DASBE: Decision-Aided Semi-Blind Equalization for MIMO Systems with Linear Precoding

José A. García-Naya, Adriana Dapena,[*] Paula M. Castro, Daniel Iglesia

Abstract: *Multiple-Input Multiple-Output* (MIMO) digital communications standards usually acquire *Channel State Information* (CSI) by means of supervised algorithms, which implies a loss of performance since pilot symbols do not convey information. We propose to obtain this CSI by using the so-called semi-blind techniques, which combine both supervised and unsupervised (blind) methods. The key idea consists in introducing a decision criterion to determine when the channel has suffered a significant change. In such a case, transmission of pilot symbols is required. The use of this criterion also allows us to determine the time instants in which CSI has to be sent to the transmitter from the receiver through a low-cost feedback channel.

Key words: *Linear Precoding, MIMO, Blind Source Separation, Hybrid Systems.*

1. Introduction

The main task when transmitting over channels with multiple antennas at the transmitter and/or at the receiver side is the separation or equalization of the transmitted data. *Linear Transmit Processing* (LTP), also termed *Linear Precoding* (LP), is a powerful method to separate signals in *Multiple-Input Multiple-Output* (MIMO) systems since it reduces computational costs and power consumption at the receiver end. Thus, the equalization task is performed at the transmitter, so the channel is pre-equalized or *precoded* before transmission with the goal of simplifying one side of the link and avoiding filter operations at the receiver. Such an operation prior to transmission is only possible for a centralized transmitter, e.g. the base-station in the downlink of a cellular system. Moreover, in case of a multiuser scenario with non-cooperative receivers, the users cannot cooperatively transform the received signals. Thus, transmit filters are necessary to separate signals for different users before transmission through a fading channel. Therefore, the advantages of carrying out this pre-equalization of channel effects at the transmitter are clear, compared to traditional receiver and equalization alternatives. Although *Wiener Filtering* (WF) for precoding has been dealt with by only a few authors [15] compared to other criteria, Wiener linear precoding is an attractive transmit optimization that minimizes the *Mean Square Error* (MSE) between the transmitted and received symbols [8, 14, 17, 19].

[*]Corresponding author: Adriana Dapena, University of A Coruña.
Postal address: Facultad de Informática, Campus de Elviña, 15071 A Coruña, Spain.
e-mail: adriana.dapena@udc.es

The design of LP schemes has been widely studied for the ideal case in which *Channel State Information* (CSI) is perfectly known at the transmitter side [8, 14, 17, 19]. However, for transmit processing the major difficulty is the availability of instantaneous CSI at the transmitter. Thus, this work is focused on determining channel changes which will allow us to update the CSI available at the transmitter side by sending appropriate information through a reverse (also termed *feedback*) channel. Most recent wireless communications standards include such a feedback channel for sending user link parameters. For example, *Worldwide Interoperability for Microwave Access* (WiMAX) standard uses this channel to send an index for selecting the most adequate code according to channel conditions [11]. However, to the knowledge of the authors, none of the current standards —even those under development— make use of such information to decide whether pilot symbols must be sent or not.

In this work, we propose a novel approach termed *Decision-Aided Semi-Blind Equalization* (DASBE), which allows us to reduce penalizations introduced by the use of pilot symbols and to get an efficient utilization of the feedback channel. The main difficulty is to detect channel variations by using a simple decision criterion. The proposed criterion compares the channel matrix estimated using an unsupervised algorithm, to the previous estimate obtained by a supervised one. Pilot symbols are required only when channel variations are significant with respect to a previously selected threshold. A similar scheme has been proposed by the authors in [7, 10], where the transmitter could send two types of frames: frames containing only pilot symbols or frames containing only user symbols. This setup differs from the frame structure used in current standards, in which frames are composed by both pilot and user symbols. For this reason, the decision criterion proposed in DASBE is used to determine the instants where pilot symbols can be eliminated (or reduced) in standards frames.

This work is organized as follows. Section 2. shows our digital communications system. Section 3. reviews some supervised and unsupervised algorithms for channel estimation and source data recovery. Section 4. proposes the DASBE approach. Representative computer simulations are presented in Section 5. and Section 6. states some concluding remarks.

Vectors and matrices are denoted by lower case bold and capital bold letters, respectively. We use $\mathrm{E}[.]$, $\mathrm{tr}(.)$, $(.)^*$, $(.)^{\mathrm{T}}$, $(.)^{\mathrm{H}}$, $\det(.)$, $\ln(.)$ and $\|.\|_2$ for expectation, trace of a matrix, complex conjugation, transposition, conjugate transposition, determinant of a matrix, natural logarithm and Euclidean norm, respectively. The i-th element of a vector $\boldsymbol{x}$ is x_i. $h(.)$ is used to denote a scalar function and $h'(.)$ and $h''(.)$ denote its first and second dereivatives.

