Depósito de investigación de la Universidad de Sevilla

https://idus.us.es/

This is an Accepted Manuscript of an article published by **IEEE Transactions on Vehicular Technology** on febrero/2012 available at: https://doi.org/10.1109/SAM.2014.6882359

"© 2011 IEEE. Personal use of this material is permitted. Permission from IEEEmust be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other Works"

Near the Cramér−Rao Bound Precoding Algorithms for OFDM Blind Channel Estimation

Francisco J. Simois, Juan José Murillo-Fuentes, Rafael Boloix-Tortosa and Luis Salamanca

***Abstract*—The authors present a blind channel estimation of cyclic prefix (CP) OFDM systems with non-redundant precoding based on second-order statistics. The study analyzes first the mean square error for the estimation of the covariance matrix of the received symbols. We prove that for high and medium signal-to-noise ratios (SNRs) the estimation error in diagonal entries of the covariance matrix exhibits a lower error than that of the off-diagonal elements. This behavior holds for SNR values in digital communication. Contrary to general belief, we prove that the diagonal of this matrix can be used for channel estimation. Hence, we develop a novel algorithm that utilizes this result. We also develop a low complexity version that provides acceptable results with reduced computational requirements. Finally, we analyze the covariance matrix and propose another new algorithm with noise suppression capabilities. Some experimental results for Rayleigh channels are included to support these conclusions. Also, they illustrate the better performance of the new methods compared to previous proposals and to the Cramér−Rao bound.**

***Index Terms*—Blind channel estimation, OFDM, non-redundant precoding, cyclic prefix (CP), variance of the estimation of a covariance matrix, Cramér−Rao bound (CRB).**

I. Introduction

Orthogonal frequency division multiplexing (OFDM) has received considerable interest in the last few years for its advantages over frequency-selective fading channels in high-rate data transmissions. These features have justified the adoption of OFDM as the transmission technique for many latest generation systems: digital audio and video broadcasting (DAB, DVB), digital subscriber lines (xDSL) or wireless local area networks (IEEE 802.11a, ETSI BRAN HIPERLAN/2). In OFDM, blocks of symbols are transmitted in parallel over several narrowband subchannels at different orthogonal subcarriers. Prior to its transmission, each block is processed by the inverse discrete Fourier transform (IDFT), and a cyclic prefix (CP), i.e., a copy of the last part of the block, is inserted. This procedure enables a very simple equalization of frequency-selective finite impulse response (FIR) channels.

Coherent detection and adaptive loading in OFDM systems require reliable channel state information. In general, the channel impulse response (CIR) is estimated by transmitting some training sequence periodically, which decreases the spectral efficiency. If we aim at avoiding this payload, we can resort to blind approaches. One of the most attractive and simplest solutions is to use non-redundant precoding at the transmitter. The result of this precoding stage is a new set of transmitted symbols that are a mixture of the original ones. The received signals are then multiplied point by point by the DFT of the channel. At this juncture, one can rely on blind source separation (BSS) techniques [1],[2], to recover the original data from the mixtures. However, such techniques do not require knowledge of the precoding matrix and are based on high-order statistics (HOS).

Many approaches have been recently proposed to estimate the CIR by using the covariance matrix of the outputs. This matrix has a known structure introduced by the precoder that can be exploited in the channel estimation. The authors in [3] proposed the extraction of the CIR from only one column of the covariance matrix. A similar work was proposed in [4] where the channel is again estimated from a single column of the crosscorrelation matrix of two consecutive received blocks; then, a direct average is applied. Later, the limited accuracy of these methods was overcome using more entries of the matrix. The authors in [5] present a new method that obtains the channel estimation from all the off-diagonal entries of the covariance matrix. The method is generalized for MIMO-OFDM systems in [6]. Also, the authors in [7] propose a singular value decomposition (SVD), and in [8] an eigenvalue decomposition of the whole covariance matrix to estimate the channel. The computational complexity of these latter methods can be reduced further by projecting the matrix into the time domain, as shown in this paper. A common feature of all these methods is that the diagonal entries of the matrix are avoided. Even in the latter methods, based on the eigenvectors of the matrix, the diagonal entries play a minor role for a large number of subcarriers. It is general belief that the estimation of the channel is more accurate if the diagonal entries of the covariance matrix are avoided, as they are corrupted by the unknown noise variance. This paper shows that this is not true for practical signal-to-noise ratios (SNR).

We analyze the mean square error for the estimation of the covariance matrix at the receiver to show that, for sufficiently high SNRs, the diagonal entries exhibit lower errors than the off-diagonal ones. Therefore, the estimation of the channel can be enhanced by using just the diagonal entries of the covariance matrix to estimate the absolute value of the DFT of the channel. This is the starting point to develop a novel algorithm for the blind estimation of the channel, called the *diagonal algorithm* (DA), which performs better than the previous ones with the same complexity. Also, a version that shows good performance and reduced complexity, similar to that of the one-column method in [3], is presented. We will

Copyright (c) 2011 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

This work was partially funded by the Spanish government (Ministerio de Educación y Ciencia TEC2009-14504-C02-02 and Consolider-Ingenio 2010 CSD2008-00010) and the European Union (FEDER).

The authors are with the Dept. Teoría de la Señal y Comunicaciones, Escuela T. S. de Ingeniería, Universidad de Sevilla, Paseo de los Descubrimientos s/n, 41092 Sevilla, Spain. E-mail: {fjsimois,murillo,rboloix,salamanca}@us.es

refer to this algorithm as the *simple diagonal algorithm* (SDA).

In some scenarios, the estimation of the channel in high-noise environment is of interest [9], [10], [11], where both the DA and SDA provide poor estimation. We avoid this problem by estimating the variance of the noise and subtracting it from the diagonal of the covariance matrix prior to the estimation of the channel with the DA or SDA. However, a quite accurate estimation of the variance of the noise is needed to achieve the same results of the DA and SDA at high SNR. The analysis of the square error of the covariance matrix proves to be useful, as discussed in this paper. Hence, we present a new algorithm, the diagonal with noise removal algorithm (DNA), which outperforms previous approaches in the literature, for all SNRs. Furthermore, we use the Cramér−Rao bound (CRB) to show that this method is close to the optimal estimation.

This paper is organized as follows. Section II presents the system model of OFDM precoded transmissions. Section III provides a general approach to the channel estimation. Section IV thoroughly studies the mean square error of the estimation of the covariance matrix. These errors are evaluated for some standard scenarios in section V. In Section VI, we develop the DA and SDA while the algorithm with noise removal capability, the DNA, is described in Section VII. In Section VIII, we describe the computational complexity of the proposed methods and compare them to previous approaches. Section IX presents various simulations to show the effectiveness of our algorithms. Section X is devoted to conclusions. The paper ends with some appendices.

The notations used in the paper are as follows. If $\mathbf{A}$ is a matrix, $\mathbf{A}_{r,q}$ is its (r,q) entry. To denote a range of values from r to q, $r\!:\!q$ is used, e.g., $\mathbf{A}_{r:q,k}$ is a column vector, the values r to q of the column k of $\mathbf{A}$. $\mathbf{A}_{:,k}$ is the kth column of $\mathbf{A}$. $\mathbf{A}^{\mathrm{T}}$ represents the transpose of $\mathbf{A}$ and $\mathbf{A}^{\mathrm{H}}$ its Hermitian. $\mathrm{tr}\,(\mathbf{A})$ is the trace of $\mathbf{A}$ and $\mathrm{diag}\,(\mathbf{A})$ is a column vector with the diagonal of $\mathbf{A}$. $\mathbf{a}_r$ is the r-th entry of vector $\mathbf{a}$. $\mathrm{diag}\,(\mathbf{a})$ is a diagonal matrix whose diagonal is vector $\mathbf{a}$. $||\mathbf{a}||$ is the L_2 norm of vector $\mathbf{a}$. $\mathbf{a}_{((r-q))_m}$ is the $((r-q) \bmod m+1)$-th entry of $\mathbf{a}$. $\mathrm{E}\{\cdot\}$ denotes statistical expectation. Finally, $\odot$ and $\oslash$ stand for the Hadamard (pointwise) product and division, respectively.

II. Nonredundant Precoding for CP-OFDM

In a CP-OFDM system with precoding, the transmitter can be described by the baseband discrete-time block equivalent model depicted in Fig. 1. The symbols first enter an S/P converter that arranges the symbols in an $m\times 1$ column vector $\mathbf{s}(k)$. Then, this vector is multiplied by $\mathbf{W}$, an $m \times m$ non-redundant precoding matrix. Later, we apply the inverse DFT (IDFT) by multiplying by $\mathbf{F}^{\mathrm{H}}$, the Hermitian of the $m \times m$ normalized FFT matrix, $\mathbf{F}$, whose (r,q)th entry is

$$\mathbf{F}_{r,q} = \frac{1}{\sqrt{m}}\, \mathrm{e}^{-j2\pi(r-1)(q-1)/m}, \tag{1}$$

for $1 \leq q,\ r \leq m$. After the IDFT and precoding stages, the last l symbols are padded at the beginning to avoid the interblock and the interchannel interferences. The symbols in the blocks are serialized and sent sequentially through the channel. Propagation is modeled as an FIR channel with

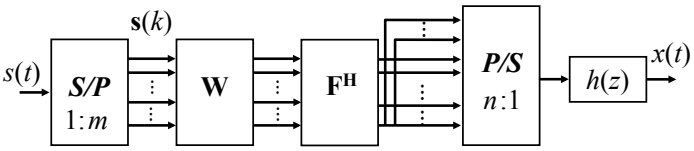

Fig. 1. Discrete-time block baseband equivalent model of precoded CP-OFDM transmitter and channel.

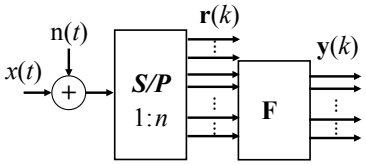

Fig. 2. Discrete-time block baseband equivalent model of CP-OFDM receiver.

impulse response $\mathbf{h} = [h_0 \cdots h_l]^{\mathrm{T}}$. It is assumed here that the maximum channel order is l (i.e., $h_i = 0, \forall i > l$).

