A Per-Survivor Processing Receiver for MIMO Transmission Systems With One Unknown Channel Order Per Output
Manuel A. Vázquez and Joaquín Míguez, Member, IEEE

Abstract—The order of a communications channel is the length of its impulse response. Recently, several works have tackled the problem of estimating the order of a frequency-selective multiple-input–multiple-output (MIMO) channel. However, all of them consider a single order, despite the fact that a MIMO channel comprises several subchannels (specifically, as many as the number of inputs times the number of outputs), each one possibly with its own order. In this paper, we introduce an algorithm for maximum-likelihood sequence detection (MLSD) in frequency- and time-selective MIMO channels that incorporates full estimation of the MIMO channel impulse response (CIR) coefficients, including one channel order per output. Simulation results following the analytical derivation of the algorithm suggest that the proposed receiver can achieve significant improvements in performance when transmitting through a MIMO channel that effectively comprises subchannels of different lengths.

Index Terms—channel estimation, dynamic programming, maximum likelihood detection, multipath channels, multiple-input–multiple-output (MIMO), signal processing algorithms, Viterbi algorithm.

I. INTRODUCTION

A COMMON assumption for the equalization of frequency-selective multiple-input–multiple-output (MIMO) channels is that the length of the channel impulse response (CIR), which is also referred to as the channel order, is known [2]. This is not true in general, and the usual approach, which consists of choosing a “high-enough” value for this parameter, i.e., an upper bound, tends to overestimate it. The reason behind this choice is to avoid the serious performance degradation that occurs when the CIR length is underestimated. However, since the complexity of maximum-likelihood sequence detection (MLSD) in frequency-selective MIMO channels grows exponentially with the channel order [3], overestimating this parameter leads to an increase in the complexity of the receiver. Moreover, an algorithm that overestimates the length of the CIR suffers a performance degradation that grows as the assumed channel order moves away from the true one [4].

The problem of detecting the order of a MIMO channel has been addressed in some recent papers. In [5], it is proposed to estimate the order of a MIMO communications channel by means of a model order selection technique known as conditional model-order estimation [6]. However, the latter is based on the assumption that the CIR is fixed for the duration of a complete data frame (typically, a few hundreds of symbols) and that the processing of the block of available observations is performed offline in batch mode. A completely different approach to the problem based on the per-survivor processing (PSP) principle [7], [8] was suggested in [4]. An advantage of this method over the previous method is that it naturally copes with time-varying channels. Related work for single-input–multiple-output (SIMO) channels can be found in [9], where a method for the estimation of the channel order based on the minimization of a combination of cost functions is introduced. However, similar to [5], this technique relies on the assumption that the channel remains constant during the transmission of a complete data frame.

In this paper, we argue that considering a single (overall) order for a MIMO channel is ultimately a simplification since a MIMO channel with $N_T$ transmitters and $N_R$ receivers spans $N_T \times N_R$ single-input–single-output (SISO) subchannels, each one with its own (1-D) CIR. Therefore, when a single order is chosen for the overall MIMO channel, we are actually estimating the maximum of the SISO subchannel orders. As a consequence, we may be overestimating (often severely) the length of many subchannels and therefore degrading the performance of the MLSD receiver.

To the best of our knowledge, not many efforts have been made to deal with the problem of estimating more than a single channel order in a MIMO system. One exception is the work in [10], which tackles the problem of estimating one channel order per transmitting antenna in a MIMO channel by means of a technique based on canonical correlation analysis (see, e.g., [11]). The algorithm proves successful in detecting the orders associated with the different SIMO channels comprised by the MIMO channel. Similar to [5] and [9], however, the channel response is assumed fixed for the duration of a data frame. A different approach is put forward in [12], where it is proposed to estimate the different channel orders associated

Manuscript received December 14, 2010; revised May 2, 2011, July 10, 2011 and September 12, 2011; accepted September 14, 2011. Date of publication October 3, 2011; date of current version December 9, 2011. This work was supported in part by the Ministry of Science and Innovation of Spain (Program Consolider-Ingenio 2010 CSD-2008-00010 COMONSENS; Project DEIPRO. Ref. TEC-2009-14504-C02-01) and in part by a joint program of the Comunidad de Madrid and the Universidad Carlos III de Madrid (CCG10-UC3M/TEC-5225 ETORS). The review of this paper was coordinated by Prof. W. A. Krzymien.

The authors are with the Departamento de Teoría de la Señal y Comunicaciones, Universidad Carlos III de Madrid, 28911 Leganés, Spain (e-mail: mvazquez@tsc.uc3m.es; jmiguez@tsc.uc3m.es).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2011.2170445
with a multiple-input–single-output (MISO) channel (i.e., one order per transmitting antenna) using higher order statistics. However, not only is this technique devised for time-invariant channels but it also does not properly handle situations in which all the orders are equal and is subject to error propagation phenomena.

In this paper, we extend the work in [4] and introduce a PSP-based method for the equalization of frequency- and time-selective MIMO channels characterized by \( N_R \) subchannel orders, one per receiving antenna.\(^1\) When the MIMO channel has indeed several different subchannel orders, the proposed method is able to outperform the algorithm in [4] without any asymptotic increase in its computational complexity. In addition, it computes maximum a posteriori (MAP) estimates of the subchannel orders and is suitable to adaptively work in nonstationary environments, i.e., when the MIMO channel response is time variant, even between consecutive symbols. We illustrate the performance of the new scheme by way of computer simulations in several scenarios. In particular, we compare the proposed receiver with PSP-based MIMO equalizers that employ one single channel order (either known or estimated), and we assess the effect of the number of survivors in the PSP sequence detector for several configurations of the subchannel orders and also show an example of application in (the particular case of) a time-invariant MIMO channel.

The remaining of this paper is organized as follows: In Section II, the discrete-time baseband equivalent signal model of a MIMO transmission system with frequency- and time-selective channel is described. The proposed algorithm is introduced in Section III. The extension of the proposed methodology to handle channels with \( N_t \) subchannel orders (one per transmitter) and \( N_t \times N_R \) subchannel orders (one per transmitter–receiver pair) is discussed in Section IV. Computer simulation results are shown in Section V, and finally, Section VI is devoted to the conclusions.

