Title	Blind MIMO channel estimation with an upper bound for channel orders
Author(s)	Zeng, Y; Ng, TS; Ma, S
Citation	Ieee International Conference On Communications, 2005, v. 3, p. 1996-2000
Issued Date	2005
URL	http://hdl.handle.net/10722/45833
Rights	©2005 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

BLIND MIMO CHANNEL ESTIMATION WITH AN UPPER BOUND FOR CHANNEL ORDERS

Yonghong Zeng, Tung Sang Ng and Shaodan Ma

Department of Electrical and Electronic Engineering
The University of Hong Kong, Pokfulam Road, Hong Kong

ABSTRACT

Many known second-order statistics based blind algorithms for MIMO channel estimation are sensitive to channel order overestimations. To overcome this problem, an algorithm is proposed in [1] for SIMO system only, and then a simple generalization of it to MIMO system is presented in [2]. In this paper, improvements and refinements on the algorithm in [2] are given, which makes the method robust to noise and round-off error. The method can give estimations of all channel impulse responses subject to a scalar matrix ambiguity when only an upper bound for all MIMO channel orders is known.

1. INTRODUCTION

Many methods have been proposed for channel estimation of multi-input multi-output (MIMO) systems [3, 4, 5, 1, 6, 7, 8, 9, 10, 11]. The subspace (SS) method [3, 8], the linear prediction (LP) approaches [4, 6, 7, 9, 10] and the outer product decomposition (OPD) [5] algorithm are most popular among them. The SS has a simple structure and achieves good performance for single-input multi-output (SIMO) system [8], but it requires precise knowledge of the channel order [12], which is very difficult to obtain due to noises and roundoff errors. Also, the extension of it to MIMO system is not so successful because it generally can only estimate the channels subject to a polynomial matrix ambiguity [3]. The LP and OPD can be used for both SIMO and MIMO system and are valid when the channel orders are overestimated, but their performance is sensitive to observation noise. It has been pointed in [13, 6] that LP's claimed robustness to channel order overestimation does not hold when the SOS contains estimation errors. It was shown in [4] that the LP algorithm can achieve acceptable performance when the assumed order equals that provided by an order detection criterion (the maximum description length (MDL) or the Akaike information criteria (AIC) [14]) overestimated by a few (one or two) taps only. This means that the LP is not fully robust to order overestimation. In MIMO systems, there are many different channels and these channels may have different channel orders. Therefore, known order detection methods such as the MDL and AIC which are valid for SIMO system cannot be used for MIMO system anymore. It is much more difficult to estimate channel orders in an MIMO system. Therefore, a practical channel estimation method must first tolerate channel overestimations.

In [1], a method robust to order overestimation is proposed for SIMO channel estimation (in the following the method is called GRDA). A simple generalization of the GRDA to MIMO system is presented in [2]. Although the generalized method is proved theoretically to tolerate order overestimation, simulations show that it is not very robust to noise and round-off errors. In this paper, two major improvements on the method, namely, using eigenvalue decomposition (EVD) to optimize an estimation and applying the MDL to detect the rank of correlation matrices, are given. The two improvements make the method not only tolerate order overestimation but also be robust to noise and round-off errors.

In the following, superscripts T, $\dagger$ and $*$ stand for transpose, Hermitian and conjugate, respectively. Symbol $\stackrel{\text{def}}{=}$ is used for introducing a new notation. $\mathbf{I}_q$ is the identity matrix of order q and $\otimes$ is the Kronecker product of matrices.

2. SYSTEM MODEL

Assume that there are P users with each user sending a symbol sequence: $s_j(n)$ $(j = 1, 2, \cdots, P)$ and M receivers (antennas) with received signal: $x_i(n)$ $(i = 1, 2, \cdots, M)$. A MIMO system can be described as

$$x_i(n) = \sum_{j=1}^{P} \sum_{k=0}^{N_{ij}} h_{ij}(k) s_j(n-k) + \eta_i(n), \qquad (1)$$

where $h_{ij}(k)$ is the channel response from user j to antenna i, N_{ij} is the order of channel $h_{ij}(k)$, and $\eta_i(n)$ is the channel noise. We assume that $M > P$. Letting $N_j \stackrel{\text{def}}{=} \max_i(N_{ij})$, $h_{ij}(k)$ are zero-padded if necessary, and

$$\begin{aligned}
\mathbf{x}(n) &\stackrel{\text{def}}{=} [x_1(n), x_2(n), \cdots, x_M(n)]^T, \\
\mathbf{h}_j(n) &\stackrel{\text{def}}{=} [h_{1j}(n), h_{2j}(n), \cdots, h_{Mj}(n)]^T, \\
\eta(n) &\stackrel{\text{def}}{=} [\eta_1(n), \eta_2(n), \cdots, \eta_M(n)]^T, \qquad (2)
\end{aligned}$$

we can express (1) into vector form as

$$\mathbf{x}(n) = \sum_{j=1}^{P} \sum_{k=0}^{N_j} \mathbf{h}_j(k) s_j(n-k) + \eta(n), \ n = 0, 1, \cdots. \quad (3)$$

