Comparison of Blind Source Separation Methods based on Iterative Algorithms

J. Rinas, K.D. Kammeyer

Department of Communications Engineering, Universität Bremen, {rinas,kammeyer}@ant.uni-bremen.de

Abstract

In this paper some approaches to detect signal streams of a multi layer transmission system are presented. We will focus on blind algorithms for the separation of the data stream and improve their performance in an iterative way in order to gain nearly the same performance as with a known channel matrix. The overall algorithm will remain blind and does not need any training data.

1 Introduction

Multi layer transmissions are an up to date technology to exploit the capacity of a spatial channel. Lately, powerful detection algorithms have been developed. These algorithms need a suitable estimation of the transmission channel. In most cases the channel estimation is obtained by including a pilot sequence in the data stream. However, this will lead to a loss of the payload rate that can be transmitted. Therefore we will introduce a new scheme that combines blind source separation techniques with an efficient detection algorithm in an iterative way. This will lead to an overall blind detection of the transmitted symbols. We will study this idea with some variants of the constant modulus algorithm (CMA) and higher order statistics (HOS) based source separation approaches.

The remainder of the paper is organized as follows: In section 2 we will introduce the transmission system that was used for all simulations. In section 3 we will present some approaches to achieve a blind separation of a multi layer transmission. We will present some variants of multiple input multiple output (MIMO) CMA algorithms and point out connections with classical blind source separation (BSS) approaches. We compare their performance using Monte Carlo simulations. A new approach to combine source separation techniques with a multi layer detection scheme is presented in section 4. This will lead to an overall blind symbol detection scheme. Section 5 introduces channel coding into the presented scheme and illustrates some problems in this context. A summary and concluding remarks can be found in section 6.

2 System description

In this paper we analyze a multi layer transmission as seen in **figure 1**. This system can be described by

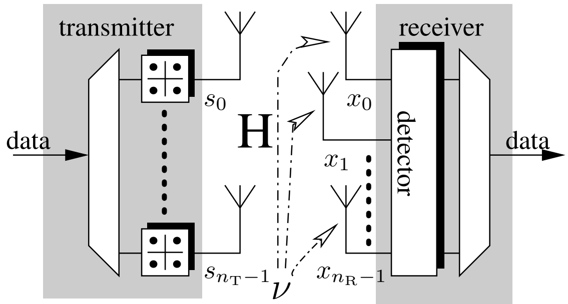

Fig. 1. transmission system

$$\mathbf{x} = \mathbf{H} \cdot \mathbf{s} + \nu \ . \tag{1}$$

Equation (1) models a transmission of signal streams s trough a spatial channel $\mathbf{H}$. Let $\mathbf{s} = [s_0(k), s_1(k), \ldots, s_{n_\mathrm{T}-1}(k)]^T$ be the vector of transmitted symbols at time instant k. We usually use a block processing of the data and use a length of $k = 0 \ldots L-1$ symbols. For a clear arrangement and short formulas we‘ll drop the index k in the following text.

The dimensions of $\mathbf{H}$ are $n_\mathrm{R} \times n_\mathrm{T}$ for n_R receive and n_T transmit antennas.

$$\mathbf{H} = \begin{bmatrix} h_{0,0} & \cdots & h_{0,n_\mathrm{T}-1} \\ \vdots & \ddots & \vdots \\ h_{n_\mathrm{R}-1,0} & \cdots & h_{n_\mathrm{R}-1,n_\mathrm{T}-1} \end{bmatrix} \tag{2}$$

The entries of $\mathbf{H}$ are assumed to be Gaussian distributed and statistical independent. They are assumed to be constant for one block of L time instants in order to model a slow fading channel. $\nu := [\nu_0(k), \nu_1(k), \ldots, \nu_{n_\mathrm{R}-1}(k)]^T$ is a vector with additive Gaussian noise that is assumed to be white.

During this paper we concentrate on a scenario with $n_\mathrm{T} = 4$ transmit and $n_\mathrm{R} = 4$ receive antennas. We transmit $L = 200$ uncoded and coded QPSK symbols. These signals have a constant magnitude if the symbol

This work was supported by the German national science foundation (DFG) within the project AKOM (project #Ka 841/9-1)

timing is perfectly known and therefore can be detected by constant modulus approaches. The energy of the transmitted signals is normalized so that the average received energy of one signal is one. (Therefore we'll only observe diversity gains and no additional array gain due to an increasing number of antennas.)

3 Blind separation approaches for communication signals

There are several methods available to separate data-streams blindly with more or less knowledge of the signals. In this paper we concentrate on two families of approaches. The classical approaches try to maximize the statistical independence of signal streams and make no use of the discrete alphabet of modulated signals, whereas the constant modulus approaches considered here use the discrete amplitude level of the QPSK signals.

