BLIND IDENTIFICATION OF FIR MIMO CHANNELS BY GROUP DECORRELATION

Senjian An[1], *Yingbo Hua*[2] *and Jonathan H. Manton*[1]

[1] Department of Electrical and Electronic Engineering
The University of Melbourne, VIC 3010, Australia
senjian, j.manton@ee.mu.oz.au

[2]Department of Electrical Engineering
University of California, Riverside, CA 92521, USA
yhua@ee.ucr.edu

ABSTRACT

In this paper, we present a new method for identification of FIR MIMO channels driven by unknown, uncorrelated and colored sources. This method, belonging to the BID (i.e., blind identification by decorrelation) family, make use of the mutual uncorrelation of the unknown sources by first decorrelating the observed signals into two uncorrelated *groups*. The two decorrelators are then used to estimate the channel matrix (i.e., MIMO channel transfer function matrix) up to a constant matrix. This constant matrix is finally determined using a BID method for instantaneous MIMO channels. This new method, named BID-G, is shown to be much more robust than the subspace method that requires the channel matrix to be irreducible and column-reduced.

1. INTRODUCTION

Blind identification of multiple-input-multiple-output (MIMO) channels is a problem arising from a wide range of applications. In this paper, we consider the case where the MIMO channel has a finite-impulse-response (FIR) and driven by unknown, uncorrelated and colored sources. The existing methods for this problem include the subspace method [1] [2] and [3]. This family of methods requires that the channel matrix is irreducible and column-reduced. Another method shown in [4] also requires the same condition. A recent method called blind identification by decorrelating subchannels (BID-S) [5] [6] has reduced the above requirement. A core step of the BID-S method is to exploit the uncorrelated sources by decorrelating subchannels. The BID-S method is a generalization of the BID-I method [7] [8] that only works on instantaneous MIMO channels. In this paper, we present a new method in the BID family, called BID-G, which represents a new progress of the development of the BID family. The BID-G method first constructs two group decorrelators and then exploits the decorrelators to estimate the MIMO channel matrix up to an unknown constant matrix. This constant matrix is then estimated using the BID-I method. The BID-G method requires that each column of the channel matrix is irreducible and the other remaining columns are irreducible and column-reduced. This is a weaker condition than that the whole channel matrix is irreducible and column-reduced. Hence, the BID-G method is more robust than the subspace method.

In the next section, we provide the problem formulation and also explain the key condition of the subspace method. In section 3, the theory of the BID-G method is presented. In section 4, we illustrate the performance of the BID-G method in a contrast to the subspace method.

2. PROBLEM FORMULATION

Consider an FIR MIMO channel described by

$$\mathbf{y}(n) = \sum_{k=0}^{q} H_k \mathbf{x}(n-k) + \mathbf{w}(n) \tag{1}$$

where $\mathbf{x}(n)$ is the vector of m input signals, $\mathbf{y}(n)$ is the vector of p output signals, $\{H_k\}_{k=0}^{q}$ is the $p \times m$ matrix sequence of the channel's impulse response with length q , and $\mathbf{w}(n)$ is a noise vector that is uncorrelated with $\mathbf{x}(n)$.

An equivalent form of (1) is $\mathbf{y}(n) = \mathbf{H}(z)\mathbf{x}(n) + \mathbf{w}(n)$ where $\mathbf{H}(z) = \sum_{k=1}^{q} H_k z^{-k}$. We assume that there are sufficient data such that the second-order-statistics can be exploited. We write the autocorrelation matrix of $\mathbf{y}(n)$ at lag τ as

$$R_{\mathbf{yy}}(\tau) \doteq \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbf{y}(n+\tau)\mathbf{y}^T(n) \tag{2}$$

and the power spectral matrix $\mathbf{S_{yy}}(z)$ of $\mathbf{y}(n)$ is defined as $\mathbf{S_{yy}}(z) = \sum_{\tau=-M}^{M} R_{\mathbf{yy}}(\tau) z^{-\tau}$ where M may be finite (or infinite) depending on the source power spectra. The power spectral matrices of the sources $\mathbf{x}(n)$ are similarly defined. The power spectrum of the white noise is $\mathbf{S_{ww}}(z) = \delta^2 \mathbf{I}$. Then from (1) and note that the noise is uncorrelated with the sources, we have

$$\mathbf{S_{yy}}(z) = \mathbf{H}(z)\mathbf{S_{xx}}(z)\mathbf{H}^T(z^{-1}) + \delta^2 \mathbf{I} \tag{3}$$

