

# Blind Equalization of Frequency Selective MIMO Systems via Statistical and Trellis-based Methods

Ansgar Scherb and Karl-Dirk Kammeyer  
Department of Communications Engineering  
University of Bremen  
Otto-Hahn-Allee, D-28359 Bremen, Germany  
Email: {scherb,kammeyer}@ant.uni-bremen.de

Tianbin Wo and Peter Adam Hoehner  
Faculty of Engineering  
University of Kiel  
Kaiserstr. 2, D-24143 Kiel, Germany  
Email: {wtb,ph}@tf.uni-kiel.de

**Abstract**—In this paper we present a composite algorithm for blind sequence estimation in frequency selective multiple-input multiple-output systems. The proposed algorithm combines methods based on second and higher order statistics with a trellis-based approach. Moreover, the computational complexity of the presented method can be reduced by pre-processing the observed data with a blindly designed linear impulse shortening filter. As verified by simulations low computational effort has not necessarily led to a deterioration of performance.

## I. INTRODUCTION

Due to its high data throughput with relatively low bandwidth consumption the research on *Multiple Input Multiple Output* (MIMO) systems has become popular in the recent years [1]. However, due to the high number of coefficients characterizing the MIMO channel the issue of channel estimation has emerged as one of the bottlenecks in coherent MIMO transmission. Hence, a large amount of pilot symbols is required for accurate MIMO channel estimates. The situation becomes more crucial if the channel is frequency selective and varies rapidly over time. Hassibi and Hochwald have shown that from the information theoretic point of view purely pilot assisted channel estimation may be highly suboptimal [2].

The use of pilot symbols can be avoided by applying blind channel estimation methods at the receiver. Depending on the degree of knowledge concerning the distribution of the data source we may distinguish different classes of blind identification approaches. The lowest level of knowledge is required by statistical methods utilizing either second or higher order statistical properties of the source, e.g. [3], [4], [5], [6]. *Second Order Statistics* (SOS) methods are only capable to determine the freq. select. MIMO channel up to an instantaneous indeterminacy, whereas the remaining spatial mixture can only be separated by using *Higher Order Statistics* (HOS) methods. Considering the finite signal alphabet (and, thus, the knowledge of the source probability distribution at the receiver) we may adopt trellis based methods for joint channel estimation and data detection. In [7] it was proposed to embed an LMS adaptation of the unknown channel within the calculation of each branch metric in *Joint Viterbi Detection* (JVD). In single-input single-output systems JVD/LMS has shown considerable performance even when starting from zero

*Channel State Information* (CSI). This method still works well in freq. select. MIMO environments if it is properly initialized, whereas it may get stuck in some local optimum without initial channel estimates. However, the computational effort of Viterbi detection in freq. select. MIMO systems may become very high and it might be desirable to decrease this effort by some sort of pre-processing as e.g. linear impulse shortening techniques.

In this paper we combine statistical methods with the JVD/LMS algorithm. This approach is promising since on the one hand side an appropriate starting point for JVD/LMS may be provided by SOS and HOS methods and on the other hand side the computational complexity may be kept low by blindly adapting a linear impulse shortening filter on the basis of SOS channel estimates.

## II. CHANNEL MODEL

We consider a frequency selective MIMO system with  $N_T$  transmit and  $N_R$  receive antennas. Using the equivalent baseband representation at symbol clock the channel output at discrete time instance  $k$  can be expressed as

$$\mathbf{r}(k) = \sum_{\ell=0}^{L_H-1} \mathbf{H}(\ell)\mathbf{b}(k-\ell) + \mathbf{n}(k), \quad (1)$$

where  $\mathbf{H}(\ell) \forall \ell = 0, \dots, L_H - 1$  is a  $(N_R \times N_T)$ -matrix characterizing the frequency selective MIMO channel with maximum delay of  $L_H$  symbol periods,  $\mathbf{b}(k)$  is the channel input at time instance  $k$  according to the  $N_T$  transmit antennas, and the length  $N_R$  vector  $\mathbf{n}(k)$  is zero mean white gaussian noise with variance  $E\{\mathbf{n}(k)\mathbf{n}^H(k)\} = \sigma_n^2 \mathbf{I}_{N_R}$ . The channel input  $\mathbf{b}(k)$  is composed of a finite alphabet  $\mathcal{A}$ , e.g.  $M$ -PSK or  $M$ -QAM and assumed to be identically independent distributed (i.i.d.) in space and time. The channel is assumed to be constant over  $K + L_H - 1$  consecutive symbol periods (block fading).