By considering L consecutive outputs and defining

$$\hat{\mathbf{x}}(n) \stackrel{\text{def}}{=} [\mathbf{x}^T(n), \mathbf{x}^T(n-1), \cdots, \mathbf{x}^T(n-L+1)]^T, \quad (4)$$

$$\hat{\eta}(n) \stackrel{\text{def}}{=} [\eta^T(n), \eta^T(n-1), \cdots, \eta^T(n-L+1)]^T, \quad (5)$$

$$\hat{\mathbf{s}}(n) \stackrel{\text{def}}{=} [s_1(n), s_1(n-1), \cdots, s_1(n-N_1-L+1), \cdots,$$
$$s_P(n), s_P(n-1), \cdots, s_P(n-N_P-L+1)]^T, \quad (6)$$

we get

$$\hat{\mathbf{x}}(n) = \mathbf{H}\hat{\mathbf{s}}(n) + \hat{\eta}(n), \quad (7)$$

where $\mathbf{H}$ is a $ML \times (N+PL)$ $(N \stackrel{\text{def}}{=} \sum_{j=1}^{P} N_j)$ matrix defined as

$$\mathbf{H} \stackrel{\text{def}}{=} [\mathbf{H}_1, \mathbf{H}_2, \cdots, \mathbf{H}_P],$$
$$\mathbf{H}_j \stackrel{\text{def}}{=}$$
$$\begin{bmatrix} \mathbf{h}_j(0) & \cdots & \cdots & \mathbf{h}_j(N_j) & \cdots & 0 \\ & \ddots & & & \ddots & \\ 0 & \cdots & \mathbf{h}_j(0) & \cdots & \cdots & \mathbf{h}_j(N_j) \end{bmatrix}.$$
$$(8)$$

3. THE ALGORITHM

The following assumptions for the statistical properties of transmitted symbols and channel noise are assumed.

(A1) Transmitted symbols are independent and identically distributed.

(A2) Noises are white and uncorrelated.

(A3) Noises and transmitted signals are uncorrelated.

(A4) A smoothing factor L has been chosen such that the matrix $\mathbf{H}$ is of full column rank.

3.1. Derivation of the algorithm

Based on the assumptions (A1-3), we can verify that

$$\mathbf{R} \stackrel{\text{def}}{=} \text{E}(\hat{\mathbf{x}}(n)\hat{\mathbf{x}}^\dagger(n)) = \sigma_s^2 \mathbf{H}\mathbf{H}^\dagger + \sigma_\eta^2 \mathbf{I}_{ML}, \quad (9)$$

$$\bar{\mathbf{R}} \stackrel{\text{def}}{=} \mathbf{R} - \sigma_\eta^2 \mathbf{I}_{ML} = \sigma_s^2 \mathbf{H}\mathbf{H}^\dagger, \quad (10)$$

$$\mathbf{Q} \stackrel{\text{def}}{=} \text{E}(\hat{\mathbf{x}}(n)\hat{\mathbf{x}}^\dagger(n-1))$$
$$= \sigma_s^2 \mathbf{H}\tilde{\mathbf{J}}_{N+PL}\mathbf{H}^\dagger + \sigma_\eta^2 (\mathbf{J}_L \otimes \mathbf{I}_M), \quad (11)$$

$$\bar{\mathbf{Q}} \stackrel{\text{def}}{=} \mathbf{Q} - \sigma_\eta^2 (\mathbf{J}_L \otimes \mathbf{I}_M) = \sigma_s^2 \mathbf{H}\tilde{\mathbf{J}}_{N+PL}\mathbf{H}^\dagger, \quad (12)$$