If $p > m$, the noise spectra can be obtained asymptotically [5]. For a simpler presentation, we will drop the noise term, and hence our theory will be based on the noiseless model

$$\mathbf{S_{yy}}(z) = \mathbf{H}(z)\mathbf{S_{xx}}(z)\mathbf{H}^T(z^{-1}) \tag{4}$$

The aim here is to estimate the channel matrix $\mathbf{H}(z)$ using the autocorrelation matrices $R_{\mathbf{yy}}(\tau)$.

THIS WORK WAS SUPPORTED IN PART BY AUSTRALIAN RESEARCH COUNCIL, THE ARC SPECIAL RESEARCH CENTER FOR ULTRA BROADBAND NETWORKING AND U.S. NATIONAL SCIENCE FOUNDATION.

Let $\mathcal{T}_l(H)$ denote the $(l+1)p \times (q+l+1)m$ generalized Sylvester matrix of $\mathbf{H}(z)$, i.e.,

$$\mathcal{T}_l(\mathbf{H}) \doteq \begin{bmatrix} H_0 & \cdots & H_q & & \\ & \ddots & & \ddots & \\ & & H_0 & \cdots & H_q \end{bmatrix}.$$

Let $\mathcal{C}_l(H)$ denote a $(q+l+1)p \times (l+1)m$ dual matrix of $\mathcal{T}_l(H)$, i.e.,

$$\mathcal{C}_l(\mathbf{H}) \doteq \begin{bmatrix} H_0 & & \\ \vdots & \ddots & \\ H_q & & H_0 \\ & \ddots & \vdots \\ & & H_q \end{bmatrix}.$$

The subspace method requires the following assumptions:

(**A1**). $\mathbf{H}(z)$ is irreducible and column reduced;

(**A2**). All columns of $\mathbf{H}(z)$ have the same degree q;

(**A3**). $l \geq \max_{1 \leq j \leq p-m} L_j^{\perp}$.

where $\{L_j^{\perp}\}_1^{p-m}$ are the dual Kronecker indices of the rational subspace spanned by the column vectors of $\mathbf{H}(z)$.

An practical upper bound of $\max_{1 \leq j \leq p-m} L_j^{\perp}$ is mq and hence one can choose $l \geq mq$ in practice.

Under the above conditions, the subspace method can obtain the channel matrix up to a unknown constant matrix Q.[2]. This matrix can be then identified up to scaling and permutation by the joint diagonalization method if the sources satisfy

(**A4**). The sources are mutually uncorrelated and of distinct power spectra.

3. THE THEORY OF THE BID-G METHOD

3.1. Group Decorrelation

Let $\mathbf{h}_i(z)$ be the ith column of $\mathbf{H}(z)$ and $\mathbf{H}_i(z)$ the remaining submatrix of $\mathbf{H}(z)$. Note that any of the above submatrices of $\mathbf{H}(z)$ is more likely to be (or closer to be) irreducible and column-reduced than $\mathbf{H}(z)$ itself. The BID-G method attempts to find the left null spaces of $\mathcal{T}_{l_1}(\mathbf{h}_i)$ and $\mathcal{T}_{l_2}(\mathbf{H}_i)$, here $l_1 \geq q$ and $l_2 \geq (m-1)q$, and then applies the subspace method to identify $\mathbf{h}_i(z)$ and $\mathbf{H}_i(z)$ up to scaling or a $(m-1) \times (m-1)$ constant matrix respectively.