We define the  $G$ -extension of a time dependent vector  $\mathbf{b}(k)$  by  $\mathbf{b}_G(k) = [\mathbf{b}^T(k), \dots, \mathbf{b}^T(k - G + 1)]^T$  and the  $G$ -convolutional extension of a freq. select. impulse response

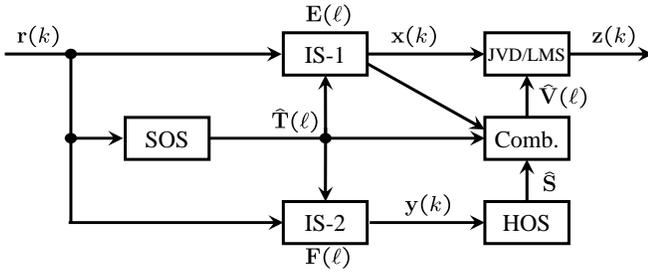


Fig. 1. Block diagram of the blind detection scheme

matrix  $\mathbf{H}(\ell)$  by the  $GN_{\text{R}} \times (G + L_{\text{H}} - 1)N_{\text{T}}$  matrix

$$\mathcal{H}_G = \begin{bmatrix} \mathbf{H}(0) & \cdots & \mathbf{H}(L_{\text{H}} - 1) & & \\ & \ddots & & \ddots & \\ & & & & \ddots \\ & & & \mathbf{H}(0) & \cdots & \mathbf{H}(L_{\text{H}} - 1) \end{bmatrix},$$

where the usage of the subscript and the calligraphic letter is a notational convention within this paper.

By the above definitions we may compactly rewrite the  $G$  extension of (1) as

$$\mathbf{r}_G(k) = \mathcal{H}_G \mathbf{b}_{G+L_{\text{H}}-1}(k) + \mathbf{n}_G(k). \quad (2)$$

### III. METHODS FOR BLIND MIMO ESTIMATION

#### A. Preliminaries

Our aim is to recover the completely unknown source signal only from the channel output without having any particular knowledge on the frequency selective MIMO channel.

A general restriction of blind multichannel identification is that the sources can only be recovered up to a complex scaling and a permutation. The complex scaling ambiguity may be reduced by exploiting the finite alphabet character of the alphabet. Thus, we call any  $\mathbf{z}(k)$  to be a proper data estimate, if its elements are from the set of admissible symbols  $\mathcal{A}$  and if it is related to the actual source by

$$\mathbf{z}(k) = \mathbf{P}\mathbf{b}(k - k_0), \quad (3)$$

where  $\mathbf{P}$  is an unknown  $N_{\text{T}} \times N_{\text{T}}$  containing only one non-zero element per each column and row and  $k_0$  is an arbitrary delay.

As mentioned above SOS methods are only able to identify the freq. select. channel up to instantaneous indeterminacy. In order to distinguish between the parts of the channel, estimated by SOS and HOS methods, we introduce the decomposition

$$\mathbf{H}(\ell) = \mathbf{T}(\ell)\mathbf{S}\mathbf{P} \quad \forall \ell = 0, \dots, L_{\text{H}} - 1 \quad (4)$$

where  $\mathbf{T}(\ell) \quad \forall \ell = 0, \dots, L_{\text{H}} - 1$  is a  $N_{\text{R}} \times N_{\text{T}}$  matrix and  $\mathbf{S}$  is a  $N_{\text{T}} \times N_{\text{T}}$  matrix not depending on the temporal parameter  $\ell$ .

#### B. Receiver Structure

The considered receiver structure is illustrated in Fig. 1. The essential part of signal detection is done in the *upper branch* consisting of the impulse shortening filter IS-1 with impulse response  $\mathbf{E}(\ell)$  and the JVD/LMS. The impulse shortening

filter IS-1 is intended for reducing the effective channel to a selectable length  $L_{\text{V}}$ , where  $\mathbf{V}(\ell)$  is the reduced length impulse response of the concatenation of the channel  $\mathbf{H}(\ell)$  and the filter  $\mathbf{E}(\ell)$ .