3.1 Classical blind source separation

The classical blind separation schemes are based on the assumption of independent components in the received signal streams. This property is fulfilled if the data carried within the symbols of the received streams is random. Established algorithms that lead to a separation are the JADE algorithm as a batch algorithm and the fastICA algorithm that solves the problem by extracting the independent components step by step.

3.1.1 JADE

The joint approximate diagonalization of eigenmatrices (JADE) algorithm [1] is a batch procedure that solves the separation problem.

The first step of this algorithm is to decorrelate the input streams $\mathbf{x}$. That is the $n_\mathrm{T} \times n_\mathrm{R}$ matrix $\mathbf{W}$ has to be calculated to fulfill

$$\mathbf{I} = \mathbf{W}\mathbf{R}_{\mathbf{xx}}\mathbf{W}^H \quad [1] \tag{3}$$

with $\mathbf{R}_{\mathbf{xx}} = \mathrm{E}\left\{\mathbf{x}\mathbf{x}^H\right\}$ being the spatial correlation matrix of the received signals. In the source separation literature [2] [3] this step is also known as principle component analysis and can be solved using a eigenvalue decomposition of $\mathbf{R}_{\mathbf{xx}}$. The decorrelation can be computed by a multiplication of the received signal streams $\mathbf{x}$ with the whitening matrix $\mathbf{W}$.

$$\mathbf{z} = \mathbf{W}\mathbf{x} \tag{4}$$

The utilization of second order information is not sufficient to obtain independent signals. Therefore the JADE algorithm additionally uses 4th order information. It maximizes some elements of the cumulant matrix $\mathbf{Q}_\mathbf{z} = \mathrm{cum}\left(e_i^\star, e_i, e_j^\star, e_l,\right)$ obtained from the extracted signals $\mathbf{e}$ defined in equation (6).

$$\max_{\mathbf{B}} \overset{!}{=} \sum_{i,j,l=0}^{n_\mathrm{T}-1} \left|\mathrm{cum}\left(e_i^\star, e_i, e_j^\star, e_l,\right)\right|^2 \tag{5}$$

This optimization problem is solved by an eigenvalue decomposition of $\mathbf{Q}_\mathbf{z}$ and a joint diagonalization of the dominant eigenvectors rearranged as matrices. This diagonalization leads to the unitary $n_\mathrm{T} \times n_\mathrm{T}$ matrix $\mathbf{B}$ and the independent data streams

$$\mathbf{e} = \mathbf{B}^H \cdot \mathbf{z} \ . \tag{6}$$

A similar approach that needs less computational effort is the SSARS algorithm presented in [4].

3.1.2 fastICA

Beside the JADE algorithm the fastICA algorithm is organized in a different way [5] [6]. The basic idea of this algorithm is to do a blind source extraction (BSE) for each component and to prevent the same signal from being extracted multiple times. It starts with the whitening of the received data in the same way as presented before in (4).

In order to extract signal number t out of the mixture $\mathbf{z}$ (4) a randomly initialized extraction vector $\mathbf{b}_t$ – a column vector of the $n_\mathrm{T} \times n_\mathrm{T}$ matrix $\mathbf{B}$ – is generated. In order to preserve the signals to remain uncorrelated $\mathbf{B}$ has to be a unitary matrix. Therefore $\mathbf{b}_t$ is constrained to form an orthonormal basis using the knowledge of the vectors $\mathbf{b}_0 \ldots \mathbf{b}_{t-1}$ obtained in former iterations. To achieve this goal matrix

$$\mathbf{B}_t = [\mathbf{b_0}, \mathbf{b_1}, \ldots \mathbf{b_{t-1}}] \tag{7}$$

containing the extraction vectors of the former iterations is build. The randomly initialized vector $\mathbf{b}_t$ is orthogonalized to the former detected ones

$$\mathbf{b}_t' = \mathbf{b}_t - \mathbf{B}_t\mathbf{B}_t^H\mathbf{b}_t \tag{8}$$

and normalized to a length of one

$$\mathbf{b}_t'' = \mathbf{b}_t' / \left|\left|\mathbf{b}_t'\right|\right| \ . \tag{9}$$

In order to determine $\mathbf{b}_t$ we choose the maximization of the kurtosis of a single signal as the criterion.

$$\begin{aligned} \max_{\mathbf{b}_t} J_{\mathrm{fastICA},t}\left(e_t\right) &= \max_{\mathbf{b}_t} \mathrm{E}\left\{\left|e_t\right|^4\right\} \\ &= \max_{\mathbf{b}_t} \mathrm{E}\left\{\left|\mathbf{b}_t^H\mathbf{z}\right|^4\right\} \end{aligned} \tag{10}$$

This can be solved using a fixed point iteration including the additional constraints (8) and (9) [6]. The resulting signal streams $\mathbf{e}$ can be extracted by multiplying the received signal with the matrix of all collected extraction vectors $\mathbf{B} = \mathbf{B}_{n_T-1}$.

$$\mathbf{e} = \mathbf{B}^H \cdot \mathbf{z} \tag{11}$$

[1] $\mathbf{W}^H$ denotes the conjugate transpose i.e. the Hermitean of $\mathbf{W}$.