2. System Model

We consider a MIMO system with N_{t} transmit antennas and N_{r} receive antennas, as plotted in Figure 1. The user symbols are expressed as $\boldsymbol{u}[n] = [u_1[n], \ldots, u_{N_{\mathrm{r}}}[n]]^{\mathrm{T}}$ and they are used by the encoder to generate the transmitted signal denoted by $\boldsymbol{x}(n) = [x_1(n), \ldots, x_{N_{\mathrm{t}}}(n)]^{\mathrm{T}}$. Suppose that $x_i(n)$ is transmitted from the i–th transmit antenna to the j–th receive antenna through the path $h_{ji}[q]$. Thus, we

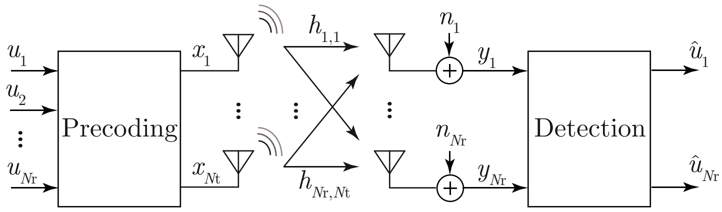

Fig. 1 *System with precoding over flat MIMO channels.*

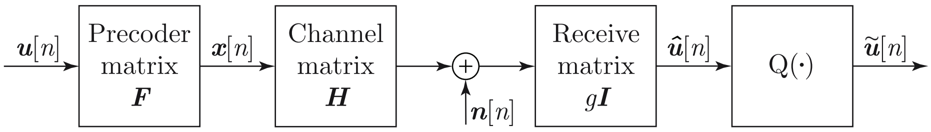

Fig. 2 *MIMO system with linear precoding.*

have that the received signals (observations) presents the form

$$\boldsymbol{y}[n] = \boldsymbol{H}[q]\boldsymbol{x}[n] + \boldsymbol{n}[n], \tag{1}$$

where $n = 0, 1, 2, \ldots$ corresponds to sample index, q denotes time-slot, and $\boldsymbol{n}(n) = [n_1(n), \ldots, n_{N_\mathrm{r}}(n)]^\mathrm{T}$ contains *Additive White Gaussian Noise* (AWGN) with covariance matrix $\boldsymbol{C_n}$. We assume that the sources are transmitted in frames of N_B symbols and that the channel remains constant during several frames, i.e. the index q is unchanged during those frames. It can be demonstrated that this discrete-time model is equivalent to the continuous-time one only if *Inter–Symbol Interference* (ISI) between samples is avoided, i.e. if the Nyquist criterion is satisfied. In that case, we are able to reconstruct the original continuous signal from samples by means of interpolation. Hereafter, we assume this channel model, known as time-varying flat block fading channel.

As it was mentioned above, in order to simplify the requirements at the receiver, the equalization task can be performed at the transmitter so the channel is precoded before transmission, as plotted in Figure 1. Such an operation —prior to transmission— is only possible when a centralized transmitter is used (e.g. the base-station for the downlink of a cellular system). The goal is to find the optimum transmit and receive filters, $\boldsymbol{F} \in \mathbb{C}^{N_\mathrm{t} \times N_\mathrm{r}}$ and $\boldsymbol{G} = g\boldsymbol{I} \in \mathbb{C}^{N_\mathrm{r} \times N_\mathrm{r}}$, respectively. Note that N_r is the number of scalar data streams. The resulting communications system is shown in Figure 2, in which the data symbols $\boldsymbol{u}[n]$ are passed through the transmit filter $\boldsymbol{F}$ to form the transmit signal $\boldsymbol{x}[n] = \boldsymbol{F}\boldsymbol{u}[n] \in \mathbb{C}^{N\mathrm{t}}$. The constraint for the transmit energy must be fulfilled, i.e.

$$E\left[||\boldsymbol{x}[n]||_2^2\right] = \mathrm{tr}\left(\boldsymbol{F}\boldsymbol{C_u}\boldsymbol{F}^\mathrm{H}\right) \leq E_\mathrm{tr},$$

where $\boldsymbol{C_u} = \mathrm{E}[\boldsymbol{u}[n]\boldsymbol{u}^\mathrm{H}[n]]$ is the correlation between the uncoded symbols and E_tr is the total transmitted energy. Thus, the received signal is given by

$$\boldsymbol{y}[n] = \boldsymbol{H}\boldsymbol{F}\boldsymbol{u}[n] + \boldsymbol{n}[n]. \tag{2}$$

After multiplying by the gain g, we get the estimated symbols

$$\hat{\boldsymbol{u}}[n] = g\boldsymbol{y}[n] = g\boldsymbol{H}\boldsymbol{F}\boldsymbol{u}[n] + g\boldsymbol{n}[n] \in \mathbb{C}^{N_\mathrm{r}}. \tag{3}$$

Therefore, the implementation of precoded systems implies very simple receivers since the observations are only multiplied by the gain factor g. Clearly, the restriction about common weights g for all the receivers is not necessary in case of decentralized receivers.