The required temporal characteristics of the unknown channel  $\mathbf{T}(\ell)$  is provided by the SOS-channel estimator (SOS) in the *mean branch*.

Since the JVD/LMS requires some initial channel estimates, we also have to identify the remaining spatial channel component  $\mathbf{S}$  (up to a permutation and a complex scaling), which is done in the *lower branch*. Therefore, the impulse shortening filter IS-2 with impulse response  $\mathbf{F}(\ell)$  is employed first, which in contrast to IS-1 is aimed at completely removing the *InterSymbol Interference* (ISI) from the observed signal. The ISI-free signal  $\mathbf{y}(k)$  is fed to the HOS channel estimator (HOS) identifying the remaining spatial part of the channel. Finally, the spatial and temporal channel estimates and the impulse shortening filter has to be combined (Comb.) in order to obtain an initial estimate of the reduced length channel  $\hat{\mathbf{V}}(\ell)$  for JVD/LMS.

Next, we will give a brief overview of the algorithms behind each constituent block in Figure 1.

#### C. SOS

Herein, we will give a brief description of the subspace approach for identification of MIMO systems developed by Meraim etc. [3]. This algorithm deals with determining an appropriate representation of the temporal channel characteristics  $\mathbf{T}(\ell)$  according to (4).

In consideration of the extended system model (2) the covariance matrix of the observation is given by

$$\Phi_G = \text{E} \{ \mathbf{r}_G(k) \mathbf{r}_G^H(k) \} = \mathcal{H}_G \mathcal{H}_G^H + \sigma_n^2 \mathbf{I}_{N_{\text{T}} \cdot M}. \quad (5)$$

If the size of the observation window  $G$  is sufficiently large, i.e.

$$G > \frac{N_{\text{T}}(L_{\text{H}} - 1)}{N_{\text{R}} - N_{\text{T}}} \quad (6)$$

(the number of rows in  $\mathcal{H}_G$  is larger than the number of columns) and  $\mathcal{H}_G$  has full column rank<sup>1</sup>, we may find an orthogonal basis of the subspace spanned by the columns of  $\mathcal{H}_G$  as well as a basis of the corresponding null space by singular value decomposition (SVD) of  $\Phi_G$ . Note that (6) does only have a finite solution if  $N_{\text{R}} > N_{\text{T}}$ . The SVD of the covariance matrix is given by

$$\Phi_G = [\mathbf{U}_{\text{S}} \quad \mathbf{U}_{\text{N}}] \begin{bmatrix} \mathbf{\Lambda} + \sigma_n^2 \mathbf{I}_S & \mathbf{0} \\ \mathbf{0} & \sigma_n^2 \mathbf{I}_N \end{bmatrix} \begin{bmatrix} \mathbf{U}_{\text{S}}^H \\ \mathbf{U}_{\text{N}}^H \end{bmatrix}, \quad (7)$$

where  $\mathbf{\Lambda}$  is a diagonal matrix containing the non-zero eigenvalues of  $\mathcal{H}_G \mathcal{H}_G^H$  on its main diagonal, the columns of  $\mathbf{U}_{\text{S}}$  and  $\mathbf{U}_{\text{N}}$  are the eigenvectors spanning the signal and noise space, respectively.  $S = N_{\text{T}} \cdot (M + L_{\text{H}} - 1)$  and  $N = N_{\text{R}} \cdot M - N_{\text{T}} \cdot (M + L_{\text{H}} - 1)$  are the dimensions of signal

<sup>1</sup>Identifiability conditions for blind SOS in terms of the channels  $Z$ -transform are extensively studied in [3].

and noise subspace, respectively. Due to the orthogonality of signal and noise subspace

$$\mathbf{U}_N^H \mathcal{H}_G = \mathbf{0} \quad (8)$$

holds. By defining  $\mathcal{U}_{L_H}$  as a convolutional matrix w.r.t. the uniformly partitioned noise space matrix  $\mathbf{U}_N = [\mathbf{U}^H(0), \dots, \mathbf{U}^H(G-1)]^H$  we may rearrange (8) as

$$\mathcal{U}_{L_H}^H \begin{bmatrix} \mathbf{H}(0) \\ \vdots \\ \mathbf{H}(L_H - 1) \end{bmatrix} = \mathbf{0}. \quad (9)$$

Thus, if  $\mathcal{U}_{L_H}$  has full row rank, we may find an admissible representation of the temporal channel characteristics by determining a basis of the right null space of  $\mathcal{U}_{L_H}^H$  in accordance to (4), e.g.  $\mathbf{T} = [\mathbf{T}^T(0), \dots, \mathbf{T}^T(L_H - 1)]^T$  satisfying  $\mathcal{U}_{L_H}^H \mathbf{T} = \mathbf{0}$ . The main steps of the subspace approach are summarized in Table 1.