3.2 MIMO CMA

The idea of the constant modulus algorithm (CMA) is to compare the amplitude of the equalized received signal with a reference amplitude [7][8][9]. This comparison expressed in mathematical terms leads to a cost function that has to be minimized. The general cost function for MIMO CMA is

$$\begin{aligned} J_{\mathrm{CMA}}(\mathbf{C}) &= \sum_{t=0}^{n_{\mathrm{T}}-1} \mathrm{E}\left\{\left(|e_{\mathrm{t}}|^2-1\right)^2\right\} \\ &= \sum_{t=0}^{n_{\mathrm{T}}-1} \mathrm{E}\left\{\left(\left|\mathbf{c}_t^H\mathbf{x}\right|^2-1\right)^2\right\} \quad , \end{aligned} \tag{12}$$

where e_{t} denotes the equalized component number t of the output vector $\mathbf{e}$. The component $\mathbf{c}_t$ of the $n_{\mathrm{R}} \times n_{\mathrm{T}}$ CMA extraction matrix $\mathbf{C} = [\mathbf{c}_0, \mathbf{c}_1, \ldots, \mathbf{c}_{n_T-1}]$ spatially filters the signal e_{t}. Separation of all signals can be done by

$$\mathbf{e} = \mathbf{C}^H \cdot \mathbf{x} \; . \tag{13}$$

In order to minimize the CMA cost function (12) the steepest descent algorithm is used.

$$\mathbf{C}^{(i+1)} = \mathbf{C}^{(i)} - \mu \frac{\partial}{\partial \mathbf{C}^{(i)}} J_{\mathrm{CMA}}\left(\mathbf{C}^{(i)}\right) \;^{2} \tag{14}$$

μ is the step factor of the gradient descent. The CMA algorithm is initialized by a whitening matrix $\mathbf{C}^0 = \mathbf{W}$ as described in equation (4) . This starts the CMA algorithm with separation vectors pointing in directions of uncorrelated sources of power and improves convergence.

The Matrix $\mathbf{C}$ can be calculated using the update equation

$$\mathbf{C}^{(i+1)} = \mathbf{C}^{(i)} - 4\mu \mathrm{E}\left\{\left(D\left\{\mathbf{e}^{(i)}\right\} - \mathbf{I}\right) \cdot \mathbf{x} \cdot \mathbf{e}^{(i)H}\right\} \; . \tag{15}$$

where the operator $D\left\{\mathbf{e}^{(i)}\right\}$ forms the diagonal matrix $D\left\{\mathbf{e}^{(i)}\right\} = \mathrm{diag}\left(\left|e_0^{(i)}\right|^2, \left|e_1^{(i)}\right|^2, \ldots, \left|e_{n_{\mathrm{T}}-1}^{(i)}\right|^2\right)$. The main problem with this algorithm is, that it is still possible to resolve the same signal multiple times. This will lead to very bad error rates of the algorithm as can be seen in **figure 3**.

Therefore we will present three approaches to avoid this problem. Two of them are already known from different sources. A new approach presented here has a slight connection to the approach of the fastICA algorithm.

3.2.1 1st approach: correlation penalty

One approach to tackle the problem of extracting the same signal multiple times was presented in [10]. The idea of this approach is to minimize the correlation of the extracted signals. This can be realized by calculating the cross correlations of all signals in every iteration step and adding the squared magnitude of the off diagonal elements to the general MIMO CMA cost function (12). This leads to the new cost function

$$J_{\mathrm{corr}}(\mathbf{C}) = J_{\mathrm{CMA}}(\mathbf{C}) + \sum_{k,l=0;\, k\neq l}^{n_{\mathrm{T}-1}} \left|\psi_{k,l}^{(i)}\right|^2 \tag{16}$$

where

$$\psi_{k,l}^{(i)} = \mathrm{E}\left\{e_k^{(i)} \cdot e_l^{(i)*}\right\} \tag{17}$$

describe the correlation of the extracted signals $e_k^{(i)}$ and $e_l^{(i)}$ as elements of a matrix

$$\mathbf{\Psi}_{\mathrm{corr}}^{(i)} = \begin{pmatrix} 0 & \psi_{0,1}^{(i)} & \cdots & \psi_{0,n_{\mathrm{T}}-1}^{(i)} \\ \psi_{1,0}^{(i)} & 0 & \cdots & \psi_{1,n_{\mathrm{T}}-1}^{(i)} \\ \vdots & & \ddots & \vdots \\ \psi_{n_{\mathrm{T}}-1,0}^{(i)} & \psi_{n_{\mathrm{T}}-1,1}^{(i)} & \cdots & 0 \; . \end{pmatrix} \tag{18}$$

Note that (18) contains no diagonal elements.

