Blind Channel Shortening for Space-Time-Frequency Block Coded MIMO-OFDM Systems

Donatella Darsena, *Member, IEEE*, Giacinto Gelli, Luigi Paura, *Member, IEEE*, Francesco Verde, *Member, IEEE*,

***Abstract*—This paper deals with multiple-input multiple-output (MIMO) broadband wireless communication systems, employing orthogonal frequency-division multiplexing (OFDM) with cyclic prefix (CP) in combination with space-time-frequency block coding (STFBC). In order to exploit the benefits of OFDM and STFBC in highly frequency-selective channels, without incurring in a significant increase of receiver complexity, a channel shortening prefilter is inserted at the receiver, aimed at reducing the length of the MIMO channel impulse response before CP removal. Two MIMO channel shorteners are proposed, both relying on linearly-constrained minimization of the mean-output-energy of the signal at the output of the channel shortener, where the linear constraints are optimally chosen so as to maximize the signal-to-noise ratio either at the output of the channel shortener or at the input of the STFBC maximum-likelihood decoder. Unlike other solutions proposed in the literature, these shortening designs are *blind* ones, i.e., a priori knowledge of the MIMO channel impulse response to be shortened is not required, and can be carried out in closed-form, i.e., iterative computation of the prefilter weights is not necessary. Numerical simulations show that, without a substantial increase in computational complexity, receivers equipped with the proposed blind channel shorteners pay only a negligible performance penalty with respect to non-blind channel shorteners, while being significantly more robust to finite sample-size effects.**

***Index Terms*—Blind channel shortening, minimum-mean-output-energy criterion, multiple-input multiple-output (MIMO) systems, orthogonal frequency-division multiplexing (OFDM), space-time-frequency block coding (STFBC).**

I. Introduction

In frequency-selective channels, the use of multiple-input multiple-output (MIMO) systems in combination with cyclic-prefixed *orthogonal frequency-division multiplexing (OFDM)* is an efficient approach to support reliable transmission at high data rates [1]. Recently, *space-time-frequency block coding (STFBC)* schemes have been proposed for MIMO-OFDM systems [2], [3], where coding is applied across multiple OFDM blocks to jointly exploit spatial, temporal, and frequency diversity. A proper design of the coding rule can assure a remarkable performance improvement over the simpler *space-time block coding (STBC)* [4] and *space-frequency block coding (SFBC)* techniques [5]–[8].

System modeling, code design, and performance analysis of MIMO-OFDM systems employing STFBC are typically developed by assuming that the length M_{cp} of the cyclic prefix (CP) is greater than the order L_h of the underlying MIMO finite-impulse response (FIR) channel, i.e., $M_{\text{cp}} \geq L_h$. In such a case, the interblock interference (IBI) between OFDM symbols can be eliminated by discarding the CP before performing an M-point Discrete Fourier Transform (DFT). However, fulfillment of condition $M_{\text{cp}} \geq L_h$ could be impractical for MIMO-OFDM systems operating over long multipath channels, since the OFDM throughput efficiency $M/(M+M_{\text{cp}})$ is significantly decreased and, at the same time, the complexity of STFBC maximum-likelihood (ML) decoding, which exponentially grows with L_h, might become very large. Such a problem can be mitigated by means of *channel shortening* (also referred to as *time-domain equalization*), a preprocessing technique that partially equalizes the underlying MIMO channel, so that the order L_{eff} of the shortened impulse response is not greater than M_{cp}, i.e., $L_{\text{eff}} \leq M_{\text{cp}} < L_h$.

In the literature, many *non-blind* channel-shortening algorithms [9]–[15] have been proposed for single-input single-output (SISO) OFDM systems, some of which have also been extended [16]–[18] to MIMO channels. These techniques require *a priori* knowledge of the impulse responses of the channels to be shortened, which can be obtained via the transmission of training sequences. The amount of training to be used increases with L_h, which, for highly dispersive channels, leads to considerable waste of resources. Furthermore, in the MIMO case, channel estimation is significantly more complicated [19] than in the SISO one. A viable alternative is to shorten the channel in *blind* mode, by deriving the parameters of the channel-shortening prefilter directly from the received data, without performing any explicit channel-estimation procedure. Blind FIR channel-shortening techniques for SISO-OFDM systems have been developed in [20]–[23], some of which have been generalized [24], [25] to MIMO-OFDM systems. These techniques are based on optimization of multimodal heuristic cost functions, which do not admit closed-form solutions in the presence of noise; hence, they must be solved by means of iterative procedures, whose local and/or

Manuscript received January 19, 2011; revised September 10, 2011; accepted November 29, 2011. The associate editor coordinating the review of this paper and approving it for publication was Prof. B. Sundar Rajan.