where σ_s^2 and σ_η^2 are the variances of the transmitted symbols and noises respectively, $\mathbf{J}_q$ is a $q \times q$ down shift matrix, that is, it is a lower triangular Toeplitz matrix with its first column being $(0, 1, 0, \cdots, 0)^T$, and $\tilde{\mathbf{J}}_{N+PL} = \text{diag}(\mathbf{J}_{N_1+L}, \mathbf{J}_{N_2+L}, \cdots, \mathbf{J}_{N_P+L})$. It is easy to verify that the rank of $\bar{\mathbf{R}}$ and $\bar{\mathbf{Q}}$ are respectively $N+PL$ and $N+P(L-1)$. Let $K = ML-(N+PL)+P$, which is the dimension of the left null space of $\bar{\mathbf{Q}}$. Let $\mathbf{u}_i$ $(i = 1, 2, \cdots, K)$ be K linear independent vectors (a basis) in the left null space of $\bar{\mathbf{Q}}$, that is,

$$\mathbf{u}_i^\dagger \bar{\mathbf{Q}} = \mathbf{u}_i^\dagger \mathbf{H}\tilde{\mathbf{J}}_{N+PL}\mathbf{H}^\dagger = 0. \quad (13)$$

It is proved in [2] that

$$\mathbf{u}_i^\dagger \mathbf{H}\hat{\mathbf{s}}(n) = \alpha_i^\dagger \mathbf{s}(n), \quad (14)$$

where $\alpha_i \stackrel{\text{def}}{=} (\alpha_{i1}, \cdots, \alpha_{iP})^T$ with α_{ij} being an arbitrary complex number, and $\mathbf{s}(n) \stackrel{\text{def}}{=} (s_1(n), \cdots, s_P(n))^T$. Use these K vectors $\mathbf{u}_i$ $(i = 1, 2, \cdots, K)$ to define a matrix $\mathbf{U}$ as

$$\mathbf{U} \stackrel{\text{def}}{=} [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_K]. \quad (15)$$

We also define a $P \times K$ matrix α by

$$\alpha \stackrel{\text{def}}{=} \begin{bmatrix} \alpha_{11} & \alpha_{21} & \cdots & \alpha_{K1} \\ \alpha_{12} & \alpha_{22} & \cdots & \alpha_{K2} \\ & \cdots & \cdots & \\ \alpha_{1P} & \alpha_{2P} & \cdots & \alpha_{KP} \end{bmatrix}. \quad (16)$$

Then we have

$$\mathbf{U}^\dagger \mathbf{H}\hat{\mathbf{s}}(n) = \alpha^\dagger \mathbf{s}(n). \quad (17)$$

From the definitions of the matrices, we have

$$\text{E}(\alpha^\dagger \mathbf{s}(n)\mathbf{s}^\dagger(n)\alpha) = \sigma_s^2 \alpha^\dagger \alpha$$
$$= \text{E}(\mathbf{U}^\dagger \mathbf{H}\hat{\mathbf{s}}(n)\hat{\mathbf{s}}^\dagger(n)\mathbf{H}^\dagger \mathbf{U})$$
$$= \mathbf{U}^\dagger \bar{\mathbf{R}}\mathbf{U}. \quad (18)$$

Note that $K = ML-(N+PL)+P > P$ and the matrix α has more columns than rows. In the following, it is assumed that α is of full row rank. Let $\mathbf{V}_1$ be a matrix of size $K \times P$ whose columns are eigenvectors corresponding to nonzero eigenvalues of $\mathbf{U}^\dagger \bar{\mathbf{R}}\mathbf{U}$. Then, it is easy to verify that there exists an invertible matrix $\mathbf{B}$ of size $P \times P$ such that $\alpha^\dagger = \mathbf{V}_1 \mathbf{B}^\dagger$. Therefore,

$$\mathbf{U}_1^\dagger \mathbf{H}\hat{\mathbf{s}}(n) = \mathbf{B}^\dagger \mathbf{s}(n). \quad (19)$$

where $\mathbf{U}_1 \stackrel{\text{def}}{=} \mathbf{U}\mathbf{V}_1$. Then we get a critical formula as follows:

$$\mathbf{s}^\dagger(n) = \hat{\mathbf{s}}^\dagger(n)\mathbf{H}^\dagger \mathbf{U}_1 \mathbf{B}^{-1} = \hat{\mathbf{z}}^\dagger(n)\mathbf{U}_1 \mathbf{B}^{-1}, \quad (20)$$

where $\hat{\mathbf{z}}(n) \stackrel{\text{def}}{=} \mathbf{H}\hat{\mathbf{s}}(n) = \hat{\mathbf{x}}(n) - \hat{\eta}(n)$. We also define $\mathbf{z}(n) \stackrel{\text{def}}{=} \mathbf{x}(n) - \eta(n)$.

