

Adapting Side Information to Transmission Conditions in Precoding Systems

Josmary Labrador, Paula M. Castro^(✉), Adriana Dapena,
and Francisco J. Vazquez-Araujo

Department of Computer Engineering, University of A Coruña,
Campus de Elviña 8, 15071 A Coruña, Spain
pcastro@udc.es

Abstract. This work proposes an hybrid precoding method in a Multiple-Input/Multiple-Output Frequency Division Duplex (MIMO FDD) system with the objective of reducing the load associated to transmit side information needed to adapt precoding matrices in both the transmitter and the receiver. The type of precoding is determined at the transmitter by using a simple rule that takes into account a receive Signal-to-Noise Ratio (SNR) estimate. The receiver computes the magnitude of the channel level fluctuations and determines the time instants when long pilot sequences are needed to estimate the precoding matrices. Using a low cost feedback channel, the receiver indicates to the transmitter both the type of precoder and transmit frames to be used.

1 Introduction

Computational Intelligence (CI)-based approaches have been widely used to solve different problems in digital communications and networking such as call admission control, management of resources and traffic, routing, multi casting, media encoding, and synchronization [5]. CI paradigms include supervised and unsupervised learning, reinforcement learning, fuzzy logic, evolutionary computation, etc. In this paper, we propose to include a decision-based learning in precoding systems to improve the transmission rate.

Precoding is an effective strategy to equalize the channel before transmission, and is included in most of the recent wireless standards with the aim of simplifying the receiver equipment by moving the equalization task to the transmitter. Linear and nonlinear precoding techniques have been widely studied in the literature. Tomlinson-Harashima Precoding (THP) is one of the best known nonlinear precoding techniques due to its adequate compromise between performance and computational complexity. THP computational complexity is still 1.6 times higher than the linear counterpart, which will be referred to as Linear Precoding (LP), as shown in [8]. In [10] we show that the combination of both precoders, which is referred to as Hybrid Precoding (HP), allows us to improve the overall system performance.

Precoding needs Channel State Information (CSI) at the transmitter, which must be obtained at the receiver by channel estimation and sent back to the

transmitter by means of a feedback channel, usually available in recent standards for Frequency Division Duplex (FDD) systems. However, this feedback channel is not necessary in Time Division Duplex (TDD) systems because the channel can be obtained by the transmitter in the uplink using reciprocity. This Partial CSI (PCSI) affects the system performance since the design of the precoding filters are based on the channel estimate and, therefore, linear and non-linear precoders will need a good channel estimate under time-varying environments [3]. Classical estimation methods are based on the sending from the transmitter of pilot symbols that are used by the receiver to obtain the channel estimate. The performance achieved with such methods, also known as *supervised learning based-methods*, is high, but the use of pilots affects throughput, spectral efficiency, and transmission energy consumption of the system [1, 12]. In this work, we propose to mitigate these limitations by using reinforcement learning in which the optimal policy will be determined to minimize the total pilot transmissions.

We assume that both the transmitter and the receiver are two individual entities with some capacity for decision, communication, and adaptation. The receiver is able to acquire channel information from the environment and then makes decisions as a consequence of these measurements. More specifically, this decision uses rules based on measurements of both the quality of the received signal (measured in terms of Signal-to-Noise Ratio (SNR)) and the channel fluctuations (measured according to an ad hoc metric also proposed in this work). The decisions will be communicated via the aforementioned low-cost feedback channel to the transmitter, which will send pilots symbols when a significant channel fluctuation is detected at the receiver. Then, with the information provided by the pilots, the receiver estimates the channel and sends this estimate to the transmitter. Both the transmitter and the receiver will adapt their precoding filters using the linear or the non-linear approaches as indicated by the receiver according to its receive SNR measurement. Therefore, we have an adaptive system that guarantees good performance with low complexity.

This work is organized as follows. Section 2 describes the signal and channel models, and shows the designs of the linear and non-linear precoders most commonly used in the literature. Section 3 briefly describes the supervised method for channel estimation used in this work. Section 4 explains the proposed scheme for minimizing the number of pilot symbols required by the receiver. Illustrative computer simulation results are presented in Sects. 5, and 6 contains the conclusions.