II. SIGNAL MODEL

We consider a MIMO communication system with \( N_t \) input data streams and \( N_R \) output observation streams. The length of the CIR between any given input and the \( t \)th output, which yields the span of the intersymbol interference for the corresponding subchannels, is modeled as a time-invariant random parameter \( m_t \) in a finite set \( \mathcal{M} \). Therefore, there are \( N_R \) (bounded but unknown) channel orders \( m_1, m_2, \ldots, m_{N_R} \). We will denote their maximum as \( m_{\max} = \max\{m_1, \ldots, m_{N_R}\} \). Obviously, \( m_{\max} \in \mathcal{M} \).

With the preceding assumptions, the discrete-time baseband equivalent model of the received signals can be written as [13]

\[
y_t = \sum_{i=0}^{m_{\max}-1} \mathbf{H}_t(i) s_{t-i} + \mathbf{g}_t, \quad t = 0, 1, \ldots \tag{1}
\]

where \( y_t \) is the \( N_R \times 1 \) vector of observations collected at the receiving antennas; \( \{ \mathbf{H}_t(i) \}_{i=0}^{m_{\max}-1} \) is the (time-varying) \( N_R \times N_t \) multidimensional CIR; \( s_t = [s_t(1), \ldots, s_t(N_t)]^\top \) is the \( N_t \times 1 \) vector of transmitted symbol streams; and \( \mathbf{g}_t \) is an \( N_R \times 1 \) vector of independent zero-mean additive white Gaussian noise (AWGN) components with variance \( \sigma_g^2 \). Subscript \( t \) denotes discrete time; hence, the CIR in (1) is both time and frequency selective when \( m_{\max} > 1 \). The different channel orders \( m_1, \ldots, m_{N_R} \), are modeled as discrete random variables in the finite set \( \mathcal{M} \), with known a priori probability mass function (pmf).

Model (1) implies that the observation vector \( y_t = [y_t(1), \ldots, y_t(N_R)]^\top \) involves \( m_{\max} \) channel matrices \( \mathbf{H}_t(i) \), \( i = 0, \ldots, m_{\max} - 1 \). However, it is possible that the subchannels on which an individual observation \( y_t(l), l \in \{1, \ldots, N_R\} \), actually depends have an order \( m_l < m_{\max} \). This means that the \( l \)th row of matrix \( \mathbf{H}_t(i) \) is null when \( i \geq m_l \).

Following the approach of [13], we model the channel variation using a first-order autoregressive (AR) process driven by white Gaussian noise, i.e.,

\[
\mathbf{H}_t(i) = \alpha \mathbf{H}_{t-1}(i) + \mathbf{V}_t(i), \quad 0 \leq i \leq m_{\max} - 1; \quad t \geq 0 \tag{2}
\]

where \( \alpha \) is the AR coefficient, and \( \mathbf{V}_t(i) \) is an \( N_R \times N_t \) matrix of independent and zero-mean Gaussian variables. Because of the different subchannel orders, some of these variables are degenerate. In particular, if \( \sigma^2_\alpha(i, i) \) is the (common) variance of the elements in the \( l \)th row of \( \mathbf{V}_t(i) \), we set \( \sigma^2_l(i, i) = \sigma^2 \) if \( i < m_l \), whereas \( \sigma^2_l(i, i) = 0 \) for \( i \geq m_l \). Model parameters \( \alpha \) and \( \sigma^2 \) are selected to fit the desired channel autocorrelation function. We further assume that the probability density function (pdf) of the channel at time \( t = -1 \) is Gaussian and known.

Equation (2) yields a suitable tool for modeling channel time variability under the wide sense uncorrelated scattering (WSUS) assumption [3], [14]. This assumption implies that the channel coefficients are mutually uncorrelated; hence, they independently evolve\(^2\) over time. This is a common modeling approach in the literature (see, e.g., [13] and [15]).

Finally, if we define the overall channel matrix as

\[
\mathbf{H}_t = [\mathbf{H}_t(m_{\max} - 1), \ldots, \mathbf{H}_t(0)] \tag{3}
\]

by putting together all the channel matrices involved in the \( t \)th observation, (1) can be rewritten in a more compact way as

\[
y_t = \mathbf{H}_t \mathbf{s}_{t;m_{\max}-1} + \mathbf{g}_t, \quad t = 0, 1, \ldots \tag{4}
\]

where the \( N_t m_{\max} \) vector \( \mathbf{s}_{t;m_{\max}+1} = [s_t(1), \ldots, s_t(N_t)]^\top \) stacks together all the symbol vectors involved in the \( t \)th observation.

III. ALGORITHM

The goal of an MLSD algorithm is to find the sequence of symbol vectors \( \mathbf{s}_{t;1} = \{s_0, \ldots, s_t\} \), with the maximum likelihood given the received sequence of observations \( y_{t;1} = \)

\(^{1}\)Note that, by assuming this model (one channel order per receiving antenna), we do not imply that it is more realistic than other possible choices, e.g., one channel order per transmitting antenna or one channel order per transmit–receive pair. The physical interpretation of the model depends on the communication system of interest.

\(^{2}\)Note that they are Gaussian distributed.
\{y_0, \ldots, y_t\}$. Therefore, we aim at the (possibly approximate) solution of the optimization problem
\[
s_{0:t}^{\text{MLSD}} = \arg\max_{s_{0:t}} p(y_{0:t} | s_{0:t})
\]
where $p(y_{0:t} | s_{0:t})$ denotes the conditional pdf of the observations $y_{0:t}$, given the sequence of symbols $s_{0:t}$ (i.e., the likelihood of $s_{0:t}$).

To search the optimal symbol sequence, we propose the use of the PSP principle. Let $P$ be the number of survivors and $s_{0:t}^{(i)}$, $i \in \{1, \ldots, P\}$, be the path (or symbol sequence) of the $i$th survivor up to time $t$. Associated to the $i$th survivor, we compute the likelihood $p(y_{0:t} | s_{0:t}^{(i)})$, and the solution of (5) is approximated as the path of the survivor with the highest likelihood, i.e.,
\[
s_{0:t}^{\text{MLSD}} = \arg\max_{s_{0:t} \in \Theta_t} p(y_{0:t} | s_{0:t})
\]
where $\Theta_t = \{s_{0:t}^{(i)}\}_{i=1}^P$.