In Fig. 2, additive white Gaussian noise (AWGN) is added in the channel. At the receiver side, a serial-to-parallel conversion is performed to obtain the $(m+l)\times 1$ vector $\mathbf{r}$. Then, the first l outputs are discarded so that a DFT can be applied later. The received block of m symbols at instant k yields [5],

$$\mathbf{y}(k) = \mathbf{F}\,[\mathbf{0}|\mathbf{I}]\,\mathbf{r}(k) = \mathbf{F}\mathbf{H}\mathbf{F}^{\mathrm{H}}\mathbf{W}\mathbf{s}(k) + \mathbf{n}(k), \tag{2}$$

where $\mathbf{n}$ is a vector whose entries are AWGN with covariance matrix $\sigma_n^2\mathbf{I}$, and $\mathbf{H}$ is the $m \times m$ circulant channel matrix with its (r,q) entry given by $\overline{\mathbf{h}}_{((r-q))_m}$. The $m \times 1$ vector $\overline{\mathbf{h}}$ is obtained by inserting $m-l-1$ zeros at the end of $\mathbf{h}$. At this point, vector $\mathbf{y}$ is used to estimate the CIR.

The model in (2) can be written in a simple and direct form as follows,

$$\mathbf{y}(k) = \tilde{\mathbf{H}}\mathbf{W}\mathbf{s}(k) + \mathbf{n}(k), \tag{3}$$

where $\tilde{\mathbf{H}} = \mathbf{F}\mathbf{H}\mathbf{F}^{\mathrm{H}} = \mathrm{diag}\left(\tilde{\mathbf{h}}\right)$ is a diagonal $m \times m$ matrix whose diagonal is the m-point DFT of the CIR,

$$\tilde{\mathbf{h}} = \sqrt{m}\mathbf{F}_{:,1:l+1}\mathbf{h}. \tag{4}$$

III. Channel Estimation

A. General Approach

The model in (3) matches that of a blind source separation (BSS) problem [2]. Matrix $\tilde{\mathbf{H}}$ is diagonal, and $\mathbf{W}$ is known. Hence, a simple decorrelation of the outputs ensures separation and estimation of the diagonal of $\tilde{\mathbf{H}}$, hereafter denoted as $\hat{\tilde{\mathbf{h}}}$. The channel estimation, up to a scale factor, can be accomplished by computing $\hat{\mathbf{h}} = (\mathbf{F}_{:,1:l+1})^{\mathrm{H}}\hat{\tilde{\mathbf{h}}}/\sqrt{m}$. This projection reduces the dimension from m to $l+1$, the number of unknowns, thereby reducing also the computational complexity.

The covariance matrix of the outputs can be written as

$$\mathbf{R}_{\mathbf{y}} = \mathrm{E}\{\mathbf{y}(k)\mathbf{y}^{\mathrm{H}}(k)\} = \tilde{\mathbf{H}}\mathbf{W}\mathbf{R}_{\mathbf{s}}\mathbf{W}^{\mathrm{H}}\tilde{\mathbf{H}}^{\mathrm{H}} + \sigma_n^2\mathbf{I}, \tag{5}$$

where $\mathbf{R_s}$ is the covariance matrix of the transmitted data. As $\tilde{\mathbf{H}}$ is diagonal, this matrix yields

$$\mathbf{R_y} = \mathbf{W}\mathbf{R_s}\mathbf{W}^{\mathrm{H}} \odot \tilde{\mathbf{h}}\tilde{\mathbf{h}}^{\mathrm{H}} + \sigma_n^2\mathbf{I}. \quad (6)$$

In the following, we assume the transmitted data to be uncorrelated, $\mathbf{R_s} = \sigma_s^2\mathbf{I}$. For the sake of simplicity we define $\mathbf{P} = \mathbf{W}\mathbf{W}^{\mathrm{H}}$ and $\mathbf{C} = \tilde{\mathbf{H}}\mathbf{W}\mathbf{W}^{\mathrm{H}}\tilde{\mathbf{H}}^{\mathrm{H}}$. It follows that

$$\mathbf{R_y} = \sigma_s^2\mathbf{C} + \sigma_n^2\mathbf{I} = \sigma_s^2\mathbf{P} \odot \tilde{\mathbf{h}}\tilde{\mathbf{h}}^{\mathrm{H}} + \sigma_n^2\mathbf{I}. \quad (7)$$

Given N received samples, $\mathbf{Y} = [\mathbf{y}(0)|\mathbf{y}(1)|...|\mathbf{y}(N-1)]$, we replace in (7) the covariance matrix by its estimation, $\hat{\mathbf{R}}_\mathbf{y} = \mathbf{Y}\mathbf{Y}^{\mathrm{H}}/N$. The channel is usually estimated mainly, or exclusively, by focusing on the off-diagonal entries of the following relation,

$$\hat{\tilde{\mathbf{h}}}\hat{\tilde{\mathbf{h}}}^{\mathrm{H}} = \frac{1}{\sigma_s^2}\hat{\mathbf{R}}_\mathbf{y} \oslash \mathbf{P} - \frac{\sigma_n^2}{\sigma_s^2}\mathbf{I} \oslash \mathbf{P}, \quad (8)$$

because diagonal elements are corrupted with the noise variance, [3], [5], [7], [8].

B. An MSE approach

An approach to solve (8) is to neglect the noise and find the vector $\hat{\tilde{\mathbf{h}}}$ that better fits the equation

$$\hat{\tilde{\mathbf{h}}}\hat{\tilde{\mathbf{h}}}^{\mathrm{H}} = \frac{1}{\sigma_s^2}\hat{\mathbf{R}}_\mathbf{y} \oslash \mathbf{P} = \mathbf{B}. \quad (9)$$

Note that for a large number of carriers, (8) and (9) only differ in a few equations, viz., the diagonal ones. Hence, using (9) we approximate (8), not taking into account the error introduced by the variance of the noise in the estimation of the covariance matrix of the outputs.

The solution to (9) can be easily computed by minimizing the Frobenius norm between $\mathbf{B}$ and $\hat{\tilde{\mathbf{h}}}\hat{\tilde{\mathbf{h}}}^{\mathrm{H}}$. Note that this loss function is the mean (averaged) value of the square errors between corresponding entries of matrices $\mathbf{B}$ and $\hat{\tilde{\mathbf{h}}}\hat{\tilde{\mathbf{h}}}^{\mathrm{H}}$. This minimization has one unique solution, up to a phase mismatch, given by the eigenvector associated to the largest eigenvalue [7]. In [8], the authors use the power method to avoid extensive computations in the decomposition. In the power method, $\tilde{\mathbf{H}}$ is estimated iteratively with a constrained learning law. Although it reduces the computational complexity compared to a standard eigendecomposition algorithm, it works on an m-dimensional space and has linear convergence. We propose to minimize the mean square error between matrices $\mathbf{B}$ and $\tilde{\mathbf{h}}\tilde{\mathbf{h}}^{\mathrm{H}}$ but project this error into the low-dimensional discrete time domain:

$$\mathrm{L} = \min_{\hat{\mathbf{h}}} \left(\mathrm{tr}\left((\mathbf{V} - \hat{\mathbf{h}}^{\mathrm{H}}\hat{\mathbf{h}}) \cdot (\mathbf{V} - \hat{\mathbf{h}}^{\mathrm{H}}\hat{\mathbf{h}})^{\mathrm{H}} \right) \right), \quad (10)$$

where $\mathbf{V} = (\mathbf{F}_{:,1:l+1})^{\mathrm{H}}\mathbf{B}\mathbf{F}_{:,1:l+1}$. To minimize it we estimate the eigenvector of $\mathbf{V}$ with dominant eigenvalue by using the shifted power method [12], [13], [14] as follows

$$\hat{\mathbf{h}}_{\mathrm{MSE}} \leftarrow \hat{\mathbf{h}}_{\mathrm{MSE}} + \mu\mathbf{V}\hat{\mathbf{h}}_{\mathrm{MSE}}, \quad (11)$$

$$\hat{\mathbf{h}}_{\mathrm{MSE}} \leftarrow \hat{\mathbf{h}}_{\mathrm{MSE}} / \| \hat{\mathbf{h}}_{\mathrm{MSE}} \|. \quad (12)$$

This algorithm is always stable regardless of the value of μ and it converges exponentially with the rate $\frac{1+\mu\rho_2}{1+\mu\rho_1}$, where ρ_1 and ρ_2 are the largest and the second largest eigenvalues of $\mathbf{V}$ [13], [14]. While the eigendecomposition approach in [8] estimates an m-dimensional vector, the proposed method estimates a vector with dimension $l+1$. Hence, we reduce the computational complexity in the optimization stage from order $\mathrm{O}(m^2)$ to $\mathrm{O}((l+1)^2)$, where m is usually much larger than $l+1$.

IV. On the estimation of the covariance matrix

As already discussed, we usually assume the covariance matrix of the transmitted data in (7)-(8) to be the identity matrix multiplied by σ_s^2. However, in practice, the covariance matrix of the outputs has to be estimated from N samples as $\hat{\mathbf{R}}_\mathbf{y} = \mathbf{Y}\mathbf{Y}^{\mathrm{H}}/N$. Hence,

$$\hat{\mathbf{R}}_\mathbf{y} = \tilde{\mathbf{H}}\mathbf{W}\hat{\mathbf{R}}_\mathbf{s}\mathbf{W}^{\mathrm{H}}\tilde{\mathbf{H}}^{\mathrm{H}} + \tilde{\mathbf{H}}\mathbf{W}\hat{\mathbf{R}}_\mathbf{sn} + \hat{\mathbf{R}}_\mathbf{ns}\mathbf{W}^{\mathrm{H}}\tilde{\mathbf{H}}^{\mathrm{H}} + \hat{\mathbf{R}}_\mathbf{n}, \quad (13)$$

where $\hat{\mathbf{R}}_\mathbf{ab} = \mathbf{A}\mathbf{B}^{\mathrm{H}}/N$, for some matrices $\mathbf{A}$ and $\mathbf{B}$. The question that arises is whether the errors in the estimation of diagonal entries of $\hat{\mathbf{R}}_\mathbf{y}$ are larger than those in the estimation of off-diagonal ones.