As mentioned before, we consider Wiener linear filtering, whose optimization consists in minimizing the MSE with a transmitted energy constraint, i.e.

$$\{\boldsymbol{F}_\mathrm{WF}, g_\mathrm{WF}\} = \arg\min_{\{\boldsymbol{F},g\}} \mathrm{E}\left[||\boldsymbol{u}[n] - \hat{\boldsymbol{u}}[n]||_2^2\right] \quad \text{s.t.: } \mathrm{tr}(\boldsymbol{F}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{F}^\mathrm{H}) \leq E_\mathrm{tr}. \tag{4}$$

Note that such a constraint is necessary to avoid the dependence of the resulting transmitted energy on the channel realization. So, the transmitted energy constraint mentioned above might be the maximum value for poor channel realizations and thus the respective precoder solution is not valid. The transmitter may also not use the whole available transmitted energy, and therefore the final quality would not be as good as possible, since it could be improved by using more transmitted energy. In [8, 17], it is shown that the solution for the linear filters designed using that MSE criterion is expressed as

$$\begin{aligned} \boldsymbol{F}_\mathrm{WF} &= g_\mathrm{WF}^{-1}\left(\boldsymbol{H}^\mathrm{H}\boldsymbol{H} + \psi\boldsymbol{I}\right)^{-1}\boldsymbol{H}^\mathrm{H}, \\ g_\mathrm{WF} &= \sqrt{\frac{\mathrm{tr}\left(\left(\boldsymbol{H}^\mathrm{H}\boldsymbol{H} + \psi\boldsymbol{I}\right)^{-2}\boldsymbol{H}^\mathrm{H}\boldsymbol{C}_{\boldsymbol{u}}\boldsymbol{H}\right)}{E_\mathrm{tr}}}, \end{aligned} \tag{5}$$

where $\psi = \frac{\mathrm{tr}(\boldsymbol{C}_{\boldsymbol{n}})}{E_\mathrm{tr}}$.

3. Source Data Recovery Methods

Current digital communications standards define a frame as a sequence of pilot and user symbols. Supervised algorithms use the pilot symbols to estimate the channel (and to recover the transmitted signals), while unsupervised (blind) approaches discard this information [9]. In particular, in order to recover transmitted signals (sources), we will use a linear system whose weight matrix $\boldsymbol{W}[n] \in \mathbb{C}^{N_\mathrm{r}\times N_\mathrm{r}}$ (termed also *recovering matrix*) will be obtained using some supervised or unsupervised algorithm. The outputs of this system are computed using

$$\boldsymbol{z}[n] = \boldsymbol{W}^\mathrm{H}[n]\boldsymbol{y}[n]. \tag{6}$$

3.1 Supervised Approach

We consider the utilization of a supervised approach to estimate the channel matrix $\boldsymbol{H}$ using as a reference the model in Eq. (1), in which $\boldsymbol{y}[n]$ and $\boldsymbol{x}[n]$ represent observations and sources, respectively.

An important family of adaptive filtering algorithms arises from the minimization of the MSE between the outputs, $\boldsymbol{z}[n]$, and the sources, $\boldsymbol{x}[n]$ [13, 18]. Mathematically, the cost function is defined as

$$\begin{aligned} J_{\mathrm{MSE}} &= \sum_{i=1}^{N_{\mathrm{B}}} \mathrm{E}\left[|z_i[n] - x_i[n]|^2\right] \\ &= \mathrm{E}\left[\mathrm{tr}\left((\boldsymbol{W}^{\mathrm{H}}[n]\boldsymbol{y}[n] - \boldsymbol{d}[n])(\boldsymbol{W}^{\mathrm{H}}[n]\boldsymbol{y}[n] - \boldsymbol{x}[n])^{\mathrm{H}}\right)\right]. \end{aligned} \tag{7}$$

Then, the recovering matrix is updated using the following gradient algorithm

$$\boldsymbol{W}[n+1] = \boldsymbol{W}[n] - \mu\ \nabla_{\boldsymbol{W}} J_{\mathrm{MSE}}[n], \tag{8}$$

where $\nabla_{\boldsymbol{W}} J_{\mathrm{MSE}}$ is the gradient of J_{MSE} with respect to $\boldsymbol{W}$, i.e.

$$\nabla_{\boldsymbol{W}} J_{\mathrm{MSE}} = \mathrm{E}\left[\boldsymbol{y}[n](\boldsymbol{W}^{\mathrm{H}}[n]\boldsymbol{y}[n] - \boldsymbol{x}[n])^{\mathrm{H}}\right]. \tag{9}$$

The classical stability analysis for gradient–based algorithms consists in finding the point in which the gradient vanishes and defining the Hessian matrix whose coefficients are given by the second derivatives of J [4]. In particular, it can be demonstrated that the stationary points of the rule defined by Eq. (8) are

$$\nabla_{\boldsymbol{W}} J_{\mathrm{MSE}} = 0 \Rightarrow \boldsymbol{W} = \boldsymbol{C}_{\boldsymbol{y}}^{-1}\boldsymbol{C}_{\boldsymbol{yx}}, \tag{10}$$

where $\boldsymbol{C}_{\boldsymbol{y}} = \mathrm{E}[\boldsymbol{y}[n]\boldsymbol{y}^{\mathrm{H}}[n]]$ is the autocorrelation of the observations and $\boldsymbol{C}_{\boldsymbol{yx}} = \mathrm{E}[\boldsymbol{y}[n]\boldsymbol{x}^{\mathrm{H}}[n]]$ is the cross-correlation between the observations and the desired signals. In practice, these desired signals are considered as known only during a finite number of instants (pilot symbols) in which the estimation is used to recover the transmitted symbols. For this reason, the performance of this type of algorithms is degraded in the presence of calibration errors.

