BLIND SOURCE SEPARATION WITH RANDOMIZED GRAM-SCHMIDT ORTHOGONALIZATION FOR SHORT BURST SYSTEMS

Constantinos B. Papadias and Alexandr M. Kuzminskiy

Bell Laboratories, Lucent Technologies
791 Holmdel-Keyport Rd. Holmdel, NJ, 07733-0400, USA, papadias@lucent.com
The Quadrant, Stonehill Green, Westlea, Swindon SN5 7DJ, UK, ak9@lucent.com

Abstract

A blind source separation problem for short burst systems is addressed by means of a constant modulus technique under orthogonal constraints. It is shown that a conventional Gram-Schmidt orthogonalization procedure normally exploited in similar applications may cause a non-uniform misadjustment distribution among the receiver outputs leading to an overall performance degradation. We propose a modified algorithm based on random reordering of the weight vectors before the orthogonalization stage and demonstrate its efficiency by means of simulations in a short burst MIMO environment.

1. INTRODUCTION

Blind source separation (BSS) of instantaneous and convolutive mixtures is a technique of interest in wireless communications and other applications. Some techniques based on higher order statistics [1-4 and others] have been proposed for BSS, including constant modulus (CM) and kurtosis based BSS approaches [3,4]. Short burst versions of some of these techniques based on iterations over the same block of data received in quasistationary conditions are addressed in [5-7 and others].

To prevent the same source signal from being extracted at different outputs it has been proposed to control crosscorrelation properties of the extracted signals by means of penalization of the conventional CM criterion [8,9] or formulation of the constrained CM or kurtosis-based criteria [3,4]. The last group of algorithms requires prewhitening of the received signal and includes a criterion optimization procedure under orthogonal constraints. A gradient search and Gram-Schmidt orthogonalization procedure are employed in [3,4] for the CM and kurtosis based versions of the technique. The asymptotic properties of this group of algorithms have been studied in [3] and their global convergence has been established.

In this paper we address a short burst application of the BSS algorithm under orthogonalization constraints. We show that a conventional application of the Gram-Schmidt orthogonalization procedure in the constrained optimization algorithm has some undesirable features under non-ideal conditions, e.g. non-ideal prewhitening stage. Particularly, we demonstrate that fixing the order of the weight vectors at the orthogonalization stage leads to different misadjustments at the different outputs causing overall performance degradation. Taking into account that any signal may be extracted at any output for a BSS technique, we accept the recovering accuracy at the output with the highest misadjustment as the performance metric.

We propose a randomized order selection of the weight vectors at the orthogonalization stage, which allows us to obtain a uniform distribution of the misadjustment error among all the extracted signals leading the overall performance improvement without any additional complexity. Simulation results for a short burst MIMO TDMA system demonstrate advantages of the randomized algorithm for variable receiving conditions.

The randomization effect can be demonstrated for different techniques based on iterative optimization of a blind criterion with orthogonal constraints. In this paper we concentrate on the CM based constrained optimization to illustrate the main idea.

The paper is organized as follows. A data model is presented in Section 2. The basic algorithm, the effect of weight vector order to the Gram-Schmidt orthogonalization stage at the CM based constrained optimization procedure and the randomized modification of the basic algorithm are given in Section 3. Some simulation results are shown in Section 4 and our conclusions are presented in Section 5.

2. DATA MODEL

We assume that M i.i.d. mutually independent zero-mean discrete-time sequences $s_m(n)$, $m = 1 \dots M$ drawn from the same constellation with the CM property $|s_m(n)|^2 = 1$, are transmitted through a $M \times K$ MIMO linear memoryless channel that introduces inter-user interference. The model of the received signal slot of N samples takes the following form

$$\mathbf{x}(n) = \mathbf{H}\mathbf{s}(n) + \mathbf{z}(n),\, n = 1 \dots N, \tag{1}$$

where $\mathbf{x}(n)$ is the $K \times 1$ vector of received signal samples, $\mathbf{s}(n) = [s_1(n) \dots s_M(n)]^{\mathrm{T}}$ is the $M \times 1$ vector of transmitted (source) signals, $\mathbf{H}$ is the $K \times M$ channel matrix, $\mathbf{z}(n)$ is the $K \times 1$ vector of additive noise samples and $(\cdot)^{\mathrm{T}}$ denotes a transpose operation. The channel matrix $\mathbf{H}$ consists of i.i.d., complex, zero-mean, unit-variance entries independent for different data slots.