This work was supported in part by Italian National Project "DRIVE IN^2".

D. Darsena is with the Department for Technologies (DIT), Parthenope University, Napoli I-80143, Italy (e-mail: darsena@uniparthenope.it).

G. Gelli, L. Paura, and F. Verde are with the Department of Biomedical, Electronic and Telecommunication Engineering (DIBET), University Federico II, Napoli I-80125, Italy (e-mail: {gelli, paura, f.verde}@unina.it).

Digital Object Identifier 10.1109/TWC.2011.110126.

Notation	Description
M	Number of subcarriers
N_T	Number of transmit antennas
N_R	Number of receive antennas
K	Time dimension (in OFDM symbols) of a codeword
B	Number of information symbols borne in a codeword
M_{cp}	CP length
P	OFDM symbol length $(= M_{\text{cp}} + M)$
T	Sampling period
T_s	OFDM symbol period
Q	Oversampling factor
L_h	Channel order
N_V	Number of virtual receiving antennas $(= Q\,N_R)$
σ_v^2	Noise variance
L_e	Channel-shortener order
L_g	Order of the channel cascade with the channel-shortener $(= L_h + L_e)$
Δ_s	Channel-shortening delay
L_{eff}	Effective order of the combined channel-prefilter impulse response
Δ_e	Signal preserving delay

TABLE I
MAIN MODEL PARAMETERS.

global convergence is not always ensured. Recently, blind channel-shortening algorithms for CP-based OFDM systems operating over single-input multiple-output (SIMO) channels (which arise either by employing multiple receiving antennas or by oversampling the received signal) have been proposed, which rely on the zero-forcing (ZF) [26] or the constrained minimum-mean-output-energy (MMOE) criterion [27] and, thereby, admit closed-form FIR solutions. In particular, in the presence of noise, the MMOE channel shortener performs significantly better than the ZF one, exhibiting also a lower computational complexity.

In this paper, we propose to extend the MMOE channel-shortening approach of [27] to a MIMO system employing STFBC. The case of a MIMO-OFDM scheme with spatial multiplexing has been considered in [28], by assuming that independent symbols are transmitted over different antennas to increase the data rate. To extend and complete the results in [28], the conditions assuring *perfect* FIR channel shortening are derived herein, which enlighten interesting relationships between the main parameters of the MIMO-OFDM system. Moreover, it is shown that, in the presence of noise, the linear constraints incorporated in the MMOE solution can be blindly optimized by means of the maximum signal-to-noise ratio (SNR) criterion, applied either at the output of the channel shortener or at the input of the STFBC ML decoder.

The paper is organized as follows. Section II describes the MIMO-OFDM system model with channel shortening. In Section III, the constrained MMOE criterion is introduced and its blind channel-shortening capabilities are analyzed in the noiseless case, whereas Section IV discusses the two different maximum-SNR blind constraint designs in the presence of noise. Monte Carlo computer simulation results, in terms of bit-error-rate (BER), are presented in Section V, followed by concluding remarks in Section VI.