A. Variance of the estimation error

We propose to study the error in the estimation of the covariance matrix in (13). The mean value of this estimator is

$$\mathrm{E}\{\hat{\mathbf{R}}_\mathbf{y}\} = \sigma_s^2\mathbf{C} + \sigma_n^2\mathbf{I}. \quad (14)$$

The expression

$$\Sigma = \mathrm{E}\{|\hat{\mathbf{R}}_\mathbf{y} - \sigma_s^2\mathbf{C} - \delta_n\sigma_n\mathbf{I}|^2\}, \quad (15)$$

where the square operator is entrywise, is the variance of the estimator for $\delta_n = 1$. By setting $\delta_n = 0$, (15) yields the mean square error between the estimation of the covariance matrix and its mean value in the absence of noise. The value of the entries of (15) is given by

$$\begin{aligned}\Sigma_{r,q} = {} & \sigma_s^4(\mathbf{C}_{r,r}\mathbf{C}_{q,q} - \Theta_{r,q})/N + \sigma_s^2\sigma_n^2(\mathbf{C}_{r,r} + \mathbf{C}_{q,q})/N \\ & + (1+\delta_{rq})\sigma_n^4/N + \delta_{rq}(1-\delta_n)\sigma_n^4. \end{aligned} \quad (16)$$

where $\delta_{rq} = 1$ if $r = q$ and 0 otherwise,

$$\Theta_{r,q} = \Phi_{r,q}|\tilde{\mathbf{h}}_r|^2|\tilde{\mathbf{h}}_q|^2, \quad (17)$$

and, in turn,

$$\Phi_{r,q} = (2 - \mu_s/\sigma_s^4)\sum_{\alpha=0}^{m-1} |\mathbf{W}_{r,\alpha}|^2 |\mathbf{W}_{q,\alpha}|^2, \quad (18)$$

where $\mu_s = \mathrm{E}\{s^4\}$ is the fourth-order moment of the inputs. This expression can be easily computed for diagonal and off-diagonal entries, by setting δ_{rq} to 1 or 0, respectively. (See Appendix A for some guidelines on the computation of (16).)

B. Crosspoint

We denote the signal-to-noise ratio (SNR) by $\gamma = \sigma_s^2/\sigma_n^2$. In this paper, we prove that, at high SNRs, the diagonal entries of the estimation of the covariance matrix exhibit a lower estimation error. The crosspoint between the variance of the error in diagonals and off-diagonal entries given by (16), as a function of γ, can be easily computed as the solution to

$$\begin{aligned}&\sigma_s^2(\sigma_s^2\mathbf{C}_{r,r}+\sigma_n^2)(\mathbf{C}_{r,r}-\mathbf{C}_{q,q})\\&-\sigma_s^4(\Theta_{r,r}-\Theta_{r,q})+(1+N-\delta_n N)\sigma_n^4=0.\end{aligned}\tag{19}$$

This yields the second-order equation

$$\begin{aligned}&\gamma^2\left[(\Theta_{r,r}-\Theta_{r,q})-\mathbf{C}_{r,r}(\mathbf{C}_{r,r}-\mathbf{C}_{q,q})\right]\\&-\gamma(\mathbf{C}_{r,r}-\mathbf{C}_{q,q})-(1+N-\delta_n N)=0.\end{aligned}\tag{20}$$

The term $(2-\mu_s/\sigma_s^4)$ in (18) can be computed for any modulation. In this paper, we focus on M-QAM modulations, for which it can be proved that

$$(2-\mu_s/\sigma_s^4)=\frac{3(M^2-1)}{5(M-1)^2}.\tag{21}$$

C. Estimation for random channels

The earlier expressions can be extended for the more general case of random channels by taking the statistical expectation of (16). If we make use of the definition of $\mathbf{C}$, it follows that

$$\begin{aligned}\Sigma_{r,q}=&\sigma_s^4(\mathbf{P}_{r,r}\mathbf{P}_{q,q}-\Phi_{r,q})\mathrm{E}\{|\tilde{\mathbf{h}}_r|^2|\tilde{\mathbf{h}}_q|^2\}/N\\&+\sigma_s^2\sigma_n^2(\mathbf{P}_{r,r}\mathrm{E}\{|\tilde{\mathbf{h}}_r|^2\}+\mathbf{P}_{q,q}\mathrm{E}\{|\tilde{\mathbf{h}}_q|^2\})/N\\&+(1+\delta_{rq})\sigma_n^4/N+\delta_{rq}(1-\delta_n)\sigma_n^4.\end{aligned}\tag{22}$$

Appendix B proves that this expression yields

$$\begin{aligned}\Sigma_{r,q}=&\sigma_s^4(\mathbf{P}_{r,r}\mathbf{P}_{q,q}-\Phi_{r,q})\\&\cdot\left[\tilde{\mathbf{e}}_1^2+|\tilde{\mathbf{e}}_{((r-q))_m}|^2-\sum_{\alpha=0}^{l}\left(2\sigma_{h_\alpha}^4-\mu_{h_\alpha}\right)\right]/N\\&+\sigma_s^2\sigma_n^2(\mathbf{P}_{r,r}+\mathbf{P}_{q,q})\tilde{\mathbf{e}}_1/N\\&+(1+\delta_{rq})\sigma_n^4/N+\delta_{rq}(1-\delta_n)\sigma_n^4,\end{aligned}\tag{23}$$

where

$$\tilde{\mathbf{e}}=\sqrt{m}\mathbf{F}_{:,1:l+1}\mathbf{e}\tag{24}$$

is the m-point DFT of the variances of the coefficients of the CIR, $\mathbf{e}=[\sigma_{h_0}^2,\sigma_{h_1}^2,...,\sigma_{h_l}^2]^{\mathrm{T}}$, $\sigma_{h_\alpha}^2$ and μ_{h_α} are the variance and the fourth-order moment of the α-th channel tap, respectively. Again, this expression can be easily computed for diagonal and off-diagonal entries, by setting δ_{rq} to 1 or 0, respectively. The crosspoint between them can be calculated by solving the expectation of (20). Again, using the definitions of $\mathbf{C}$ and Φ, (20) yields

$$\begin{aligned}&\gamma^2\left[(\mathbf{P}_{r,r}\mathbf{P}_{q,q}-\Phi_{r,q})\,|\tilde{\mathbf{h}}_r|^2|\tilde{\mathbf{h}}_q|^2+\left(\Phi_{r,r}-\mathbf{P}_{r,r}^2\right)|\tilde{\mathbf{h}}_r|^4\right]\\&-\gamma\left(\mathbf{P}_{r,r}|\tilde{\mathbf{h}}_r|^2-\mathbf{P}_{q,q}|\tilde{\mathbf{h}}_q|^2\right)-(1+N-N\delta_n)=0,\end{aligned}\tag{25}$$

and after computing its expectation,

$$\begin{aligned}&\gamma^2\left[(\mathbf{P}_{r,r}\mathbf{P}_{q,q}-\Phi_{r,q})(\tilde{\mathbf{e}}_1^2+|\tilde{\mathbf{e}}_{((r-q))_m}|^2)\right.\\&-\left(\mathbf{P}_{r,r}^2-\Phi_{r,r}\right)2\tilde{\mathbf{e}}_1^2-\left(\mathbf{P}_{r,r}\mathbf{P}_{q,q}-\mathbf{P}_{r,r}^2+\Phi_{r,r}-\Phi_{r,q}\right)\\&\left.\cdot\sum_{\alpha=0}^{l}\left(2\sigma_{h_\alpha}^4-\mu_{h\alpha}\right)\right]-\gamma(\mathbf{P}_{r,r}-\mathbf{P}_{q,q})\tilde{\mathbf{e}}_1\\&-(1+N-N\delta_n)=0.\end{aligned}\tag{26}$$

V. Estimation errors in some scenarios

The previous section showed that the values $\mathbf{C}_{r,r}$ and $\Theta_{r,q}$ depend on the channel and the precoding matrix. In the following, we analyze some standard scenarios. As discussed in Section III-A, noise is usually ignored in the estimation of the channel. Hence, we set $\delta_n=0$ to estimate the mean square error for the estimation of the covariance matrix.

A. AWGN channel and no precoding

Assume we transmit an M-QAM, the channel is AWGN, in (5) we have no precoding, i.e. $\mathbf{W}=\tilde{\mathbf{H}}=\mathbf{C}=\mathbf{I}$, and the covariance matrix of the transmitted data is $\mathbf{R_s}=\sigma_s^2\mathbf{I}$. Then, at reception the covariance yields $\mathbf{R_y}=\sigma_s^2\mathbf{I}+\sigma_n^2\mathbf{I}$. We compute the mean square error of the estimation of the covariance matrix of the outputs as $\Sigma=\mathrm{E}\{|\hat{\mathbf{R}}_\mathbf{s}-\sigma_s^2\mathbf{I}|^2\}$. In this scenario, with the identity matrix as precoder, (17) yields

$$\begin{aligned}\Theta_{r,r}^{(\mathrm{I})}&=2-\mu_s/\sigma_s^4,\\\Theta_{r,q}^{(\mathrm{I})}&=0,\ r\neq q.\end{aligned}\tag{27}$$

and equation (16),

$$\Sigma_{r,r}=(\sigma_s^2+\sigma_n^2)^2/N-\sigma_s^4\Theta_{r,r}^{(\mathrm{I})}/N+\sigma_n^4/N+\sigma_n^4,\tag{28}$$

$$\Sigma_{r,q}=(\sigma_s^2+\sigma_n^2)^2/N,\ \ q\neq r.\tag{29}$$

As the SNR increases, we have

$$\lim_{\mathrm{SNR}\to\infty}\Sigma_{r,r}=(1-\Theta_{r,r}^{(\mathrm{I})})\sigma_s^4/N,\tag{30}$$

$$\lim_{\mathrm{SNR}\to\infty}\Sigma_{r,q}=\sigma_s^4/N,\ \ q\neq r.\tag{31}$$

Hence, at suffciently high SNR, the difference between the estimation error of the diagonal entries and the off-diagonal ones is $(2\sigma_s^4-\mu_s)/N$. Besides, it follows that the quotient of errors is

$$\lim_{\mathrm{SNR}\to\infty}\frac{\Sigma_{r,q}}{\Sigma_{r,r}}=\left(1-\Theta_{r,r}^{(\mathrm{I})}\right)^{-1}.\tag{32}$$