The received signal $\mathbf{x}(n)$ is subsequently filtered by a $K \times M$ "matrix equalizer" $\mathbf{W}(n) = [\mathbf{w}_1(n) \dots \mathbf{w}_M(n)]$ that produces the $M \times 1$ vector of output signals $\mathbf{y}(n) = [y_1(n) \dots y_M(n)]^{\mathrm{T}}$. The receiver output can be represented as

$$\mathbf{y}(n) = \mathbf{W}^{\mathrm{T}}(n)\mathbf{x}(n) = \mathbf{G}^{\mathrm{T}}(n)\mathbf{s}(n) + \mathbf{z}'(n), \tag{2}$$

where $\mathbf{G}(n) = \mathbf{H}^{\mathrm{T}}\mathbf{W}(n)$ is the $M \times M$ overall response matrix and $\mathbf{z}'(n) = \mathbf{W}^{\mathrm{T}}(n)\mathbf{z}(n)$ is the spatially colored noise at the receiver output. The problem is to find the weight matrix $\mathbf{W}(n)$ that matches the receiver outputs $y_m(n)$, $m = 1 \dots M$ with all the transmitted signals $s_m(n)$, $m = 1 \dots M$ up to some rotation and permutation factors.

3. MULTIUSER CONSTRAINED CM ALGORITHM

3.1. Basic algorithm

Similarly to [3,4] the multi-user CM criterion can be formulated for sub-Gaussian signals as

$$\min_{\mathbf{W}} \sum_{m=1}^{M} E\{(|\mathbf{w}_m^{\mathrm{T}}\mathbf{x}(n)|^2 - 1)^2\} \tag{3}$$

$$\text{subject to: } \mathbf{G}^*\mathbf{G} = \mathbf{I}_M, \tag{4}$$

where $\mathbf{I}_M$ is the $M \times M$ identity matrix and $(\cdot)^*$ denotes complex conjugate transpose operation and $E\{\cdot\}$ is a mathematical expectation operator. If the channel matrix is unitary, i.e. $\mathbf{H}^*\mathbf{H} = \mathbf{I}_M$, the constraint (4) can be reduced to $\mathbf{W}^*\mathbf{W} = \mathbf{I}_M$.

Then, the basic algorithm can be formulated as follows:

Step 1. Prewhitening:

$$\tilde{\mathbf{x}}(n) = \hat{\mathbf{L}}\mathbf{x}(n),\, n = 1 \dots N \tag{5}$$

Step 2: Initialization

$$\mathbf{W}^{(1)}(1) = \mathbf{W}_0 \tag{6}$$

Step 3: Calculate for $j = 1 \dots J$

Step 4: Calculate for $n = 1 \dots N$

Step 5: Calculate for $m = 1 \dots M$

$$\mathbf{v}_m^{(j)}(n+1) = \mathbf{w}_m^{(j)} + \mu e_m^{(j)}(n)\tilde{\mathbf{x}}(n)^{\mathrm{C}}, \tag{7}$$

$$e_m^{(j)}(n) = \frac{y_m^{(j)}(n)}{|y_m^{(j)}(n)| - 1},\; y_m^{(j)}(n) = \mathbf{w}_m^{(j)\mathrm{T}}(n)\tilde{\mathbf{x}}(n) \tag{8}$$

Step 6: end of step 5

Step 7: Orthogonalization

$$\mathbf{W}^{(j)}(n+1) = \mathrm{GS}\{\mathbf{V}^{(j)}(n+1)\} \tag{9}$$

Step 8: end of step 4

Step 9: $\mathbf{W}^{(j+1)}(1) = \mathbf{W}^{(j)}(N)$

Step 10: end of step 3.

The following notations are used: $\tilde{\mathbf{x}}(n)$ is the $M \times 1$ input vector after prewhitening, $\hat{\mathbf{L}}$ is the $M \times K$ prewitening matrix (will be specified in Section 4), $\mathbf{W}^{(j)}(n) = [\mathbf{w}_1^{(j)}(n) \dots \mathbf{w}_M^{(j)}(n)]$ and $\mathbf{V}^{(j)}(n) = [\mathbf{v}_1^{(j)}(n) \dots \mathbf{v}_M^{(j)}(n)]$ are the $M \times M$ weight matrices at the j-th burst iteration, J is the number of burst iterations, μ is the adaptation coefficient in the gradient search procedure, $\mathrm{GS}\{\cdot\}$ denotes a Gram-Schmidt orthogonalization operation and $(\cdot)^{\mathrm{C}}$ denotes a complex conjugate operation.