2 System Model

Figure 1 shows a MIMO system with N_t TX antennas and N_r RX antennas. In this paper we will assume $N_t = N_r = N$. We can model the received observations as

$$\mathbf{y}[n] = \mathbf{H}[q]\mathbf{x}[n] + \eta[n], \quad (1)$$

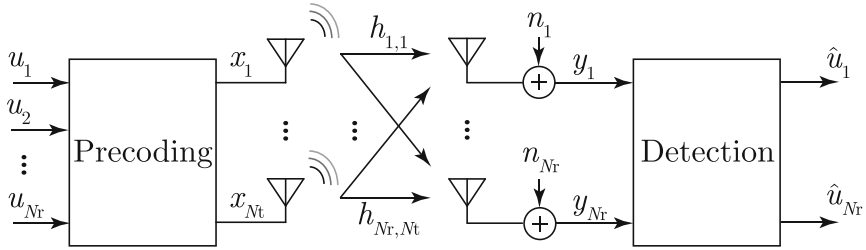


Fig. 1. Scheme of a MIMO system with precoding.

where $n = 0, 1, 2, \dots$ corresponds to the sample index. Given that the channel remains constant during several frames of N_B symbols, we use $\mathbf{H}[q]$ to denote the time-varying flat block fading channel.

The equalization task can be performed at the TX and thus the channel is pre-equalized or *precoded* before the transmission with the goal of simplifying the requirements at the RX. Such an operation is only possible when a centralized TX is employed (e.g. the base-station of the downlink of a cellular system). We have considered the linear and non-linear precoding schemes more commonly used in the literature: Wiener LP and THP, respectively.

2.1 Linear Precoder

We assume hereinafter that the RX filter is an identity matrix (multiplied by a scalar $\beta[q]$, with $\beta[q] \in \mathbb{C}$), which allows the use of decentralized RX (see, for instance, [6]). Clearly, the restriction that all receivers apply the same scalar weight $\beta[q]$ is not necessary for decentralized receivers, but it ensures closed-form solutions for the design of the filters. The goal is to find the optimum TX filter $\mathbf{F}[q] \in \mathbb{C}^{N \times N}$ and the RX filter $\mathbf{G}[q] = \beta[q]\mathbf{I} \in \mathbb{C}^{N \times N}$. The data symbols $\mathbf{u}[n]$ are passed through the transmit filter $\mathbf{F}[q]$ to form the transmitted signal $\mathbf{x}[n] = \mathbf{F}[q]\mathbf{u}[n] \in \mathbb{C}^N$. Note that the constraint for the transmitted energy must be fulfilled, $\mathbb{E}[\|\mathbf{x}[n]\|_2^2] \leq E_{\text{tx}}$, where E_{tx} is the fixed total transmitted energy. The received signal is thus given by

$$\mathbf{y}[n] = \mathbf{H}[q]\mathbf{F}[q]\mathbf{u}[n] + \boldsymbol{\eta}[n], \quad (2)$$

where $\mathbf{y}[n] \in \mathbb{C}^N$, $\mathbf{H}[q] \in \mathbb{C}^{N \times N}$, and $\boldsymbol{\eta}[n] \in \mathbb{C}^N$ is the Additive White Gaussian Noise (AWGN). After multiplying by the receive gain $\beta[q]$, we get the estimated symbols $\hat{\mathbf{u}}[n] = \beta[q]\mathbf{H}[q]\mathbf{F}[q]\mathbf{u}[n] + g\boldsymbol{\eta}[n]$, where $\hat{\mathbf{u}}[n] \in \mathbb{C}^N$.

Wiener Filtering (WF) is a very powerful transmit optimization that minimizes the Mean Square Error (MSE) with a transmit energy constraint [2, 7, 11], and therefore the linear precoders of our proposal will be obtained according to that optimization.