Assume that, at time $t - 1$, we have computed the set of survivor paths $\Theta_{t-1}$. For each path, $s_{0:t-1}^{(i)} \in \Theta_{t-1}$, its likelihood $p(y_{0:t-1} | s_{0:t-1}^{(i)})$ is available, and so are the channel order probabilities $p(m_l | s_{0:t-1}^{(i)}, y_{0:t-1})$, with $l = 1, \ldots, N_R$ and $m_l \in \mathcal{M}$. If we now let $\mathcal{S}_{N_R}$ denote the $N_R$-dimensional alphabet, where vector $s_t$ can take values, the proposed PSP algorithm consists of four steps.

1) Extend each survivor $s_{0:t-1}^{(i)}$ into $\{s_{0:t-1}^{(i)}, s_t^{(j)}\}_{j=1}^P [\mathcal{S}_{N_R}]$, with one for each possible vector element $s_t \in \mathcal{S}_{N_R}$.

2) For each sequence in the set of prospective paths $\Theta_t^* = \{s_{0:t-1}^{(i)}, s_t^{(j)}\}_{i=1, j=1}^P [\mathcal{S}_{N_R}]$, compute the likelihood $p(y_{0:t} | s_{0:t-1}^{(i)}, s_t^{(j)})$.

3) Build the new survivor set $\Theta_t = \{s_{0:t}^{(i)}\}_{i=1}^P$ by selecting the $P$ paths with the highest likelihoods in $\Theta_t^*$.

4) For each survivor $s_{0:t}^{(i)}$ in $\Theta_t$, update the channel order probabilities to obtain $p(m_l | s_{0:t}^{(i)}, y_{0:t})$ for $l = 1, \ldots, N_R$, and every possible value of each $m_l \in \mathcal{M}$.

Note that the survivors are globally selected from set $\Theta_t^*$ and not on a per-state basis as in conventional PSP algorithms. Simulations show that this strategy can lead to an improved performance.

Steps 1) and 3) of the proposed algorithm are self-explanatory. Details about the recursive computation of the channel order posterior probabilities $p(m_l | s_{0:t}^{(i)}, y_{0:t})$ and the likelihood $p(y_{0:t} | s_{0:t-1}^{(i)}, s_t^{(j)})$ can be found in Section III-A. In Section III-B, we show that the MAP estimation of $N_R$ sub-channel orders does not significantly increase the complexity of the MIMO CIR estimation algorithm compared to the case in which only $m_{\max}$ is estimated (as, e.g., in [4]).

### A. Recursive Update

We address the update of the likelihood first. In particular, a straightforward application of the chain rule yields
\[
p(y_{0:t} | s_{0:t-1}^{(i)}, s_t^{(j)}) = p(y_t | s_{0:t-1}^{(i)}, s_t^{(j)}, y_{0:t-1}) p(y_{0:t-1} | s_{0:t-1}^{(i)})
\]
where we have used the fact that $y_{0:t-1}$ is independent of $s_t$.

The WSUS assumption implies that the coefficients in a certain row of the channel matrix are independent of those contained in any other row, and as a consequence, so are the corresponding observations obtained for these coefficients (conditional on the transmitted symbols). Hence, the first term on the right-hand side of (7) can be factorized as
\[
p(y_t | s_{0:t-1}^{(i)}, s_t^{(j)}, y_{0:t-1}) = \prod_{l=1}^{N_R} p(y_t | s_{0:t-1}^{(i)}, s_t^{(j)}, y_{0:t-1}(l))
\]
(See Appendix A for a formal proof of this result.) Each of the marginal likelihoods on the right-hand side of (8) can be computed using a Kalman filter (KF) when the order of the corresponding channel is known. Indeed, since any channel order can only take values in the finite set $\mathcal{M}$, we can carry out this computation by means of the law of total probability, which yields
\[
p(y_t | s_{0:t-1}^{(i)}, s_t^{(j)}, y_{0:t-1}(l)) = \sum_{m_l \in \mathcal{M}} p(y_t | m_l, s_{0:t-1}^{(i)}, s_t^{(j)}, y_{0:t-1}(l))
\]
Details on the computation of the marginal likelihoods on the right-hand side of (9) (which are conditional on the channel order) are given in Appendix B.

The symbols transmitted at time $t$ do not contribute with any useful information to the estimation of the $l$th channel order $m_l$ if the corresponding observation $y_t$ is unknown; hence,

5This equation still holds when there is spatial correlation among the subchannels associated with a given receiving antenna (i.e., correlation among the coefficients in a given row of the channel matrix). However, the equation does not hold if there is correlation among channels associated to different receiving antennas (i.e., correlation among coefficients in different rows of the channel matrix).
where 

\[ p(m_l|s_{0:t-1}^{(i)}, s_l^{(j)}, y_{0:t-1}(l)) = p(m_l|s_{0:t-1}^{(i)}, y_{0:t-1}(l)) \],

so that (9) can be rewritten as

\[
p \left( y_{l}(l)|s_{0:t-1}^{(i)}, s_l^{(j)}, y_{0:t-1}(l) \right) \]

\[ = \sum_{m_{i} \in \mathcal{M}} p \left( y_{l}(l)|m_{i}, s_{0:t-1}^{(i)}, s_l^{(j)}, y_{0:t-1}(l) \right) \]

\[ \times p \left( m_l|s_{0:t-1}^{(i)}, y_{0:t-1}(l) \right). \] (10)

As for the a posteriori probability for the lth channel order, it can be easily updated by means of the Bayes’ theorem as

\[
p \left( m_l|s_{0:t-1}^{(i)}, y_{0:t-1}(l) \right) \propto p \left( y_{l-1}(l)|m_{l}, s_{0:t-1}^{(i)}, y_{0:t-2}(l) \right) \]

\[ \times p \left( m_l|s_{0:t-2}^{(i)}, y_{0:t-2}(l) \right) \] (11)

for \( l = 1, \ldots, N_R \). Since the channel orders take values in a discrete and finite set \( \mathcal{M} \), it is straightforward to compute the proportionality constants for the probabilities in (11).

### B. Complexity

Just like any other algorithm performing MLSD in a MIMO channel, the proposed scheme has an exponential complexity on the number of input streams \( N_T \), \( N_R \). In particular, at each time step, the likelihood of every possible symbol vector must be computed for each survivor and every possible channel order. Therefore, the complexity of the data detection procedure in the proposed scheme is \( \mathcal{O}(P|\mathcal{M}|S^{N_T}) \). Notice that, in this case and as opposed to classical trellis-search algorithms such as the PSP or the Viterbi algorithms, the complexity can be ultimately controlled through the number of overall survivors considered, regardless of the channel order (which is unknown).

