A CHANNEL ORDER INDEPENDENT METHOD FOR BLIND EQUALIZATION OF MIMO SYSTEMS

João Xavier and Victor Barroso

Instituto Superior Técnico – Instituto de Sistemas e Robótica
Av. Rovisco Pais, 1096 Lisboa Codex, Portugal
{jxavier,vab}@isr.ist.utl.pt

ABSTRACT

We study blind equalization of noisy MIMO-FIR systems driven by white sources. We present a new second order statistics (SOS) based approach which does not require the knowledge of the channel order. This technique blindly transforms a convolutive mixture of users into an instantaneous one. Thus, in the special case of a single user (SIMO systems), an estimate of the input signal is readily obtained. Computer simulations results illustrate the promising performance of the proposed technique. We compare our method with the multistep prediction (MSP) approach (in the context of SIMO systems), and evaluate the algorithm capability in globally nulling the intersymbol interference (ISI) for MIMO systems.

1. INTRODUCTION

The paper by Tong *et al.* [1] was a major breakthrough in the area of blind channel equalization/identification of SIMO systems. It showed the feasibility of blind equalization/identification schemes based only on second order statistics (SOS); see also [2, 10, 11]. The performance of these subspace-based approaches tends to drop down significantly when the channel order is wrongly estimated. The methods in [3, 4], based on linear prediction theory, are formulated for the single-user case and offer robustness to overdetermination of the channel order. In this paper, we present a SOS based blind equalization technique for the general case of multiple users (MIMO systems), which is robust to either under or over evaluations of the channel order. The paper is organized as follows. Section 2 introduces the data model and states the main assumptions. Section 3 describes the proposed blind equalization technique. Section 4 shows simulation results assessing the performance of our method for SIMO and MIMO systems. Section 5 presents the conclusions of this work.

Notation. Matrices (capital) and vectors are in boldface type. $\mathbb{C}^{n \times m}$ is the set of $m \times n$ matrices with complex entries. $\mathcal{R}(\mathbf{A})$ and $\mathcal{N}(\mathbf{A})$ denote, respectively, the range and null-space of matrix $\mathbf{A}$. The notations $(\cdot)^T$ and $(\cdot)^*$ stand for the transpose and Hermitian operator, respectively. The symbols $\mathbf{I}_n$ and $\mathbf{0}_{n \times m}$ stand for the $n \times n$ identity and the $n \times m$ all-zero matrices, respectively. $\|\cdot\|$ denotes the Schatten 2-norm and $\delta(t)$ is the Kronecker delta. $\mathcal{G}_n$ and $\mathcal{U}_n$ denote the general linear group and the unitary group of $\mathbb{C}^{n \times n}$, respectively.

2. SYSTEM MODEL

Consider an M_0th order P-input/L-output noisy linear FIR-MIMO system, described by the equation

$$\mathbf{x}(t) = \sum_{m=0}^{M_0} \mathbf{H}(m)\mathbf{s}(t - m) + \mathbf{b}(t). \qquad (1)$$

Here, $\mathbf{x}(t) \in \mathbb{C}^L$ is the vector of system outputs, $\mathbf{b}(t)$ represents additive noise, and $\mathbf{s}(t) \equiv [\, s_1(t)\ s_2(t)\ \cdots\ s_P(t) \,]^T$ contains the P input scalar signals; $\mathbf{H}(m) \in \mathbb{C}^{L \times P}$, $m = 0, 1, \ldots, M_0$ denote the matricial filter coefficients. By stacking $N \geq M_0$ successive samples according to

$$\mathbf{x}_N(t) \equiv \left[\, \mathbf{x}(t)^T\ \mathbf{x}(t-1)^T\ \cdots\ \mathbf{x}(t-N+1)^T \,\right]^T$$
$$\mathbf{b}_N(t) \equiv \left[\, \mathbf{b}(t)^T\ \mathbf{b}(t-1)^T\ \cdots\ \mathbf{b}(t-N+1)^T \,\right]^T,$$

we have the equivalent model

$$\mathbf{x}_N(t) = \mathbf{H}_N \mathbf{s}_M(t) + \mathbf{b}_N(t); \qquad (2)$$

$\mathbf{s}_M(t) \equiv \left[\, \mathbf{s}(t)^T\ \mathbf{s}(t-1)^T\ \cdots\ \mathbf{s}(t-M+1)^T \,\right]^T$, $(M \equiv M_0 + N)$ and