Now we show that the left null spaces of $\mathcal{T}_{l_1}(\mathbf{h}_i)$ and $\mathcal{T}_{l_2}(\mathbf{H}_i)$ can be found by decorrelating the output signals into two uncorrelated groups. Let $L_1 = p(l_1+1) - (l_1+1+q)$, $L_2 = p(l_2+1) - (m-1)(l_2+q+1)$, $\mathbf{G}_1(z) = \sum_{i=0}^{l_1} G_{1i} z^{-i} \in \mathbb{R}[z]^{L_1 \times p}$, $\mathbf{G}_2(z) = \sum_{i=0}^{l_2} G_{2i} z^{-i} \in \mathbb{R}[z]^{L_2 \times p}$ with

$$\mathcal{T}_0(\mathbf{G}_1) = [G_{11}, G_{12}, \cdots, G_{1l_1}] \in \mathbb{R}^{L_1 \times p(l_1+1)},$$

$$\mathcal{T}_0(\mathbf{G}_2) = [G_{21}, G_{22}, \cdots, G_{2l_2}] \in \mathbb{R}^{L_2 \times p(l_2+1)}$$

having full row ranks. Here $\mathbb{R}, \mathbb{R}[z]$ represent the real number field and polynomial matrix set with real coefficient matrices.

The main technical result of this paper is the following

Theorem 1. *Suppose $\mathbf{G}_1(z)$ and $\mathbf{G}_2(z)$ satisfy the conditions discussed above. Also assume that $\mathbf{H}(z), \mathbf{S}_{\mathbf{xx}}(z)$ meet (**A1-A4**) with (**A3**) replaced by*

- (**A3'**). $l_1 \geq q, l_2 \geq (m-1)q$.

Then, by constructing two decorrelators $\mathbf{G}_1(z)$ and $\mathbf{G}_2(z)$ such that the output of first decorrelator is uncorrelated with that of the second, i.e.,

$$\mathbf{G}_1(z)\mathbf{S}_{\mathbf{yy}}(z)\mathbf{G}_2^T(z^{-1}) = 0 \quad (5)$$

we achieve that the output signals of the two decorrelators are due to two distinct groups of the input signals of the MIMO channel, i.e.,

$$\exists i \in \{1, 2, \cdots, m\}, \mathbf{G}_1(z)\mathbf{h}_i(z) = 0, \mathbf{G}_2(z)\mathbf{H}_i(z) = 0 \quad (6)$$

which is equivalent to

$$\exists i \in \{1, 2, \cdots, m\}, \mathcal{T}_0(\mathbf{G}_1)\mathcal{T}_{l_1}(\mathbf{h}_i) = 0, \mathcal{T}_0(\mathbf{G}_1)\mathcal{T}_{l_2}(\mathbf{H}_i) = 0.$$

where $\mathbf{h}_i(z)$ is the ith column of $\mathbf{H}(z)$ and $\mathbf{H}_i(z)$ denotes the remaining submatrix of $\mathbf{H}(z)$.

The proof is omitted due to space limitations.

3.2. Channel Estimation

Note that

$$\mathcal{T}_0(\mathbf{G}_1)\mathcal{T}_{l_1}(\mathbf{h}_i) = 0 \Longleftrightarrow \mathcal{C}_q(\mathbf{G}_1)\mathcal{C}_0(\mathbf{h}_i) = 0$$
$$\mathcal{T}_0(\mathbf{G}_2)\mathcal{T}_{l_2}(\mathbf{H}_i) = 0 \Longleftrightarrow \mathcal{C}_q(\mathbf{G}_2)\mathcal{C}_0(\mathbf{H}_i) = 0$$

The coefficient matrices of $\mathbf{h}_i(z)$ and $\mathbf{H}_i(z)$, namely $\mathcal{C}_0(\mathbf{h}_i)$ and $\mathcal{C}_0(\mathbf{H}_i)$ in the above equations, can be found up to scaling or a constant matrix respectively by performing singular value decomposition on $\mathcal{C}_q(\mathbf{G}_1)$ and $\mathcal{C}_q(\mathbf{G}_2)$ and getting their right null space.