This ratio can be read as the achieved gain for large SNR when the diagonal entries are used instead of the off-diagonal ones for estimating a given parameter. For an M-QAM, the quotient of errors can be easily computed by introducing (21) into (32). For $M=4$ the gain is infinite, and for $M=64$ it is 4.2 dB. The solution to (20) is the crosspoint between (28) and (29). In this scenario, the linear term in (20) vanishes and the solution is

$$\gamma=\sqrt{\frac{5}{3}\frac{(M-1)^2}{M^2-1}(1+N)}.\tag{33}$$

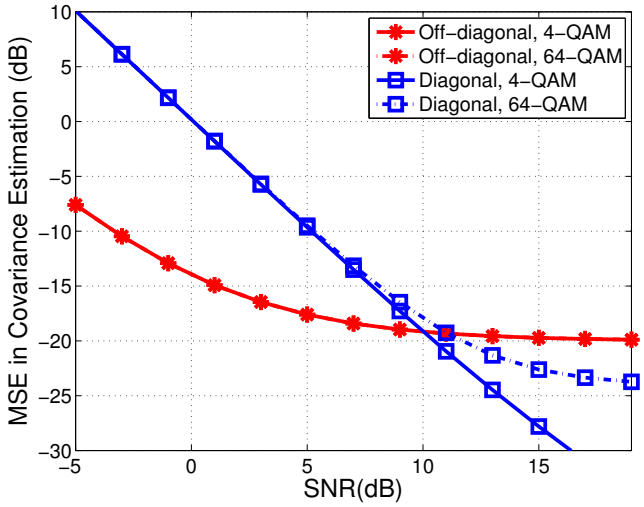

Fig. 3. Mean square error for the estimation of the diagonal (□) and the off-diagonal (∗) entries for the covariance matrix of the outputs for a memoryless channel, no precoding, and a 4-QAM modulation (solid) and a 64-QAM modulation (dot-dashed) for $N = 100$ samples.

The crosspoint in (33) yields 10.02 dB and 11.06 dB for $M = 4$ and $M = 64$, respectively and $N = 100$. For a large number of samples, e.g., $N = 1000$, we need an SNR > 16 dB to have better estimations in the diagonal entries. For $N = 10$, the threshold reduces to SNR > 6.3 dB. These crosspoints are near the values of the SNR where the optimal detectors provide useful bit error rates (BER), e.g. for a 4-QAM we need SNR ≥ 8 dB to get BER$\leq 10^{-3}$ in AWGN channels.

Fig. 3 displays the mean square error of the estimation of a diagonal (□) and an off-diagonal (∗) entry of $\hat{\mathbf{R}}_{\mathbf{y}}$ in (28)-(29). The transmitted signals were randomly generated for a 4-QAM (solid) and a 64-QAM (dot-dashed) along different signal-to-noise ratios (SNR) for $N = 100$ samples (blocks). For an SNR beyond the crosspoints, the error is lower in the diagonal entries. For the 4-QAM, a constant modulus modulation, this error approaches zero as the noise vanishes.

B. AWGN channel and Precoding

If we use a precoding matrix, the estimation of the covariance of the received symbols is given by (13) with $\tilde{\mathbf{H}} = \mathbf{I}$. The errors in the estimation of an entry of $\mathbf{R}_{\mathbf{s}}$ spread over the entries of $\mathbf{R}_{\mathbf{y}}$. This effect depends on the structure of the precoding matrix, as analyzed in this section. For large SNR, we prove that for the precoding matrix proposed in [5] the errors in the diagonal entries are lower than in the off-diagonal ones.

In the precoding matrix, $\mathbf{W}$, proposed in [5] by Gao and Nallanathan, the diagonal entries of $\mathbf{P} = \mathbf{W}\mathbf{W}^{\mathrm{H}}$ are equal to 1 and the off-diagonal entries are $p < 1$. It can be easily proved that $\mathbf{W}_{r,r} = w_1 \ \forall r$, and $\mathbf{W}_{r,q} = w_2 \ \forall q \neq r$, where $0 < w_2 < w_1 < 1$. The k-th precoded signal is the k-th input signal multiplied by w_1 plus the sum of all the other signals multiplied by w_2. Values of p close to one yield a good estimation of the channel and poor BER, while near-zero values involve poor estimates of the channel [5].

We compute the mean square error of the estimation of the covariance matrix of the outputs as $\Sigma = \mathrm{E}\{|\hat{\mathbf{R}}_{\mathbf{y}} - \sigma_s^2\mathbf{C}|^2\}$. Assume the channel is AWGN. In this scenario, $\mathbf{C} = \mathbf{P}$ and $\Theta = \Phi$. For the precoding in [5], Φ in (18) yields

$$\Phi_{r,r}^{(\mathrm{Gao})} = (2\sigma_s^4 - \mu_s)(w_1^4 + (m-1)w_2^4)/\sigma_s^4, \tag{34}$$

$$\Phi_{r,q}^{(\mathrm{Gao})} = (2\sigma_s^4 - \mu_s)(2w_1^2w_2^2 + (m-2)w_2^4)/\sigma_s^4, \ r \neq q, \tag{35}$$

and (16) simplifies to

$$\Sigma_{r,r} = (\sigma_s^2 + \sigma_n^2)^2/N - \sigma_s^4\Theta_{r,r}^{(\mathrm{Gao})}/N + \sigma_n^4/N + \sigma_n^4, \tag{36}$$

$$\Sigma_{r,q} = (\sigma_s^2 + \sigma_n^2)^2/N - \sigma_s^4\Theta_{r,q}^{(\mathrm{Gao})}/N, \ \ q \neq r. \tag{37}$$

It follows that, as the SNR increases, the gain obtained by the diagonal entries can be computed as

$$\lim_{\mathrm{SNR}\to\infty} \frac{\Sigma_{r,q}}{\Sigma_{r,r}} = \frac{1 - \Theta_{r,q}^{(\mathrm{Gao})}}{1 - \Theta_{r,r}^{(\mathrm{Gao})}}. \tag{38}$$

Hence, at sufficiently high SNR, the difference between estimation errors for the off-diagonal and the diagonal entries is

$$(\Theta_{r,r}^{(\mathrm{Gao})} - \Theta_{r,q}^{(\mathrm{Gao})})\sigma_s^4/N = \frac{3(M^2-1)}{5(M-1)^2}(w_1^2 - w_2^2)^2\sigma_s^4/N, \tag{39}$$

where the term on the right is the value particularized for an M-QAM modulation. If $0.8 \geq p \geq 0.2$, the gain in (38) is in the range $[0.9, 14.9]$ dB for $M = 4$ and $[0.6, 6.5]$ dB for $M = 64$. As the SNR decreases, the difference between these errors in (36)-(37) reduces, to cancel at some point. For SNR below this crosspoint, the off-diagonal entries are better estimated. The solution to (20) is this crosspoint. In this scenario, the linear term in (20) vanishes, as $\mathbf{C}_{r,r} = \mathbf{C}_{q,q}$, and the solution yields

$$\gamma = \sqrt{\frac{5}{3}\frac{(M-1)^2}{M^2-1}\frac{1+N}{(w_1^2 - w_2^2)^2}}, \tag{40}$$

where we make use of (39).

Fig. 4 depicts the mean square error in the estimation of a diagonal (□) entry and an off-diagonal (∗) entry of $\hat{\mathbf{R}}_{\mathbf{y}}$ in (36) and (37). The transmitted signals were randomly generated for a 4-QAM (solid) and a 64-QAM (dot-dashed) along different signal-to-noise ratios (SNR) for $N = 100$ samples (blocks). The crosspoint in (40) yields 12.4 dB and 13.5 dB for $M = 4$ and $M = 64$, respectively, for the value $p = 0.54$ proposed in [5]. For $N = 10$ and $N = 1000$, the crosspoints shift left and right, respectively, 5 dB. These crosspoints are below the values of the SNR where the optimal detectors provide useful BER in fading channels.

C. Rayleigh channel and precoding

In [5], $\mathbf{P}_{r,r} = \mathbf{P}_{q,q} = 1$, and (23) yields

$$\Sigma_{r,q} = \sigma_s^4(1 - \Phi_{r,q}^{(\mathrm{Gao})}) \cdot \left[\tilde{\mathbf{e}}_1^2 + |\tilde{\mathbf{e}}_{((r-q))_m}|^2 - \sum_{\alpha=0}^{l}\left(2\sigma_{h_\alpha}^4 - \mu_{h_\alpha}\right)\right]/N + 2\sigma_s^2\sigma_n^2\tilde{\mathbf{e}}_1/N + (1+\delta_{rq})\sigma_n^4/N + \delta_{rq}\sigma_n^4, \tag{41}$$

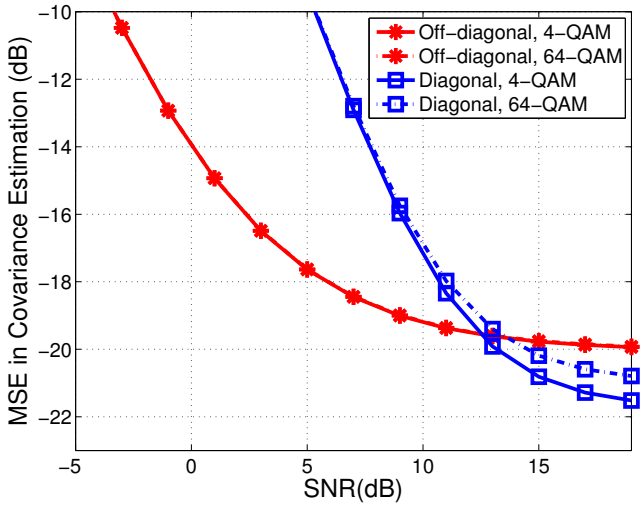

Fig. 4. Mean square error for the estimation error in the diagonal (□) and the off-diagonal (∗) entries for the covariance matrix of outputs for an AWGN channel and precoding, and a 4-QAM modulation (solid) and a 64-QAM modulation (dot-dashed) for $N = 100$ samples.