$$\mathbf{H}_N \equiv \begin{bmatrix} \mathbf{H}(0) & \cdots & \mathbf{H}(M_0) & \mathbf{0}_{L \times P} & \cdots & \mathbf{0}_{L \times P} \\ \mathbf{0}_{L \times P} & \mathbf{H}(0) & \cdots & \mathbf{H}(M_0) & \cdots & \mathbf{0}_{L \times P} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0}_{L \times P} & \cdots & \mathbf{0}_{L \times P} & \mathbf{H}(0) & \cdots & \mathbf{H}(M_0) \end{bmatrix}.$$

Equation (2) can also be rewritten as

$$\mathbf{x}_N(t) = \sum_{m=0}^{M} \mathbf{H}_N(m)\mathbf{s}(t - m) + \mathbf{b}_N(t), \qquad (3)$$

where $\mathbf{H}_N(m) \in \mathbb{C}^{NL \times P}$ is the submatrix of the generalized Sylvester matrix $\mathbf{H}_N$ obtained by retaining the columns $mP + 1$ to $(m + 1)P$. We assume that: **(A1)** the number P of users is known and that $\mathbf{H}_N$ is full column-rank; **(A2)** $s_p(t)$ is a zero-mean white unit-variance sequence and the P users are uncorrelated with each other, i.e., $\mathbf{R}_\mathbf{s}(\tau) = \mathrm{E}\{\mathbf{s}(t)\mathbf{s}(t - \tau)^*\} = \mathbf{I}_P \delta(\tau)$; **(A3)** the noise process $\mathbf{b}(t)$ is zero-mean, wide sense stationary, and uncorrelated with $\mathbf{s}(t)$, with known autocorrelation matrices $\mathbf{R}_\mathbf{b}(\tau)$.

3. BLIND EQUALIZATION ALGORITHM

Overview. We describe our blind equalization technique. This is theoretically insensitive to under/over evaluations of the system order M_0 or, equivalently, of rank$(\mathbf{H}_N)$. Consider the data model in (3) and let $\widehat{M}$ be an arbitrary estimate of M (obtained by the MDL or AIC criteria [5]). In our approach, we blindly compute $\widehat{M} + 1$ minimum variance distortionless response (MVDR) beamformers, $\mathbf{G}_m \in \mathbb{C}^{LN \times P}$, $m = 0, 1, \ldots, \widehat{M} + 1$; the mth beamformer aims at extracting the signal $\mathbf{s}(t - m)$ from the observations $\mathbf{x}_N(t)$. For the computation of each beamformer, the knowledge of the channel order is not needed. The penalty for order under-estimation ($\widehat{M} < M$) is that the last $M - \widehat{M}$ delayed replicas of $\mathbf{s(t)}$ in (3) are not retrieved. This is sub-optimal since we may be discarding the best replica, in terms of signal-to-noise ratio (SNR). The penalty for order over-estimation ($\widehat{M} > M$) is the unnecessary extra computation; in fact, the last $\widehat{M} - M$ beamformers are meaningless. However, in both cases, $\min(M, \widehat{M}) \geq 1$ valid instantaneous mixtures of the P users are obtained for blind source separation (BSS).

Algorithm. Our technique relies only on the SOS of $\mathbf{x}_N(t)$ and exploits the nested range property of the sucessive denoised correlation matrices $\mathbf{R}(\tau) \equiv \mathbf{R}_{\mathbf{x}_N}(\tau) - \mathbf{R}_{\mathbf{b}_N}(\tau)$, $\tau \geq 0$. We obtain the $\widehat{M} + 1$ beamformers in sequence. We assume that we are at the mth stage ($m \leq M$), which aims at computing the MVDR beamformer

$$\mathbf{G}_m = \underset{\mathbf{G}^* \mathbf{H}_N(m) = \mathbf{I}_P}{\arg \min} \, \mathrm{E}\left\{\|\mathbf{G}^* \mathbf{x}_N(t)\|^2\right\}.$$

This is given by

$$\mathbf{G}_m = \mathbf{R}_{\mathbf{x}_N}^{-1}(0)\,\mathbf{H}_N(m)\left[\mathbf{H}_N(m)^* \mathbf{R}_{\mathbf{x}_N}^{-1}(0)\,\mathbf{H}_N(m)\right]^{-1}. \quad (4)$$

We describe the steps involved in the computation of $\mathbf{G}_m$.