---

**Algorithm 1** Blind SOS Channel Estimation

---

- 1: Estimate the covariance matrices  $\Phi_G$
  - 2: Find an orthogonal basis  $\mathbf{U}_N$  of the noise subspace by SVD on  $\Phi_G$
  - 3: Rearrange  $\mathbf{U}_N$  to  $\mathcal{U}_{L_H}$
  - 4: Estimate  $\mathbf{T}(\ell)$  by determining the right null space of  $\mathcal{U}_{L_H}$
- 

#### D. Impulse Shortening

In the previous section we have discussed a method for estimation estimate of the temporal channel characteristics  $\mathbf{T}(\ell)$  without using pilots symbols. Substituting  $\mathbf{T}(\ell)\mathbf{S} = \mathbf{H}(\ell)$  we may rewrite (1) as

$$\mathbf{r}(k) = \sum_{\ell=0}^{L_H-1} \mathbf{T}(\ell) \underbrace{\mathbf{S}\mathbf{b}(k-\ell)}_{\mathbf{x}(k-\ell)} + \mathbf{n}(k). \quad (10)$$

Considering a Rayleigh block fading channel model, we might assume an i.i.d. character of the coefficients of the spatial component  $\mathbf{S}$ , i.e.  $\mathbb{E}\{\mathbf{S}\mathbf{S}^H\} = \mathbf{I}_{N_T}$ . Thus, knowing  $\mathcal{T}_G$  the covariance matrix can be expressed as

$$\Phi_G = \mathcal{T}_G \mathcal{T}_G^H + \sigma_n^2 \mathbf{I}_{N_T \cdot M}. \quad (11)$$

Given this knowledge our aim is to design a linear filter  $\mathbf{E}(\ell)$  of length  $L_E$  such that the concatenation of filter and channel  $\sum_{\ell'} \mathbf{E}(\ell') \mathbf{T}(\ell - \ell')$  has less effective memory than the channel itself. Due to the lack of space in this section we only discuss the general principle of linear space-time impulse shortening without gdetailed discussion. A comprehensive overview in this field can be found in [8] and a detailed derivation of the herein proposed impulse shortening technique is given in [9].

Let

$$\mathbf{V}(\ell) = \sum_{\ell'} \mathbf{E}(\ell') \mathbf{T}(\ell - \ell' + k_0) \quad \forall 0 \leq \ell < L_V \quad (12)$$

be the reduced length target impulse response (TIR) parameterized by its length  $L_V < L_H$ , an arbitrary delay of  $k_0$

symbol periods, and the number of parallel outputs, which we fixed to  $N_T$  for the sake of convenience. Given the above parameters the TIR  $\mathbf{V}(\ell)$  still has several degrees of freedom. The ability of  $\mathbf{E}(\ell)$  to maximize the total *Signal to Interference and Noise Ratio* (SINR) at the equalizer output significantly depends on the realization of  $\mathbf{V}(\ell)$ . Thus, we have to optimize  $\mathbf{E}(\ell)$  and  $\mathbf{V}(\ell)$  in parallel. We may seek to completely remove ISI choosing  $L_V = 1$ . However, better results in terms of SINR might be obtained by only partially removing the ISI as done in case of  $L_V > 1$ . Thus, the output of TIR can be used as reference while defining the error between reference and the equalizer output by

$$\epsilon(k) = \mathcal{E}_1 \mathbf{r}_{L_E}(k) - \mathcal{V}_1 \mathbf{x}_{L_V}(k - k_0). \quad (13)$$

In terms of minimum mean squared error (MMSE) our goal is to determine the pair

$$(\mathcal{E}_1, \mathcal{V}_1) = \arg \min_{\mathcal{E}_1, \mathcal{V}_1} \mathbb{E} \{ \epsilon^H(k) \epsilon(k) \} \quad \text{s.t. } \mathcal{V}_1 \mathcal{V}_1^H = \mathbf{I}_{N_T}. \quad (14)$$

