^T \quad (6)$$

where $\mathbf{W}_1$ and $\mathbf{W}_2$ are weighting matrices and a specific choice of them leads to different algorithms [1]. For example,

$$\mathbf{W}_1 = \mathbf{I}_{iL},\ \mathbf{W}_2 = \mathbf{I}_j \quad (7)$$

and

$$\mathbf{W}_1 = \mathbf{I}_{iL},\ \mathbf{W}_2 = (\mathbf{U}_f^\perp)^T(\mathbf{U}_f^\perp(\mathbf{U}_f^\perp)^T)^\dagger\mathbf{U}_f^\perp, \quad (8)$$

result in the popular N4SID and MOESP algorithms, respectively, with † denoting Moore-Penrose pseudo-inverse.

Define the extended observability matrix and state sequence as

$$\Gamma_i \overset{\text{def}}{=} \begin{pmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{i-1} \end{pmatrix}$$

$$\mathbf{X}_i \overset{\text{def}}{=} (\mathbf{x}_i\mathbf{x}_{i+1}, \ldots, \mathbf{x}_{i+j-1}) \in R^{N\times j}.$$

The second step of a subspace algorithm is to find estimates of the extended observability matrix and state sequences via

$$\Gamma_i = \mathbf{W}_1^{-1}\mathbf{U}_1\mathbf{\Sigma}_1^{1/2}$$

$$\mathbf{X}_i = \Gamma_i^\dagger\mathbf{O}_i, \quad \mathbf{X}_{i+1} = \Gamma_{i-1}^\dagger\mathbf{O}_{i+1}$$

The final step is to determine the system matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ (up to within a similarity transformation) by

$$\min_{\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D}} \left\| \begin{pmatrix} \mathbf{X}_{i+1} \\ \mathbf{Y}_{i/i} \end{pmatrix} - \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}\begin{pmatrix} \mathbf{X}_i \\ \mathbf{U}_{i/i} \end{pmatrix} \right\|_F^2. \quad (9)$$

Note that the sizes of matrices involved in the subspace methods are dependant on i. It will be shown in Section 4 that while the order can be estimated reliably across a broad range of i's as long as the condition $i > N$ is satisfied, the quality of subsequent coefficient estimates depends on i in a non-monotonous fashion.

3. The proposed LRL1

As mentioned previously, subspace methods are inefficient in dealing with sparse transfer function coefficients. That is, for a system with sparse transfer function coefficients, the outcome of subspace methods is one of many (non-sparse) state space realizations ($\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$). In order to exploit the sparseness in the system, system (1) is represented using transfer function coefficients. In particular, for the lth output the following linear regression (LR) relationship holds:

$$\begin{aligned} y_l(k+N) &= \sum_{n=1}^{N} a_n y_l(k+N-n) \\ &+ \sum_{m=1}^{M}\sum_{n=0}^{N} b_{mn,l}u_m(k+N-n), \quad k = 1,\ldots,S-N, l = 1,\ldots,L, \end{aligned} \quad (10)$$

Where$\{a_n, b_{mn,\ l}\}$ are the system transfer function coefficients. Equation (10) can be re-written into a vector form:

$$y_l(k+N) = \mathbf{y}_{k,l}^T\mathbf{a} + \sum_{m=1}^{M}\mathbf{b}_{m,l}^T\mathbf{u}_{mk}, \quad k = 1,\ldots,S-N, \quad (11)$$