Step 1: Determination of $\mathbf{B}_m$. At the mth stage, we have available the matrices $\mathbf{T}_i = \mathbf{H}_N(i)\,\mathbf{Q}_i$, for $i = 0, 1, \ldots, m - 1$, where $\mathbf{Q}_i \in \mathcal{U}_P$ (see step 4 below for how this stage generates $\mathbf{T}_m$). Also, from **(A2)** it follows that the range of $\mathbf{R}(m)$ is given by

$$\mathcal{R}(\mathbf{R}(m)) = \text{span}\{\mathbf{H}_N(m), \mathbf{H}_N(m+1), \ldots, \mathbf{H}_N(M)\}. \quad (5)$$

This is the nested range property we mentioned above. We define the matrix

$$\mathbf{B}_m \equiv \sum_{i=0}^{m-1} \mathbf{T}_i \mathbf{T}_i^* + \mathbf{R}(m+1)\,\mathbf{R}^*(m+1).$$

It is readily seen that $\mathcal{R}(\mathbf{B}_m)$ equals

$$\text{span}\{\mathbf{H}_N(0), \ldots, \mathbf{H}_N(m-1), \mathbf{H}_N(m+1), \ldots, \mathbf{H}_N(M)\}. \quad (6)$$

Notice that only $\mathbf{H}_N(m)$ is missing in (6).

Step 2: Determination of $\mathbf{W}_m$. We compute a matrix $\mathbf{W}_m \in \mathbb{C}^{LN \times P}$, which satisfies the *zero-forcing* equation

$$\mathbf{W}_m^* \mathbf{H}_N(i) = \mathbf{0}_{P \times P}, \quad (7)$$

for $i \neq m$. For this, we exploit the following theorem.

Theorem 1 *Let $\mathbf{A}, \mathbf{B} \in \mathcal{C}^{n \times n}$ be positive semidefinite Hermitian matrices. Let $\mathcal{R}(\mathbf{B}) \subset \mathcal{R}(\mathbf{A})$, and $\dim \mathcal{R}(\mathbf{A}) = \dim \mathcal{R}(\mathbf{B}) + l$, where l denotes a positive integer. For $\epsilon > 0$, define $\mathbf{C}_\epsilon = \epsilon(\mathbf{B} + \epsilon \mathbf{I}_n)^{-1}\mathbf{A}$. Then,*

$$\mathbf{C} \equiv \lim_{\epsilon \to 0} \mathbf{C}_\epsilon = \mathbf{U}\mathbf{V}^*,$$

where $\mathbf{U}, \mathbf{V} \in \mathcal{C}^{n \times l}$ are full column-rank matrices, and $\mathbf{U}$ spans $\mathcal{R}(\mathbf{A}) \overset{\perp}{\ominus} \mathcal{R}(\mathbf{B})$, where $\overset{\perp}{\ominus}$ denotes the relative orthogonal complement.
Proof: See Appendix A.

We use the theorem with $\mathbf{A} = \mathbf{R}(0)$ and $\mathbf{B} = \mathbf{B}_m$; $\mathbf{C}$ is evaluated in practice as $\mathbf{C} \simeq \mathbf{C}_\epsilon$, with $\epsilon \ll 1$. In our case, rank$(\mathbf{C}) = P$ (recall equations (5) and (6)). Thus, through a singular value decomposition (SVD) of $\mathbf{C}$, we obtain an isometry $\mathbf{W}_m \in \mathbb{C}^{NL \times P}$ ($\mathbf{W}_m^* \mathbf{W}_m = \mathbf{I}_P$) which verifies $\mathcal{R}(\mathbf{W}_m) = \mathcal{R}(\mathbf{C})$ (the P columns of $\mathbf{W}_m$ are the P dominant left-singular vectors of $\mathbf{C}$). Notice that, since $\mathcal{R}(\mathbf{C}) \perp \mathcal{R}(\mathbf{B})$, $\mathbf{W}_m$ satisfies indeed equation (7). Also, after some algebra, one proves that $\mathbf{K}_m \equiv \mathbf{W}_m^* \mathbf{H}_N(m) \in \mathcal{G}_P$.