Using this cost function leads to the update equation

$$\begin{aligned} &\mathbf{C}^{(i+1)} = \mathbf{C}^{(i)} \\ &-\mu\left(4\mathrm{E}\left\{\left(D\left\{\mathbf{e}^{(i)}\right\} - \mathbf{I} + \mathbf{\Psi}_{\mathrm{corr}}^{(i)\,H}\right) \cdot \mathbf{x} \cdot \mathbf{e}^{(i)H}\right\}\right) \; . \end{aligned} \tag{19}$$

The computational effort of this approach can be immense, because the correlation matrix $\mathbf{\Psi}_{\mathrm{corr}}^{i}$ has to be estimated in every iteration.

3.2.2 2nd approach: determinant penalty

Another idea to prevent the CMA algorithm to extract the same signal twice can be found in [11]. The basic idea of this approach is to minimize the linear dependency of the separation vectors contained in $\mathbf{C}$. The linear dependency is measured here in terms of the logarithm of the magnitude of the determinant of $\mathbf{C}$. This leads to the cost function

$$J_{\mathrm{det}}(\mathbf{C}) = J_{\mathrm{CMA}}(\mathbf{C}) - \ln\left|\det \mathbf{C}^H\right| \; . \tag{20}$$

For nearly linear dependent components in $\mathbf{C}$ the right hand side term of (20) will grow to very high values.

The corresponding update equation can be calculated to be

$$\begin{aligned} &\mathbf{C}^{(i+1)} = \mathbf{C}^{(i)} \\ &-\mu\left(4\mathrm{E}\left\{\left(D\left\{\mathbf{e}^{(i)}\right\} - \mathbf{I}\right) \cdot \mathbf{x} \cdot \mathbf{e}^{(i)H}\right\} - \left(\mathbf{C}^{(i)H}\right)^{-1}\right) . \end{aligned} \tag{21}$$

This drawback from the computational side of view is that this update equation contains a matrix inversion in every iteration.

[2]The superscript (i) denotes the iteration step.

3.2.3 3rd approach: subspace limitation

Motivated by the CMA approaches presented above and the idea to successively separate signals we present a new scheme to separate constant modulus signals from a MIMO system.

As a first step we decorrelate the received signal streams as described by equation (4). In order to extract the signal number t from our whitened stream $\mathbf{z}$ we start with a $n_{\mathrm{R}} \times 1$ vector $\mathbf{c}_t$ that is zero except for a one at row t. This vector is orthogonalized to the former vectors as done by the fastICA algorithm in equations (8) and (9).

Then we adjust $\mathbf{c}_t$ by the steepest descent algorithm using the CMA cost function (12) trimmed down to update only $\mathbf{c}_t$.

This leads to the update equation

$$\mathbf{c}_t^{(i+1)} = \mathbf{c}_t^{(i)} - 4\mu \mathrm{E}\left\{ \left(\left| e_t^{(i)} \right|^2 - 1 \right) \cdot \mathbf{z} \cdot e_t^{(i)H} \right\} . \tag{22}$$

After separating one signal we consider the known separation vector using equations (8) and (9) to select the next initialization in order to gain the next signal.

Compared to the update equation of the general MIMO CMA (15) we need no additional computations inside the iteration loop.

3.3 performance comparison

All separation methods presented so far will result in separated data streams $\mathbf{e}$. In order to compare their performance in terms of bit error rates (BER) we use a set up as depicted in **figure 2**.

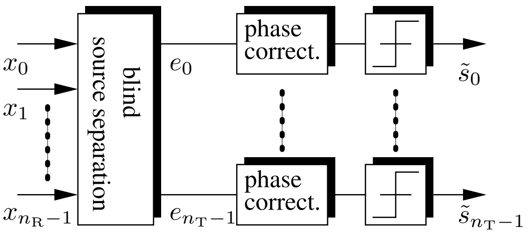

Fig. 2. source separation set up for BER measurements

Since the phase of every data stream e_t has not been taken into account during the separation, every data stream is rotated by a random phase factor and permutated.

$$e_{t,\mathrm{derot}} = e_t \cdot e^{-\left(\arg \mathrm{E}\left\{e_t^4\right\}\right)/4} \tag{23}$$