A. Notations

The fields of complex, real, and integer numbers are denoted with $\mathbb{C}$, $\mathbb{R}$, and $\mathbb{Z}$, respectively; matrices [vectors] are denoted with upper [lower] case boldface letters (e.g., $\mathbf{A}$ or $\mathbf{a}$); the field of $m\times n$ complex [real] matrices is denoted as $\mathbb{C}^{m\times n}$ [$\mathbb{R}^{m\times n}$], with $\mathbb{C}^m$ [$\mathbb{R}^m$] used as a shorthand for $\mathbb{C}^{m\times 1}$ [$\mathbb{R}^{m\times 1}$]; the superscripts $*$, T, H, -1, and $\dagger$ denote the conjugate, the transpose, the conjugate transpose, the inverse, and the Moore-Penrose generalized inverse [30] of a matrix, respectively; $\{\mathbf{A}\}_{ij}$ indicates the $(i+1,j+1)$th element of $\mathbf{A}\in\mathbb{C}^{m\times n}$, with $i\in\{0,1,\ldots,m-1\}$ and $j\in\{0,1,\ldots,n-1\}$; $\mathbf{0}_m\in\mathbb{R}^m$, $\mathbf{O}_{m\times n}\in\mathbb{R}^{m\times n}$, and $\mathbf{I}_m\in\mathbb{R}^{m\times m}$ denote the null vector, the null matrix, and the identity matrix, respectively; trace$(\mathbf{A})$ denotes the trace of $\mathbf{A}\in\mathbb{C}^{n\times n}$; rank$(\mathbf{A})$, $\mathcal{R}(\mathbf{A})$, and $\mathcal{R}^{\perp}(\mathbf{A})$ denote the rank, the range (column space), and the orthogonal complement of the column space of $\mathbf{A}\in\mathbb{C}^{m\times n}$ in $\mathbb{C}^m$; $\|\mathbf{A}\|^2 \triangleq \text{trace}(\mathbf{A}\,\mathbf{A}^H)$ is the squared (Frobenius) norm of $\mathbf{A}\in\mathbb{C}^{m\times n}$; finally, $j\triangleq\sqrt{-1}$ is the imaginary unit and the operator $\text{E}[\,\cdot\,]$ denotes ensemble averaging.

II. THE MIMO-OFDM TRANSCEIVER MODEL WITH CHANNEL SHORTENING

Let us consider the baseband equivalent model of a MIMO-OFDM system, whose main parameters are summarized in Table I. The system employs channel-independent STFBC of rate $B/(KM)$, wherein the ith block of B information symbols is encoded over N_T transmit antennas, K consecutive OFDM symbols, and M subcarriers into the matrix codeword $\mathbf{C}(i)\in\mathbb{C}^{(KM)\times N_T}$, whose $(k+1,\alpha)$th block $\mathbf{c}_\alpha^{(k)}(i)\in\mathbb{C}^M$ is the code vector transmitted by the αth antenna in the $(iK+k)$th OFDM symbol interval, with $\alpha\in\{1,2,\ldots,N_T\}$ and $k\in\{0,1,\ldots,K-1\}$. Such a framework is general and subsumes different transmit diversity rules: for instance, the full-rate Alamouti code [29] corresponds to setting $B=2\,M$, $N_T=2$, $K=2$, $\mathbf{c}_2^{(1)}(i)=\mathbf{c}_1^{(0)}(i)^*$, and $\mathbf{c}_1^{(1)}(i)=-\mathbf{c}_2^{(0)}(i)^*$; moreover, simple spatial multiplexing, wherein $M\,N_T$ data

symbols are directly mapped into different antennas and subcarriers, can be obtained as a limiting case for $K = 1$.

Each code vector $\mathbf{c}_\alpha^{(k)}(i)$ is subject to an M-point Inverse Discrete Fourier Transform (IDFT), followed by the insertion of a CP of length M_{cp}, thus obtaining the vector $\mathbf{u}_\alpha^{(k)}(i) \triangleq [u_\alpha^{(k,0)}(i), u_\alpha^{(k,1)}(i), \ldots, u_\alpha^{(k,P-1)}(i)]^T \in \mathbb{C}^P$, with $P \triangleq M + M_{\text{cp}}$. The sequence $u_\alpha^{(k,p)}(i)$ feeds a digital-to-analog (D/A) converter working at rate $1/T = P/T_s$, where T and T_s denote the sampling and the OFDM symbol period, respectively, followed by up-conversion and transmission by the αth antenna. The *composite* channel (encompassing the cascade of the D/A interpolator, the pulse shaping filter, the physical channel, and the A/D antialiasing filter) between the αth transmit antenna ($\alpha \in \{1, 2, \ldots, N_T\}$) and the βth receive antenna ($\beta \in \{1, 2, \ldots, N_R\}$) is modeled as a linear time-invariant system, whose impulse response $h_{\beta,\alpha}(\tau)$ spans a maximum of L_h sampling periods, that is $h_{\beta,\alpha}(\tau) \equiv 0$ for $\tau \notin [0, L_h T]$. Assuming perfect timing and frequency synchronization, the received signal at each antenna is fractionally sampled at rate Q/T, with $Q \geq 1$ denoting the *oversampling* factor. By stacking all the received samples in the vector $\mathbf{r}(n) \in \mathbb{C}^{N_V}$, where $N_V \triangleq Q\, N_R$ is interpreted as the number of *virtual* receiving antennas, one obtains the MIMO model