Step 3: Determination of $\mathbf{Y}_m = \mathbf{K}_m \mathbf{U}_m$. By noticing that $\mathbf{R}(0) = \sum_{i=0}^{M} \mathbf{H}_N(i)\,\mathbf{H}_N^*(i)$ and exploiting equation (7), it follows that

$$\mathbf{S}_m \equiv \mathbf{W}_m^* \mathbf{R}(0)\,\mathbf{W}_m = \mathbf{K}_m \mathbf{K}_m^*.$$

Thus, the positive square root of $\mathbf{S}_m$, say $\mathbf{Y}_m \equiv \mathbf{S}_m^{1/2}$, verifies $\mathbf{Y}_m = \mathbf{K}_m \mathbf{U}_m$, where $\mathbf{U}_m \in \mathcal{U}_P$. Also, $\mathbf{Y}_m \in \mathcal{G}_P$.

Step 4: Determination of $\mathbf{T}_m = \mathbf{H}_N(m)\,\mathbf{Q}_m$. We define the $NL \times NL$ matrix

$$\mathbf{Z}_m \equiv \mathbf{R}(0)\,\mathbf{W}_m \mathbf{Y}_m^{-*} \mathbf{Y}_m^{-1} \mathbf{W}_m^* \mathbf{R}(0) = \mathbf{H}_N(m)\,\mathbf{H}_N^*(m). \quad (8)$$

The outer-product in (8) permits to determine $\mathbf{H}_N(m)$ up to a rotation ambiguity. Let $\mathbf{Z}_m = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^*$ be an eigenvalue decomposition (EVD) of $\mathbf{Z}_m$, where $\mathbf{V} \in \mathcal{U}_{NL}$ and $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{NL})$, with $\lambda_1 \geq \cdots \geq \lambda_P > \lambda_{P+1} = \cdots \lambda_{LN} = 0$. Let $\mathbf{V} = [\mathbf{V}_1\ \mathbf{V}_2]$ $(\mathbf{V}_1 \in \mathbb{C}^{NL \times P})$ and $\boldsymbol{\Lambda}_1 = \text{diag}(\lambda_1, \ldots, \lambda_P)$. Then, $\mathbf{T}_m \equiv \mathbf{V}_1 \boldsymbol{\Lambda}_1^{1/2}$ satisfies $\mathbf{T}_m = \mathbf{H}_N(m)\,\mathbf{Q}_m$, where $\mathbf{Q}_m \in \mathcal{U}_P$.

Step 5. Determination of the MVDR beamformer $\widetilde{\mathbf{G}}_m = \mathbf{G}_m \mathbf{V}_m$. Step 4 determines $\mathbf{H}_N(m)$ up to an (unknown) unitary matrix $\mathbf{Q}_m$. This permits to determine the MVDR beamformer $\mathbf{G}_m$ in (4) up to a rotation, i.e.

$$\widetilde{\mathbf{G}}_m \equiv \mathbf{R}_{\mathbf{x}_N}^{-1}(0)\,\mathbf{T}_m\left(\mathbf{T}_m^* \mathbf{R}_{\mathbf{x}_N}^{-1}(0)\,\mathbf{T}_m\right)^{-1} = \mathbf{G}_m \mathbf{Q}. \quad (9)$$

Thus, the output of the beamformer in (9) is given by $\mathbf{z}_m(t) \equiv \widetilde{\mathbf{G}}_m^* \mathbf{x}_N(t) = \mathbf{Q}_m^* \mathbf{s}(t - m) + \mathbf{n}(t)$, where $\mathbf{n}(t)$ accounts for the residual ISI and noise. Notice that since $\mathbf{Q}_m$ is unitary, usage of $\widetilde{\mathbf{G}}_m$ instead of $\mathbf{G}_m$ does not result in loss of SNR. To recover the emitted symbols $\mathbf{s}(t - m)$ in the samples $\mathbf{z}_m(t)$, a BSS algorithm must be invoked, e.g., [6, 7, 8, 9]. In [11], a BSS algorithm exploiting the fact that the mixing matrix is unitary is proposed.