Now we have in hand an estimate of the channel, namely $\widehat{\mathbf{H}}(z) = [\widehat{\mathbf{h}}_i(z), \widehat{\mathbf{H}}_i(z)]$, which is equal to $\mathbf{H}(z)$ up to a right invertible constant matrix, i.e., there is a nonsingular matrix Q such that $\mathbf{H}(z) = \widehat{\mathbf{H}}(z)Q$. Since $\mathbf{H}(z)$ is irreducible, so is $\widehat{\mathbf{H}}(z)$. We can compute an equalizer $\mathbf{E}_H(z)$ satisfying $\mathbf{E}_H(z)\widehat{\mathbf{H}}(z) = I$. However, the computation of the equalizer may be very sensitive to estimation errors if (**A1**) is not well satisfied. In this paper, we use the generalized equalizer $\mathbf{G}_H(z)$ satisfying $\mathbf{G}_H(z)\widehat{\mathbf{H}}(z) = \alpha(z)I$ where $\alpha(z)$ can be any polynomial.

The output of the generalized equalizer, say $\mathbf{v}(n) = \mathbf{G}_H(z)\mathbf{y}(n)$, satisfies

$$\mathbf{S}_{\mathbf{vv}}(z) = \mathbf{G}_H(z)\mathbf{S}_{\mathbf{yy}}(z)\mathbf{G}_H^T(z^{-1})$$

and hence $\mathbf{S}_{\mathbf{vv}}(z) = Q\alpha(z)\alpha(z^{-1})\mathbf{S}_{\mathbf{xx}}(z)Q^T$. The matrix Q is then obtained by the diagonalization method [8].

3.3. The Cost Function of BID-G

Note that (5) is equivalent to

$$\mathcal{T}_0(\mathbf{G}_1)P(\tau)\mathcal{T}_0(\mathbf{G}_2)^T = 0, -M < \tau < M \quad (7)$$

where $\mathcal{T}_0(\mathbf{G}_1) \in \mathbf{R}^{L_1 \times p(l_1+1)}$, $\mathcal{T}_0(\mathbf{G}_2) \in \mathbf{R}^{L_2 \times p(l_2+1)}$,and

$$P(\tau) = \begin{bmatrix} R_{yy}(\tau) & \cdots & R_{yy}(\tau+l_2) \\ \vdots & \ddots & \vdots \\ R_{yy}(\tau-l_1) & \cdots & R_{yy}(\tau+l_2-l_1) \end{bmatrix}$$

We can choose the following cost function

$$\mathcal{J}_1(G_1, G_2) = trace\left\{ \sum_{\tau=-M_1}^{M_1} G_1 P(\tau) G_2^T G_2 P^T(\tau) G_1^T \right\} \quad (8)$$

where G_1 and G_2 are the coefficient matrices, denoted by $\mathcal{T}_0(\mathbf{G}_1)$ and $\mathcal{T}_0(\mathbf{G}_2)$ in (7), of $\mathbf{G}_1(z)$ and $\mathbf{G}_2(z)$ respectively. Here, M_1 is the lag length we used. It may be smaller than M since some part of the equations in (7) may be enough (in theory) to imply all the other equations. In order to make the solution unique, up to a left unitary matrix, and the minimum of the cost function is close to zero, we need some extra constraints on $G_1$and G_2 and recondition the matrix $P(\tau)$.

Let

$$Y_l \doteq \begin{bmatrix} R_{yy}(0) & \cdots & R_{yy}(l) \\ \vdots & \ddots & \vdots \\ R_{yy}(-l) & \cdots & R_{yy}(0) \end{bmatrix}, l = l_1, l_2.$$

If Y_{l_1} and Y_{l_2} are nonsingular, we can add constraints on $\mathbf{G}_1(z)$ and $\mathbf{G}_2(z)$ such that

$$\begin{aligned} \mathcal{T}_0(\mathbf{G}_1) Y_{l_1} \mathcal{T}_0(\mathbf{G}_1)^T &= I_{L_1} \\ \mathcal{T}_0(\mathbf{G}_2) Y_{l_2} \mathcal{T}_0(\mathbf{G}_2)^T &= I_{L_2} \end{aligned}$$

This idea is very similar as the prewhitening technique for blind source separation [8]. The difference is the prewhitening is performed for the left decorrelaor and the right decorrelaor separately.