The orthogonality constraint (ONC) in (14) was introduced, in order to avoid the trivial solution  $\mathcal{E}_1 = \mathbf{0}$ . Indeed, also other constraints were proposed in past [8], though ONC seems to be the most reasonable choice in regard to SINR. Substituting (13) after little algebra the squared mean error in (14) can be expressed as

$$\begin{aligned} \mathbb{E} \{ \epsilon^H(k) \epsilon(k) \} &= \| (\mathcal{E}_1 \Phi_{L_E} - \mathcal{V}_1 \Delta_{k_0} \mathcal{T}_G^H) \Phi_{L_E}^{-1/2} \|_F^2 \\ &+ \| \mathcal{V}_1 \Delta_{k_0} (\mathcal{T}_{L_E}^H \mathcal{T}_{L_E} + \sigma_n^2 \mathbf{I}_{N_T \cdot L_E})^{-1} \Delta_{k_0}^H \mathcal{V}_1^H \|_F^2, \end{aligned} \quad (15)$$

where  $\Delta_{k_0} = [\mathbf{0}_{L_V N_T \times k_0 N_T} \quad \mathbf{I}_{L_V N_T} \quad \mathbf{0}_{L_V N_T \times \delta N_T}]^2$ . Apparently, the first line in (15) can be forced to zero by

$$\mathcal{E}_1 = \mathcal{V}_1 \Delta_{k_0} \mathcal{T}_G^H \Phi_{L_E}^{-1}, \quad (16)$$

whereas the second line represents the remaining error only depending on  $\mathcal{V}_1$ . An appropriate TIR  $\mathcal{V}_1$  minimizing this term subject to the ONC can be obtained by calculating the right eigenvectors of  $\Theta = \Delta_{k_0} (\mathcal{T}_{L_E}^H \mathcal{T}_{L_E} + \sigma_n^2 \mathbf{I}_{N_T \cdot L_E})^{-1} \Delta_{k_0}^H$  corresponding to the  $N_T$  smallest eigenvalues. The ONC design of the impulse shortening filter is summarized in Table 2.

---

**Algorithm 2** ONC design of the impulse shortening

---

- 1: Determine  $\Theta$
  - 2: Determine an appropriate TIR  $\mathbf{V}(\ell)$  by EVD on  $\Theta$
  - 3: Determine the equalizer  $\mathbf{E}(\ell)$  by (16)
- 

#### E. HOS

Whereas the former two section have dealt with the estimation and equalization of the temporal channel characteristics, in this section we focus only on identifying the spatial part. Starting from the decomposition (4), we have

$$\mathbf{x} = \mathbf{S}\mathbf{b} \quad (17)$$

$${}^2\delta = L_E + L_H - L_V + k_0$$

where  $\mathbf{b}^3$  is a vector with  $N_T$  statistically independent random variables as entries,  $\mathbf{S}$  is an  $N_T \times N_T$  invertible spatial mixing matrix and  $\tilde{\mathbf{n}}$  is noise.

The general approach of independent component analysis (ICA) consists of the two steps at first whitening the observed signal  $\mathbf{x}$ , e.g.  $\tilde{\mathbf{x}} = \mathbf{C}\mathbf{x}$  with  $\mathbf{E}\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}^H\} = \mathbf{I}_{N_T}$  and afterwards finding an unitary separation matrix  $\mathbf{A}$  such that the components of  $\mathbf{y} = \mathbf{A}\tilde{\mathbf{x}}$  are mutually independent. While the whitening matrix  $\mathbf{C}$  might be easily found by the inverse Cholesky factors of the covariance matrix  $\mathbf{E}\{\mathbf{x}\mathbf{x}^H\}$ , we have to establish a statistical measurement of the independence for the second part of ICA. With  $\mathbf{b}$  being an independent non gaussian random vector (as stated above), we call an empirical function  $f(\mathbf{b})$  to be an appropriate contrast for ICA, if  $f(\mathbf{b}) < f(\mathbf{P}\mathbf{b})$  is satisfied for any matrix  $\mathbf{P}$  with equality only if  $\mathbf{z} = \mathbf{P}\mathbf{b}$  is a permutation of  $\mathbf{b}$  according to the definition in (3). In other words the contrast becomes minimum, if there is no statistical connection among the single elements of  $\mathbf{z}$ .