4. COMPUTER SIMULATIONS

To illustrate the performance of our technique, which we refer to as the blind MVDR (BMVDR) approach, we present the results of numerical simulations in two distinct scenarios: **(i)** the single-user case (SIMO), and **(ii)** the multi-user case (MIMO). In both cases, the noise $\mathbf{b}(t)$ in the data model (1) is taken to be spatio-temporal white Gaussian noise; SNR $\equiv \mathrm{E}\left\{\|\mathbf{y}(t)\|^2\right\}/\mathrm{E}\left\{\|\mathbf{b}(t)\|^2\right\}$, where $\mathbf{y}(t) \equiv \sum_{m=0}^{M_0} \mathbf{H}(m)\mathbf{s}(t-m)$ is the noiseless output of the system in (1). Also, the matrix $\mathbf{C}$ in Theorem 1, is computed as $\mathbf{C} \simeq \mathbf{C}_\epsilon$ with $\epsilon = 0.001$.

Single-user. We consider a $M_0 = 4$th order SIMO system with $L = 2$ outputs. The filter coefficients are given by $\mathbf{H}^T(0) = [-1.25 - 1.29j \ , \ -0.30 - 0.02j]^T$, $\mathbf{H}^T(1) = [2.54 + 1.46j \ , \ -1.21 + 1.52j]^T$, $\mathbf{H}^T(2) = [0.28 + 0.03j \ , \ 0.13 - 0.60j]^T$, $\mathbf{H}^T(3) = [0.55 - 0.34j \ , \ -0.79 - 0.82j]^T$ and $\mathbf{H}^T(4) = [0.74 + 1.43j \ , \ -0.41 - 0.83j]^T$. The SIMO system is driven by a QPSK sequence and $N = 7$ sucessive observations $\mathbf{x}(t)$ are stacked to form the samples $\mathbf{x}_N(t)$. The number of observations is $T = 400$. The input signal is estimated as the output of the $m = 4$th MVDR beamformer $\widetilde{\mathbf{G}}_m$ – recall equation (9). The channel order is overestimated as $\widehat{M} = 6$. We compare our technique with the multistep prediction (MSP) approach in [4]. Moreover, for the MSP approach we use the *exact* prediction filters, whereas for the BMVDR approach we estimate the correlation matrices from the available samples. The SNR is varied between $\mathrm{SNR}_{min} = 10$dB and $\mathrm{SNR}_{max} = 25$dB, in steps of $\Delta\mathrm{SNR} = 5$dB. For each SNR, $K = 100$ independent Monte-Carlo simulations were performed. At the end of the kth experiment, the variance at the output of each equalizer, i.e., $\sigma^2_{k;msp}$ and $\sigma^2_{k;bmvdr}$, is computed. This permits to obtain an average variance for each SNR, $\sigma^2_{msp} = \text{mean}\left\{\sigma^2_{k;msp}\right\}$ and $\sigma^2_{bmvr} = \text{mean}\left\{\sigma^2_{k;bmvdr}\right\}$; mean $\{\cdot\}$ is the sample mean operator. The probability of symbol misclassification is then computed from σ^2_{msp} and σ^2_{bmvdr}. In figure 1, we plot the probabilities of error thus obtained against the SNR. As
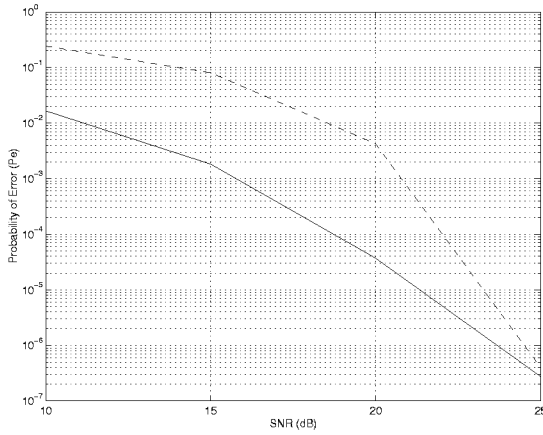
seen, the BMVDR approach yields better symbol estimates, specially at the lower SNRs considered.

Multi-user. We consider a $M = 2$th order MIMO system, with $L = 10$ outputs, driven by $P = 3$ BPSK users. The MIMO filter coefficients are samples of a zero-mean complex Gaussian random variable with power $\sigma^2 = 2$. The input multi-user signal $\mathbf{s}(t)$ is estimated as the output of the $m = 2$th MVDR beamformer, i.e., $\hat{\mathbf{s}}(t) = \widetilde{\mathbf{G}}_2^* \mathbf{x}_N(t)$, where $N = 1$. Each user scalar signal $s_p(t)$ is then extracted from $\hat{\mathbf{s}}(t)$ by a MVDR beamformer based on the exact mixing matrix. We consider $T = 800$ observations. Figures 2 and 3 show the result of a typical run of our scheme under $SNR = 15$dB. Figure 2 displays the output of the first unequalized channel (i.e., the first component of the multichannel vector $\mathbf{x}(t)$), and figure 3 shows the signal estimate for the first user (similar results hold for the remaining users). As seen, the eye of the output constellation is significantly opened.