As for channel estimation, one KF per output and channel order must be kept and updated for each survivor. This yields \( |\mathcal{M}|N_R \) KFs per survivor, with each one keeping track of the channel state information related to a specific output and a specific channel order. Let \( m_{\mathcal{M}} \) denote the upper bound of any possible channel order, i.e., \( m_{\mathcal{M}} = \max \{ \mathcal{M} \} \). (Note that \( m_{\mathcal{M}} \geq m_{\max} \) and, normally, \( m_{\mathcal{M}} > m_{\max} \)) The KFs considering \( m_{\mathcal{M}} \) as the channel order (one per output) determine the complexity of the channel estimation because their associated state vectors are the longest ones: they comprise \( N_T m_{\mathcal{M}} \) elements. (The length of the state vector of a KF tracking a subchannel of order \( m \) is exactly \( N_T m_{m} \).) Then, taking into account that the complexity of a KF is cubic on the length of its state vector, it is easy to see that the complexity of the algorithm due to the channel estimation procedure is \( \mathcal{O}(P|\mathcal{M}|N_R N_T^2 m_{\mathcal{M}}^3) \). This stems from \( N_T^2 m_{\mathcal{M}}^3 \) being the complexity of the most expensive KF, \( |\mathcal{M}|N_R \) being the number of KFs kept per survivor, and \( P \) being the number of survivors.

Let us remark that, if we were to estimate a single order for the entire channel (the maximum among all outputs \( m_{\max} \)), the computational burden would not be lesser. In such case, there would be no need to use (8) to compute the likelihood of \( Y_1 \) as the product of the individual observations, and hence, \( |\mathcal{M}| \) KFs per survivor (each one estimating all the channel coefficients under a different channel order assumption) would be enough. However, the state vector associated to the KF considering \( m_{\mathcal{M}} \) as the channel order would have \( N_R N_T m_{\mathcal{M}} \) elements, and by the same argument as before, that would yield an overall complexity \( \mathcal{O}(P|\mathcal{M}|N_R^2 N_T^3 m_{\mathcal{M}}^3) \) for the channel estimation algorithm.

Therefore, the proposed algorithm does not incur any (asymptotic) computational penalty due to the estimation of one channel order per output, when compared with a similar algorithm estimating a single channel order. However, performing channel order estimation always entails a computational burden. One way of alleviating it is to switch to an algorithm that fixes the different channel orders once these have been reliably estimated, as suggested in [4].

### IV. Alternative Channel Models

A natural question to ask is whether the results of Section III still hold when we allow for a different channel order associated with each pair of transmitting–receiving antennas. In such case, the factorization in (8), as well as its proof in Appendix A, is still valid. However, the likelihoods 

\[ p(y_{l}(l)|s_{0:t-1}^{(i)}, s_l^{(j)}, y_{0:t-1}(l)) \]

with \( l = 1 \ldots N_R \), can no longer be computed using (9). The reason is that a single observation \( y_{l}(l) \) recorded at time \( t \) in the \( l \)th receiving antenna depends now on \( N_T \) channel orders (rather than a single one).

To be specific, let \( m_{i,k} \in \mathcal{M} \) denote the order of the channel between the \( k \)th transmitting antenna and the \( l \)th receiving antenna. Computing the likelihood of a single observation, following the same approach we used before (i.e., by means of the law of total probability), yields

\[
p \left( y_{l}(l)|s_{0:t-1}^{(i)}, s_l^{(j)}, y_{0:t-1}(l) \right) \]

\[ = \sum_{m_{i,1}} \cdots \sum_{m_{i,N_T}} p \left( y_{l}(l)|m_{i,1}, \ldots, m_{i,N_T}, s_{0:t-1}^{(i)}, s_l^{(j)}, y_{0:t-1}(l) \right) \]

\[ = \sum_{m_{i,1}} \cdots \sum_{m_{i,N_T}} p \left( y_{l}(l)|m_{i,1}, \ldots, m_{i,N_T}, s_{0:t-1}^{(i)}, s_l^{(j)}, y_{0:t-1}(l) \right) \]

\[ \times p \left( m_{i,1}, \ldots, m_{i,N_T}|s_{0:t-1}^{(i)}, s_l^{(j)}, y_{0:t-1}(l) \right). \] (12)

From (12), it becomes clear that the complexity associated with the computation of 

\[ p(y_{l}(l)|s_{0:t-1}^{(i)}, s_l^{(j)}, y_{0:t-1}(l)) \]

is now \( \mathcal{O}(|\mathcal{M}|N_T^3) \), i.e., exponential on the number of transmitting antennas \( N_T \).

The same argument applies if we allow for one channel order per transmitting antenna, instead of per receiving antenna as in Sections II and III. Let us denote by \( m_{k,i} \), with \( k = 1, \ldots, N_T \), the channel order associated with the \( k \)th transmitting antenna. In this case, the single observation \( y_{l}(l) \) still depends on \( N_T \) channel orders; hence, the law of total probability yields

\[
p \left( y_{l}(l)|s_{0:t-1}^{(i)}, s_l^{(j)}, y_{0:t-1}(l) \right) \]

\[ = \sum_{m_{i,1}} \cdots \sum_{m_{N_T}} p \left( y_{l}(l)|m_{i,1}, \ldots, m_{N_T}, s_{0:t-1}^{(i)}, s_l^{(j)}, y_{0:t-1}(l) \right) \]

\[ \times p \left( m_{i,1}, \ldots, m_{N_T}|s_{0:t-1}^{(i)}, s_l^{(j)}, y_{0:t-1}(l) \right). \] (13)
Equation (13) shows that the complexity of computing the marginal likelihoods on the right-hand side of (8) is exponential on the number of transmitting antennas, i.e., $O(|\mathcal{M}|^{N_T})$.

From the previous discussion, it follows that the proposed method is only advantageous when considering one channel order per receiving antenna. When that is not the case, the method can still be applied, but the complexity due to data detection procedure is doubly exponential on the number of input streams. Specifically, Section III-B shows that the complexity of the algorithm is, in any case, exponential on the number of input streams. However, if one channel order per receiving antenna is considered, the complexity due to data detection becomes $O(P|M| |S^{N_T}|)$, whereas if either one channel order per transmitting antenna or one channel order per each transmitting–receiving pair of antennas is allowed, the complexity (due to data detection) becomes $O(P|M|^{N_T} |S^{N_T}|)$.