An efficient batch algorithm referred to as *Joint Approximate Diagonalization of Eigen-matrices* (JADE) was proposed by Cardoso [5]. This method is essentially based on the fourth order cumulant contrast

$$f^{\text{JADE}}(\mathbf{y}) = \sum_{ijkl \neq iikl} |\text{cum}(y_i, y_j^*, y_k, y_l^*)|^2. \quad (18)$$

Cumulants are statistical parameters describing the particular distribution of random variables. The fourth order cumulant of zero mean and symmetrically distributed random variables (as given in our case) is defined by

$$\text{cum}(y_i, y_j^*, y_k, y_l^*) \quad (19)$$

$$= \mathbf{E}\{y_i y_j^* y_k y_l^*\} - \mathbf{E}\{y_i y_j^*\} \mathbf{E}\{y_k y_l^*\} \quad (20)$$

$$- \mathbf{E}\{y_i y_k\} \mathbf{E}\{y_j^* y_l^*\} - \mathbf{E}\{y_i y_l^*\} \mathbf{E}\{y_k y_j^*\}. \quad (21)$$

Defining the cumulant matrix

$$\mathbf{Q}_{ij} = \mathbf{E}\{x_i x_j^* \mathbf{x}\mathbf{x}^H\} - \mathbf{E}\{x_i x_j^*\} \mathbf{E}\{\mathbf{x}\mathbf{x}^H\} \quad (22)$$

$$- \mathbf{E}\{x_i \mathbf{x}\} \mathbf{E}\{x_j^* \mathbf{x}^H\} - \mathbf{E}\{x_j^* \mathbf{x}\} \mathbf{E}\{x_i \mathbf{x}^H\} \quad (23)$$

we may express the desired cumulants of  $\mathbf{y} = \mathbf{A}\mathbf{x}$  with  $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_{N_T}]$  being an arbitrary unitary matrix as  $\text{cum}(y_i, y_j^*, y_k, y_l^*) = \mathbf{a}_k^H \mathbf{Q}_{ij} \mathbf{a}_l$ . Thus, the above contrast is minimized by diagonalizing jointly the set of all cumulant matrices  $\{\mathbf{Q}_{ij} | 0 < i, j < N_T\}$ . A common technique for finding those left and right multipliers jointly diagonalizing a set of Hermitic matrices is the Jacoby algorithm. The obtained matrix  $\mathbf{A}$  is the desired separation matrix, whereas  $\mathbf{C}^{-1} \mathbf{A}^H = \hat{\mathbf{S}}$  is an estimate of spatial mixing matrix. The steps of JADE are summarized in 3.

### F. JVD/LMS

Starting from the reduced length channel model

$$\mathbf{r}(k) = \mathcal{W}_1 \mathbf{z}_{L_W}(k) + \tilde{\mathbf{n}}(k) \quad (24)$$

<sup>3</sup>For the sake of readability we have dropped the time index  $k$  in the following.

### Algorithm 3 JADE

- 1: Determine the pre-whitening matrix  $\mathbf{C}$  by Cholesky decomposition
- 2: Estimate the cumulant matrix set  $\{\mathbf{Q}_{ij} | 0 < i, j < N_T\}$
- 3: Determine the separation matrix  $\mathbf{A}$  by jointly diagonalizing  $\{\mathbf{Q}_{ij} | 0 < i, j < N_T\}$  via Jacoby algorithm

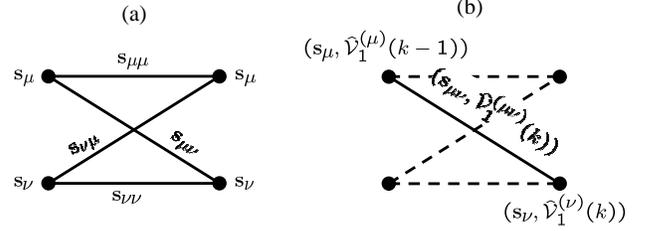


Fig. 2. Trellis segment in JVD (a) and JVD/LMS (b)

where  $\mathbf{W}(\ell) = \sum_{\ell'} \mathbf{E}(\ell') \mathbf{T}(\ell - \ell' + k_0) \mathbf{A}$ ,  $L_W = L_V$  and  $\tilde{\mathbf{n}}(k) = \mathcal{E}_1 \mathbf{n}_{L_E}(k - k_0)$  is white noise, the maximum likelihood sequence estimator is given by

$$[\hat{\mathbf{z}}(0), \dots, \hat{\mathbf{z}}(K)] = \arg \min \sum_k \|\mathbf{r}(k) - \mathcal{W}_1 \mathbf{z}_{L_W}(k)\|^2. \quad (25)$$