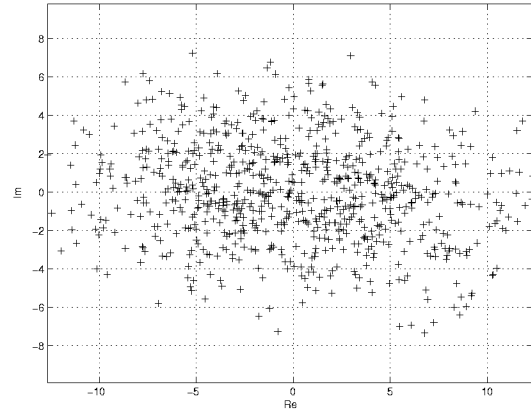To assess more precisely the performance of the proposed

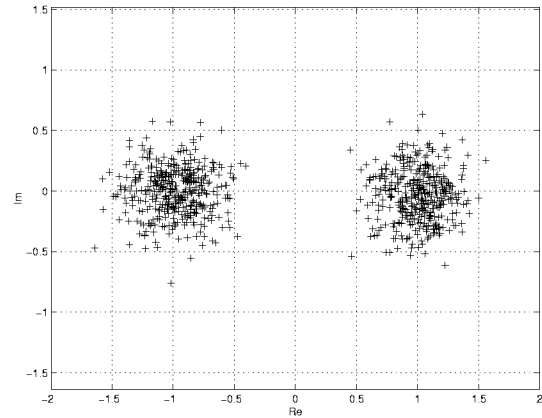Figure 2: Output of the first unequalized channel

Figure 1: Probability of error for the MSP (dashed) and the BMVDR (solid) approaches

Figure 3: Signal estimate for the first user

scheme, we varied the SNR between $\mathrm{SNR}_{min} = 10$dB and

$SNR_{max} = 25$dB, in steps of $\Delta SNR = 2.5$dB. For each SNR, $K = 100$ Monte-Carlo trials were performed. The bit-error rate (BER) is then evaluated for each user as in the SIMO case (see above). Figure 4 presents the results obtained. The signal replicas are clearly separated, and the ISI per user well rejected.
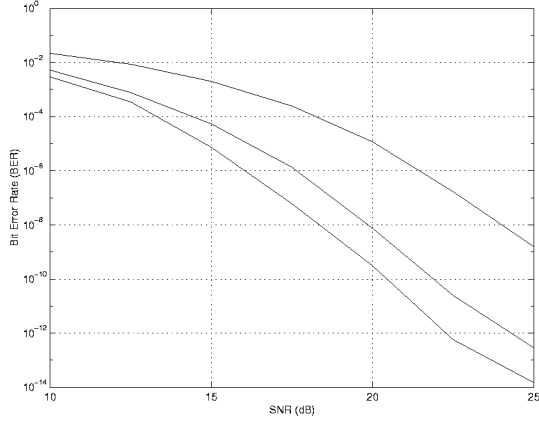

Figure 4: Bit error rate (BER) for user p=1,2 and 3

5. CONCLUSIONS

We proposed a blind equalization technique for MIMO systems, driven by white sources. It relies only on the SOS of the system outputs, and does not need an estimate of the channel order. We blindly compute several MVDR beamformers, each one matched to a distinct multichannel filter tap. The penalty for under/over estimation of the channel order is fewer delayed replicas of the input signal to process or some meaningless beamformers, respectively. In any case, the algorithm always yields valid instantaneous mixtures of the users' signals. Computer simulations illustrated the good performance of the proposed technique, either in the context of a single-user (SIMO) (compared to the MSP approach) or in the case of multi-users (MIMO).