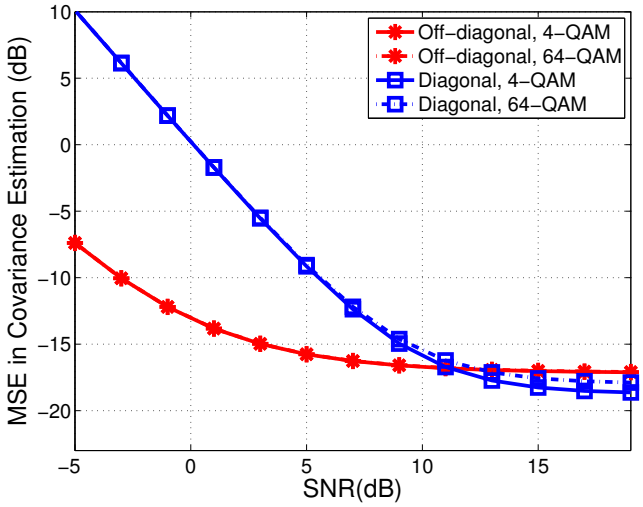

Fig. 5. Mean square error for the estimation error in the diagonal (□) and the off-diagonal (∗) entries for the covariance matrix of the outputs for a fading channel and precoding, for a 4-QAM modulation (solid) and a 64-QAM modulation (dot-dashed) for $N = 100$ samples averaged over 1000 Rayleigh channels.

where $\Phi^{(\mathrm{Gao})}$ is given in (34)-(35). For channel taps with Rayleigh-distributed envelope, $\mu_{h_r} = 2\sigma_{h_r}^4$, and (41) yields, for diagonal and off-diagonal entries,

$$\Sigma_{r,r} = \sigma_s^4(1-\Phi_{r,r}^{(\mathrm{Gao})})\cdot 2\tilde{\mathbf{e}}_1^2/N + 2\sigma_s^2\sigma_n^2\tilde{\mathbf{e}}_1/N + 2\sigma_n^4/N + \sigma_n^4, \quad (42)$$

$$\Sigma_{r,q} = \sigma_s^4(1-\Phi_{r,q}^{(\mathrm{Gao})})\cdot\left[\tilde{\mathbf{e}}_1^2 + |\tilde{\mathbf{e}}_{((r-q))_m}|^2\right]/N + 2\sigma_s^2\sigma_n^2\tilde{\mathbf{e}}_1/N + \sigma_n^4/N. \quad (43)$$

Both errors are equal at the crosspoint given in (26), which depends on the channel. In this scenario, it follows that

$$\gamma^2\left[(1-\Phi_{r,q}^{(\mathrm{Gao})})(\tilde{\mathbf{e}}_1^2 + |\tilde{\mathbf{e}}_{((r-q))_m}|^2) - (1-\Phi_{r,r}^{(\mathrm{Gao})})2\tilde{\mathbf{e}}_1^2\right] - (1+N) = 0. \quad (44)$$

Fig. 5 includes the mean square error of the estimation of a diagonal (□) entry and an off-diagonal (∗) entry of $\hat{\mathbf{R}}_{\mathbf{y}}$ in (42)-(43), with the precoding in [5] with $p = 0.54$. We model the channel by randomly generating the coefficients of the filter with a Rayleigh distribution, scaled with an exponential power-delay profile for every path as in [15], [5],

$$\sigma_{h_r}^2 = \mathrm{E}\{|h_r|^2\} = \exp(-r/10), \quad r = 0, ..., l \quad (45)$$

and the phase of each coefficient, $\angle h_r$, is uniformly distributed over $[0, 2\pi)$. We include (42) and (43) for $r = 1$ and $q = 2$ averaged over 1000 different and randomly generated channels with $l = 8$. The transmitted signals were also randomly obtained with a 4-QAM (solid) and a 64-QAM (dot-dashed) along different signal-to-noise ratios (SNR), and the covariance matrix was estimated with $N = 100$ samples (blocks). We observe that the variances of the error in both diagonal and off-diagonal entries are similar to those in the previous subsection but with a vertical shift due to the average over the Rayleigh channels. By solving (44) we have a similar crosspoint, at 11.31 dB and 12.47 dB for 4-QAM and 64-QAM, respectively, as observed in Fig. 5.

VI. The Diagonal of the Covariance Matrix in the CIR Estimation

A. The diagonal algorithm

Bearing the earlier results in mind, we conclude that the diagonal of the covariance matrix is of major importance in the estimation of the channel. Therefore, we estimate the CIR with the following method, denoted by the diagonal algorithm (DA). The diagonal of the covariance matrix provides the square of the absolute values of the DFT of the channel, (see (9)),

$$|\hat{\tilde{\mathbf{h}}}_{\mathrm{DA}\,i}| = \sqrt{\mathbf{B}_{i,i}}\ . \quad (46)$$

The angle may be estimated using a robust method such as the earlier MSE approach,

$$\angle\hat{\tilde{\mathbf{h}}}_{\mathrm{DA}} = \angle\left(\sqrt{m}\mathbf{F}_{:,1:l+1}\hat{\mathbf{h}}_{\mathrm{MSE}}\right). \quad (47)$$

Finally, the channel can be estimated as $\hat{\mathbf{h}}_{\mathrm{DA}} = (\mathbf{F}_{:,1:l+1})^{\mathrm{H}}\hat{\tilde{\mathbf{h}}}_{\mathrm{DA}}/\sqrt{m}$. This method can be improved by subtracting the noise variance, if any estimation is available, as discussed later in this paper.

B. Simple Diagonal Algorithm

The previous diagonal approach showed a good performance. However, its computational complexity can be reduced, as the estimation of the angles with the MSE method involves the whole covariance matrix. The angle of the DFT of the channel can only be estimated with off-diagonal entries, but we can use just one column of the covariance matrix to compute it. The estimation is not as accurate as the one using the MSE or the method by Gao and Nallathanan [5], but only m products and divisions are needed in its computation. The estimation of the absolute value of the DFT of the channel yields

$$|\hat{\tilde{\mathbf{h}}}_{\mathrm{SDA}\,i}| = \sqrt{\mathbf{B}_{i,i}}\ . \quad (48)$$

The angle may be estimated by the simple one-column method in [3]. If we use any column of $\mathbf{B}$, e.g., the r-th, the angle yields

$$\angle\hat{\tilde{\mathbf{h}}}_{\mathrm{SDA}\,i} = \angle\Big(\mathbf{B}_{i,r}/|\hat{\tilde{\mathbf{h}}}_{\mathrm{SDA}\,r}|\Big). \tag{49}$$

Finally, the channel can be estimated as $\hat{\mathbf{h}}_{\mathrm{SDA}} = (\mathbf{F}_{:,1:l+1})^{\mathrm{H}}\hat{\tilde{\mathbf{h}}}_{\mathrm{SDA}}/\sqrt{m}$.

VII. Noise removal

In the estimation of the covariance matrix of the outputs, noise is present in the diagonal entries. This fact has motivated other authors to estimate the absolute value of the DFT of the channel by using off-diagonal entries [7], [8], [3], [5]. On the other hand, we have proved that the estimation of the diagonal entries exhibits a lower square error for large and medium SNR. As discussed in Section V, the SNR values for which the diagonal entries are better estimated are usually the values needed for useful BER. Hence, the proposed approaches in Section VI achieve better results in this range of SNR, as illustrated in the experimental section later in this paper. However, we may aim at obtaining a better performance also in high-noise scenarios [9], [10], [11]. To improve the estimation, we propose removal of the noise from the diagonal of the covariance matrix, by previously estimating it. Because the variance of the noise is suppressed from the diagonal, we expect to have a significant improvement in the estimation error of the diagonal entries. Note that the performance of this method can be easily analyzed by studying the variance of the estimator instead of the mean square error, i.e., by setting $\delta_n = 1$ in (16) and (20). For example, for the AWGN channel with precoding and a 4-QAM in Subsection V-B, the crosspoint in (20) yields $\gamma = w_1^2 - w_2^2 < 1$, and we have a better estimation of the diagonal entries for the whole useful range of SNR.

The variance of the noise can be estimated from the diagonal of the covariance matrix. With a proper estimation of the channel, we may compute the product $\sigma_s^2\mathbf{W}\mathbf{W}^{\mathrm{H}} \odot \tilde{\mathbf{h}}\tilde{\mathbf{h}}^{\mathrm{H}}$ and estimate the noise variance by subtracting it from $\hat{\mathbf{R}}_{\mathbf{y}}$ for any diagonal entry, e.g., the rth,

$$\sigma_n^2 \approx \left(\hat{\mathbf{R}}_{\mathbf{y}} - (\sigma_s^2\mathbf{W}\mathbf{W}^{\mathrm{H}}) \odot \hat{\tilde{\mathbf{h}}}\hat{\tilde{\mathbf{h}}}^{\mathrm{H}}\right)_{r,r}. \tag{50}$$

As the entries of the estimated covariance matrix show errors, this estimation could be improved by the whole set of diagonal entries, e.g., averaging them. However, this yields poor results. Some diagonal entries are badly estimated and distort the computation of the noise. A poorly estimated entry, e.g. the r-th, can be detected by checking its corresponding square error, in (16). The square error of the estimation is highly dependent on $\mathbf{C}_{r,r}$. Therefore, we discard the estimations (50) with the highest values of $\mathbf{C}_{r,r}$. The algorithm, also called *diagonal noise removal algorithm* (DNA), is explained as follows.