where

$$\mathbf{a} = (a_1 \ldots a_N)^T, \quad (12)$$

$$\mathbf{y}_{k,l}^T = \left(y_l(k+N-1), y_l(k+N-2), \ldots, y_l(k)\right),$$

$$\mathbf{u}_{mk}^T = \left(u_m(k+N), u_m(k+N-1), \ldots, u_m(k)\right),$$

and the contribution of the mth input to the lth output is presented by

$$\mathbf{b}_{m,l} = (b_{m0,l}, \ldots, b_{mN,l})^T \quad (13)$$

By stacking up the output samples in (11), one can have

$$\mathbf{y}_l = [\mathbf{Y}_l \ \mathbf{U}]\,\boldsymbol{\gamma}_l, \tag{14}$$

where

$$\mathbf{y}_l = [\gamma_l(N+1)\gamma_l(N+2)\ldots\gamma_l(S)]^T,$$

$$\mathbf{U} = (\mathbf{U}_1 \ldots \mathbf{U}_M),$$

$$\mathbf{Y}_l = \begin{pmatrix} \mathbf{y}_{1,l}^T \\ \mathbf{y}_{2,l}^T \\ \vdots \\ \mathbf{y}_{S-N,l}^T \end{pmatrix}, \quad \mathbf{U}_m = \begin{pmatrix} \mathbf{u}_{m1}^T \\ \mathbf{u}_{m2}^T \\ \vdots \\ \mathbf{u}_{m(S-N)}^T \end{pmatrix},$$

$$\boldsymbol{\gamma}_l = (\mathbf{a}^T \ \boldsymbol{\beta}_l^T)^T,$$

$$\boldsymbol{\beta}_l = (\mathbf{b}_{1,l}^T \ldots \mathbf{b}_{M,l}^T)^T,$$

Note that it is a common practice to identify a multi-input single-output model for each output separately and then combine the individual models into a final MIMO model. However, if there are common or correlated parameters among models for different output variables and/or correlated noise, then performing identification on all outputs simultaneously can lead to better and more robust models [18]. By combining all outputs of (14), we have

$$\mathbf{y} = \mathbf{H}\,\boldsymbol{\beta} \tag{15}$$

where

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_L \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \mathbf{a} \\ \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_L \end{pmatrix},$$

$$\mathbf{H} = \begin{pmatrix} \mathbf{Y}_1 & \mathbf{U} & \mathbf{0} & \cdots & \cdots & \mathbf{0} \\ \mathbf{Y}_2 & \mathbf{0} & \mathbf{U} & \mathbf{0} & \cdots & \mathbf{0} \\ \cdots & & \cdots & & & \\ \mathbf{Y}_L & \mathbf{0} & \cdots & & \mathbf{0} & \mathbf{U} \end{pmatrix}.$$

We now propose a two-step algorithm which first identifies the system order in state space format. This allows the formation of a MIMO model in terms of transfer function coefficients. L1-norm optimization is then exploited to estimate the coefficients of the model. One of the key factors leading to recent advances in compressed sensing (CS) is that L1-norm optimization promotes sparsity. That is, as compared with its popular L2-norm counterpart, L1-norm optimization produces better results when the parameters to be estimated are sparse [9]. In our case, the data model (15) allows $\boldsymbol{\beta}$ to be obtained by L1-norm minimization with guaranteed convergence since, unlike the cases in CS, this is strict convex optimization. The minimization generally takes the following form [19]:

$$\hat{\boldsymbol{\beta}} = \arg\ \min_{\mathbf{x}} \left(||\mathbf{x}||_1 + \lambda||\mathbf{y} - \mathbf{H}\mathbf{x}||_2^2\right) \tag{16}$$

Where λ balances the sparsity of the solution with the fidelity to the data and should be inversely proportional to the noise level. Efficient algorithms for solving (16) have been developed [9,19].

The proposed algorithm makes further use of the SVD in (6), where matrix $\Sigma_1 = \mathrm{diag}(\lambda_1,\ldots,\lambda_N)$ contains the principal singular-values and $\Sigma_2 = \mathrm{diag}(\lambda_{N+1},\ldots,\lambda_{iL})$ contains noise singular-values. When noise is absent, Σ_2 is an all-zero matrix. Otherwise, its diagonal entries will be non-zero positive numbers. Although it is impossible here to accurately estimate the noise variance, these diagonal entries do reflect the noise level and can be represented by

$$\bar{\sigma} = \sqrt{\left(\sum_{k=N+1}^{iL} \lambda_k^2\right) / (iL - N)} \tag{17}$$

In our proposed method, we will set the parameter in (16) to $\lambda = \delta/\bar{\sigma}$ where δ is a positive constant. The proposed LRL1 algorithm can now be summarized as follows:

Step 1: Subspace identification of system order using (5) and (6), and estimation of noise level using (17). This step is a systematic order selection procedure with reliable performance.

Step 2: Estimation of system transfer function coefficients using (16). This step is simultaneous optimization over all outputs with robust sparse solutions.