6. REFERENCES

[1] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second order statistics: a time-domain approach," *IEEE Transactions on Information Theory*, vol. 40, no. 2, pp. 340–349, March 1994

[2] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Transactions on Signal Processing*, vol. 43, no. 2, pp. 516–526, February 1995

[3] K. Abed Meraim, P. Duhamel, D. Gesbert, P. Loubaton, S. Mayrargue, E. Moulines, and D. Slock, "Prediction error methods for time-domain blind identification of multichannel FIR filters," in *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'95)*, pp. 1968–1971, 1995

[4] D. Gesbert and P. Duhamel, "Robust blind channel identification and equalization based on multi-step predictors," in *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'97)*, pp. 3621–3624, 1997

[5] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 2, pp. 387–392, April 1985

[6] S. Talwar, M. Viberg, and A. Paulraj, "Blind estimation of multiple co-channel digital signals using an antenna array," *IEEE Signal Processing Letters*, vol. 1, no. 2, pp. 29–31, February 1994

[7] C. Papadias and A. Paulraj, "A space-time constant modulus algorithm for SDMA systems," in *Proceedings IEEE Vehicular Technology Conference*, pp. 86–90, Atlanta, Georgia, USA, 1996

[8] J. Xavier and V. Barroso, "Blind SDMA with a signal space geometry matched algorithm," in *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'97)*, pp. 3435–3440, 1997

[9] V. Barroso, J. M. F. Moura, and J. Xavier, "Blind array channel division multiple access (AChDMA) for mobile communications," *IEEE Transactions on Signal Processing*, vol. 46, pp. 737–752, March 1998

[10] J. Xavier, V. Barroso, and J. M. F. Moura, "Closed-form blind channel identification and source separation in SDMA systems through correlative coding," to appear in *Journal of Selected Areas on Communications*, special issue on *Signal Processing for Wireless Communications*

[11] J. Xavier and V. Barroso, "Blind source separation, ISI cancellation and carrier phase recovery in SDMA systems for mobile communications," to appear in *Wireless Personal Communications*, special issue on *Wireless Broadband Communications*

A. PROOF OF THEOREM 1

Suppose that $\dim \mathcal{R}(\mathbf{B}) = r$. Let $\mathbf{A} = \mathbf{QR}$ and $\mathbf{B} = \mathbf{S\Lambda S}^*$ be truncated QR and eigenvalue decompositions of $\mathbf{A}$ and $\mathbf{B}$, respectively; here, $\mathbf{Q} \in \mathbb{C}^{n \times (r+l)}$ and $\mathbf{S} \in \mathbb{C}^{n \times r}$ are isometries, $\mathbf{R} \in \mathcal{G}_{r+l}$ and $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_r)$, with $\lambda_i > 0$ for $i = 1, \ldots, r$. Let the columns of $\mathbf{U} \in \mathbb{C}^{n \times l}$ be an orthonormal basis for $\mathcal{R}(\mathbf{A}) \overset{\perp}{\ominus} \mathcal{R}(\mathbf{B})$, i.e., $\widetilde{\mathbf{Q}} = [\,\mathbf{S}\,\mathbf{U}\,]$ spans $\mathcal{R}(\mathbf{A})$, and $\mathbf{U}^*\mathbf{S} = \mathbf{0}_{l \times r}$. Since $\mathbf{Q}$ and $\widetilde{\mathbf{Q}}$ are isometries spanning the same space, there exists $\mathbf{Z} \in \mathcal{U}_{r+l}$ such that $\mathbf{Q} = \widetilde{\mathbf{Q}}\mathbf{Z}$. Write $\mathbf{Y} \equiv \mathbf{ZR} \in \mathcal{G}_{r+l}$ as

$$\mathbf{Y} = \left[\begin{array}{c} \mathbf{W}^* \\ \mathbf{V}^* \end{array} \right],$$

where $\mathbf{W} \in \mathbb{C}^{(r+l) \times r}$ and $\mathbf{V} \in \mathbb{C}^{(r+l) \times l}$. Both $\mathbf{W}$ and $\mathbf{V}$ are full column-rank matrices, and $\mathbf{A} = \mathbf{SW}^* + \mathbf{UV}^*$. Finally,

$$
\begin{aligned}
\mathbf{C} &= \lim_{\epsilon \to 0} \epsilon \left(\mathbf{B} + \epsilon \mathbf{I}_n\right)^{-1} \mathbf{A} \\
&= \left(\mathbf{I}_n - \mathbf{SS}^*\right)\left(\mathbf{SW}^* + \mathbf{UV}^*\right) \\
&= \mathbf{UV}^*.
\end{aligned}
$$