However, generally, Y_{l_1} and Y_{l_2} may be singular and the ranks may be difficult to estimate correctly in practice. In this case, it means that there exist a polynomial vector $\mathbf{g}(z)$ with degree l_1 or l_2 respectively satisfying $\mathbf{g}^T(z)\mathbf{H}(z) = 0$. In our algorithm, we modify the "prewhitening" algorithm in the following way.

Denote the singular value decompositions of Y_{l_1} and Y_{l_2} by $Y_{l_1} = U_1 S_1 U_1^T$ and $Y_{l_2} = U_2 S_2 U_2^T$. (Note that $R_{yy}(-k) = R_{yy}^T(k)$ and hence Y_l is symmetric). Set

$$\begin{aligned} \widehat{S}_1(k,k) &= \begin{cases} S_1(k,k), \text{ if } \frac{S_1(k,k)}{S_1(1,1)} > \varepsilon \\ 1, \text{otherwise} \end{cases}, \\ \widehat{S}_2(i,i) &= \begin{cases} S_2(i,i), \text{ if } \frac{S_2(i,i)}{S_2(1,1)} > \varepsilon \\ 1, \text{otherwise} \end{cases}, \end{aligned}$$

for $1 \le k \le p(l_1+1), 1 \le i \le p(l_2+1)$. The off diagonals of $\widehat{S}_1$ and $\widehat{S}_2$ are zero. Then we replace $P(\tau)$ by

$$\overline{P}(\tau) = \widehat{S}_1^{-\frac{1}{2}} U_1^T P(\tau) U_2 \widehat{S}_2^{-\frac{1}{2}}$$

and use the following new cost function

$$\mathcal{J}(X_1, X_2) = trace\left\{ \sum_{\tau=-M_1}^{M_1} X_1 \overline{P}(\tau) X_2^T X_2 \overline{P}^T(\tau) X_1^T \right\} \quad (9)$$

We find $G_i = X_i \widehat{S}_i^{-\frac{1}{2}} U_i^T$ by minimizing this cost function under the constraint$X_i X_i^T = I_{L_i}, i = 1, 2$.

3.4. Alternating Projection

Since the cost function is nonlinear, it is difficult to get the global minimum generally. However, for a fixed G_1 (or G_2), the cost function is a quadratic function of G_2 (or G_1 respectively). We can minimize the cost function with respect to G_1 or G_2 in an alternating fashion. We refer to this procedure as alternating projection(AP).

Since there are m possible group decorrelators, it is not always necessary to get the global minimum. We only need to find one of the possible group decorrelators.

Let (G_1, G_2, ε) , where ε is the value of the cost function, be a local minima achieved by alternating projection from a randomly selected initial point. In simulations for $m \ge 3$, we observe that (G_1, G_2) may not be a group decorrelator, but G_2 is very likely to be a part of the left null space of one column of the channel, i.e., $G_2 \mathcal{T}_{l_2}(\mathbf{h}_i) \approx 0$. Based on this observation, we estimate one polynomial vector, say $\mathbf{h}(z)$, from $G_2 \mathcal{T}_{l_2}(\mathbf{h}) \approx 0$, and use the left null space of $\mathcal{T}_{l_1}(\mathbf{h})$, say $\widehat{G_1}$, as the initial point of G_1 for the next round of alternating projection. We now outline the proposed algorithm in the following