4. Simulation

We evaluate the performances of the proposed LRL1 method against that of the subspace method N4SID [1]. The first MIMO system under consideration is generated by modifying the SISO system in [[1], p. 155]. In particular, the system order and eigenvalues (i.e., the vector $[1\ \mathbf{a}^T]^T$) are kept unchanged and one more input and one more output are added to the original system. Then, some elements in each of the transfer function vector $\mathbf{b}_{m,l}$ (see (13)) are set to zero to create the desired sparsity. The resulting 2-input 2-output system of order $N = 4$ has the following transfer function coefficients:

System 1:

$$\mathbf{a} = (0\ 0\ 0\ \ 0.5288)^T,$$

$$\begin{pmatrix} \mathbf{b}_{1,1}^T \\ \mathbf{b}_{2,1}^T \\ \mathbf{b}_{1,2}^T \\ \mathbf{b}_{2,2}^T \end{pmatrix} = \begin{pmatrix} -0.7139\ 0\ 1.2080\ 0\ -4.0361 \\ 0\ 2.9505\ -2.7061\ 0\ 0.5238 \\ 0\ 0.0526\ -0.0239\ -0.6024\ 0 \\ 0\ -0.2686\ -0.2374\ 0.1585\ 0 \end{pmatrix}.$$

In the simulation, the data size is chosen to be S = 200 and the inputs are sequences of Gaussian variables with zero mean and unit variance. A white noise vector $\mathbf{v}_k$ is added to the output vector $\mathbf{y}_k$ (see (1b)) and the SNR as defined below is kept constant for each k,

$$\text{SNR} = \frac{||\mathbf{y}_k||^2}{||\mathbf{v}_k||^2}.$$

100 runs are conducted for each simulated scenario, where for each run the input and noise vectors are independently generated. The root mean square error (RMSE) for the estimates of the $T = N + M(N + 1)L$ parameters in $\boldsymbol{\beta}$ is measured as

$$\text{RMSE} = \sqrt{\frac{1}{100T}\sum_{r=1}^{100} ||\boldsymbol{\beta}(r) - \boldsymbol{\beta}||_2^2}$$

where $\boldsymbol{\beta}(r)$ denotes the estimates in the rth run. In the simulation, δ = 0.5 is fixed for all the scenarios and the convex programming package from [20] is used in solving (16).

As mentioned in Section 2, the subspace approach identifies the system order (N) by examining the iL diagonal entries of $\boldsymbol{\Sigma}$ in (6), i.e., the singular values (λ_1,..., λ_{iL}) of the system, where i is the number of block rows of the Hankel data matrices (see (2)) and L = 2 is the number of outputs. For SNR = 30 and i changing from 4 to 10, we obtain iL singular values for each particular choice of i. Figure 1 shows (λ_1,..., λ_{iL}, 0,..., 0) where zero padding is made for cases with i < 10.

Checking for the number of principal (dominant) singular values clearly indicates that the system order is N = 4. This result verifies that the identification of system order is reliable regardless of the value of i as long as it is chosen to be larger than N. This feature is very attractive in practice since it is relatively easy to select an i which is larger than the largest possible value the system order might have.

Having identified the system order, we now follow the subsequent steps of the subspace and proposed LRL1 methods to estimate model coefficients. Figure 2 shows the RMSE of the subspace method when i changes from 5 to 10 at SNR = 30. It can be seen here that the performance of the subspace method depends on i in a non-

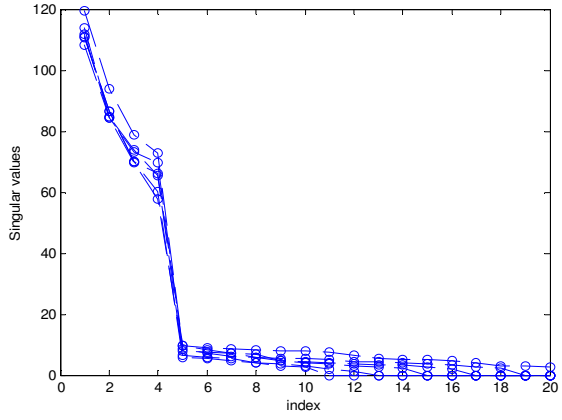

Figure 1 System singular values (System 1).

monotonous fashion. That is, a larger i does not necessarily leads to better coefficient estimates. This is due to the fact that only finite datasets are available; increasing the number of rows of the data matrices will lead to a reduction in the number of columns. In the subsequent simulations, i = 9 is used.