A conventional implementation of the Gram-Schmidt procedure, e.g. the one exploited in [3,4] in the similar environment, can be expressed as

$$\mathbf{w}_1 = \mathbf{v}_1/||\mathbf{v}_1||, \tag{10}$$

$$\mathbf{w}_m = \frac{\mathbf{v}_m - \sum_{l=1}^{m-1}(\mathbf{w}_l^*\mathbf{v}_m)\mathbf{w}_l}{||\mathbf{v}_m - \sum_{l=1}^{m-1}(\mathbf{w}_l^*\mathbf{v}_m)\mathbf{w}_l||},\, m = 2 \dots M, \tag{11}$$

where the indexes j and $n+1$ are omitted for simplicity.

3.2. Misadjustment effect

Let us assume that some misadjustment exists at the $(n+1)$-th adaptation step in (7):

$$\mathbf{v}_m(n+1) = \mathbf{v}_{0m} + \delta\mathbf{v}_m(n+1),\, m = 1 \dots M, \tag{12}$$

where $\mathbf{v}_{0m}$ is the ideal weight vector for the m-th output and $\delta\mathbf{v}_m(n+1)$ is the overall error (misadjustment) vector. Some possible sources of misadjustment after convergence are non-ideal prewhitening, noise and non-zero adaptation coefficient.

Substituting equation (12) into equations (10) and (11) we obtain the following expression for the weight vectors after the orthogonalization step:

$$\mathbf{w}_m(n+1) = f_m(\delta\mathbf{v}_1(n+1), \dots, \delta\mathbf{v}_m(n+1)), \tag{13}$$

for $m = 1 \dots M$, where f_m some function.

An important observation can be made from equation (13): the vector order at the orthogonalization step makes a difference in terms of misadjustment effects. Indeed, the first weight vector depends only on its own misadjustment, i.e. $\mathbf{w}_1(n+1) = f_1(\delta\mathbf{v}_1(n+1))$, the second vector depends on its own error as well as on the first's vector error, i.e. $\mathbf{w}_2(n+1) = f_2(\delta\mathbf{v}_1(n+1), \delta\mathbf{v}_2(n+1))$ and the last vector depends on all the misadjustments, i.e. $\mathbf{w}_M(n+1) = f_M(\delta\mathbf{v}_1(n+1), \ldots, \delta\mathbf{v}_M(n+1))$, reflecting the algorithm's "deflation" nature.

An immediate consequence of this observation is that the signal extraction accuracy depends on the weight vector order at the orthogonalization step. A signal extracted at the first output is expected to have the lowest error level while the signal recovered with the last weight vector may have the highest error level. Taking into account that a blind approach does not allow us to control the signal extraction order, i.e. any signal may be extracted at any output up to some error and phase rotation, this effect may cause an overall performance degradation.

3.3. Modified algorithm

To make the misadjustment error uniformly distributed (on average) over all the extracted signals we propose random selection of the vectors order at the orthogonalization step independently for the consecutive iterations. In this case, step 7 at the basic algorithm in Section 3.1 can be modified as follows:

Step 7a: Random order selection

$$\tilde{\mathbf{V}}^{(j)}(n+1) = \text{ORDER}\{\mathbf{V}^{(j)}(n+1)\} \tag{14}$$

Step 7b: Orthogonalization

$$\tilde{\mathbf{W}}^{(j)}(n+1) = \text{GS}\{\tilde{\mathbf{V}}^{(j)}(n+1)\} \tag{15}$$

Step 7c: Order recovery

$$\mathbf{W}^{(j)}(n+1) = \text{IORDER}\{\tilde{\mathbf{W}}^{(j)}(n+1)\}, \tag{16}$$

where ORDER$\{\mathbf{A}\}$ is the operator, which randomly reorders the columns of matrix $\mathbf{A}$ and IORDER is the inverse operator, i.e. $\mathbf{A} = \text{IORDER}\{\text{ORDER}\{\mathbf{A}\}\}$.

We refer to the basic algorithm in Section 3.1 as the multiuser CM algorithm (MUC) and the modified algorithm as randomized MUC (RMUC). It is worth emphasising that there is almost no additional computational complexity involved in RMUC compared to MUC.