Due to the finiteness of the signal alphabet and the convolutional structure of  $\mathcal{W}_{L_W}$  the desired sequence can be efficiently found by trellis search techniques, e.g. the joint Viterbi detector (JVD). In Fig. 2 (a) a segment of a trellis diagram is illustrated.

Each branch in Fig. 2 connecting two states is associated to a hypothesis for  $\mathbf{z}_{L_W}(k)$ , where  $\mathbf{z}_{\mu \rightarrow \nu}$  denotes the hypothesis corresponding to the branch from the  $\mu$ -th state to the  $\nu$ -th state. The standard Viterbi algorithm calculates (25) by recursively updating the metric

$$m_\nu(k) = \min_\mu (m_\mu(k-1) + \|\mathbf{r}(k) - \mathcal{W}_1 \mathbf{z}_{\mu \rightarrow \nu}\|^2). \quad (26)$$

In order to deal with non or imperfect CSI, the JVD/LMS approach embeds a LMS channel adaptation within the metric calculations. As illustrated in Fig. 2 (b) the estimate  $\hat{\mathcal{V}}_1^{(\mu)}(k)$  is assigned to the  $\mu$ -th node at instance  $k$  and estimate  $\hat{\mathcal{V}}_1^{(\mu\nu)}(k)$  is assigned to the branch connecting the  $\mu$ -th node at instance  $k-1$  and the  $\nu$ -th node at instance  $k$ . According to the particular node and the instance  $k$  each of the considered estimates corresponds to a different hypothetical signal sequence. Given the previously obtained estimate  $\hat{\mathcal{V}}_1^{(\mu)}(k-1)$  and the step size  $\Delta$  the LMS adaptation yields

$$\begin{aligned} \hat{\mathcal{W}}_1^{(\mu \rightarrow \nu)}(k) & \quad (27) \\ &= \hat{\mathcal{V}}_1^{(\mu)}(k-1) + \Delta \left( \mathbf{r}(k) - \hat{\mathcal{V}}_1^{(\mu)}(k-1) \mathbf{z}_{\mu \rightarrow \nu} \right) \mathbf{z}_{\mu \rightarrow \nu}^H. \end{aligned}$$

Replacing  $\mathcal{V}_1$  by the estimate the metric calculations becomes

$$m_\nu(k) = \min_\mu (m_\mu(k-1) + \|\mathbf{r}(k) - \hat{\mathcal{V}}_1^{(\mu \rightarrow \nu)}(k) \mathbf{z}_{\mu \rightarrow \nu}\|^2) \quad (28)$$

and with  $\mu_{\min}$  being the index minimizing (28) the channel estimate associated to the  $\nu$ -th state is given by

$$\hat{\mathcal{W}}_1^\nu(k) = \hat{\mathcal{V}}_1^{\mu_{\min} \rightarrow \nu}(k). \quad (29)$$

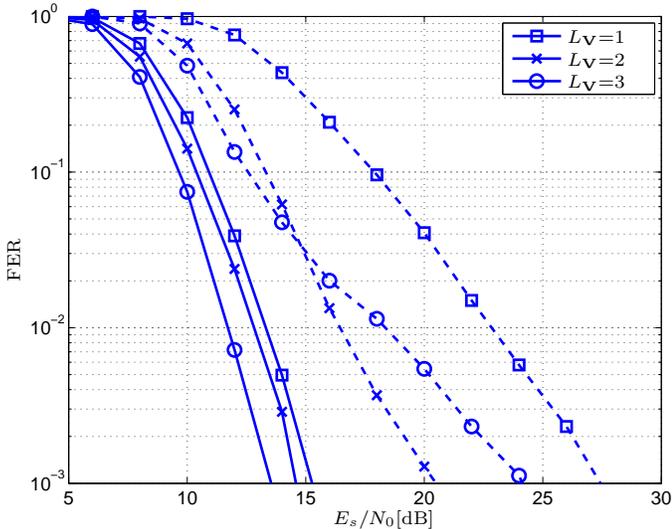


Fig. 3. BER vs. SNR with perfect CSI (solid line) and blind (dashed line); system parameter: BPSK, ( $N_R = 6 \times N_T = 4$ ) MIMO channel of length  $L = 3$  with i.i.d. gaussian channel gains, frame length  $K = 200$

The JVD/LMS algorithm consists of the following steps summarized in Alg. 4.