2.2 Tomlinson-Harashima Precoder

In this subsection, we will briefly describe the Tomlinson-Harashima (TH) non-linear precoder, which will be used in this paper. This precoder employs two filters: one, denoted by $\mathbf{F}[q]$, placed at the transmitter to suppress parts of the interference linearly, and another one, given by $\mathbf{I} - \mathbf{B}[q]$, inside a feedback loop and also at the transmitter to subtract the remaining interferences non-linearly, with \mathbf{B} being strictly lower triangular to ensure the causality of the feedback process. Since the order of precoding has an important effect on performance, the data signal $\mathbf{u}[n]$ is reordered by means of the permutation filter $\mathbf{P}[q] = \sum_{i=1}^N \mathbf{e}_i \mathbf{e}_{n_i}^T$, where \mathbf{e}_i is the i -th column of the $N \times N$ identity matrix and n_i is the index of the i -th data stream to be precoded [8]. The signal $\mathbf{P}[q]\mathbf{u}[n]$ is first passed through the feedback loop to get the output $\mathbf{v}[n]$. The nonlinear modulo operator $\mathbf{M}(\bullet)$ of the feedback loop limits the amplitude of $\mathbf{v}[n]$ and thus, the power of the transmit signal $\mathbf{x}[n]$. The received signal is expressed as

$$\mathbf{y}[n] = \mathbf{M}(g[q]\mathbf{H}[q]\mathbf{F}[q]\mathbf{v}[n] + g\eta[n]), \quad (3)$$

because the modulo operator is applied again at the receiver to invert its effect at the transmitter [11]. The receive weight $g[q]$ directly follows from the transmit energy constraint. The resulting estimate of $\mathbf{u}[n]$ is denoted again by $\hat{\mathbf{u}}[n]$.

The Wiener THP for flat fading channels results from the minimization of the MSE and the restriction of a spatially causal feedback filtering. The filters obtained from that minimization are determined column by column [8, 9, 11], and each column requires one matrix inverse which results in a total complexity order of $O(N^4)$. With the decomposition described in [4, 13], the complexity is reduced to $O(N^3)$. In addition, some heuristic ordering strategies can be applied as described in [8].

3 Supervised Channel Estimation

Channel estimation is crucial in wireless communication systems. In this work CSI is acquired at the receiver and sent back to the transmitter via a feedback channel so that the precoding filters can be updated at both link sides. This channel estimation can be performed by means of pilot symbols, also called *training sequences*.

When pilots are employed, the received signal $\mathbf{Y}[q]$ is a linear combination of the transmitted signals $\mathbf{S}[q]$ as follows

$$\mathbf{Y}[q] = \mathbf{S}[q]\mathbf{H}[q] + \eta[q] \in \mathbb{C}^{K \times N}, \quad (4)$$

where K is the length of the pilot sequence. The matrix $\eta[q] \in \mathbb{C}^{K \times N}$ is the AWGN with covariance matrix denoted as \mathbf{C}_η . Thus, the channel estimate is obtained as

$$\hat{\mathbf{H}}[q] = \mathbf{W}[q]\mathbf{Y}[q], \quad (5)$$

where $\mathbf{W}[q] \in \mathbb{C}^{N \times K}$ is the matrix that calculates the estimate from the observations.

The Linear Minimum Mean Square Error (LMMSE) channel estimation minimizes the average MSE between the channel and its estimate, which leads to the final expression for the MMSE linear filter

$$\mathbf{W}[q] = \mathbf{C}_{\mathbf{H}\mathbf{Y}}\mathbf{C}_{\mathbf{Y}}^{-1}, \quad (6)$$

where $\mathbf{C}_{\mathbf{H}\mathbf{Y}} = \mathbf{C}_{\mathbf{H}}\mathbf{S}^{\mathbf{H}}[q]$ and $\mathbf{C}_{\mathbf{Y}} = \mathbf{S}[q]\mathbf{C}_{\mathbf{H}}\mathbf{S}^{\mathbf{H}}[q] + \mathbf{C}_{\eta}$, being $\mathbf{C}_{\mathbf{H}} = N\mathbf{I}$. Therefore, the channel estimate can be obtained as

$$\hat{\mathbf{H}}[q] = \mathbf{C}_{\mathbf{H}}\mathbf{S}^{\mathbf{H}}[q](\mathbf{S}[q]\mathbf{C}_{\mathbf{H}}\mathbf{S}^{\mathbf{H}}[q] + \mathbf{C}_{\eta})^{-1}\mathbf{Y}[q]. \quad (7)$$