First, estimate the channel using, mainly or exclusively, the off-diagonal entries of the estimation of the covariance matrix of the outputs. The MSE approach in Section III-B, or the Gao and Nallathanan algorithm [5] is a suitable method. By avoiding the diagonal entries, we get a noise-free estimation of the channel. Second, build a noise-free estimation of the covariance matrix as follows

$$\sigma_s^2\mathbf{C} = \sigma_s^2\mathbf{W}\mathbf{W}^{\mathrm{H}} \odot \hat{\tilde{\mathbf{h}}}\hat{\tilde{\mathbf{h}}}^{\mathrm{H}}. \tag{51}$$

Third, estimate the noise from every diagonal entry as

$$\mathbf{\Xi} = \mathrm{diag}\left(\hat{\mathbf{R}}_{\mathbf{y}} - \sigma_s^2\mathbf{C}\right), \tag{52}$$

where $\mathbf{\Xi}$ is a vector whose rth entry is the estimation of the noise using the rth diagonal entry of $\hat{\mathbf{R}}_{\mathbf{y}}$. Next, find the R entries $\mathcal{Q} = \{q_1, q_2, ..., q_r, ..., q_R\}$ of $\mathbf{\Xi}$ with the lowest corresponding values of $\mathbf{C}_{q_r,q_r}$. Fifth, estimate the variance of the noise as

$$\hat{\sigma_n^2} = 1/R \sum_{q_r \in \mathcal{Q}} \mathbf{\Xi}_{q_r}. \tag{53}$$

Finally, obtain the estimate of the channel as:

$$|\hat{\tilde{\mathbf{h}}}_{\mathrm{DNA}\,i}| = \sqrt{\left(\mathbf{B} - \frac{\hat{\sigma_n^2}}{\sigma_s^2}\mathbf{I} \oslash \mathbf{P}\right)_{i,i}}. \tag{54}$$

The angle may be estimated with any approach, e.g. the method used in the first step to estimate $\mathbf{C}$, the MSE,

$$\angle\hat{\tilde{\mathbf{h}}}_{\mathrm{DNA}} = \angle\Big(\sqrt{m}\mathbf{F}_{:,1:l+1}\hat{\mathbf{h}}_{\mathrm{MSE}}\Big). \tag{55}$$

The estimation of the CIR yields $\hat{\mathbf{h}}_{\mathrm{DNA}} = (\mathbf{F}_{:,1:l+1})^{\mathrm{H}}\hat{\tilde{\mathbf{h}}}_{\mathrm{DNA}}/\sqrt{m}$.

VIII. Computational complexity

Once we have computed (9), the order of the computational complexity of the methods, in ascending order, is calculated as follows. For the one-column method and the SDA just divide, entry by entry, one $m \times 1$ vector by another. Also, a de-noising stage (IFFT and FFT) of order $\mathrm{O}(m\log(m))$ is run, which yields the computational complexity of these methods. The MSE approach has a complexity $\mathrm{O}(m(l+1)\log m)$ given by the projection stage, $\mathbf{V} = (\mathbf{F}_{:,1:l+1})^{\mathrm{H}}\mathbf{B}\mathbf{F}_{:,1:l+1}$. Note that the updated rule, of $\mathrm{O}((l+1)^2)$, is performed a few times. The DA has the same complexity as the MSE, which is the method used in this paper to estimate the phases. The power method by Seo and Chung in [8] requires the multiplication of matrix $\mathbf{B}$ by a vector several times; hence, it is of order $\mathrm{O}(m^2)$. The DNA first estimates the channel with some method, in this paper we use the MSE. Then, any of the operations to estimate the noise variance yields at most $\mathrm{O}(m^2)$. Hence, the DNA has $\mathrm{O}(m^2)$. Finally, the method by Gao and Nallanathan gives $\mathrm{O}(m^2(l+1))$, as described in [5]. Table I shows a comparison of the computational complexity for all these approaches.

TABLE I
Computational complexity.

Method	Complexity order
One-column	$\mathrm{O}(m\log(m))$
SDA	$\mathrm{O}(m\log(m))$
DA	$\mathrm{O}(m(l+1)\log m)$
DNA	$\mathrm{O}(m^2)$
Gao and Nallanathan	$\mathrm{O}(m^2(l+1))$

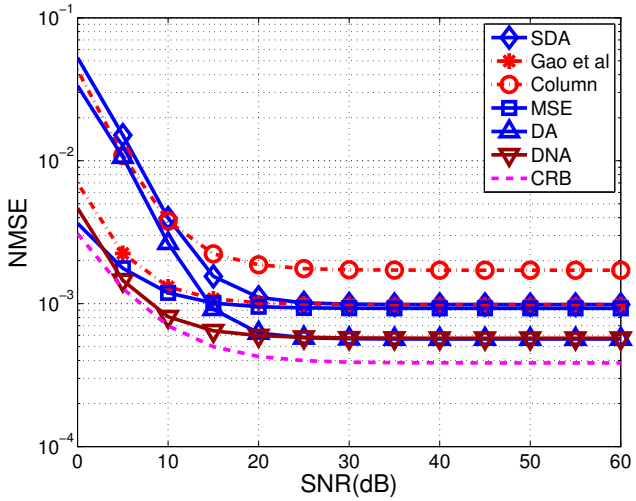

Fig. 6. Averaged NMSE for $m = 64$ subcarriers, a 4-QAM modulation, $N = 100$, and Rayleigh channel with $l+1 = 9$ taps. The channel is estimated using the SDA, the Gao and Nallanathan, the one-column, the MSE, the DA, and the DNA methods.

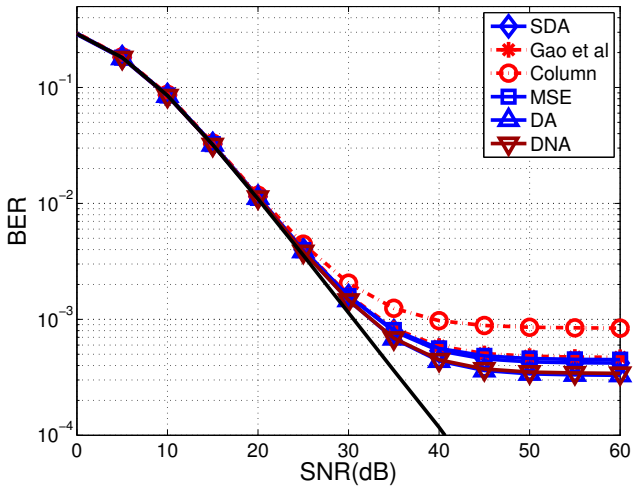

Fig. 7. Averaged BER for a ZF equalizer for $m = 64$ subcarriers, a 4-QAM, $N = 100$, and Rayleigh channel with $l + 1 = 9$ taps. The channel is estimated using the SDA, the Gao and Nallanathan, the one-column, the MSE, the DA, and the DNA methods.

IX. Experimental results

To compare the performance of the above-mentioned methods with other approaches, we use the normalized mean square error (NMSE). As the channel can be estimated up to a constant [16], the scalar ambiguity is solved by computing

$$\alpha_{min} = \min_{\alpha} ||\tilde{\mathbf{h}} - \alpha\hat{\tilde{\mathbf{h}}}||^2. \tag{56}$$

Then, the NMSE is defined as

$$NMSE(\hat{\tilde{\mathbf{h}}}) = \frac{1}{t}\sum_{i=1}^{t}\frac{||\tilde{\mathbf{h}} - \alpha\hat{\tilde{\mathbf{h}}}(i)||^2}{||\tilde{\mathbf{h}}||^2}, \tag{57}$$

where $\hat{\tilde{\mathbf{h}}}(i)$ and $\tilde{\mathbf{h}}$ are the estimations at run i and the true DFT of the channel impulse response, respectively, and t is the number of runs or estimations.

We compute a lower bound on the variance of estimators of the channel taps. As, we are analyzing blind algorithms, we resort to the stochastic Cramér−Rao bound (CRB) [17], [5]:

$$\sum_i \mathbf{CRB}_{\mathbf{h}\,i} \leq NMSE(\hat{\tilde{\mathbf{h}}}), \quad \forall\hat{\tilde{\mathbf{h}}}. \tag{58}$$

We examine the performance of the proposed methods compared to other previous proposals. The scenario is similar to that used in [5] and [8]. The block length is set to $m = 64$ according to the IEEE 802.11a standard. We simulate 5-tap ($l + 1 = 5$) and 9-tap channels, a similar and a larger value compared to those used in [5] and [3]. The precoding matrix is the one proposed in [5] with $p = 0.54$. All results are averaged over 1000 Monte Carlo runs. In every run the channel, noise, and signals are generated randomly. The SNR in the experiments is the ratio $\gamma = \sigma_s^2/\sigma_n^2$. Because the channel is normalized to have a unit norm, this value approximates $E_s/N_0 = \sigma_y^2/\sigma_n^2$, where only the precoding matrix introduces some deviation. The MSE approach in (11) used in these experiments is initialized with the one-column method in [3]

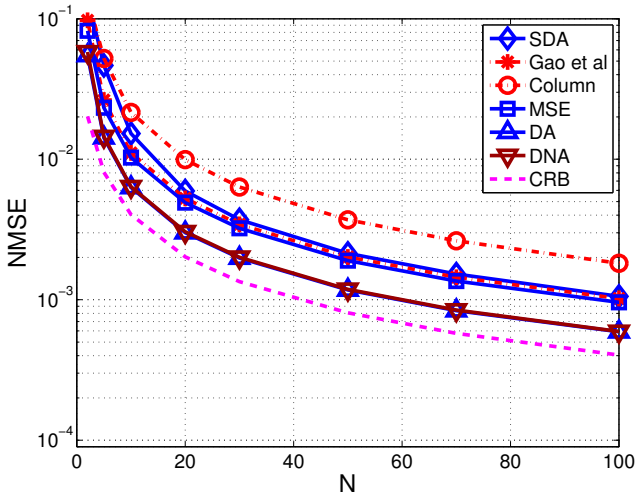

Fig. 8. Averaged NMSE for $m = 64$ subcarriers, a 4-QAM, SNR $= 25$dB, and Rayleigh channel with $l + 1 = 9$ taps. The channel is estimated using the SDA, the Gao and Nallanathan, the one-column, the MSE, the DA, and the DNA methods.

and is run for 20 iterations with $\mu = 0.01$ in (12). Parameter R of the DNA is set to 10% of the number of subcarriers.