3.2 Unsupervised Approach

The transmission of pilot symbols and the prior knowledge about channel matrices can be avoided by using *Blind Source Separation* (BSS) algorithms [5, 9, 16]. BSS methods simultaneously estimate the mixing matrix and the realizations of the source vector. In particular, we consider the model given by Eq. (2), where $\boldsymbol{y}[n]$ and $\boldsymbol{u}[n]$ represent observations and sources, respectively. The joint matrix $\boldsymbol{HF}$ is the matrix to be estimated.

One of the best known BSS algorithms has been approached by Bell and Sejnowski [3]. The idea proposed by these authors is to obtain the weighted coefficients of a Artificial Neural Network, $\boldsymbol{W}[n]$, in order to maximize the mutual information (MI) between the outputs before the activation function $\boldsymbol{h}(\boldsymbol{z}[n]) = \boldsymbol{h}(\boldsymbol{W}^{\mathrm{H}}[n]\boldsymbol{y}[n])$, where $h(.)$ is the activation function, and the inputs $\boldsymbol{y}[n]$. The resulting cost function is given by

$$J_{\mathrm{MI}}(\boldsymbol{W}[n]) = \ln(\det(\boldsymbol{W}^{\mathrm{H}}[n])) + \sum_{i=1}^{N_t} \mathrm{E}[\ln(h_i'(z_i[n]))], \tag{11}$$

where h_i is the i–th element of the vector $\boldsymbol{h}(\boldsymbol{z}[n])$, and $'$ denotes the first derivative. The maximum of this cost function can be obtained using a relative gradient algorithm [1, 2], which gives

$$\begin{aligned}\boldsymbol{W}[n+1] &= \boldsymbol{W}[n] + \mu \boldsymbol{W}[n]\boldsymbol{W}^{\mathrm{H}}[n]\left(\boldsymbol{z}[n]\ \boldsymbol{f}^{\mathrm{H}}(\boldsymbol{y}[n]) - \boldsymbol{W}^{-\mathrm{H}}[n]\right)\\ &= \boldsymbol{W}[n] + \mu \boldsymbol{W}[n]\left(\boldsymbol{z}[n]\boldsymbol{f}^{\mathrm{H}}(\boldsymbol{z}[n]) - \boldsymbol{I}\right). \end{aligned} \tag{12}$$

where $\boldsymbol{f}(\boldsymbol{z}) = [-h''(z_1)/h'(z_1), ..., -h''(z_{N_r})/h'(z_{N_r})]^T$. The expression in Eq. (12) admits an interesting interpretation by means of the use of the non-linear function $f(z) = z^*(|z|^2 - 1)$. In this case, Castedo and Macchi [6] have shown that the Bell and Sejnowski rule can be interpreted as an extension of the *Constant Modulus Algorithm* (CMA) proposed by Godard [12].

4. Decision-Aided Semi-Blind Equalization (DASBE)

Recent digital communications standards include a low-cost feedback channel which can be used to send estimates obtained using a supervised approach. Using this information, the transmitter adapts the precoding matrix $\boldsymbol{F}$ according to existing channel conditions. This approach has several limitations: firstly, transmission of pilot symbols penalizes throughput and secondly, as a consequence, overhead of the feedback channel appears in case of CSI must be sent from the receiver each time a new frame is acquired. In addition, a large number of pilot symbols is needed to guarantee the convergence of the adaptive algorithm in Eq. (8) or to ensure that the matrix $\boldsymbol{C_y}$ in Eq. (10) is not singular.

In this section, we present the novel DASBE approach, which combines supervised and unsupervised techniques to mitigate the limitations found in classical approaches. We denote by $\boldsymbol{W_u}[n]$ and $\boldsymbol{W_s}[n]$ the respective matrices for the unsupervised and supervised modules.

We consider two frames types: firstly, classical frames formed by pilots and user symbols, and secondly, user frames containing only user symbols. The following procedure is performed at the receiver side each time a classical frame is received:

- First, the supervised algorithm estimates the channel matrix $\boldsymbol{H}$ from pilot symbols and, subsequently, computes the gain parameter g_{WF} and the precoding matrix $\boldsymbol{F}$ according to Eq. (5).
- The joint matrix $\boldsymbol{HF}$ (denoted by $\hat{\boldsymbol{HF}}$) is computed and the unsupervised algorithm is initialized so that $\boldsymbol{W_u}[n] = \hat{\boldsymbol{HF}}^{-\mathrm{H}}$.
- The channel matrix $\boldsymbol{H}$ is sent to the transmitter through the feedback channel allowing the transmitter to update the precoding matrix $\boldsymbol{F}$ as given by Eq. (5).

On the contrary, when user frames are received, the unsupervised algorithm (see Eq. (12)) is adapted and the decision criterion is evaluated after processing all the frame symbols. An “alarm” is sent to the transmitter through the feedback channel when that decision criterion indicates that a significant channel variation has occurred. The user symbols included in both types of frames are recovered using $\hat{\boldsymbol{u}}[n] = g_{\mathrm{WF}}\boldsymbol{y}[n]$.