Equation (23) points out an approach feasible for QPSK signals. The average of the data signal taken to the power of four will remove the data. The result can be taken to estimate the phase rotation whereby a discrete phase ambiguity (quadrant ambiguity) will remain. We can take the derotated signals to decide the symbol positions $\tilde{s}_t$. In order to do the bit demapping of the datastreams we have to solve the quadrant ambiguity and the permutation problem. In this paper we neglect this problem and calculate a propper estimation of the permutation matrix $\hat{\mathbf{P}}$ by using all transmitted data. (The quadrant ambiguity problem can be solved by e.g. using differential coding, while the permutation problem can be used by including addresses that can also be used by higher layers.)

$$\hat{\mathbf{P}} = \mathrm{E}\left\{ \tilde{\mathbf{s}} \cdot \mathbf{s}^H \right\} \tag{24}$$

Using $\hat{\mathbf{P}}$ we get the assignments by successively taking the maximum absolute values. The corresponding phase factors can be used to determine the quadrant ambiguity.

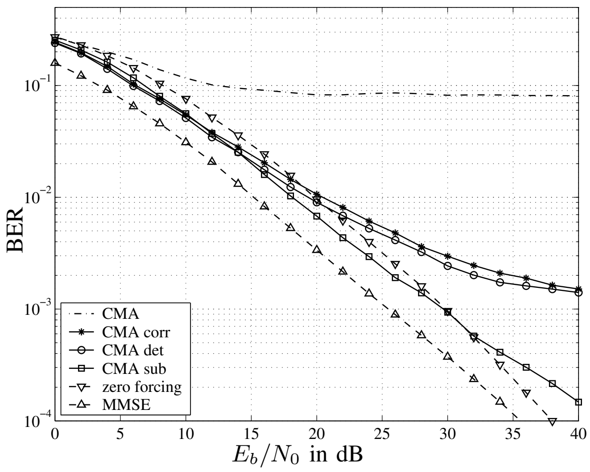

Fig. 3. BER error rates of all CMA approaches

Figure 3 depicts the simulation results of the three CMA approaches as well as two reference curves. One reference curve is the linear zero forcing filter and one is the MMSE solution with ideally known channel matrix to equalize the spatial mixture. For the CMA approaches we use a step factor of $\mu = 0.01$ and used 300 iterations (not optimized).

From **figure 3** (CMA curve) we can state that it is absolutely necessary to prevent the CMA from detecting one signal multiple times. The other 3 CMA approaches work well but suffer from an error floor due to the steepest descent algorithm; We reach BER regions where an application of channel coding should be possible. As expected the performance of the algorithms is between the zero forcing and the MMSE solution for the equalization of the spatial channel. The new subspace limitation approach has the best performance of all CMAs because it includes a very hard constraint to avoid same signals.

Figure 4 shows bit error curves of the classical source separation algorithms. They gain comparable performance to the CMA with subspace limitation, but do not show an error floor, because these algorithms don't use a steepest descent.

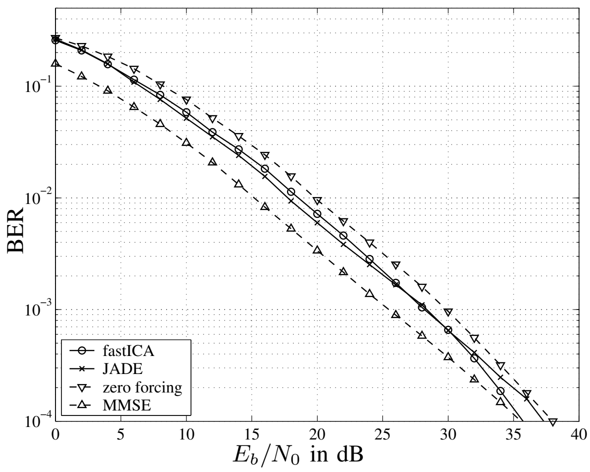

Fig. 4. BER error rates of the source separation approaches using HOS

4 Application of Iteration Techniques - uncoded

The algorithms presented until this point only approximated linear spatial filters in order to separate the data streams. In this section we try to improve the detection performance by applying a cancellation scheme. This will utilize the finite symbol alphabet that was only used marginally till now.