Figure 3 compares the estimation errors of the subspace and LRL1 methods for different SNR values. The results show that LRL1 is more robust to noise in dealing with this sparse system.

Further test is conducted by modifying *System 1* to create an even sparser system. In particular, some of the transfer function coefficients of *System 1* are set to zero, which results in

System 2:

$$\mathbf{a} = (0\ 0\ 0\ \ 0.5288)^T,$$

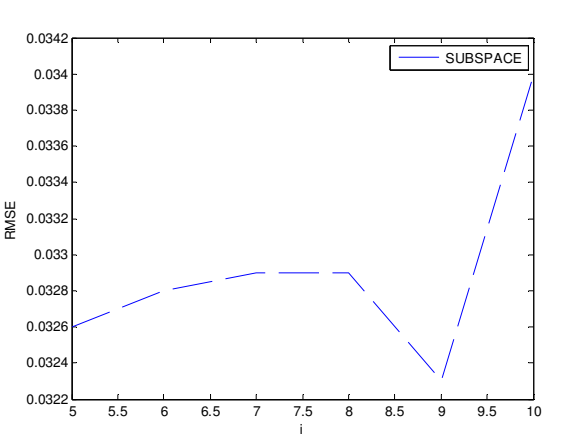

Figure 2 Estimation error of the subspace method (System 1).

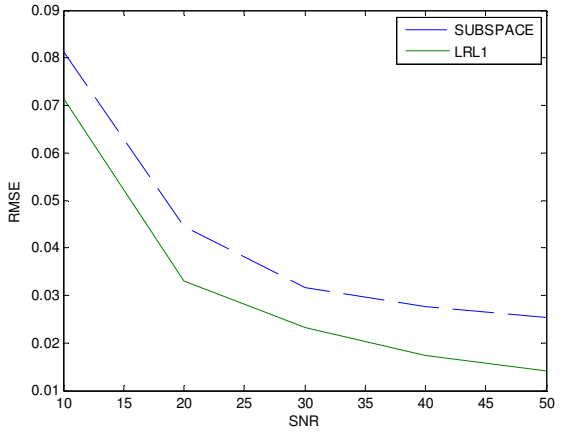

Figure 3 Performance comparison (System 1).

$$\begin{pmatrix} \mathbf{b}_{1,1}^T \\ \mathbf{b}_{2,1}^T \\ \mathbf{b}_{1,2}^T \\ \mathbf{b}_{2,2}^T \end{pmatrix} = \begin{pmatrix} 0\ 0 & 1.2080 & 0 & -4.0361 \\ 0\ 0 & -2.7061 & 0 & 0 \\ 0\ 0.0526 & -0.0239 & 0 & 0 \\ 0\ 0 & 0 & 0.1585 & 0 \end{pmatrix}.$$

Figure 4 shows the performances of the two methods under the same conditions as in Figure 3. The results demonstrate that the improvement of LRL1 over the original subspace method increases when the system to be identified becomes sparser.

5. Conclusion

A noise-robust algorithm for the identification of MIMO systems has been presented. The proposed method leverages reliable system order identification of subspace principle and exploits L1-norm optimization to achieve high effectiveness for identifying systems with sparse transfer function coefficients. While retaining good features of original subspace methods such as the convenience for multivariable systems, the proposed method is shown to be able to significantly improve estimation accuracy for sparse systems.