An important question is how to design the decision module in order to detect such channel variations. By combining Eqs. (2) and (6), the output $\boldsymbol{z}[n]$ can be rewritten as a linear combination of the sources

$$\boldsymbol{z}[n] = \boldsymbol{\Gamma}[n]\boldsymbol{u}[n], \tag{13}$$

where $\boldsymbol{\Gamma}[n] = {\boldsymbol{W}_{\boldsymbol{u}}}^{\mathrm{H}}[n]\boldsymbol{H}\boldsymbol{F}$ represents the overall mixing/separating system. Sources are optimally recovered in case of selecting the matrix $\boldsymbol{W}_{\boldsymbol{u}}[n]$ such that every output extracts a different single source. This occurs when the matrix $\boldsymbol{\Gamma}[n]$ has the form

$$\boldsymbol{\Gamma}[n] = \boldsymbol{D}\boldsymbol{P}, \tag{14}$$

where $\boldsymbol{D}$ is a diagonal invertible matrix and $\boldsymbol{P}$ is a permutation matrix. An interesting consequence of using a linear precoder is that the permutation ambiguity associated to unsupervised algorithms is avoided because of the initialization $\boldsymbol{W}_{\boldsymbol{u}}[n] = (\hat{\boldsymbol{H}\boldsymbol{F}})^{-\mathrm{H}}$. This implies that the data sources are recovered in the same order as they were transmitted. Therefore, taking Eq. (14) into account, the optimum separation matrix produces a diagonal matrix $\boldsymbol{\Gamma}[n]$ and thus, the mismatch of $\boldsymbol{\Gamma}[n]$ with respect to a diagonal matrix allows us to measure channel variations.

Although channel matrices are unknown, we can use as a reference the estimation $\hat{\boldsymbol{H}\boldsymbol{F}}$ obtained by means of the supervised approach. Thus, we compute $\boldsymbol{\Gamma}[n] = {\boldsymbol{W}_{\boldsymbol{u}}}^{\mathrm{H}}[n]\hat{\boldsymbol{H}\boldsymbol{F}}$ after processing the symbols in a frame. Consequently, that distance with respect to a diagonal matrix is measured using the following "error" criterion:

$$\mathrm{Error}[n] = \sum_{i=1}^{N_{\mathrm{t}}} \sum_{j=1, j\neq i}^{N_{\mathrm{t}}} \left(\frac{|\gamma_{ij}[n]|^2}{|\gamma_{ii}[n]|^2} + \frac{|\gamma_{ji}[n]|^2}{|\gamma_{ii}[n]|^2} \right), \tag{15}$$

where $\gamma_{ii}[n]$ denotes the i–th diagonal element of the matrix $\boldsymbol{\Gamma}[n]$. A possibility for determining when the channel has changed significantly is to compare the above error value to some fixed threshold value (denoted by t), i.e. $\mathrm{Error}[n] > t$ would mean that a classical frame (i.e. a frame with pilot and user symbols) is required.

5. Simulation Results

In order to show the performance achieved with the proposed DASBE approach, we present results obtained by several computer simulations performed considering that 10 000 QPSK symbols have been transmitted through a MIMO system in blocks of 200 symbols each one (i.e. 50 frames). The system consists of four transmit and four receive antennas. The channel matrix changes each 10 frames according to the following model

$$\boldsymbol{H} = (1-\alpha)\boldsymbol{H} + \alpha\boldsymbol{H}_{\mathrm{new}},$$

where $\boldsymbol{H}_{\mathrm{new}}$ is a 4×4 complex matrix randomly generated according to a Gaussian distribution. The rest of parameters used for DASBE has been: threshold of $t = 0.1$ and initial step-size parameter of $\mu = 0.001$ for the unsupervised algorithm. The

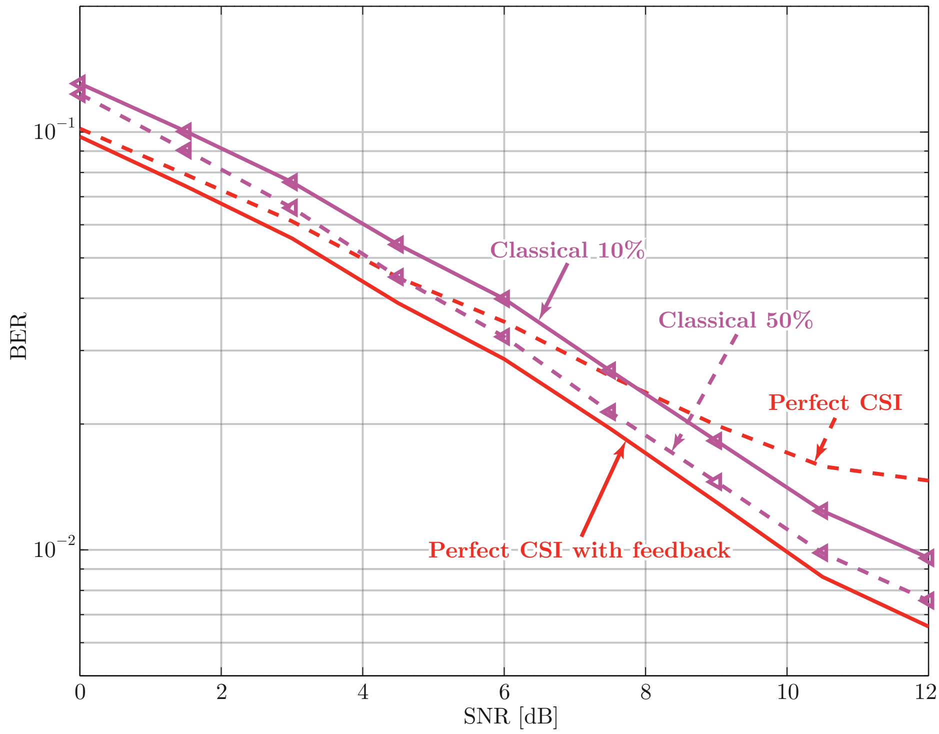

Fig. 3 *BER versus SNR obtained using a classical approach.*

following results have been obtained by averaging 1 000 independent realizations, varying both channels and transmitted symbols.