Now we are ready to estimate the channels. From (3) we see that $\mathbf{h}_j(k) = \text{E}(\mathbf{x}(n)s_j^*(n-k))/\sigma_s^2$. Let

$$\mathbf{h}(k) \stackrel{\text{def}}{=} [\mathbf{h}_1(k), \mathbf{h}_2(k), \cdots, \mathbf{h}_P(k)] \quad (21)$$

be the MIMO channel matrix. Similar to [2], it can be verified that

$$\mathbf{h}(k)\sigma_s^2 = \text{E}(\mathbf{z}(n)\hat{\mathbf{z}}^\dagger(n-k))\mathbf{U}_1 \mathbf{B}^{-1}. \quad (22)$$

By defining an $M \times M$ matrix $\mathbf{r}_z(k)$ as

$$\mathbf{r}_z(k) \stackrel{\text{def}}{=} \text{E}(\mathbf{z}(n)\mathbf{z}^\dagger(n-k)), \quad (23)$$

which can be computed from $\mathbf{r}_x(k) \stackrel{\text{def}}{=} \text{E}(\mathbf{x}(n)\mathbf{x}^\dagger(n-k))$ by the formula

$$\mathbf{r}_z(k) = \mathbf{r}_x(k) - \mathbf{r}_\eta(k) = \begin{cases} \mathbf{r}_x(k), & k \neq 0 \\ \mathbf{r}_x(k) - \sigma_\eta^2 \mathbf{I}_M, & k = 0 \end{cases},$$

we finally get

$$\mathbf{h}(k) = [\mathbf{r}_z(k), \mathbf{r}_z(k+1), \cdots, \mathbf{r}_{\mathbf{z}}(k+L-1)]\mathbf{U}_1 \mathbf{B}^{-1}\sigma_s^{-2}. \quad (24)$$

In (24), only matrix $\mathbf{B}^{-1}$ cannot be determined from the statistics of outputs of the MIMO system. That is, the MIMO channels are estimated from its output second order statistics (SOS) with ambiguity of a $P \times P$ matrix. This ambiguity matrix is intrinsic for any SOS-based blind channel estimation method.

The right null space of $\bar{\mathbf{Q}}$ is also used by the GRDA method. However, it is difficult to use the right null space for MIMO system identification because the characterization of it needs the unknown channel orders, and, unlike a SIMO system, a MIMO system usually has many different channel orders.

3.2. Practical algorithm

We assume that an upper bound for all the channel orders, that is, a number N_{upp} such that $N_j \leqslant N_{upp}$ ($j = 1, 2, \cdots, P$), is known or estimated. Choose a smoothing factor L such that $ML > PN_{upp} + PL$. The number K, that is, the dimension of the left null space of $\bar{\mathbf{Q}}$, is unknown because the sum of all the channel orders, N, is unknown. To overcome this difficulty, we use the MDL [14] to estimate it. The practical algorithm is summarized in the following.

Algorithm Blind MIMO Channel Estimation
Step 1. Compute $\mathbf{R} = \frac{1}{L_s} \sum_{n=L}^{L+L_s-1} \hat{\mathbf{x}}(n)\hat{\mathbf{x}}^\dagger(n)$, where L_s is the number of samples used. Compute the EVD of $\mathbf{R}$. Use the MDL to estimate the rank of $\mathbf{H}$. Let the estimated rank be r. Average the smallest $ML - r$ eigenvalues of $\mathbf{R}$ to get an estimation ($\bar{\sigma}_\eta^2$) for the noise variance. Let $\bar{\mathbf{R}} = \mathbf{R} - \bar{\sigma}_\eta^2 \mathbf{I}_{ML}$.
Step 2. Compute $\mathbf{Q} = \frac{1}{L_s} \sum_{n=L}^{L+L_s-1} \hat{\mathbf{x}}(n)\hat{\mathbf{x}}^\dagger(n-1)$ and $\bar{\mathbf{Q}} = \mathbf{Q} - \bar{\sigma}_\eta^2 (\mathbf{J}_L \otimes \mathbf{I}_M)$. Then compute the singular value decomposition (SVD) of $\bar{\mathbf{Q}}$. Let $K_1 = ML - r + P$. Choose K_1 vectors $\mathbf{u}_i$ ($i = 1, 2, \cdots, K_1$) corresponding to the K_1 smallest left singular values of $\bar{\mathbf{Q}}$ and denote a matrix $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{K_1}]$.
Step 3. Compute $\mathbf{W} = \mathbf{U}^\dagger \bar{\mathbf{R}} \mathbf{U}$ and the EVD of $\mathbf{W}$. Let $\mathbf{V}_1$ be the matrix of size $K_1 \times P$ whose columns are eigenvectors corresponding to nonzero eigenvalues of $\mathbf{W}$, and $\mathbf{U}_1 = \mathbf{U}\mathbf{V}_1$.
Step 4. For $k = 0, 1, \cdots, N_{upp} + L - 1$, compute