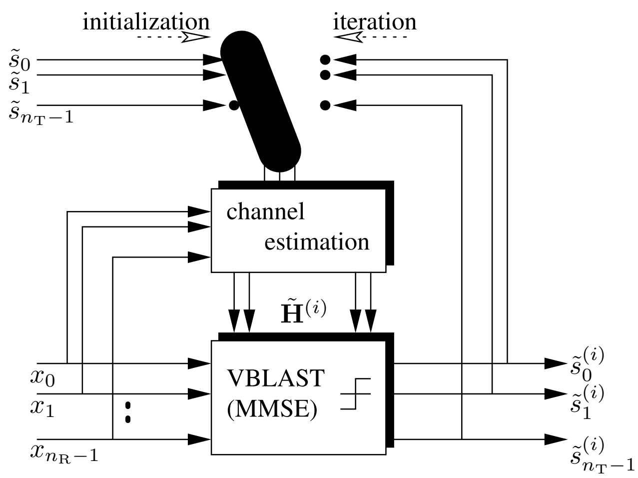

Fig. 5. improving the channel estimation / detection

We will use a system as depicted in **figure 5**. We start with coarsely decided symbols $\tilde{s}_0 \ldots \tilde{s}_{n_\mathrm{T}-1}$ that were obtained using a blind separation method as depicted in **figure 2**. Using this data we produce a first channel estimation $\tilde{\mathbf{H}}^{(0)}$. This channel estimation will be used to detect the symbols once more using the VBLAST detection algorithm as presented in [12]. We use the MMSE variant of the VBLAST algorithm but also other multi layer detection schemes are possible e.g. SQRD [13] [14].
Using the output of the VBLAST detector for improved channel estimation in combination with a new detection of the data will iteratively lead to better results.

Since this detection loop is initialized in a blind way and the VBLAST algorithm only decides symbol positions the whole detection scheme remains blind.

Because of quadrant ambiguities of the initially used symbols $\tilde{s}_0 \ldots \tilde{s}_{n_\mathrm{T}-1}$ we have to prove that the proposed scheme works. Therefore we define symbol matrices from the vectors used so far. The transmitted data streams are arranged in matrix $\mathbf{S}$

$$\mathbf{S} = [\mathbf{s}(0), \mathbf{s}(1), \ldots, \mathbf{s}(L-1)] \tag{25}$$

and the received and coarsely decided data streams in matrix $\tilde{\mathbf{S}}$

$$\tilde{\mathbf{S}} = [\tilde{\mathbf{s}}(0), \tilde{\mathbf{s}}(1), \ldots, \tilde{\mathbf{s}}(L-1)] \ . \tag{26}$$

$$\tilde{\mathbf{S}} \approx \mathbf{\Phi} \mathbf{P} \mathbf{S} \tag{27}$$

We can formulate the quadrant ambiguity of the decision by equation (27). Whereby $\mathbf{P}$ is a random permutation matrix and $\Phi = \mathrm{diag}\,[\phi_0, \phi_1, \ldots \phi_{n_T-1}]$ is a diagonal matrix modelling the discrete quadrant ambiguities. The equal sign is only valid, if there are no decision errors.

In order to use the coarse decision results $\tilde{\mathbf{S}}$ for channel estimation we calculate

$$\tilde{\mathbf{H}} = \mathbf{X} \cdot \tilde{\mathbf{S}}^{+\ 3} \ . \tag{28}$$

Assuming no noise in the system will lead to

$$\begin{aligned} \tilde{\mathbf{H}} &= \mathbf{H}\mathbf{S} \cdot \tilde{\mathbf{S}}^{+} \\ &= \mathbf{H}\mathbf{S} \cdot \tilde{\mathbf{S}}^{H} \left(\tilde{\mathbf{S}} \tilde{\mathbf{S}}^{H} \right)^{-1} \ . \end{aligned} \tag{29}$$

If we introduce equation (27) and assume correct decisions[4] we get

$$\tilde{\mathbf{H}} = \mathbf{H}\mathbf{S} \cdot \mathbf{S}^H \mathbf{P}^H \mathbf{\Phi}^H \left(\mathbf{S}\mathbf{P}\mathbf{\Phi}\mathbf{\Phi}^H \mathbf{P}^H \mathbf{S}^H \right)^{-1} \ . \tag{30}$$

Because $\mathbf{\Phi}$ and $\mathbf{P}$ are unity matrices ($\mathbf{\Phi}\mathbf{\Phi}^H = \mathbf{I}$ and $\mathbf{P}\mathbf{P}^H = \mathbf{I}$) we can simplify the expression to

$$\tilde{\mathbf{H}} = \mathbf{H}\mathbf{S} \cdot \mathbf{S}^H \mathbf{P}^H \mathbf{\Phi}^H \left(\mathbf{S}\mathbf{S}^H \right)^{-1} \ . \tag{31}$$

If we assume that $\mathbf{S}$ contains uncorrelated signal streams of sufficient length L, the terms $\mathbf{S} \cdot \mathbf{S}^H$ will be approximately diagonal matrices. Therefore we can further simplify to

$$\tilde{\mathbf{H}} = \mathbf{H} \cdot \mathbf{P}^H \mathbf{\Phi}^H \ . \tag{32}$$

This leads to an estimation of the channel matrix $\tilde{\mathbf{H}}$ in a permutated form where every column contains a quadrant error. If we apply this channel estimation in the VBLAST algorithm we get symbols $\tilde{s}_0^{(i)}, \ldots \tilde{s}_{n_\mathrm{T}-1}^{(i)}$ (**figure 5**) with the corresponding discrete phase ambiguities $\phi_0, \ldots, \phi_{n_\mathrm{T}-1}$, but this will not influence our further detection and cancellation process, as long as we only want to decide the symbol positions.