Using a classical supervised approach, Figure 3 shows the performance in terms of *Bit Error Rate* (BER) versus *Signal–to–Noise Ratio* (SNR) for a channel updating parameter $\alpha = 0.05$ and different percentages of pilot symbols per frame. Specifically, we select 10%, which means that 20 symbols per frame are dedicated to pilot symbols, and 50%, which corresponds to 100 pilots per frame. As a performance bounds, the following curves are also plotted in Figure 3:

- BER curve when both perfect CSI and feedback channel are available between the receiver and the transmitter side (labeled as *Perfect CSI with feedback*).
- BER curve without feedback channel (labeled as *Perfect CSI*). In such a case, the precoding matrix is never updated, which leads to a loss in performance with respect to the previous situation with existing feedback channel.

Notice that the utilization of the feedback channel produces a considerable improvement in terms of BER and SNR (in fact, for the SNR plotted in this figure, the system without feedback is not able to achieve a BER of 10^{-2}). It is also apparent that the classical approach needs 50% of pilot symbols to obtain a performance close to the Perfect CSI with feedback.

Figure 4 plots the results obtained with the DASBE approach considering both

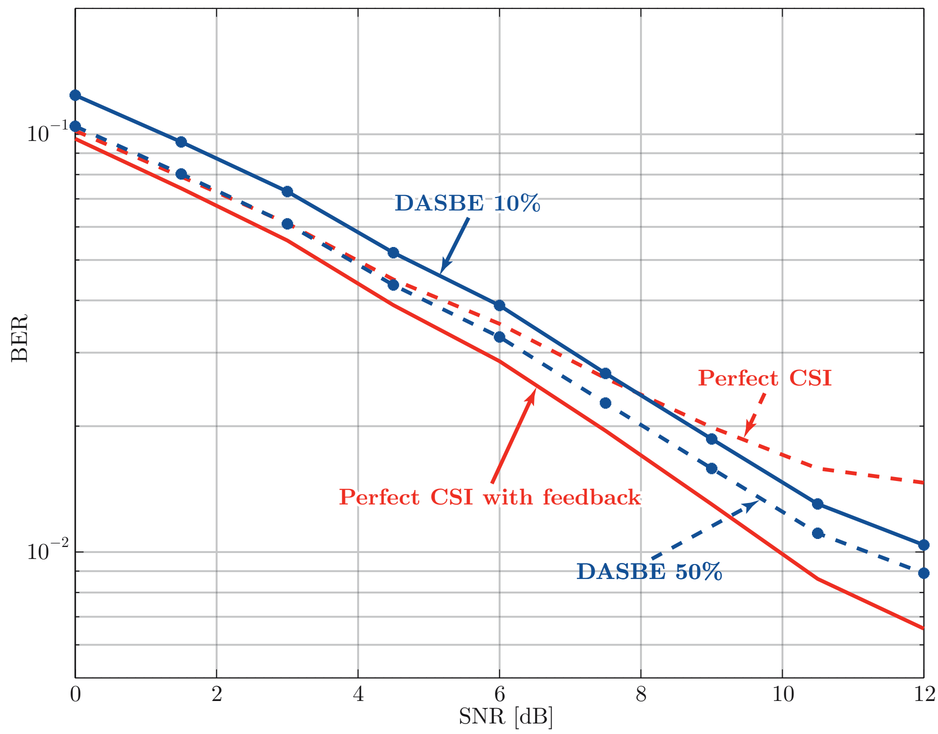

Fig. 4 *BER versus SNR obtained using DASBE.*

10% and 50% percentages of pilot symbols per frame. Notice that using the DASBE approach only the classical frames carry pilot symbols, while the user frames exclusively contain data symbols. From Figure 4 it is apparent that the performance is similar to that offered by the classical approach (see Figure 3), but with the advantages of reducing the feedback channel overhead (see Figure 5) as well as the amount of needed pilot symbols (see Figure 6).

Figure 5 presents the utilization of the feedback channel depending on the approach used to track channel variations. Note that the classical approach transmits through the feedback channel each time a new frame is received, i.e. 50 times in total (independently of the pilot symbols percentage). However, the channel utilization for DASBE depends only on the decision criterion. It can be observed from Figure 5 that the feedback channel utilization is considerably reduced in case of implementing DASBE.

Finally, Figure 6 shows another important advantage of DASBE with respect to the classical approach, which consists in a considerable reduction in the number of needed pilot symbols. This is because pilots are included only when the degradation of the channel estimates is too large (according to the previously fixed threshold). Also note that for the DASBE approach, Figure 6 plots the mean number of pilot symbols per frame considering the two frame types (i.e. classical and user frames) required to transmit 50 frames.