$$\mathbf{r}_z(k) = \mathbf{r}_x(k) - \mathbf{r}_\eta(k) = \begin{cases} \mathbf{r}_x(k), & k \neq 0 \\ \mathbf{r}_x(k) - \bar{\sigma}_\eta^2 \mathbf{I}_M, & k = 0 \end{cases},$$

where $\mathbf{r}_x(k)$ is estimated by output samples.
Step 5. Form a matrix

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}^T(0) & \mathbf{G}^T(1) & \cdots & \mathbf{G}^T(N_{upp}) \end{bmatrix}^T =$$

$$\begin{bmatrix} \mathbf{r}_z(0) & \mathbf{r}_z(1) & \cdots & \mathbf{r}_z(L-1)) \\ \mathbf{r}_z(1) & \mathbf{r}_z(2) & \cdots & \mathbf{r}_z(L)) \\ & \cdots & \cdots & \\ \mathbf{r}_z(N_{upp}) & \mathbf{r}_z(N_{upp}+1) & \cdots & \mathbf{r}_z(N_{upp}+L-1) \end{bmatrix} \mathbf{U}_1.$$

The MIMO channel matrix is then $\mathbf{h}(k) = \mathbf{G}(k)\beta$, where β is a $P \times P$ matrix to be determined and

$$\mathbf{h}(k) = \begin{bmatrix} h_{11}(k) & \cdots & h_{1P}(k) \\ \vdots & \cdots & \vdots \\ h_{M1}(k) & \cdots & h_{MP}(k) \end{bmatrix}.$$

4. DISCUSSIONS ON THE METHOD

There are two major improvements over the method in [2], namely, applying the MDL to detect the rank of correlation matrices (Step 1) and using EVD to optimize an estimation for $\mathbf{s}(n)$ (Step 3). The two improvements make the method not only tolerate order overestimation but also be robust to noise and round-off errors.

The tolerance to noise and round-off errors of the algorithm is closely related to the property of matrix α. Since $K_1 = ML - r + P > P$ (in MDL, $r < ML$), α has more columns than rows (this is different from [2] where α has only P columns). Simulations show that larger K_1 (if $K_1 \leqslant K$) gives better performance to α (larger Pth singular value). This justifies using the MDL to estimate K because it gives a relatively precise estimation. Alternatively, we can use $(M - P)L - PN_{upp} + P$ to approximate

K (that is, replace K_1 by this number in the algorithm), which is a direct generalization of that used in GRDA. If the channel orders are overestimated, this estimation may be much smaller than the actual value and therefore the algorithm's performance will be bad. For a SIMO system, the estimation of K is equivalent to the estimation of the channel order. However, for a MIMO system, they are different because a MIMO system has many different channel orders and therefore the orders cannot be obtained from K.

For a SIMO system, the algorithm is the GRDA with an amendment. The amendment is that the rank of $\mathbf{H}$ is estimated from the eigenvalues of $\mathbf{R}$ by the MDL. For simplicity, in the following, the method is called modified GRDA (MGRDA). Using the EVD to optimize an estimation for $\mathbf{s}(n)$ (see Step 3) is equivalent to the EPC criteria in GRDA. The computational complexities of the two algorithms are nearly the same, because the estimation of the rank of $\mathbf{H}$ by the MDL is very simple if the eigenvalues of $\mathbf{R}$ are known. The eigenvalues of $\mathbf{R}$ are also required in GRDA for estimating the noise variance. Since the MDL usually gives a more precise estimation of K than the direct estimation in GRDA, the amendment usually gives rise to better performances both in accuracy and robustness to order overestimation, which is supported by simulations.

5. SIMULATIONS

In the following, signal-noise-ratio (SNR) means the ratio of the average received signal power to the average noise power SNR $\stackrel{\text{def}}{=} \frac{\mathrm{E}(||\mathbf{x}(n)-\eta(n)||^2)}{\mathrm{E}(||\eta(n)||^2)}$. The mean square error (MSE) between the estimated and true channel responses is defined as MSE $\stackrel{\text{def}}{=} \min_\beta \frac{\sum_{l=0}^{N_{upp}} ||\mathbf{h}(l)-\bar{\mathbf{h}}(l)\beta||_F^2}{\sum_{l=0}^{N_{upp}} ||\mathbf{h}(l)||_F^2}$, where $\bar{\mathbf{h}}(l)$ and $\mathbf{h}(l)$ are the estimated and true channel responses respectively, while the true channel responses are zero-padded, β is a matrix with size $P \times P$, and $|| \cdot ||_F$ means the Frobenius norm for matrix. It is easy to prove that the optimal $\beta = \left(\sum_{l=0}^{N_{upp}} \bar{\mathbf{h}}^\dagger(l)\bar{\mathbf{h}}(l) \right)^{-1} \left(\sum_{l=0}^{N_{upp}} \bar{\mathbf{h}}^\dagger(l)\mathbf{h}(l) \right)$.