---

#### Algorithm 4 JVD/LMS

---

- 1: Initialize  $m_\mu(0) = 0$  and  $\hat{W}_1^{(\mu)}(0)$  (by zeros or an initial estimate) for any  $\mu$ .
  - 2: **for**  $k=1:K+L-1$  **do**
  - 3: Perform a LMS adaption of the channel estimates by (27) for any admissible state transition  $\mu \rightarrow \nu$
  - 4: Calculate the metric  $m_\mu(k)$  by (28) for any  $\mu$
  - 5: Update channel estimates by (29)
  - 6: **end for**
- 

#### IV. NUMERICAL RESULTS

As already mentioned in the previous section, blind data detection always suffers from a permutation and phase ambiguity. In order to fade out these effects, we have fitted the final detection result in the best way in terms of possible permutations and complex scalings. Figure 3 shows the frame error rate vs. the SNR for different target impulse lengths  $L_V$ . A reference for the proposed blind algorithm is given by the solid curves, where we presumed perfect CSI. The dashed curves exhibit the results of the herein proposed blind receiver. The case  $L_V = 1$  involves that within the JVD/LMS step instantaneous maximum likelihood detection and LMS channel adaption alternates. Comparing the FER of  $L_V = 1$  and  $L_V = 2$  an apparent gain by approximately 8 dB can be observed, whereas contrary to the reference curves the performance in case of  $L_V = 3$  might become worse than in case of  $L_V = 2$ . Obviously, JVD/LMS is more sensitive to an initial channel estimation error when the considered effective channel order is large. Therefore, the prepended impulse shortening filter IS-1 may not only provide lower

computational complexity but also deliver better performance in terms of FER.

#### V. CONCLUSION

The proposed scheme combining statistical and trellis-based methods for equalizing and estimating blindly frequency selective MIMO systems has shown considerable performance. In fact, purely statistical methods suffer from the disregard of the a-priori known source statistics (finite alphabet character), whereas purely trellis-based approaches cannot be applied in MIMO environment without appropriate initialization. Furthermore, the computational effort consumed by Viterbi detection can be drastically reduced by applying a blindly designed linear impulse shortening filter. Thus, our method provides low computational complexity as well as high quality data estimates.

#### REFERENCES

- [1] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," in *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585-595, 1999.
- [2] B. Hassibi, and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?," in *IEEE Trans. on Information Theory*, vol. 49, no. 4, pp. 951-963, April 2003.
- [3] K. A. Meraim, P. Loubaton, and E. Moulines, "A subspace algorithm for certain blind identification problems," in *IEEE Trans. on Information Theory*, vol. 43, no. 2, pp. 499-511, March 1997.
- [4] L. Tong, and S. Perreau, "Multichannel blind identification: From subspace to maximum likelihood method," in *Proceedings of the IEEE*, October 1998.
- [5] J.-F. Cardoso, "High-Order Contrasts for Independent Component Analysis," in *Neural Computations*, vol. 11, pp. 157-192, 1999.
- [6] A. Hyvrinen and E. Oja, "Independent Component Analysis: A Tutorial," in *Neural Computations*, vol. 13, pp. 411-430, 2000.
- [7] X.-M. Chen, and P. A. Hoeher, "Trellis-based iterative adaptive blind sequence estimation for uncoded/coded systems with differential precoding," in *EURASIP Journal on Applied Signal Processing*, vol. 2005, no. 6, pp. 828-843, May 2005.
- [8] D. Wuebben, and K. D. Kammeyer, "Impulse Shortening and Equalization of Frequency-Selective MIMO Channels with Respect to Layered Space-Time Architectures," in *EURASIP Signal Processing, Volume 83, Issue 8*, pp. 1643-1659, August 2003
- [9] N. Al-Dhahir, "FIR channel-shortening equalizers for MIMO ISI channels," in *IEEE Trans. Commun.*, vol. 49, no. 2, pp. 213-218, 2001.