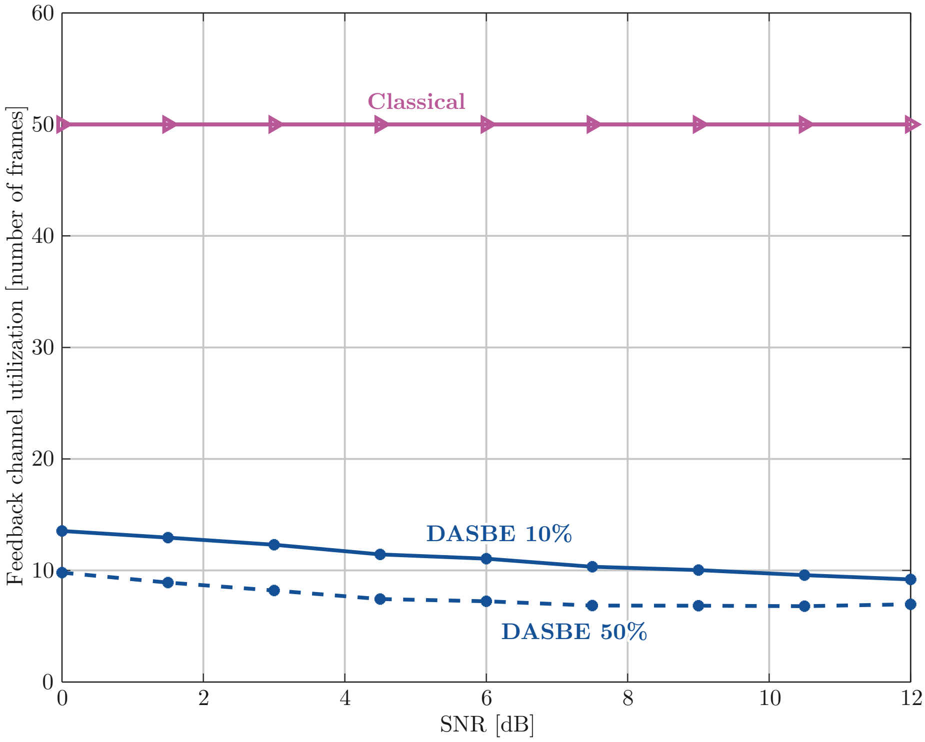

Fig. 5 *Utilization of the feedback channel versus SNR obtained using a classical approach and DASBE.*

5.1 Remarkable Comments

It is important to note that in the case of the supervised estimation in Eq. (10), the matrix $\boldsymbol{C_y}$ may be singular. When this occurs, we have decided to consider the previous channel estimate. Moreover, for those frames in which the unsupervised algorithm diverges, we have reduced the step-size parameter to $\mu = \mu/10$ and initialized the algorithm to the matrix $\boldsymbol{W_u}[n]$ given by the previous frame.

Moreover, note that the BSS problem assumes that the observations are linear mixtures of the sources. From Eq. (5) it is easy to verify that for LP systems, assuming perfect CSI at the transmitter side, the joint matrix $\boldsymbol{HF}$ is diagonal when ψ is close to zero or, equivalently, when SNR is large. In that case, BSS methods are not justified. However, under realistic transmission scenarios, SNR is usually constrained to the interval $[5\,\text{dB}, 15\,\text{dB}]$ and perfect CSI is not available at the transmitter, which produces a non–diagonal matrix $\boldsymbol{HF}$ that allows us to use BSS algorithms.