Intuitively, the reason for this method being only efficient for MIMO channels in which there is one channel order for receiving antennas is that any individual observation is inherently affected by the communication channels established between itself and every transmitting antenna. Thus, the channel order associated with every transmitting antenna affects the likelihood of any given observation and must be integrated out when unknown. Assuming that one channel order per receiving antenna entails that, for a specific receiving antenna, all the transmitting antennas have the same channel order, only one channel order needs to be integrated out.

V. SIMULATION RESULTS

We have carried out computer simulations to assess the proposed algorithm in several scenarios. In every case, we have considered a system with $N_s = 3$ transmitting antennas and $N_r = 3$ receiving antennas. The modulation format is binary phase-shift keying, and transmission is carried out in frames of $K = 300$ symbol vectors (i.e., 900 binary symbols overall), including a short training sequence of $T = 15$ pilot symbols from each transmitter. The set of possible channel orders is $\mathcal{M} = \{1, \ldots, 4\}$; hence, $m_M = 4$. In the simulations, each data frame is generated independently of all others (including the transmitted data, the MIMO channel realization, and the noise terms).

Since a training sequence is being employed, the proposed method is here used in a semiblind context. Note that no assumption is made that prevents its application in a fully blind manner, however, by completely removing the training sequence. In practice, however, it can be expected that a blind implementation of the method would demand a larger number of survivors to attain the desired performance. The resulting increase in the computational load may well be more relevant than the gain in the efficiency of the channel usage.

Although, for the purpose of algorithm design, we have assumed the multidimensional AR process (2) to model the channel dynamics, to assess the performance of the proposed algorithm in a more realistic environment, we have carried out our computer simulations, assuming CIRs generated using the classical Clarke autocorrelation function [17]. In particular, the channel coefficients are uncorrelated (as we abide by the WSUS assumption), and the Doppler spectrum of each coefficient is set to be the classical Jake’s spectrum [18, Section 5.4], which yields the Clarke’s autocorrelation function

$$
\rho(k) = J_0 \left(2\pi \frac{v_m}{c} F \cdot k T_s \right)
$$

where $k$ is the delay in symbol periods, $J_0$ is the Bessel function of the first kind and zeroth order, $v_m = 50$ m/s is the relative speed of the transmitter with respect to the receiver, $c$ is the speed of light, $F_c = 2$ GHz is the carrier frequency, and $T_s = 1/(500 \cdot 10^3)$ s is the symbol period.

Within this simulation setup, we have compared the following;

1) The MLSD method introduced in this paper, which estimates one channel order per output, which is labeled “MLSD-$m$.”

2) A standard PSP algorithm (labeled “PSP”) that estimates the CIR using a KF, considering $m_M = \max\{|\mathcal{M}| = 4$ as the channel order.

3) A modified version of the standard PSP algorithm (labeled “G-PSP”) that selects the survivors globally rather than on a per-survivor basis. The channel order for the method is assumed to be $m_M = 4$.

4) A genie-aided maximum-likelihood sequence detector (labeled “MLSD”) implemented using the Viterbi algorithm with perfect knowledge of the CIR (including all of its associated channel orders).

The PSP and G-PSP algorithms perform a search over a typical trellis diagram with $|S^{N_T (m_M - 1)}| = 2^9 = 512$ states. Unless otherwise explicitly stated, the number of survivors is, in both cases, $P = 1024$ (which amounts to keeping two survivors per state in the standard PSP technique), and this is also the number of survivors employed by the MLSD-$m$.

We evaluate the performance of the receivers in terms of the bit error rate (BER) and the mean square error (MSE) of the channel estimates. To measure the MSE, the best survivor is selected at the end of the frame, and for that survivor, the channel order for the method is assumed to be $m_M = 4$. To compute the normalized MSE of a channel estimate at time $t$, we group its coefficients in a vector $\hat{\mathbf{h}}_t$ and then obtain the normalized MSE according to

$$
\text{MSE}_t = \left( \frac{\mathbf{h}_t - \hat{\mathbf{h}}_t}{\mathbf{h}_t} \right)^H \left( \frac{\mathbf{h}_t - \hat{\mathbf{h}}_t}{\mathbf{h}_t} \right).
$$

The overall normalized MSE is then obtained by averaging over all time steps

$$
\text{MSE} = \frac{1}{K} \sum_{t=0}^{K-1} \text{MSE}_t.
$$

Three different scenarios have been investigated, each one considering a different choice of channel orders.

6Note that zero padding might be needed to compare the channel estimates produced by the proposed algorithm with the actual channel vector (so that the dimensions match).
Fig. 1. BER for several values of the SNR (in decibels) when \( m_1 = 2, m_2 = 2, \) and \( m_3 = 1 \).

Fig. 2. MSE for several values of the SNR (in decibels) when \( m_1 = 2, m_2 = 2, \) and \( m_3 = 1 \).

A. Scenarios \( m_1 = 2, m_2 = 2, \) and \( m_3 = 1 \):

time-varying channel

Fig. 1 shows the BER attained by the four algorithms, for several values of signal-to-noise ratio (SNR), when the channel order of the first two outputs is \( m_1 = m_2 = 2 \) and that of the remaining one is \( m_3 = 1 \). It can be seen that the MLSD-\( m \) algorithm performs around 2 dB worse than the genie-aided MLSD receiver (for a BER of \( 10^{-4} \)) but also 2 dB better (for the same BER) than the G-PSP algorithm, whose relative performance degrades at high SNRs. As of the standard PSP, its BER curve is always above that of the G-PSP receiver.

In detecting the channel order of the different outputs, the proposed algorithm reduces the effective number of channel coefficients that must be estimated, which, in turn, improves the quality of the channel estimates. This is illustrated in Fig. 2, which plots the MSE of the channel estimates produced by the algorithms against the SNR.

As for the channel order probabilities computed by the MLSD-\( m \) algorithm, Table I shows the relative frequency of channel order misdetection for every receiving antenna and several values of the SNR. It can be seen that the proposed algorithm is (almost always) able to accurately detect the channel orders of the different outputs.

The proposed algorithm computes a posteriori channel order probabilities for every receiving antenna at each time instant [see (11)]. Fig. 3 shows the evolution of the a posteriori pmf for the first subchannel order \( m_1 \) during the first 50 time instants. It can be seen that the a posteriori probability of the true channel order is above 0.9 after a few time instants (i.e., already during the training stage).