Fig. 6 displays the NMSE along the signal-to-noise ratio (SNR) for the one-column by Petropulu *et al* in [3] ($\circ$), Gao and Nallanathan in [5] ($*$), the MSE in (11) ($\square$), the simple diagonal (SDA) ($\diamond$) in (48)-(49), the diagonal (DA) ($\triangle$) in (46)-(47), and the diagonal noise removal (DNA) ($\triangledown$) in (54)-(55) methods for $N = 100$ samples. The Cramér−Rao bound is also included (dashed) as reference. The signal constellation is 4-QAM. Fig. 7 shows the corresponding BER for the same scenario and methods, by a zero forcing (ZF) equalizer with the estimated channels. We also depict the BER for the ZF equalizer with perfect knowledge of the channel as reference (solid). For SNR ≥ 14 dB, the one-column method exhibits the worst NMSE, the DA and DNA the lowest estimation

error, and the other approaches have a similar performance. For $\mathrm{SNR} \leq 14$ dB, the estimation error of the diagonal elements in the covariance matrix is largely distorted by noise, as discussed in Sections V-B, V-C and illustrated in Figs. 4 and 5, and the one-column, DA and SDA exhibit a worse performance. The DNA has the best NMSE along the whole range of SNR. The 14-dB threshold depends on the precoding matrix and the number of samples. For a number of samples in the range $[10, 1000]$ we have a maximum deviation of about 5 dB around the previous 14-dB threshold. It is important to remark that, as illustrated in Fig. 7, the BER for $\mathrm{SNR} \leq 20$ dB is the same for all approaches, close to the BER estimated with the true channel. Therefore, in digital communications scenarios, as the one simulated here, good results in the estimation of the channel at low SNR are not useful and simple and accurate approaches such as the SDA are good estimators.

Fig. 8 depicts the NMSE and the CRB along the number of samples or blocks, N, used in the estimation of the covariance matrix for an $\mathrm{SNR} = 25$ dB. The DA and DNA provide good estimations for a low number of samples and a curve near to that of the CRB. The other methods, including the SDA, exhibit a similar performance for large enough numbers of samples, except for the one-column approach. For a low number of samples, the SDA is not as accurate as the MSE. The information of all the off-diagonal elements is relevant to better estimate the phase of channel.

It may be concluded that if the computational complexity is to be reduced and the coherence time of the channel is not too short, the SDA is a good option, as it provides the same BER as the Gao and Nallanathan, and MSE methods with the same low complexity as the one-column method. If the best estimation is needed, or the coherence time is short, the DA or DNA may to be used.

The experimental results in the previous example are similar to the ones obtained for other constellations; only the BER changes according to the characteristics of the modulation. Fig. 9 also shows the results for the NMSE for a 64-QAM and $N = 100$ samples. Fig. 4 proves that the crosspoint between the curves for the error in the diagonal ($\square$) and the off-diagonal entries ($\circ$) in a 64-QAM (dash-dotted) is similar to the crosspoint of the 4-QAM (solid) curves. The crosspoints of the DA and the MSE approach in Figs. 6 and 9 are also similar. The DA and SDA present similar or better NMSE than the other methods for $\mathrm{SNR} \geq 15$ dB. The DNA exhibits a good performance, near to the Cramér−Rao bound, for all SNR. Besides, as in Rayleigh channels the SNR needed for a BER higher than 10^{-2} for 64-QAM is larger than this 15-dB value (see Fig. 10) the DA is a good choice. We may resort to the SDA when computationally limited.

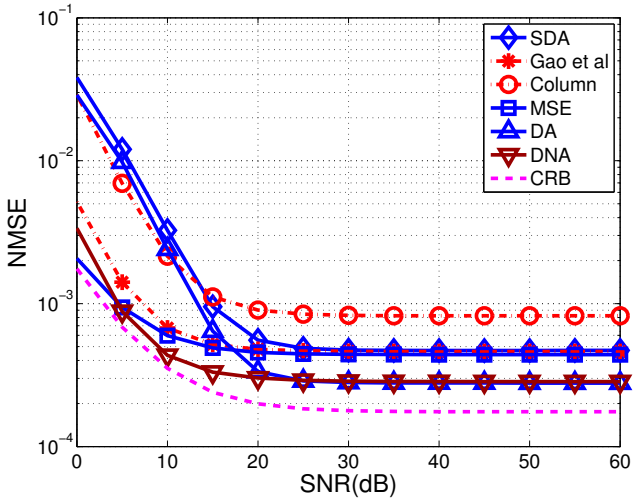

Fig. 9. Averaged NMSE for $m = 64$ subcarriers, a 64-QAM modulation, $N = 100$, and Rayleigh channel with $l+1 = 5$ taps. The channel is estimated using the SDA, the Gao and Nallanathan, the one-column, the MSE, the DA and the DNA methods.

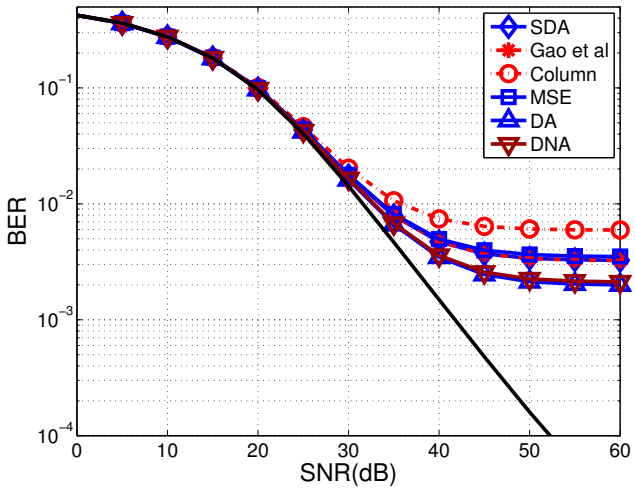

Fig. 10. Averaged BER for a ZF equalizer for $m = 64$ subcarriers, a 64-QAM modulation, $N = 100$, and Rayleigh channel with $l+1 = 5$ taps. The channel is estimated using the SDA, the Gao and Nallanathan, the one-column, the MSE, the DA and the DNA methods.

X. Conclusions

This paper revisits the estimation of the CIR in OFDM transmissions using non-redundant precoding and second-order statistics. We prove that for large and medium SNR the estimation of the diagonal entries of the covariance matrix of the outputs exhibits lower errors than that of the estimation of the off-diagonal ones. This result holds for typical SNR in fading channels. Accordingly, and contrary to previous approaches, we propose a method that uses the diagonal entries of the covariance matrix of the outputs to estimate the absolute value of the DFT of the CIR. If the channel is not fast time-variant, and we have adequate number of output samples, the phases of the DFT of the CIR can be estimated by resorting to the one-column method, thereby quite reducing the computational complexity. We also analyze the mean square error of the estimation matrix to derive a novel approach with noise suppression capabilities. In Rayleigh channels, these novel proposals exhibit a good performance, close to the CRB for typical SNR values in digital communications, and are an improvement to previous methods.

XI. Appendix A

The variance or the mean square error of the estimator for an entry (r,q) of the covariance matrix of the outputs is defined as

$$\Sigma_{r,q} = E\left\{\left|\frac{1}{N}\sum_{k=0}^{N-1}\mathbf{y}_r(k)\mathbf{y}_q^*(k) - \sigma_s^2\mathbf{C}_{r,q} - \delta_{rq}\delta_n\sigma_n^2\right|^2\right\}. \tag{59}$$

Expanding the square error, it follows that

$$\begin{aligned}\Sigma_{r,q} &= \frac{1}{N^2}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1} E\left\{\mathbf{y}_r(k)\mathbf{y}_q^*(k)\mathbf{y}_r^*(l)\mathbf{y}_q(l)\right\} \\ &- \frac{1}{N}(\sigma_s^2\mathbf{C}_{r,q} + \delta_{rq}\delta_n\sigma_n^2)\sum_{l=0}^{N-1} E\left\{\mathbf{y}_r^*(l)\mathbf{y}_q(l)\right\} \\ &- \frac{1}{N}(\sigma_s^2\mathbf{C}_{r,q}^* + \delta_{rq}\delta_n\sigma_n^2)\sum_{k=0}^{N-1} E\left\{\mathbf{y}_r(k)\mathbf{y}_q^*(k)\right\} + \sigma_s^4\left|\mathbf{C}_{r,q}\right|^2 \\ &+ \delta_{rq}\delta_n\sigma_s^2\sigma_n^2(\mathbf{C}_{r,q} + \mathbf{C}_{r,q}^*) + \delta_{rq}^2\delta_n^2\sigma_n^4. \end{aligned}\tag{60}$$

In addition,

$$\begin{aligned}&E\left\{\mathbf{y}_r(k)\mathbf{y}_q^*(k)\mathbf{y}_r^*(l)\mathbf{y}_q(l)\right\} = E\left\{\mathbf{y}_r(k)\mathbf{y}_q^*(k)\right\}E\left\{\mathbf{y}_r^*(l)\mathbf{y}_q(l)\right\} \\ &= \mathbf{R}_{\mathbf{y}\,r,q}\mathbf{R}_{\mathbf{y}\,r,q}^* = \left|\mathbf{R}_{\mathbf{y}\,r,q}\right|^2, \quad k \neq l \end{aligned}\tag{61}$$

and

$$E\left\{\mathbf{y}_r(k)\mathbf{y}_q^*(k)\mathbf{y}_r^*(l)\mathbf{y}_q(l)\right\} = E\left\{\left|\mathbf{y}_r(k)\right|^2\left|\mathbf{y}_q(k)\right|^2\right\}, \quad k = l. \tag{62}$$

Because the samples are i.i.d, $\mathbf{R}_{\mathbf{y}\,r,q} = \sigma_s^2\mathbf{C}_{r,q} + \delta_{rq}\sigma_n^2$ and $\delta^2 = \delta$, expression (59) can be simplified to

$$\begin{aligned}\Sigma_{r,q} &= \frac{1}{N}E\left\{\left|\mathbf{y}_r\right|^2\left|\mathbf{y}_q\right|^2\right\} - \frac{1}{N}(\sigma_s^2\mathbf{C}_{r,q} + \delta_{rq}\sigma_n^2) \\ &\cdot(\sigma_s^2\mathbf{C}_{r,q}^* + \delta_{rq}\sigma_n^2) + \delta_{rq}(1-\delta_n)\sigma_n^4, \end{aligned}\tag{63}$$

where

$$\begin{aligned}E\left\{\left|\mathbf{y}_r\right|^2\left|\mathbf{y}_q\right|^2\right\} &= \sigma_s^4(\mathbf{C}_{r,r}\mathbf{C}_{q,q} + \left|\mathbf{C}_{r,q}\right|^2 - \Theta_{r,q}) \\ &+ \sigma_s^2\sigma_n^2(\mathbf{C}_{r,r} + \mathbf{C}_{q,q} + \delta_{rq}(\mathbf{C}_{r,q} + \mathbf{C}_{r,q}^*)) \end{aligned}\tag{64}$$

where we used the relation $\mu_n/\sigma_n^4 - 3 = 0$ between the fourth-order moment, μ_n, and the standard deviation, σ_n, of Gaussian noise. Introducing (64) into (63) yields (16).