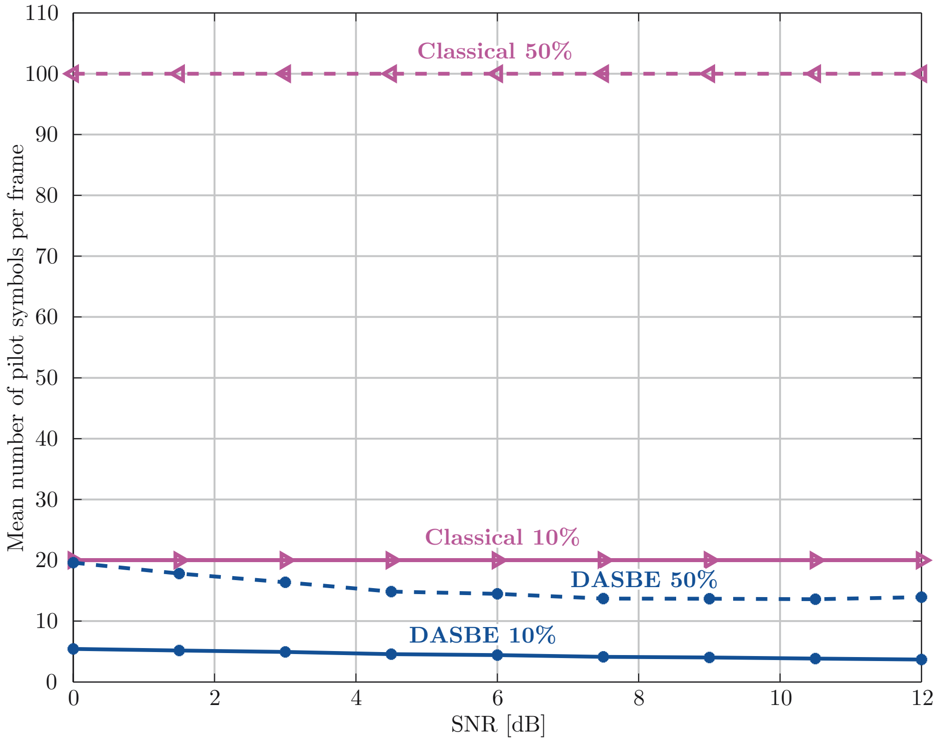

Fig. 6 *Pilot symbols versus SNR obtained using a classical approach and DASBE.*

6. Conclusions

Given a communications system in which a block flat fading channel is considered, we proposed an intuitive as well as simple method to detect channel variations. This decision criterion is used to develop a novel hybrid approach which combines both supervised and unsupervised algorithms. In case of significant channel variations, our system utilizes a supervised approach to estimate the channel coefficients, which are sent to the transmitter through a low–cost feedback channel. Otherwise, an unsupervised adaptive algorithm is used to track those channel variations. Simulation results have shown that the proposed approach is an attractive solution for wireless systems since it provides an adequate BER with a low overhead caused by transmitted pilot symbols and with reduced feedback channel occupancy.

Acknowledgments

This work has been funded by the Ministerio de Ciencia e Innovación of Spain, and the FEDER funds of the European Union under the grants 09TIC008105PR, TEC2007-68020-C04-01, and CSD2008-00010.

References

[1] S.-I. Amari. Gradient Learning in Structured Parameter Spaces: Adaptive Blind Separation of Signal Sources. In *Proc. WCNN'96*, pages 951–956. San Diego, 1996.

[2] S.-I. Amari, T.-P. Chen, and A. Cichocki. Stability Analysis of Learning Algorithms for Blind Source Separation. *Neural Networks*, 10(no. 8):1345–1351, August 1997.

[3] A. Bell and T. Sejnowski. An Information-Maximization Approach to Blind Separation and Blind Deconvolution. *Neural Computation*, 7(no. 6):1129–1159, November 1995.

[4] A. Benveniste, M. Metivier, and P. Priourent. *Adaptive Algorithms and Stochastic Approximations*. Springer-Verlag, New York, 1990.

[5] J.-F. Cardoso. Blind Signal Separation: Statistical Principles. *Proceedings of IEEE*, 86(no. 10):2009–2025, October 1998.

[6] L. Castedo and O. Macchi. Maximizing the Information Transfer for Adaptive Unsupervised Source Separation. In *Proc. SPAWC'97*, pages 65–68. Paris, France, April 1997.

[7] P. M. Castro, J. A. García-Naya, D. Iglesia, and A. Dapena. A Novel Hybrid Approach to Improve Performance on Frequency Division Duplex Systems with Linear Precoding. In *Lecture Notes on Computer Science*, number 6077, vol. II, pages 248–255, 2010.

[8] R. L. Choi and R. D. Murch. New Transmit Schemes and Simplified Receiver for MIMO Wireless Communication Systems. *IEEE Transactions on Wireless Communications*, 2(6):1217–1230, November 2003.

[9] P. Comon and C. Jutten. *Handbook of Blind Source Separation, Independent Component Analysis and Applications*. Academic Press, 2010.

[10] A. Dapena, P. M. Castro, and J. Labrador. Combination of Supervised and Unsupervised Algorithms for Communication Systems with Linear Precoding. In *Proc. of WCCI'2010*, pages 1265–1272, July, 2010.

[11] M. Ergen. *Mobile Broadband: Including WiMAX and LTE*. Springer Verlag, 2009.

[12] D. Godard. Self-recovering Equalization and Carrier Tracking in Two–Dimensional Data Communication Systems. *IEEE Transactions on Communications*, 28:1867–1875, November 1980.

[13] S. Haykin. *Neural Networks: A Comprehensive Foundation*. Macmillan College Publishing Company, New York, 1994.

[14] M. Joham. *Optimization of Linear and Nonlinear Transmit Signal Processing. PhD dissertation*. Munich University of Technology, 2004.

[15] M. Joham, K. Kusume, M. H. Gzara, W. Utschick, and J. A. Nossek. Transmit Wiener Filter for the Downlink of TDD DS-CDMA Systems. In *Proc. ISSSTA*, volume 1, pages 9–13, September 2002.

[16] C. Jutten and J. Herault. Blind Separation of Sources, Part I: An Adaptive Algorithm Based on Neuromimetic Architecture. *Signal Processing*, 24:1–10, July 1991.

[17] H. R. Karimi, M. Sandell, and J. Salz. Comparison between Transmitter and Receiver Array Processing to Achieve Interference Nulling and Diversity. In *Proc. PIMRC*, volume 3, pages 997–1001, September 1999.

[18] O. Macchi. *Adaptive Processing. The Least Mean Squares Approach with Applications in Transmission*. John Wiley and Sons, 1995.

[19] J. A. Nossek, M. Joham, and W. Utschick. Transmit Processing in MIMO Wireless Systems. In *Proc. of the 6th IEEE Circuits and Systems Symposium on Emerging Technologies: Frontiers of Mobile and Wireless Communication*, pages I–18 – I–23, May/June 2004. Shanghai, China.