B. Scenarios \( m_1 = 4, m_2 = 1, \) and \( m_3 = 1 \):

Time-Varying Channel

We now test the algorithm in a setup in which the channel order of the first output is \( m_1 = 4 \) (hence, \( m_1 = m_{\text{max}} = m_M \)), and those of the two remaining ones are \( m_2 = m_3 = 1 \). Fig. 4(a) shows the BER for the different SNRs. In this scenario, the MLSD-\( m \) algorithm still performs close to the genie-aided MLSD receiver (around 2 dB worse for a BER of \( 10^{-4} \)) but only slightly better (around 1 dB) than the G-PSP algorithm. The gap in performance between the MLSD-\( m \) and G-PSP becomes narrower, compared to Fig. 1, because the (single) channel order assumed by the G-PSP method \( m_M = 4 \) coincides with the maximum of the actual orders, i.e., in this case, \( m_M = \max\{m_1, m_2, m_3\} = m_{\text{max}} \). However, the performance of the G-PSP receiver severely degrades for high SNRs (i.e., for SNR \( \geq 18 \) dB). Note that, for this method, the assumed channel

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
\textbf{SNR} & \textbf{output 1} & \textbf{output 2} & \textbf{output 3} \\
\hline
3 & 1.47e-02 & 7.37e-03 & 1.01e-02 \\
6 & 4.60e-03 & 6.45e-03 & 4.60e-03 \\
9 & 0.00e+00 & 0.00e+00 & 9.21e-04 \\
12 & 0.00e+00 & 0.00e+00 & 0.00e+00 \\
\hline
\end{tabular}
\caption{Relative frequency of channel order misdetection for each output and different values of the SNR (scenario \( m_1 = 2, m_2 = 2, \) and \( m_3 = 1 \). The results are averaged over 1086 data frames.)}
\end{table}

\footnote{The results in Table I were obtained by simply choosing at the end of the frame (for every receiving antenna) the channel order for which the a posteriori probability computed by the algorithm is maximum.}
order is $m_M = 4$; hence, $N_T N_R m_M = 36$ channel coefficients have to be estimated and then tracked over the entire frame. In addition, recall that the receivers are initialized with a training sequence of only $T = 15$ symbols per transmitter. For high SNRs, the result is an initial estimate that is overfitted, which leads to poor detection of the subsequent data symbols. The standard PSP receiver is more robust to this problem because the allocation of a fixed number of survivors to each state in the trellis graph guarantees better diversity of the candidate paths.

The BER attained by the G-PSP can be greatly improved by simply increasing the number of training symbols. Fig. 4(b) shows the BER of this receiver when initialized with $T = 15$ and $T = 30$ training data per transmitter. The improved initialization removes the problem of misconvergence in this SNR range, but increasing the length of the training sequence reduces the efficiency of the communications system since less information symbols are transmitted per frame.

For low and medium SNRs, the Gaussian noise is dominant and “regularizes” the initial estimate. For this scenario, the MLSD-$m$ receiver was again almost always able to accurately detect the channel orders of the different outputs. Table II shows the relative frequency of channel order misdetection for every receiving antenna and several values of the SNR.

### C. Scenarios $m_1 = 4$, $m_2 = 1$, and $m_3 = 1$: Time-Invariant Channel

If a time-invariant channel is considered, the results noticeably differ. Fig. 6 shows the BER achieved by the four algorithms when the CIR is constant along the entire frame. The range of SNRs in which we evaluate the performance of the algorithms is here different because an SNR of 15 dB is now enough for the algorithms to reach very low BERs (below $10^{-5}$ for all the algorithms, except the conventional PSP algorithm). In this case, the difference between the performance of the MLSD-$m$ and G-PSP algorithms is very small for low-to-medium SNRs (about 1 dB for a BER of $10^{-2}$) and almost negligible for high SNRs (above 12 dB). However, note that, for SNR = 15, both algorithms are less than 0.5 dB away from the (optimal) Viterbi algorithm.

In terms of channel estimation, the MLSD-$m$ algorithm is however still clearly superior to the G-PSP, as shown in Fig. 7.
These previous results suggest that taking into account several subchannel orders is especially advantageous when the channel is time varying.

**D. Scenarios $m_1 = 4$, $m_2 = 4$, and $m_3 = 1$: Time-Varying Channel**

Fig. 8 shows the BER when two of the outputs have the maximum channel order $m_1 = m_2 = 4$, and hence perfectly fit those assumed by the PSP and the G-PSP algorithms, whereas the remaining one has no memory at all, i.e., $m_3 = 1$, meaning that it is severely overestimated by the PSP and G-PSP methods. Both the G-PSP and the MLSD-$m$ receivers perform about 2 dB worse than the genie-aided MLSD receiver. Therefore, overestimating the channel order of a single output does not have a remarkable impact on the performance of the G-PSP algorithm. It is also worth mentioning that, in this particular case, the difference between the standard PSP and the G-PSP is more blatant (about 2 dB for a BER of $10^{-4}$).

As for the channel estimation, the MSE of the different algorithms is shown in Fig. 9. It can be seen that the proposed algorithm performs slightly better than the G-PSP, but the gap between the curves is very small because, in this scenario, both algorithms are estimating a similar number of coefficients (as opposed to the scenario in Section V-A, in which the number of coefficients estimated by the MLSD-$m$ algorithm was small, compared with that of the G-PSP).

In this scenario, all the algorithms exhibit performance improvement in terms of BER due to the use of a channel with more diversity. For the same reason, the detection of the channel orders in the MLSD-$m$ algorithm is now more accurate. This is illustrated by Table III, which shows the relative frequency of channel order misdetection for every receiving antenna and three values of the SNR. (For SNR $\geq 9$ dB, all the frequencies are zero.)

**E. Estimating a Single Channel Order Versus Estimating One Channel Order Per Output**

It is interesting to study what kind of performance gain can be expected when estimating one channel order per output, instead
of a single overall channel order.\textsuperscript{9} With that purpose, we have compared the algorithm in [4] for estimating a single overall channel order (labeled “MLSD-$m$ (single channel order)” here) with the MLSD-$m$ algorithm proposed in this paper, which estimates one order per receiving antenna.