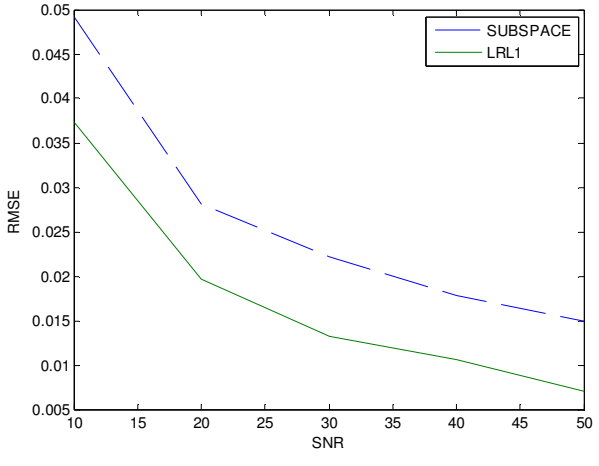

Figure 4 Performance comparison (System 2).

Competing interests
The authors declare that they have no competing interests.

Received: 9 January 2012 Accepted: 9 May 2012 Published: 9 May 2012

References

1. P Vanoverschee, B DeMoor, *Subspace Identification for Linear Systems: Theory-Implementation-Applications* (Springer, 1996).
2. P Young, *Recursive Estimation and Time-series Analysis: An Introduction* (Springer, 1984)
3. H Akaike, A new look at the statistical model identification. IEEE Trans Autom Control. **19**(6), 716–723 (1974)
4. S Ouyang, Y Hua, Bi-iterative least square method for subspace tracking. IEEE Trans Signal Process. **53**(8), 2984–2996 (2005)
5. W Qiu, E Skafidas, Robust estimation of GCD with sparse coefficients. Signal Process. **90**(3), 972–976 (2010)
6. H Andrews, C Patterson, Singular value decompositions and digital image processing. IEEE Trans Acoust Speech Signal Process. **24**(1), 26–53 (2003)
7. KA Meraim, W Qiu, Y Hua, Blind system identification. Proc IEEE. **85**(8), 1310–1322 (1997)
8. W Qiu, SK Saleem, M Pham, Blind identification of multichannel systems driven by impulsive signals. Digital Signal Process. **20**(3), 736–742 (2010)
9. E Candès, M Wakin, An introduction to compressive sampling. IEEE Signal Process Mag. **25**(2), 21–30 (2008)
10. SF Cotter, BD Rao, Sparse channel estimation via matching pursuit with application to equalization. IEEE Trans Commun. **50**(3), 374–377 (2002)
11. WF Schreiber, Advanced television systems for terrestrial broadcasting: some problems and some proposed solutions. Proc IEEE. **83**, 958–981 (1995)
12. IJ Fevrier, SB Gelfand, MP Fitz, Reduced complexity decision feedback equalization for multipath channels with large delay spreads. IEEE Trans Commun. **47**, 927–937 (1999)
13. S Ariyavisitakul, NR Sollenberger, LJ Greenstein, Tap selectable decision-feedback equalization. IEEE Trans Commun. **45**, 1497–1500 (1997)
14. M Kocic, D Brady, M Stojanovic, Sparse equalization for real-time digital underwater acoustic communications, in *Proc OCEANS95*, San Diego, CA, pp. 1417–1422 (October 1995)
15. Y Chen, Y Gu, AO Hero III, Sparse LMS for system identification, in *Proceedings of ICASSP*, Taipei, Taiwan, pp. 3125–3128 (19-24 April 2009)
16. M Sharp, A Scaglione, Application of sparse signal recovery to pilot-assisted channel estimation, in *Proc of Intl Conf on Acoustics, Speech and Signal Proc*, Las Vegas, NV, (April 2008)
17. M Sharp, A Scaglione, Estimation of sparse multipath channels, in *Proceedings of IEEE Military Communications Conference*, San Diego, CA, pp. 1–7 (17-19 November 2008)
18. BS Dayal, JF MacGregor, Multi-output process identification. J Process Control Elsevier Sci. **7**(4), 269–282 (1997)
19. E Candès, T Tao, The Dantzig selector: statistical estimation when *p* is much larger than n. Ann Stat. **35**, 2313–2351 (2005)
20. M Grant, S Boyd, CVX: Matlab software for disciplined convex programming (web page and software). http://stanford.edu/~boyd/cvx. Accessed 12 January 2012

doi:10.1186/1687-6180-2012-104
Cite this article as: Qiu *et al.*: Identification of MIMO systems with sparse transfer function coefficients. *EURASIP Journal on Advances in Signal Processing* 2012 **2012**:104.