Let $N_{max} \stackrel{\text{def}}{=} \max_j N_j$. Only an upper bound for all the orders is assumed known in our simulations, that is, a number N_{upp} is known such that $N_{upp} \geq N_{max}$. To verify the robustness to channel order overestimation, we test four cases of the channel order upper bound (N_{upp}): N_{max} (exact upper bound), $N_{max} + 2$, $N_{max} + 4$ and $N_{max} + 6$, respectively. In the following, a Monte Carlo realization means: ① a random Rayleigh fading channel (tap coefficients are random complex numbers with Gaussian distribution) is created; ② a data packet with 1000 random 4-QAM input samples is transmitted through the wireless channel with random Gaussian noises. Extensive simulations show that the algorithm is truly robust to channel order overestimation, noise and round-off errors.

5.1. MIMO channel estimation

Consider a 2-user 4-antenna system ($M = 4$, $P = 2$) with $N_1 = 2$ and $N_2 = 4$. The smoothing factor is chosen as $L = \lfloor PN_{upp}/(M-P) \rfloor + 1$. The MSE versus SNR is shown in Figure 1, where 500 samples are used for estimating the statistical correlation matrices. In Figure 2, the SNR is fixed to 20dB, while the sample size used is changed from 300 to 900. All the results are averaged over 100 Monte Carlo realizations. From the figures, we see that the algorithm works well when only an upper bound for the channel orders is known. Furthermore, the assumed upper bound
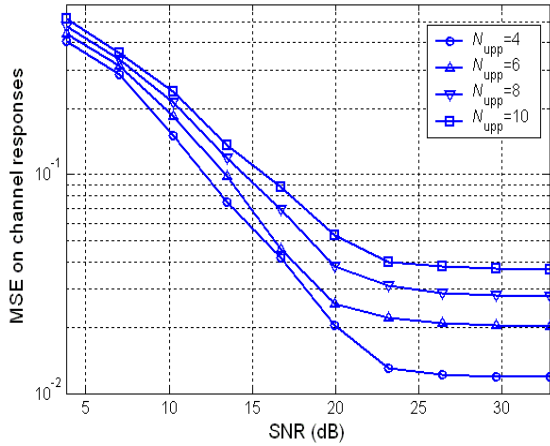
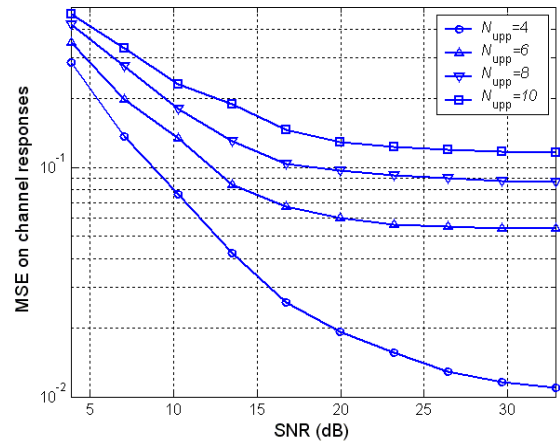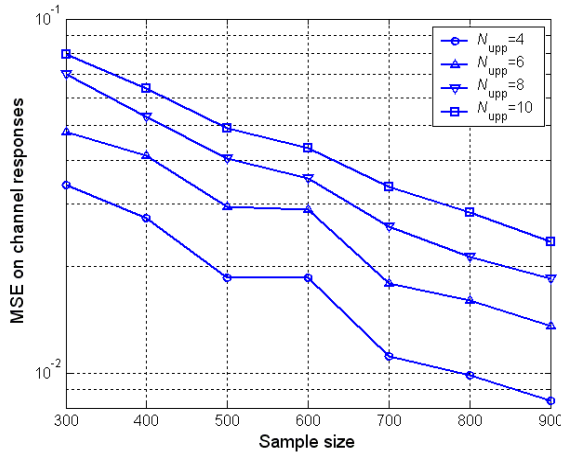Figure 1: MSE versus SNR (500 samples)