4. SIMULATION RESULTS

We simulate a 4×6 MIMO system ($M = 4$, $K = 6$). The channel matrix $\mathbf{H}$ is chosen from a complex Gaussian distribution of zero mean and unit variance $h_{ij} \sim \mathcal{N}(0,1)$. Each user input is a QPSK signal. On the receiver side additive white Gaussian noise of SNR=20 dB is added to each of six received signals. A Schur algorithm is applied to estimate the prewhitening matrix $\hat{\mathbf{L}}$ similarly to [3]. The estimation is performed over the whole data slot of N samples. The fixed adaptation coefficient $\mu = 0.03$ and initialization $\mathbf{W}_0 = \mathbf{I}_M$ are applied at all simulations. Variable slot size is considered $N = [50, 100, 200, 400]$ leading to different prewhitening errors. The total number of iterations is 4000 for all simulations, i.e. $J = 4000/N$.

Figure 1 quantitatively illustrates non-uniform misadjustment distribution for MUC and the randomization effect for RMUC. The first row of the constellation pictures shows typical outputs of MUC for $N = 100$. The second row of pictures presents the output signals of RMUC for the same data slot. One can see the expected non-uniform behaviour of MUC and randomization effect of RMUC.

Figures 2 - 5 present the learning curves (mean square error (MSE) after phase correction versus the current block iteration number j) averaged over 500 independent trials for both MUC and RMUC for different N. The first row of pictures corresponds to the particular user without specification of which output this user is extracted at. The second row of pictures corresponds to the users extracted at the indicated outputs. One can see that although the randomization at the orthogonalization stage leads to some accuracy degradation for the signal extracted first, the proposed algorithm demonstrates an overall performance improvement especially for the short burst scenarios ($N = 50$ and $N = 100$). Particularly, the improvement can be observed for the signal recovered last (last plot in the second row of pictures in Figures 2 - 5), which is our performance metric as formulated in Section 1, as well as for average MSE for all users (first row of pictures in Figures 2 - 5). The improvement decreases with growing N because of the better accuracy at the prewhitening step.

5. CONCLUSIONS

The BSS problem for short burst systems has been addressed by means of a CM technique under orthogonal constraints. It has been shown that the conventional implementation of the Gram-Schmidt orthogonalization procedure may cause a non-uniform misadjustment distribution among the receiver outputs leading to an overall performance degradation. A modified algorithm has been proposed, which is based on random reordering of the weight vectors before the orthogonalization stage. Its efficiency has been demonstrated by means of simulations in a short burst MIMO environment.

6. REFERENCES

[1] S.Haykin, "Unsupervised adaptive filtering, volume 1: blind source separation", John Wiley and Sons, Inc., 2000.

[2] Y.Li, K.J.R.Liu, "Adaptive blind source separation and equalization for multiple-input/multiple-output systems", IEEE Trans. Inf. Theory, vol. 44, pp. 2864-2876, 1998.

[3] C.B.Papadias, "Globally convergent blind source separation based on a multiuser kurtosis maximization criterion", IEEE Trans Sig. Proc., vol. 48, n. 12, pp. 3508-3519, 2000.

[4] P.Sansrimahachai, D.B.Ward, A.G.Constantinides, "Blind source separation for BLAST", in Proc. 14th Int. Conf. DSP, pp. 139-142, 2002.

[5] B.Kim, D.C.Cox, "Blind equalization for short burst communications over frequency selective wireless channels", in Proc. IEEE 47th VTC, pp. 544-548, Phoenix, 1997.

[6] Y.Chen, J.C-I Chuang, K.B.Letaief, "Blind equalization for short burst TDMA systems in wireless communications", in Proc. IEEE 47th VTC, pp. 535-538, Phoenix, 1997.

[7] A.M.Kuzminskiy, "Finite amount of data effects in spatio-temporal filtering for equalisation and interference rejection in short burst wireless communications", Signal Processing, Elsevier, vol. 80, n. 10, pp. 1987-1997, 2000.

[8] C.B.Papadias, A.Paulraj, "A constant modulus algorithm for multi-user signal separation in presence of delay spread using antennas arrays", IEEE Signal Processing Letters, vol. 4, n. 6, pp. 178-181, 1997.

[9] C.B.Papadias, H.Huang, "Linear space-time multiuser detection for multipath CDMA channels", IEEE Journal on Selected Areas in Communications (J-SAC), vol. 19, n. 2, pp. 254-265, 2001.