To summarize: We found an iterative estimation and detection scheme that utilizes the finite symbol alphabet and remains completely blind.

[3] The $+$ sign denotes the Moore-Penrose pseudo inverse.
[4] The terms in (27) are equal.

4.1 performance of the proposed iterative scheme

In order to show the feasibility of our detection approach we present some BER results of the transmission system presented in **figure 1**. As an initialization we exemplary use the output of the JADE algorithm. The permutation problem was solved in the same way as above (24).

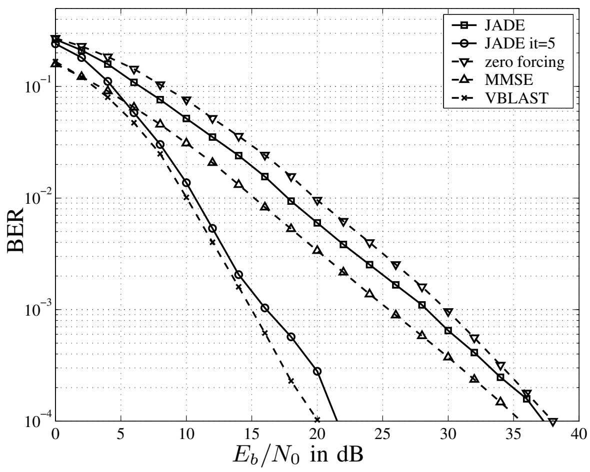

Fig. 6. BER of JADE with iterative VBLAST detection

Figure 6 depicts the results of our simulations. Beside the reference curves of the MMSE and zero forcing detection the raw VBLAST detection with ideal known channel matrix was introduced as a further reference.
Using our new iterative scheme (**figure 5**) we can observe a gain of about $10\ \mathrm{dB}$ at a bit error rate of 10^{-3} (JADE it=5) compared to the classical source separation using only the JADE algorithm. For this improvement we need only 5 iterations of detections and channel estimations. We nearly reach the curve of the VBLAST algorithm with known channel matrix $\mathbf{H}$.[5] We have to emphasize that the whole detection scheme remains blind since no reference data is used to gain the symbol decisions.

5 Inclusion of channel coding

In order to improve the overall robustness against noise and to improve the iterative estimation loop we introduce channel coding in our system. For channel coding we use the $(5,7)_8$ convolutional code with tail bits in each stream as depicted in **figure 7**.

On the receiving side we replace every decision device with a Viterbi decoder. But a simple replacement will lead to a problem: The input signal of the decision device in **figure 2** has to cope with a quadrant ambiguity of the signal. This can not work with a convolutional

[5]We can even decrease the gap to the detection with ideally known channel matrix if we increase the length of the data block L.

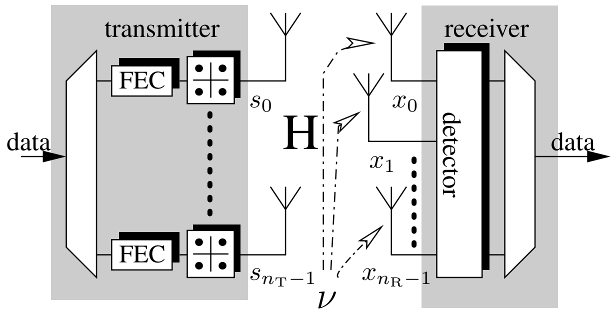

Fig. 7. transmission system with channel coding

code, because a phase rotated QPSK signal will lead to an invalid codeword. Therefore we introduce a further phase comparison with the transmitted data (after the hard decision of $\tilde{s}_0 \ldots \tilde{s}_{n_T-1}$) in order to make the Viterbi decoding process possible.
The decision device inside the iterations using the VBLAST algorithm can be replaced by a Viterbi detector without any problems.