Figure 2: MSE versus sample size (SNR=20dB)

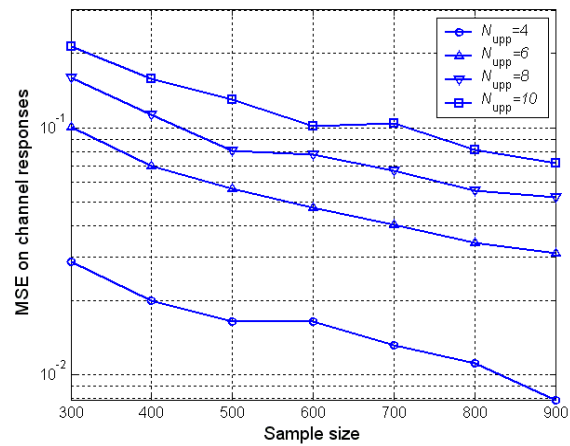Figure 3: MSE versus SNR (method in [2], 500 samples)

Figure 4: MSE versus sample size (method in [2], SNR=20dB)

can be much larger than the exact upper bound. When the channel orders are overestimated, there are additional channel coefficients at the tails to be estimated. Errors are inevitably introduced to these coefficients (ideally should be zeros), which causes the MSE to become higher. Other simulations show that the smoothing factor does not affect the MSE significantly.

For comparison, simulation results are given in Figure 3 to 4 when the method in [2] is used, where the parameters are the same as those in Figure 1 to 2, respectively, except that the smoothing factor is chosen as $L = \lfloor PN_{\text{upp}}/(M - P) \rfloor + 7$ to obtain the best performance. Obviously, the results are much worse. Here the smoothing factor do have significant impact on the performance.

5.2. SIMO channel estimation

The MGRDA and GRDA (by EPC criteria) are tested on the same conditions. A single-user 4-antenna system ($M = 4$, $P = 1$) is considered. First, the channel in [1] (p. 1455) is used for the simulation. The smoothing factor is chosen as $L = N_{upp} + 2$, which is recommended by the GRDA. We test the cases for N_{upp} being 4 (true order), 6, 8 and 10, respectively. The MSE versus SNR is shown in Figure
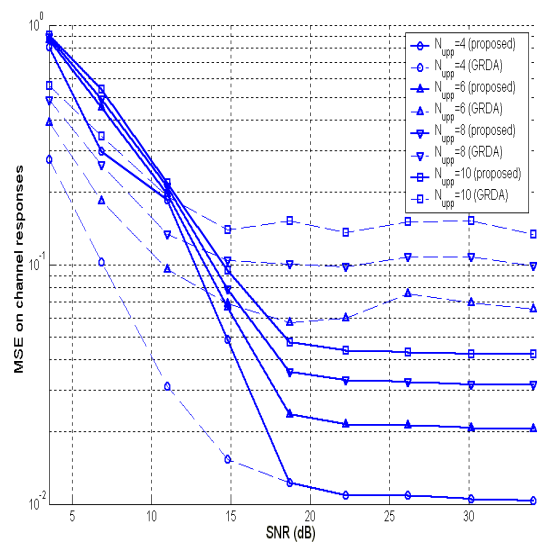

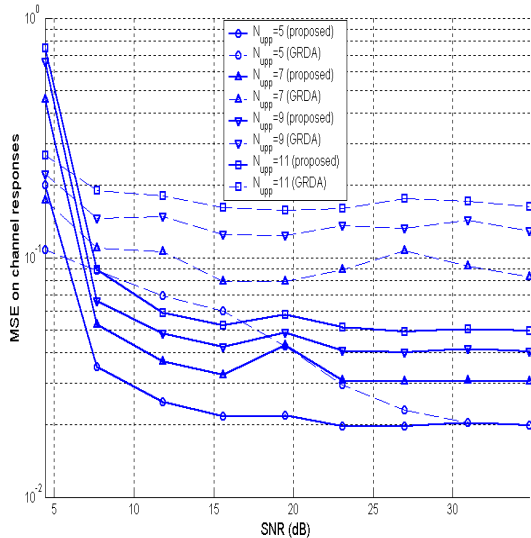Figure 5: MSE versus SNR (200 samples)

Figure 6: MSE versus SNR (bad channel, 200 samples)

5, where 200 samples are used. Secondly, the channel is constructed by oversampling a three-ray, long delay (delays at 0, 0.5, and 3 baud periods) multipath channel. The channel, with order 5, is given in [1] (p. 1456). This channel is very ill-conditioned because the least nonzero singular value of $\bar{\mathbf{R}}$ is very small [1]. The smoothing factor is again chosen as $L = N_{upp} + 2$. We test the cases for N_{upp} being 5 (true order), 7, 9 and 11, respectively. Figure 6 shows the MSE versus SNR, where 200 samples are used.