The simulation setup is the same as before (cf. Section V for details), except for the number of survivors considered, which has now been reduced to $P = 128$. Regarding the subchannel orders, two of the previous scenarios are separately analyzed.

Fig. 10(a) shows the BER achieved by both algorithms when the first two receiving antennas have channel order $m_1 = m_2 = 2$, whereas the third one has channel order $m_3 = 1$. It can be seen that, in this case, the gap between the curves of the MLSD-$m$ (single channel order) and the MLSD-$m$ algorithms is negligible (only for very high SNRs the MLSD-$m$ exhibits some advantage), and hence, estimating several channel orders rather than a single one does not pose any noticeable advantage. The reason is that both algorithms are estimating a similar number of coefficients. Indeed, assuming that the two of them achieve their goals (the MLSD-$m$ (single channel order) estimates an overall channel order $\hat{m}_1 = 2$ and the MLSD-$m$ estimates $\hat{m}_1 = 2$, $\hat{m}_2 = 2$, and $\hat{m}_3 = 1$), the number of channel coefficients that need to be estimated by the MLSD-$m$ is $(2 + 2 + 1) \times N_T = 15$, and the number of coefficients that need to be estimated by the MLSD-$m$ (single channel order) is $N_R \times 2 \times N_T = 18$.

The BER of the algorithms for the different SNRs when $m_1 = 4$ and $m_2 = m_3 = 1$ is shown in Fig. 10(b). In this scenario, the performance exhibited by the MLSD-$m$ algorithm is significantly better than that of the MLSD-$m$ (single channel order) for low-medium SNRs, i.e., $\sim 1$ dB of improvement when BER $\approx 10^{-4}$. Since $m_{\text{max}} = 4$, the latter estimates $N_R \times 4 \times N_T = 36$ coefficients, whereas the former only needs to estimate $(4 + 1 + 1) \times N_T = 18$ coefficients. That allows the MLSD-$m$ algorithm to take advantage of more accurate estimates of the channel coefficients. In this scenario, the benefits of estimating one channel order per receiving antenna, instead of a single (overall) channel order, become clear.

\textbf{F. Impact of the Number of Survivors on the Performance}

To assess the impact of the number of survivors in the performance of our proposed algorithm, we have conducted experiments to study the evolution of the BER as the number of survivors increases for different sets of subchannel orders. The results from the simulations are shown in two separate pictures:

\textsuperscript{9} Recall from Section III-B that this improvement comes at no cost since estimating one channel order per output does not entail any penalty on the computational complexity with respect to estimating a single channel order.
be estimated. This is shown in Fig. 11(a) when the number of survivors is $P = 2$; in such case, the BER of the proposed algorithm is higher for $m_1 = 4$, $m_2 = 1$, and $m_3 = 4$ than it is for $m_1 = 4$, $m_2 = 1$, and $m_3 = 2$.

When the maximum order among the subchannel orders $(m_{\text{max}})$ changes, the results are quite similar. Fig. 11(b) shows the BER of the MLSD-$m$ algorithm for different sets of subchannel orders and different numbers of survivors. Here, the first subchannel order $m_1$ takes values on the set $\{2, 3, 4\}$, whereas the other subchannel orders are fixed to $m_2 = m_3 = 1$. Again, if the number of survivors is large enough ($P \geq 8$ in the figure), then the algorithm behaves better (in terms of BER) with high diversity channels (many coefficients) than with low diversity channels. See, e.g., the BERs for $P = 64$ in the figure. For a small number of survivors, the opposite behavior is observed. This is, again, justified by the need to obtain accurate channel estimates.

VI. CONCLUSION

In this paper, we have introduced a method based on the PSP principle for the equalization of frequency- and time-selective MIMO channels with one (unknown) channel order per output. The computational complexity of the proposed scheme is not increased, compared to previously proposed receivers that estimate a single channel order in a similar framework. We have numerically assessed the performance of the new scheme in several scenarios and observed that it clearly outperforms the standard PSP algorithm that considers a single channel order (equal to the maximum possible value) for the entire MIMO channel.

Unlike the conventional PSP algorithm, where a given number of survivors is maintained for every possible state of the trellis, our approach selects the survivors globally (that is, regardless of the state they entail). To determine the extent to which the performance gain of the proposed algorithm comes from the use of this selection strategy, we have also implemented a modified version of the classical PSP algorithm. The modified technique also assumes a single channel order equal to the maximum possible channel order, but the survivors are selected globally, instead of per state. Our simulation results show that, even though the performance of this algorithm (termed G-PSP) is usually better than that of the standard PSP, it is still inferior to that of the proposed method, which accounts for the different subchannel orders that actually exist in the model. Moreover, we have shown that the performance of the G-PSP can severely deteriorate if the algorithm is not appropriately initialized.

We have also numerically verified that estimating one channel order per receiving antenna can bring a performance improvement over methods that estimate a single order for the complete MIMO channel. Furthermore, this improvement is, in our simulations, larger when the CIR is time varying than in scenarios where the channel remains fixed.

APPENDIX A

In this appendix, we formally prove the validity of (8). We start on the left-hand side of that equation by integrating out the unknown channel to obtain

$$p(y_t|s_{0:t}, y_{0:t-1}) = \int p(y_t|H_t, s_{0:t}, y_{0:t-1})dH_t = \int p(y_t|H_t, s_{0:t}, y_{0:t-1})p(H_t|s_{0:t}, y_{0:t-1})dH_t. \quad (16)$$

From (4), it is clear that the $l$th element of vector $y_t$ only depends on the $l$th row of matrix $H_t$, which we will denote by $h(t,l)$, rather than on the entire matrix. Therefore, we can write

$$p(y_t|H_t, s_{0:t}, y_{0:t-1}) = \prod_{l=1}^{N_R} p(y_t(l)|h(t,l), s_{0:t}). \quad (17)$$

Conversely, only observations coming from the $l$th output contain information about $h(t,l)$, and since the channel coefficients are a priori independent (see Section II), we arrive at the factorization

$$p(H_t|s_{0:t}, y_{0:t-1}) = \prod_{l=1}^{N_R} p(h(t,l)|s_{0:t}, y_{0:t-1}(l)). \quad (18)$$
Substituting (17) and (18) in (16) yields

\[ p(y_t | s_{0:t}, y_{0:t-1}) = \prod_{t=1}^{N_T} p(y_t | h_t(l), s_{0:t}) \prod_{t=1}^{N_T} p(h_t(l) | s_{0:t}, y_{0:t-1}(l)) \, dH_t \]

\[ = \prod_{t=1}^{N_R} \int_{y_{0:t-1}(l)} p(y_t | h_t(l), s_{0:t}) p(h_t(l) | s_{0:t}, y_{0:t-1}(l)) \, dh_t(l) \]

\[ = \prod_{t=1}^{N_R} p(y_t | s_{0:t}, y_{0:t-1}(l)) . \quad (19) \]