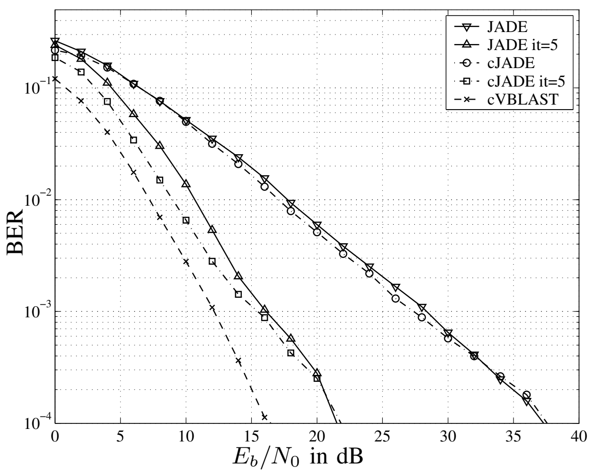

Fig. 8. BER or coded multi layer transmission

Figure 8 depicts the results of a Monte Carlo simulation using such a set up. We can find only a small coding gain between the uncoded (JADE) and the coded (cJADE) scheme. This is due to the hard decision input of the Viterbi decoder. When using VBLAST iterations in order to improve the channel estimation, we can only see a coding gain at low SNRs. This is a common behavior of coded multi layer systems with VBLAST. Since the SNR gets better, the interference cancellation process inside the VBLAST can't be improved by a coder. Based on this first analysis we can state that it is a promising approach to introduce coding in the system. But we have to cope the problem of quadrant ambiguity of the incoming signal in the input of the decision devices. Therefore further work will be done on tackling this problem.

6 Summary and Conclusions

Blind source separations techniques are promising in connection with MIMO communication systems. Their strength is that they can extract signal streams utilizing only some statistical properties of the signals. They do not need any reference data.
In this paper we presented in detail some separation methods that are based on the MIMO constant modulus property. We presented a new computational efficient CMA approach using a subspace limitation for MIMO systems.
We used the blind separation methods to initialize MIMO detection schemes and to improve channel estimation and data detection in an iterative way. Therefore we pointed out that these iterations can be done without reference data, so that the overall symbol detection algorithm remains blind. We have shown that we gain about $10\ \mathrm{dB}$ compared to the classical separation approaches and reach nearly the performance of the VBLAST detection algorithm with known channel matrix.
We presented a first step to combine the proposed iterative scheme with channel coding and pointed out directions to further improvements.

References

[1] J.-F. Cardoso and A. Souloumiac, "Blind beamforming for non Gaussian signals," *IEE Proceedings-F*, vol. 140, no. 6, pp. 362–370, December 1993.

[2] A. Cichocki and S. Amari, *Adaptive Blind Signal and Image Processing*. Wiley, 2002.

[3] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. New York: Wiley, 2001.

[4] M. Feng and K. Kammeyer, "Blind Source Separation for Communication Signals Using Antenna Arrays," in *Proc. IEEE Int. Conference on Personal Wireless Communications (ICUPC-98)*, Florence, Italy, October 1998.

[5] A. Hyvärinen and E. Oja., "A fast fixed-point algorithm for independent component analysis," *Neural Computation*, vol. 9, no. 7, pp. 1483–1492, 1997.

[6] E. Bingham and A. Hyvärinen, "A fast fixed-point algorithm for independent component analysis of complex-valued signals," *Int. J. of Neural Systems*, vol. 10, no. 1, pp. 1–8, 2000.

[7] K. D. Kammeyer, *Nachrichtenübertragung*, 2nd ed. Stuttgart: Teubner, 1996.

[8] R. Mann Pelz, W. Tobergte, and K. Kammeyer, "Applications of Constant Modulus Algorithms for Adaptive Equalization of Time-Varying Multipath Channels," in *International Conference on Digital Signal Processing*, Florence, Italy, January 1987, pp. 421–425.

[9] R. Mann Pelz and K. Kammeyer, "A Pole-Zero-Tracking Constant Modulus Algorithm," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Proc.*, Glasgow, Great Britain, January 1989, pp. 1227–1230.

[10] C. Papadias and A. Paulraj, "A constant modulus algorithm for multiuser signal separation in presence of delay spread using antenna arrays," *IEEE Signal Processing Letters*, vol. 4, no. 6, pp. 178–181, June 1997.

[11] C. Mejuto, A. Dapena, and L. Castedo, "Frequency-Domain Infomax for Blind Separation of Convolutive Mixtures," in *Second International Workshop on Independent Component Analysis and Blind Signal Separation*, June 2000, pp. 315–320.

[12] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An Architecture for Realizing Very High Data Rates Over the Rich-Scattering Wireless Channel," *invited paper, Proc. ISSSE-98*, 29. September 1998.

[13] D. Wübben, R. Böhnke, J. Rinas, V. Kühn, and K. Kammeyer, "Efficient algorithm for decoding layered space-time codes," *IEE Electronic Letters*, vol. 37, no. 22, pp. 1348–1350, November 2001.

[14] D. Wübben, J. Rinas, R. Böhnke, V. Kühn, and K. Kammeyer, "Efficient Algorithm for Detecting Layered Space-Time Codes," in *4th International ITG Conference on Source and Channel Coding*, Berlin, Germany, January 2002.

[15] A. Stuard and K. Ord, *Kendalls Advanced Theory of Statistics - Volume I: Distribution Theory*, 6th ed. New York: Wiley, 2001.