4 Decision-Aided Precoding System

In this section we propose a MIMO system with decision-aided precoding that requires the updating of the precoding filters depending on the channel fluctuations. This system will be referred to as Decision-aided Precoding (DP) in the following. The goal of this solution is to reduce the computational complexity of the overall system without penalizing in a significant way its performance. In standard systems, the pilots are transmitted in all the frames, which produces a strong degradation of performance, spectral efficiency, and transmit energy. With our approach we will be able to minimize the loss of effective transmission rate or the channel overload produced by the sending of pilot symbols.

We will consider that the transmitter sends two types of frames: *classic* and *user frames*. The *classic frames* contain a long pilot sequence and user data symbols. The *user frames* contain a short pilot sequence and user data symbols.

For determining if the channel variations are important enough to request the sending of classic frames and the updating of the precoding filters, we propose a metric that compares the estimate of the channel matrix corresponding to the current frame, denoted by $\hat{\mathbf{H}}[n]$, and that estimated in the previous frame, denoted by $\hat{\mathbf{H}}[n-1]$. Both estimates are obtained by the receiver using the short pilot sequence of the user frames, which will calculate for each transmit frame the matrix $\mathbf{\Gamma}[n] = (\hat{\mathbf{H}}[n])^{-1}\hat{\mathbf{H}}[n-1]$. In particular, we will use the error measurement, denoted as ϵ_{CSI} , as follows

$$\epsilon_{\text{CSI}} = \frac{1}{N} \sum_{i=1}^N \sum_{j=1, j \neq i}^N \left(\frac{|\gamma_{ij}[n]|^2}{|\gamma_{ii}[n]|^2} + \frac{|\gamma_{ji}[n]|^2}{|\gamma_{ii}[n]|^2} \right), \quad (8)$$

where $\gamma_{ii}[n]$ is the i -th diagonal entry of the matrix $\mathbf{\Gamma}[n]$. Thus, this value, that shows the distance between $\mathbf{\Gamma}[n]$ and the identity matrix, gives us a measurement of the channel time variations. If ϵ_{CSI} is high, the channel is suffering from significant fluctuations and therefore, the receiver will request a classic frame including a long pilot sequence to the transmitter. The receiver will estimate the channel from pilots using LMMSE and the updated coefficients will be sent to the transmitter using the feedback channel. Transmit and receive precoding filters will be updated at both link sides. Otherwise, if ϵ_{CSI} is low, the precoding filters remain unchanged as had been used in the previous frame.

Moreover, in [10] we have demonstrated that LP is better than THP for low SNRs, and vice versa. Therefore, we propose to include a decision rule at the receiver to determine the action to be performed by our system, as follows

State 1: A user frame is received

Actions:

Estimate $\hat{\mathbf{H}}[q]$ using the pilot sequence in the supervised algorithm
 Compute $\epsilon_{CSI}[q]$, estimate the receive SNR, and determine $p_{i,\text{SNR}}$
If $\epsilon_{CSI}[q] \geq p_{i,\text{SNR}}$, require a classic frame
otherwise, require an user frame.

State 2: A classic frame is received

Actions:

Estimate $\hat{\mathbf{H}}[q]$ using the pilot symbols in the supervised algorithm
 Estimate the receive SNR
If $\text{SNR} > \text{SNR}_l$, update THP using $\hat{\mathbf{H}}[q]$;
otherwise, update LP using $\hat{\mathbf{H}}[q]$
 Send $\hat{\mathbf{H}}[q]$ to the transmitter and require an user frame

Notice that $p_{i,\text{SNR}}$ and SNR_l are the two thresholds of our decision-aided algorithm that will be determined in a training step prior to real transmission.