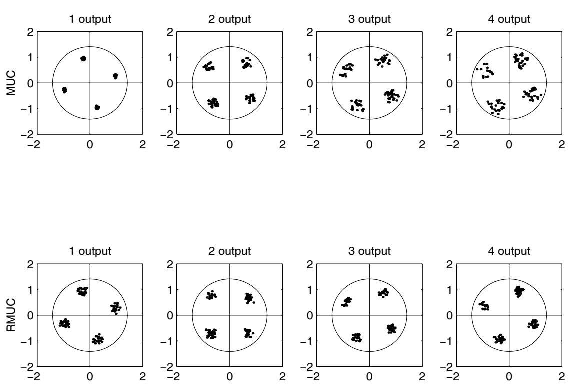

Fig. 1. Typical constellation pictures for $N = 100$

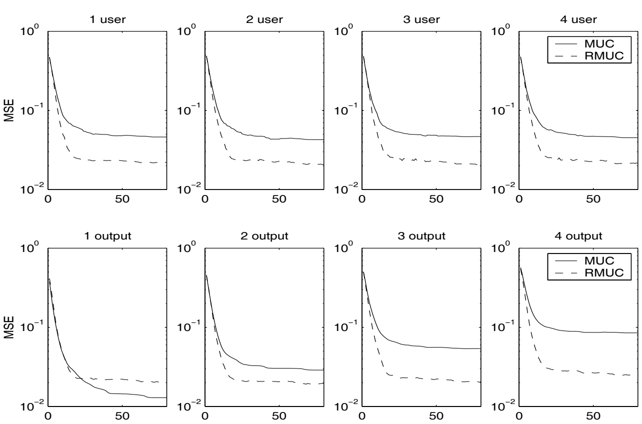

Fig. 2. Learning curves for $N = 50$: MSE after phase correction versus block iteration number

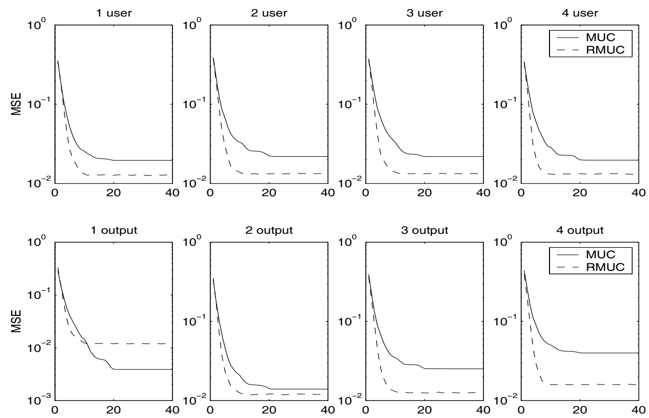

Fig. 3. Learning curves for $N = 100$: MSE after phase correction versus block iteration number

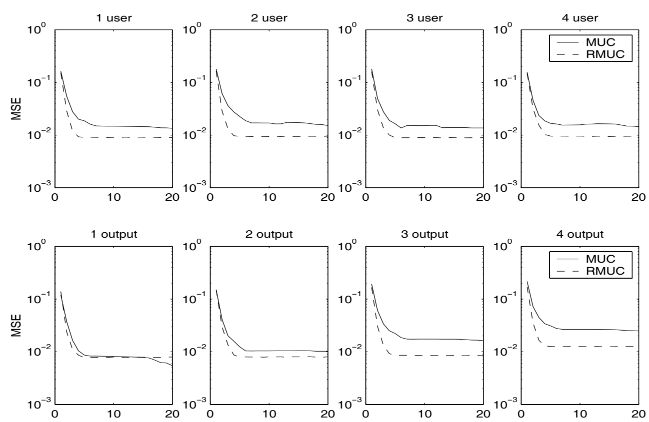

Fig. 4. Learning curves for $N = 200$: MSE after phase correction versus block iteration number

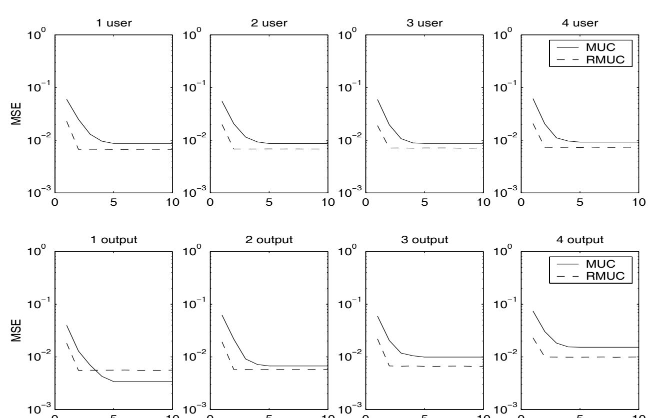

Fig. 5. Learning curves for $N = 400$: MSE after phase correction versus block iteration number