1. Find a local minima (G_1, G_2, ε) using AP from any initial point. Set $G_1^1 = G_1, G_2^1 = G_2, \varepsilon_1 = \varepsilon, k = 1$;
2. Compute the minimal right singular vector, denoted by h, of $\mathcal{C}_q(\mathbf{G}_2^k(z))$ where $\mathbf{G}_2^k(z)$ is a polynomial matrix with $\mathcal{T}_0(\mathbf{G}_2^k) = G_2^k$;
3. Set G_1 to be an orthogonal basis of the left null space of $\mathcal{T}_{l_1}(\mathbf{h})$ where $\mathbf{h}(z)$ is a polynomial vector with $\mathcal{C}_0(\mathbf{h}) = h$;
4. Using G_1 as the initial point and apply the AP algorithm, one get another local minima (G_1, G_2, ε).
5. Check if this local minima has been encountered before. If it does not repeat the previous ones, set $k = k+1, G_1^k = G_1, G_2^k = G_2, \varepsilon_k = \varepsilon$ and go to step 2. Otherwise, find the best one with minimal cost function among this chain of local minimas.

Our simulations show that this algorithm is rather effective to overcome the local minima problem. In the simulations performed for this paper, we perform each group decorrelation with randomly selected initial points.

4. SIMULATIONS

In this section, some numerical simulations are performed on the proposed blind identification algorithm based on group decorrelation to compare with subspace method presented in [2].

Example 1. We randomly selected 100 channels $\mathbf{H}(z)$ with dimension 4×3 and of degree 3 and three sources with randomly selected power spectra of degree 8. The SNR is chosen to be 20dB and the number of samples is 10000. In order to show how many channels are well estimated by the BID-G method and the subspace method, the channels are ordered with increasing NMSE for BID-G in the upper part of Figure 1. Here, NMSE stands for normalized mean square error of the channel estimation. In the lower part of Figure 1, the channels are ordered with increasing NMSE for the subspace method. The advantage of the BID-G method is clearly shown in Fig. 1. The BID-G performs well (NMSE less than -20dB) for about 80% of all channels while the subspace method for about 15%. Moreover, for most channels, the BID-G performs better than the subspace method.

Example 2. In this simulation, we chose a relatively well-conditioned channel from 100 channels randomly selected and tested the performance of the two methods versus signal noise rate (SNR). The minimal singular value of $\mathcal{T}_l(\mathbf{H})$ for this channel is 0.0947 where $l = mq = 9$. The power spectra of the sources are of degree 8 and randomly selected for each run. The NMSE is the averaged among 60 runs. The number of samples was chosen to be 10000.

Example 3. Fig. 3 demonstrates the performance of the two methods for a relatively ill-conditioned channel. The minimal singular value of $\mathcal{T}_l(\mathbf{H})$ for this channel is 0.0020 where $l = 9$. The other parameters were selected in the same way as Example 2.