XII. Appendix B

Equation (22) differs from (16) in the expectation of the CIR: $\mathrm{E}\{|\tilde{\mathbf{h}}_r|^2\}$ and $\mathrm{E}\{|\tilde{\mathbf{h}}_r|^2|\tilde{\mathbf{h}}_q|^2\}$. The first one can be written as

$$\mathrm{E}\{|\tilde{\mathbf{h}}_r|^2\} = \sum_{\alpha_1=0}^{l}\sum_{\alpha_2=0}^{l}\mathrm{E}\{h_{\alpha_1}h_{\alpha_2}^*\}e^{-j\frac{2\pi}{m}r(\alpha_1-\alpha_2)}, \tag{65}$$

where $\mathrm{E}\{h_{\alpha_1}h_{\alpha_2}^*\} = \mathrm{E}\{|h_{\alpha_1}|^2\} = \sigma_{h_{\alpha_1}}^2$ for $\alpha_1{=}\alpha_2$ and zero otherwise, provided the taps of the CIR are zero mean and independent. Hence,

$$\mathrm{E}\{|\tilde{\mathbf{h}}_r|^2\} = \sum_{\alpha=0}^{l}\sigma_{h_\alpha}^2 = \tilde{\mathbf{e}}_1, \tag{66}$$

where $\tilde{\mathbf{e}}$ is defined in (24). On the other hand,

$$\begin{aligned}&\mathrm{E}\{|\tilde{\mathbf{h}}_r|^2|\tilde{\mathbf{h}}_q|^2\} = \\ &\sum_{\alpha_1=0}^{l}\sum_{\alpha_2=0}^{l}\sum_{\alpha_3=0}^{l}\sum_{\alpha_4=0}^{l}\mathrm{E}\{h_{\alpha_1}h_{\alpha_2}^*h_{\alpha_3}h_{\alpha_4}^*\}e^{-j\frac{2\pi}{m}[r(\alpha_1-\alpha_2)+q(\alpha_3-\alpha_4)]} \\ &= -\sum_{\alpha=0}^{l}\left(2\sigma_{h_\alpha}^4 - \mu_{h_\alpha}\right) + \left[\sum_{\alpha=0}^{l}\sigma_{h_\alpha}^2\right]^2 + \left|\sum_{\alpha=0}^{l}\sigma_{h_\alpha}^2 e^{-j\frac{2\pi}{m}(r-q)}\right|^2 \end{aligned}\tag{67}$$

Both expressions, (67) and (66), written as a function of $\tilde{\mathbf{e}}$ and introduced into (22), yield (23).

References

[1] Chiu Shun Wong, Drangan Obradovic, and Nilesh Madhu, "Independent component analysis (ICA) for blind equalization of frequency selective channels," in *IEEE 13th Workshop on Neural Networks for Signal Processing,*, 17-19 Sept. 2003, pp. 419 – 428.

[2] R. Boloix-Tortosa and J. J. Murillo-Fuentes, "Blind source separation in the adaptive reduction of inter-channel interference for OFDM," *Lecture Notes on Computer Sciences 3195: Independent Component Analysis and Blind Signal Separation.*, pp. 1142–1149, Granada, Sept. 2004.

[3] A. Petropulu, Ruifeng Zhang, and R. Lin, "Blind OFDM channel estimation through simple linear precoding," *IEEE Trans. on Wireless Communications*, vol. 3, no. 2, pp. 647–655, March 2004.

[4] R. Lin and A.P. Petropulu, "Linear precoding assisted blind channel estimation for OFDM systems," *IEEE Trans. on Vehicular Technology*, vol. 54, no. 3, pp. 983 – 995, May 2005.

[5] F. Gao and A. Nallanathan, "Blind channel estimation for OFDM systems via a generalized precoding," *IEEE Trans. on Vehicular Technology*, vol. 56, no. 3, pp. 1155–1164, May 2007.

[6] F. Gao and A. Nallanathan, "Blind channel estimation for MIMO OFDM systems via nonredundant linear precoding," *IEEE Trans. on Signal Processing*, vol. 55, no. 2, pp. 784 –789, Jan 2007.

[7] Ruifeng Zhang, "Blind channel estimation for precoded OFDM system," in *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, March 2005, vol. 3, pp. 469–472.

[8] Bangwon Seo and Hyun Gyu Chung, "Improved precoding-assisted blind channel estimation scheme for OFDM systems," in *IEEE Vehicular Technology Conference*, Oct 2007, pp. 1260–1264.

[9] Hung Nguyen-Le and Tho Le-Ngoc, "Pilot-aided joint CFO and doubly-selective channel estimation for OFDM transmissions," *IEEE Trans. on Broadcasting*, vol. 56, no. 4, pp. 514 –522, Dec. 2010.

[10] M. Tran, A. Nix, and A. Doufexi, "Mobile WiMAX: Impact of channel estimation error on the performance of limited feedback linear precoding," in *IEEE 21st International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, Sep. 2010, pp. 1066 – 1070.

[11] Zhongjun Wang, Yan Xin, G. Mathew, and Xiaodong Wang, "Efficient phase-error suppression for multiband OFDM-based UWB systems," *IEEE Trans. on Vehicular Technology*, vol. 59, no. 2, pp. 766 –778, Feb. 2010.

[12] G.W. Stewart, *Matrix Algorithms. Volume II: Eigensystems*, SIAM, 2001.

[13] J.-F. Yang and M. Kaveh, "Adaptive eigensubspace algorithms for direction or frequency estimation and tracking," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 36, no. 2, pp. 241 –251, Feb. 1988.

[14] Fan Xu and Jr. Willson, A.N., "Fixed-point analysis and realization of an adaptive eigenvector algorithm for use in sensor networks," in *Conference Record of the Thirty-Third Asilomar Conference on Signals, Systems, and Computers*, 1999, vol. 1, pp. 668 –672.

[15] Michelle Morelli and Umberto Mengali, "A comparison of pilot-aided channel estimation methods for OFDM systems," *IEEE Trans. on Signal Processing*, vol. 49, no. 12, pp. 3065–3073, 1999.

[16] E. De Carvalho and D.T.M. Slock, "Cramer-Rao bounds for semi-blind, blind and training sequence based channel estimation," in *First IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications*, Apr. 1997, pp. 129 –132.

[17] P. Stoica, E.G. Larsson, and A.B. Gershman, "The stochastic CRB for array processing: a textbook derivation," *IEEE Signal Processing Letters*, vol. 8, no. 5, pp. 148 –150, May. 2001.

Francisco J. Simois (S'99, M'05) was born in Sevilla, Spain, in 1974. He received his B.Sc. in Telecommunications Engineering from the Universidad de Sevilla, Spain, in 1997 and his Ph.D. from the same university in 2005. Currently he is an associate lecturer at the University of Sevilla, teaching both undergraduate and postgraduate courses related to signal processing and communications. His research fields are digital signal processing and digital communications.

Juan José Murillo-Fuentes was born in Sevilla in 1973. He received his Telecommunications Engineering degree in 1996 from the Universidad de Sevilla, Spain, and his Ph.D. degree in Telecommunication Engineering in 2001 from the Universidad Carlos III de Madrid, Spain. He is currently an Associate Professor at the Department of Signal Theory and Communication, Universidad de Sevilla. His research interests lie in algorithm development for machine learning, and their application to digital communications. Juan J. Murillo-Fuentes has been teaching in Bs, Ms and Ph.D. courses for ten years now. He is a lecturer in radiation and radiocommunications at the University of Seville since 2001, teaching fundamentals in wireless communications and wireless systems. He is also a professor of the Vodafone Master and doctoral courses. Since July 2010 he is a Senior member of the IEEE.

Rafael Boloix-Tortosa graduated in Telecommunications Engineering from the Universidad de Sevilla, Spain, in 2000, and joined the Department of Electronic Engineering at the School of Engineering as a Research Assistant. He received the Ph.D. degree in 2005. He is currently an Assistant Professor with the Department of Signal Theory and Communications, Universidad de Sevilla. His research interests include statistical communication signal processing, blind source separation, and array processing, and their application to digital communications.

Luis Salamanca (IEEE Student Member) is a Ph.D. student in the Universidad de Sevilla. He was born in Badajoz, Spain, in 1985. He obtained the B.Sc. in Telecommunication Engineering and the Ms.C. in Electronics and Signal Processing from the Universidad de Sevilla in 2009 and 2010, respectively. He has been a lecturer in diverse topics such as signal processing, radio propagation and communication theory in the degree of Telecommunication Engineering. His main research interests are in the field of signal processing and communications, as well as Bayesian statistics and machine learning applied to digital communications.