### APPENDIX B

Assuming a specific survivor path \( s^{(i)}_{0:t} \) and a particular value of \( m_1 \in M \), a linear-Gaussian random dynamic system in state-space form can be associated with the \( l \)th receiving antenna according to (2) and (1), i.e.,

\[ h_t(l) = a h_{t-1}(l) + v_t \quad \text{and} \quad y_t(l) = s^{(i)}_{m_1,t} h_t(l) + g_t(l) \quad (20) \]

where \( h_t(l) \) is an \( N_T m_1 \times 1 \) vector comprising the first \( N_T m_1 \) coefficients in the \( l \)th row of matrix \( H_t \). \( v_t \) is a Gaussian random vector with pdf \( p(v_t) = N \left( 0, \sigma_v^2 I_{N_T m_1} \right) \), and

\[ s^{(i)}_{m_1,t} = \begin{bmatrix} s^{(i)}_{t-1,m_1+1} & \cdots & s^{(i)}_{t-1,m_1+N_T-1} \end{bmatrix}^{\top} . \quad (21) \]

Assuming \( p(h_{t-1}(l)) \) to be Gaussian and known, the posterior pdf \( p(h_t(l) | m_1, y_{0:t}(l), s^{(i)}_{0:t}) \) is also Gaussian and can be recursively computed using the KF [19]. In particular, if

\[ p \left( h_{t-1}(l) | m_1, y_{0:t-1}(l), s^{(i)}_{0:t-1} \right) = N \left( h_{t-1}(l), C_{t-1}^{(i)}(l) \right) \quad (22) \]

then

\[ p \left( h_t(l) | m_1, y_{0:t-1}(l), s^{(i)}_{0:t-1} \right) = N \left( h_t(l), C_{t}^{(i)}(l) \right) \]

where

\[ \tilde{h}_{t-1}^{(i)}(l) = \alpha \tilde{h}_{t-1}^{(i)}(l) \]

\[ C_{t-1}^{(i)}(l) = \alpha^2 C_{t-1}^{(i)}(l) + \sigma_v^2 I_{N_T m_1} \]

and

\[ p \left( h_t(l) | m_1, y_{0:t}, s^{(i)}_{0:t} \right) = N \left( h_t(l), C_{t}^{(i)}(l) \right) , \]

where

\[ \tilde{h}_{t}^{(i)}(l) = \tilde{h}_{t-1}^{(i)}(l) + k_t^{(i)}(l) \left( y_t(l) - s^{(i)}_{m_1,t} h_{t-1}(l) \right) \]

\[ C_{t}^{(i)}(l) = \left( I_{N_T m_1} - k_t^{(i)}(l) s^{(i)}_{m_1,t}^{\top} \right) C_{t-1}^{(i)}(l) \]

Notice that, according to the signal model given in Section II, the last \( N_T (m_{\text{max}} - m_1) \) coefficients in the \( l \)th row of matrix \( H_t \) (those associated with the channel matrices \( H_t(i) \) with \( i \geq m_1 \)) are zero.

with the Kalman gain vector being

\[ k_t^{(i)}(l) = C_{t-1}^{(i)}(l) \left( s^{(i)}_{m_1,t}^{\top} C_{t-1}^{(i)}(l) s^{(i)}_{m_1,t} + \sigma_{v,t}^2 \right)^{-1} \]

The marginal likelihood

\[ p \left( y_t(l) | m_1, y_{0:t-1}(l), s^{(i)}_{0:t} \right) = \int p \left( y_t(l) | h_t(l), s^{(i)}_{m_1,t} \right) \times p \left( h_t(l) | m_1, y_{0:t-1}(l), s^{(i)}_{0:t-1} \right) \, dh_t(l) \quad (27) \]

is the integral of the product of two Gaussian pdfs and, as a consequence, is itself Gaussian, i.e.,

\[ p \left( y_t(l) | m_1, y_{0:t-1}(l), s^{(i)}_{0:t} \right) = \mathcal{N} \left( y_t(l) | \tilde{h}_{t-1}^{(i)}(l) \alpha v_t, \tilde{h}_{t}^{(i)}(l) S^{(i)}_{m_1,t} + \sigma_{v,t}^2 \right) . \quad (28) \]

The latter result enables the computation of the likelihoods on the right-hand side of (9).

### REFERENCES


Manuel A. Vázquez received the M.Sc. degree and the Ph.D. degree in computer engineering from the University of A Coruña, A Coruña, Spain, in 2003 and 2008, respectively. From April 2008 to March 2009, he was a Researcher with Tampere University of Technology, Tampere, Finland. From February to August 2011, he was a Visiting Researcher with Nanyang Technological University, Singapore. He is currently a Ph.D. Assistant Professor with the Departamento de Teoría de la Señal y Comunicaciones, Universidad Carlos III de Madrid, Leganés, Spain. His research interests include particle filtering, multiple-input–multiple-output systems, dynamic programming, positioning, and electroencephalogram signal processing.

Joaquín Míguez (S’98–A’00–M’03) received the M.Sc. and Ph.D. degrees in computer engineering from the University of A Coruña, A Coruña, Spain, in 1997 and 2000, respectively. Late in 2000, he joined the Department of Electronics and Systems, University of A Coruña, where he became an Associate Professor in July 2003. From April to December 2001, he was a Visiting Scholar with the Department of Electrical and Computer Engineering, State University of New York at Stony Brook, Stony Brook. In September 2004, he moved to the Departamento de Teoría de la Señal y Comunicaciones, Universidad Carlos III de Madrid, Leganés, Spain, as an Associate Professor. He has coauthored more than 30 international journal papers in the areas of electrical engineering, mathematical physics, and statistics. His research interests are statistical signal processing, Bayesian analysis, dynamical systems, and the theory and applications of Monte Carlo methods. Dr. Míguez was a co-recipient of the IEEE Signal Processing Magazine Best Paper Award 2007.