From the simulation results, it is obvious that in most cases the MGRDA achieves better performance in both accuracy and robustness to channel order overestimation than the GRDA. For low SNR, the MGRDA is sometimes worse than the GRDA because the MDL usually gives an overestimation of K. It is seen from the figures that even if N_{upp} equals to the exact channel order, the two methods still achieve different performances. The reason is that the MGRDA does not assume that the exact order is known and still uses the MDL to estimate the dimension of the null space, while the GRDA acknowledges the information and therefore obtains the true dimension of the null space. For a good conditioned channel, knowing the true dimension can obtain the best performance (Figure 5), but for a bad conditioned channel, it may not (Figure 6, the reason is that the smallest nonzero eigenvalue is very small and an overestimation of K is justified). Furthermore, the GRDA method recommends the smooth factor L larger than the estimated channel order (preferably $N_{upp}+2$). In MGRDA, L can be much smaller, usually $L = \lfloor N_{upp}/3 \rfloor + 1$. Noticing that the size of the matrix $\mathbf{R}$ is $ML \times ML$, we know that the computational complexity increase rapidly with larger L.

6. CONCLUSIONS

It has been shown that the proposed SOS-based blind method can accurately estimate the channels subject to a scalar matrix ambiguity when only an upper bound for all the MIMO channel orders is known, that is, it tolerates channel order overestimation. Furthermore, simulations have shown that the method is robust to noise and round-off errors.

7. REFERENCES

[1] H. Gazzah, P. A. Regalia, J. P. Delmas, and K. Abed-Meraim, "A blind multichannel identification algorithm robust to order overestimation," IEEE Trans. Signal Processing, vol. 50, no. 6, pp. 1449–1458, 2002.

[2] Y. H. Zeng and T. S. Ng, "A blind MIMO channel estimation method robust to order overestimation," Signal Processing, vol. 84, no. 2, pp. 435–439, 2004.

[3] K. Abed-Meraim, P. Loubaton, and E. Moulines, "A subspace algorithm for certain blind identification problem," IEEE Trans. Information Theory, vol. 32, no. 2, pp. 499–511, 1997.

[4] K. Abed-Meraim, P. Loubaton, and E. Moulines, "Prediction error method for second-order blind identification," IEEE Trans. Signal Processing, vol. 45, no. 3, pp. 694–705, 1997.

[5] Z. Ding, "Matrix outer product decomposition method for blind multiple channel identification," IEEE Trans. Signal Processing, vol. 45, no. 12, pp. 3053–3061, 1997.

[6] A. Gorokhov and P. Loubaton, "Blind identification of MIMO-FIR systems: a generalized prediction approach," Signal Processing, vol. 73, pp. 105–124, 1999.

[7] R. Lopez-Valcarce and S. Dasgupta, "Blind channel equalization with colored sources based on second-order statistics: A linear prediction approach," IEEE Trans. Signal processing, vol. 49, no. 9, pp. 2050–2059, 2001.

[8] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," IEEE Trans. Signal Processing, vol. 43, no. 2, pp. 516–525, 1995.

[9] D. Slock, "Blind fractionally-spaced equalization, perfect reconstruction filterbanks, and multilinear prediction," in Proceedings of IEEE ICASSP Conf., (Adalaide, Australia), 1994.

[10] J. K. Tugnait, "On linear predictors for MIMO channels and related blind identification and equalization," IEEE Signal processing Letters, vol. 5, no. 11, pp. 289–291, 1998.

[11] Y. H. Zeng and T. S. Ng, "A semi-blind channel estimation method for multi-user multi-antenna OFDM systems," IEEE Trans. Signal Processing, vol. 52, no. 5, pp. 1419–1429, 2004.

[12] W. Qiu and Y. Hua, "Performance analysis of the subspace method for blind channel identification," Signal Processing, vol. 50, pp. 71–81, 1996.

[13] A. P. Liavas, P. A. Regalia, and J. P. Delmas, "On the robustness of the linear prediction method for blind channel identification with respect to effective channel undermodeling/overmodeling," IEEE Trans. Signal Processing, vol. 48, no. 5, pp. 1477–1481, 2000.

[14] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," IEEE Trans. Acoust., Speech, Signal Processing, vol. 32, pp. 387–392, Apr. 1985.