5 Simulation Results

In this section we will show some results obtained from computer simulations. First, the time-varying channel will be modeled as follows

$$\mathbf{H}[q] = \begin{cases} \frac{(1-\alpha)\mathbf{H}[q-1] + \alpha\mathbf{H}_R[q]}{\sqrt{(1-\alpha)^2 + \alpha^2}} & \text{if } q = bF, b = 1, 2, \dots \\ \mathbf{H}[q-1] & \text{otherwise,} \end{cases} \quad (9)$$

where F is the number of frames in which the channel remains unchanged. $\mathbf{H}_R[q]$ is randomly generated following a Rayleigh distribution. The α parameter determines the speed in channel variations. If $\alpha = 0$ the channel is constant, whereas for $\alpha = 1$ the channel changes randomly from one block to another.

Additionally, the following simulation parameters are considered: $N = 4$ transmit and receive antennas; 1000 independent experiments; 128 channel realizations in each experiment; 512 frames of 128 symbols; $F = 4$ frames in which the channel remains unchanged; $P_l = 12$ QPSK pilot symbols per long pilot sequence; $P_s = 4$ QPSK per short pilot sequence; LMMSE channel estimation, and $\alpha = 0.2$ in (9).

5.1 Training Step

In a training step prior to transmission we have evaluated the distance between the performance, evaluated in terms of Bit Error Rate (BER), obtained with

both precoders, LP and THP, when the channel information is partially known at the transmitter. Classic frames are transmitted so that the long pilot sequence included in these frames are used to obtain the CSI via LMMSE estimation.

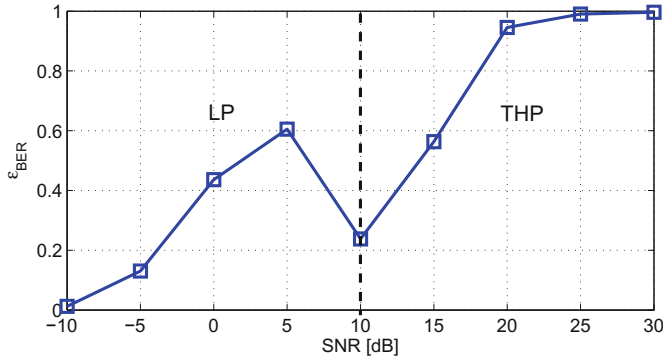


Fig. 2. Training step: SNR_i for using LP or THP.

Therefore, the range of application of each type of precoder is determined by using the following distance measurement

$$\epsilon_{\text{BER}} = \frac{|\text{BER}_{\text{LP}} - \text{BER}_{\text{THP}}|}{\text{BER}_{\text{LP}}}. \quad (10)$$

Figure 2 shows this merit figure as a function of the receive SNR. Taking into account these results, we have decided to consider an SNR threshold, denoted as SNR_1 , of 10 dB in (4), so that LP is used for SNR values equal or less than 10 dB and THP for SNR higher than that value.

Otherwise, we need to calculate the threshold values $p_{i,\text{SNR}}$ of (4) to decide if pilot symbols are required or not for channel estimation. For this purpose, for each SNR we consider only the values of $\epsilon_{\text{CSI}}[q]$ obtained every 4 frames, i.e. when the channel changes. Then, we calculate the threshold as the i -th percentile, where the $i\%$ of those values are lower than this threshold, and the $100 - i\%$ are greater. Table 1 shows the threshold values $p_{i,\text{SNR}}$ as a function of receive SNR. We have selected the percentiles 1, 2, and 5 to illustrate the performance of our decision-aided system. These percentiles will be respectively denoted as $p_{1,\text{SNR}}$, $p_{2,\text{SNR}}$, and $p_{5,\text{SNR}}$.

5.2 Transmission Step

In our experiment we will consider an $\text{SNR}_1 = 10$ dB and the values of $p_{i,\text{SNR}}$ in Table 1, obtained during the training step.

Figure 3 (top) shows the performance in terms of BER of the proposed DP scheme for $p_{1,\text{SNR}}$, $p_{2,\text{SNR}}$ and $p_{5,\text{SNR}}$. Notice that the floor effect is produced

Table 1. $p_{i,\text{SNR}}$ thresholds.