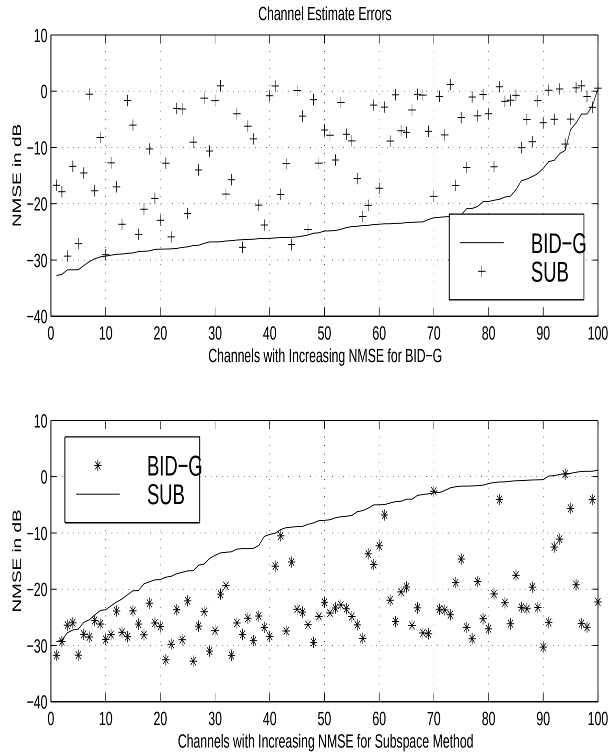

Fig. 1. A comparison of the BID-G and subspace methods

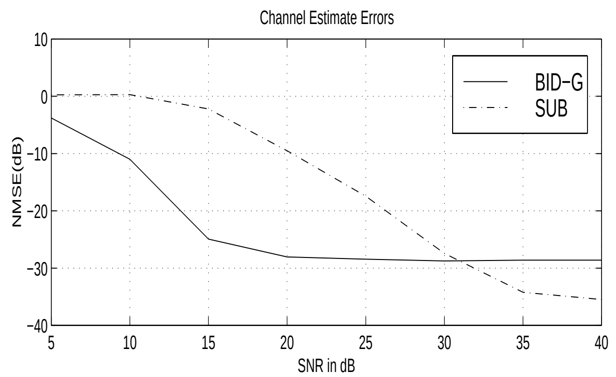

Fig. 2. Performance comparison of the BID-G and subspace methods for a well-conditioned channel

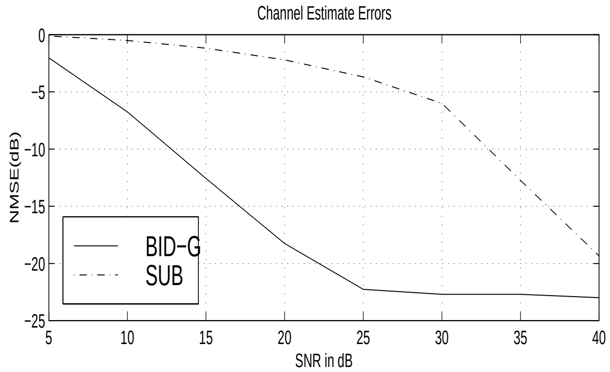

Fig. 3. Performance comparison of the BID-G and subspace methods for an ill-conditioned channel

5. CONCLUSION

In this paper, we have presented a new method in the BID family for blind identification of FIR MIMO channels driven by unknown, uncorrelated and colored signals. This new method, BID-G, is shown to be much more robust than the subspace method. Further study of the BID-G method in comparison to the BID-S method is currently underway.

6. REFERENCES

[1] A. Gorokhov and P. Loubaton, "Subspace-based techniques for blind separation of convolutive mixtures with temporally correlated sources," *IEEE Trans. Circuits and Systems-I*, vol. 44, no. 9, pp. 813–820, 1997.

[2] P. Loubaton, E. Moulines, and P. Regalia, "Subspace method for blind identification and deconvolution," in *Signal processing advances in wireless and mobile communications, Vol. 1: trends in channel estimation and equalization*, G. B. Giannakis, Y. Hua, P. Stoica, and L. Tong, Eds., chapter 3, pp. 63–112. Prentice-Hall, 2001.

[3] K. Abed-Meraim and Y. Hua, "Blind identification of Multi-Input Multi-Output system using minimum noise subspace," *IEEE Trans. Signal Processing*, vol. 45, no. 1, pp. 254–258, Jan. 1997.

[4] T. Ma, Z. Ding, and S. F. Yau, "A two-stage algorithm for mimo blind deconvolution of colored input signals," *IEEE Trans. Signal Processing*, vol. 48, no. 4, pp. 1187–1192, 2000.

[5] Y. Hua, S. An, and X. Yong, "Blind identification and equalization of FIR MIMO channels by BIDS," in *Proc. ICASSP*, May 2001, vol. 4, pp. 2157–2160.

[6] Y. Hua, S. An, and X. Yong, "Blind identification of FIR MIMO channels by decorrelating subchannels," *IEEE Trans. Signal Processing*, To appear.

[7] L. Tong, V. C. Soon, R. Liu, and Y. Huang, "AMUSE: a new blind identification algorithm," in *Proc. ISCAS*, 1990, New Orleans, LA.

[8] A. Belouchrani, K. Abed-Meraim, J. F. Cardoso, and E. Moulines, "A blind source separation technique using second-order statistics," *IEEE Trans. Signal Processing*, vol. 45, no. 2, pp. 434–443, 1997.