SNR [dB]	-10	-5	0	5	10	15	20	25	30
$p_{1,\text{SNR}}$	0.6285	0.5537	0.3292	0.1418	0.0570	0.0295	0.0210	0.0179	0.0169
$p_{2,\text{SNR}}$	0.7290	0.6493	0.3933	0.1680	0.0682	0.0350	0.0249	0.0211	0.0201
$p_{5,\text{SNR}}$	0.9180	0.8200	0.5228	0.2241	0.0903	0.0464	0.0328	0.0276	0.0264

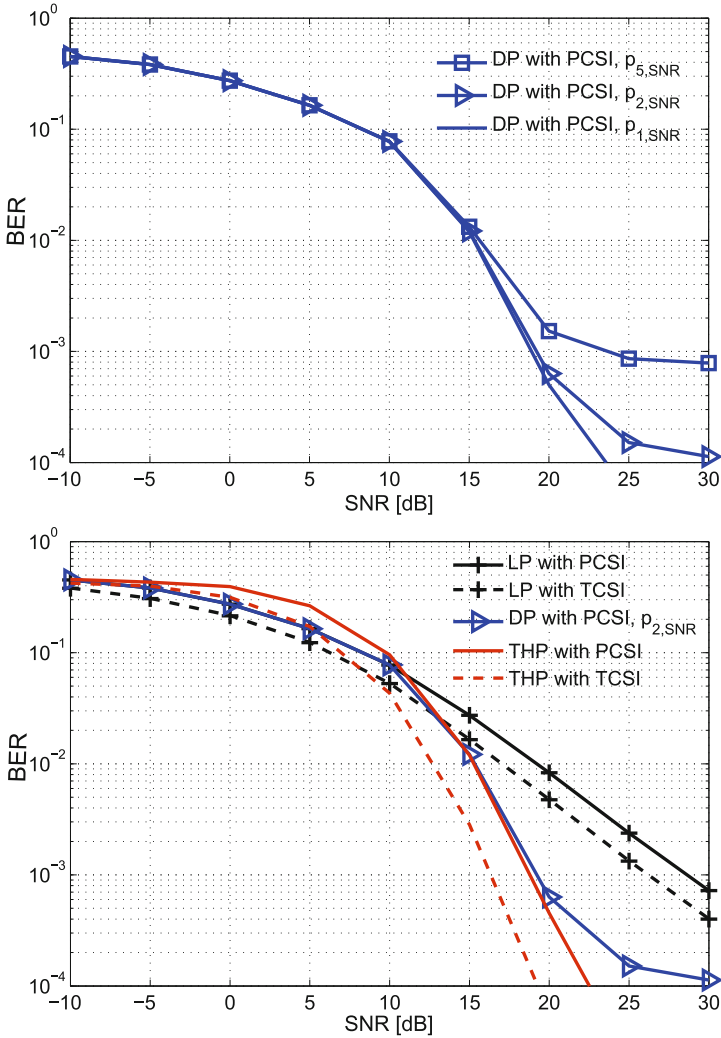


Fig. 3. BER vs. SNR for LP, THP, and DP.

by the use of a precoder that is not adapted to the actual channel state due to the channel fluctuations not being strong enough to trigger a filter update. The curve corresponding to DP for $p_{2,\text{SNR}}$ exhibits a medium performance with a floor

Table 2. Percentage of precoder updates and pilot reduction as a function of receive SNR in dB.

SNR [dB]	-10	-5	0	5	10	15	20	25	30
Precoder updates [in %]	98.05	97.86	97.75	97.09	94.77	85.28	62.41	40.41	30.21
ϵ_{pilot} [in %]	34.30	34.43	34.50	34.94	36.49	42.81	58.06	72.73	79.53

effect for SNRs higher than 30 dB. In Fig. 3 (bottom), we compare the results for LP, THP, and the proposed DP. As boundary cases, we have included the curves for LP and THP with Total CSI (TCSI), i.e. perfect CSI at the transmitter. The curve corresponding to DP for $p_{2,\text{SNR}}$ exhibits a medium performance, close to that achieved with LP for low SNR and to that obtained with THP for high SNR, according to the decision rule of (4).

Considering a threshold $p_{2,\text{SNR}}$, Table 2 shows the percentage of filter updates and the reduction in pilot symbols computed using the following expression

$$\epsilon_{\text{pilot}} = \left(1 - \frac{N_u P_s + N_c P_l}{(N_u + N_c) P_l} \right) \times 100, \quad (11)$$

where N_u and N_c are the number of user frames and classic frames, respectively. We can see that the reduction of pilot symbols is higher than 34% for all SNRs. The reduction is considerable for SNR higher than 15 dB. In addition, for SNR higher than 15 dB, the THP filter is updated in a reduced number of times which implies a considerable improvement in terms of computational load.

6 Conclusions

In this paper a decision-aided MIMO hybrid precoding system with partial transmit CSI is proposed. The system increases the effective data rate (or spectral efficiency) by minimizing the overhead caused by the transmission of pilot symbols. This is achieved by means of limiting the number of updates of the precoding filters to the time instants in which the channel significantly varies according to a given threshold, which is fixed prior to transmission in a training step. As shown with simulation results, the loss in performance is not very significant, especially if adequate decision thresholds are selected.

Acknowledgments. This work has been funded by the Galician Government under grants ED431C 2016-045 and ED341D R2016/012 as well as by the Spanish Government under grants TEC2013-47141-C4-1-R (RACHEL project) and TEC2016-75067-C4-1-R (CARMEN project).

References

1. Castro, P.M., Garca-Naya, J.A., Dapena, A., Iglesia, D.: Channel estimation techniques for linear precoded systems: supervised, unsupervised, and hybrid approaches. *Sig. Process.* **91**(7), 1578–1588 (2011). <http://www.sciencedirect.com/science/article/pii/S0165168411000028>

2. Choi, R.L., Murch, R.D.: New transmit schemes and simplified receiver for MIMO wireless communication systems. *IEEE Trans. Wirel. Commun.* **2**(6), 1217–1230 (2003)
3. Fischer, R.F.H.: *Precoding and Signal Shaping for Digital Transmission*. Wiley, Hoboken (2002)
4. Golub, G.H., Van Loan, C.F.: *Matrix Computations*, vol. 3. JHU Press, Baltimore (2012)
5. Guturu, P.: Computational intelligence in multimedia networking and communications: trends and future directions. In: Hassaniien, A.-E., Abraham, A., Kacprzyk, J. (eds.) *Computational Intelligence in Multimedia Processing Recent Advances*, pp. 51–76. Springer, Heidelberg (2008)
6. Hunger, R., Joham, M., Utschick, W.: Extension of linear and nonlinear transmit filters for decentralized receivers. In: *European Wireless 2005*, vol. 1, pp. 40–46, April 2005
7. Karimi, H.R., Sandell, M., Salz, J.: Comparison between transmitter and receiver array processing to achieve interference nulling and diversity. In: *Proceedings of the PIMRC*, vol. 3, pp. 997–1001 (1999)
8. Kusume, K., Joham, M., Utschick, W., Bauch, G.: Efficient Tomlinson-Harashima precoding for spatial multiplexing on flat MIMO channel. In: *Proceedings of the International Conference on Communications*, Seoul, Korea, vol. 3, pp. 2021–2025, May 2005
9. Kusume, K., Joham, M., Utschick, W., Bauch, G.: Cholesky factorization with symmetric permutation applied to detecting and precoding spatially multiplexed data streams. *IEEE Trans. Signal Process.* **55**(6), 3089–3103 (2007)
10. Labrador, J., Castro, P.M., Vazquez-Araujo, F.J., Dapena, A.: Hybrid precoding scheme with partial CSI at the transmitter. In: *16th International Conference on Knowledge-Based and Intelligent Information and Engineering Systems (KES 2012)*, September 2012
11. Nossek, J.A., Joham, M., Utschick, W.: Transmit processing in MIMO wireless systems. In: *Proceedings of the 6th IEEE Circuits and Systems Symposium on Emerging Technologies: Frontiers of Mobile and Wireless Communication*, Shanghai, China, pp. I-18 - I-23, May/June 2004
12. Scharf, L.L.: *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Addison-Wesley, Boston (1991)
13. Schnorr, C.P., Euchner, M.: Lattice basis reduction: improved practical algorithms and solving subset sum problems. *Math. Program.* **66**(1), 181–199 (1994). <http://dx.doi.org/10.1007/BF01581144>