$$\mathbf{r}(n) = \sum_{\ell=0}^{L_h} \mathbf{H}(\ell)\, \mathbf{u}(n-\ell) + \mathbf{v}(n) \tag{1}$$

where $\mathbf{H}(\ell) \in \mathbb{C}^{N_V \times N_T}$ is the matrix representation of the oversampled MIMO channel, the entries $u_\alpha(n)$ of the vector $\mathbf{u}(\ell) \triangleq [u_1(\ell), u_2(\ell), \ldots, u_{N_T}(\ell)]^T \in \mathbb{C}^{N_T}$ are such that $u_\alpha[(iK+k)P+p] = u_\alpha^{(k,p)}(i)$, and $\mathbf{v}(n) \in \mathbb{C}^{N_V}$ is noise.

We assume that the CP length is insufficient to counteract the multipath-channel effects, i.e., $M_{\text{cp}} < L_h$, which is likely to happen in highly dispersive channels, and, thus, a MIMO channel shortener is employed to jointly shorten all the channel impulse responses. Let L_e be the FIR channel-shortener order and $\widetilde{\mathbf{y}}(n) = [\widetilde{y}_1(n), \widetilde{y}_2(n), \ldots, \widetilde{y}_{N_R}(n)]^T \in \mathbb{C}^{N_R}$, the input-output relationship of the MIMO channel shortener is $\widetilde{\mathbf{y}}(n) = \widetilde{\mathbf{F}}\, \widetilde{\mathbf{r}}(n)$, where $\widetilde{\mathbf{F}} \in \mathbb{C}^{N_R \times N_V(L_e+1)}$ is the channel-shortening matrix (to be designed according to a certain optimization criterion), and the augmented vector $\widetilde{\mathbf{r}}(n) \triangleq [\mathbf{r}^T(n), \mathbf{r}^T(n-1), \ldots, \mathbf{r}^T(n-L_e)]^T \in \mathbb{C}^{N_V(L_e+1)}$ can be expressed as

$$\widetilde{\mathbf{r}}(n) = \widetilde{\mathbf{H}}\, \widetilde{\mathbf{u}}(n) + \widetilde{\mathbf{v}}(n) \tag{2}$$

where the block Toeplitz channel matrix $\widetilde{\mathbf{H}}$ is shown at the top of the next page, with $L_g \triangleq L_e + L_h$, whereas $\widetilde{\mathbf{u}}(n) \triangleq [\mathbf{u}^T(n), \mathbf{u}^T(n-1), \ldots, \mathbf{u}^T(n-L_g)]^T \in \mathbb{C}^{N_T(L_g+1)}$ and $\widetilde{\mathbf{v}}(n) \triangleq [\mathbf{v}^T(n), \mathbf{v}^T(n-1), \ldots, \mathbf{v}^T(n-L_e)]^T \in \mathbb{C}^{N_V(L_e+1)}$. By (2), the channel-shortener output becomes

$$\widetilde{\mathbf{y}}(n) = \widetilde{\mathbf{G}}\, \widetilde{\mathbf{u}}(n) + \widetilde{\mathbf{d}}(n) \tag{4}$$

where the matrix $\widetilde{\mathbf{G}} \triangleq \widetilde{\mathbf{F}}\, \widetilde{\mathbf{H}} \in \mathbb{C}^{N_R \times N_T(L_g+1)}$ collects all the taps of the *combined* MIMO channel-prefilter response, and $\widetilde{\mathbf{d}}(n) = [\widetilde{d}_1(n), \widetilde{d}_2(n), \ldots, \widetilde{d}_{N_R}(n)]^T \triangleq \widetilde{\mathbf{F}}\, \widetilde{\mathbf{v}}(n) \in \mathbb{C}^{N_R}$ is the noise contribution at the channel-shortener output. *Perfect* [31] (or ideal) channel shortening amounts to design the MIMO channel shortener such that the shortened impulse response is FIR of order $L_{\text{eff}} \leq M_{\text{cp}} < L_h$, which assures perfect IBI cancellation through CP removal. In matrix terms, this goal is obtained by choosing $\widetilde{\mathbf{F}}$ such that $\widetilde{\mathbf{G}}$ has the *target* form

$$\widetilde{\mathbf{G}}_{\text{target}} \triangleq [\mathbf{O}_{N_R \times (N_T \Delta_s)}, \widetilde{\mathbf{G}}_{\text{win}}, \mathbf{O}_{N_R \times N_T(L_g - \Delta_s - L_{\text{eff}})}] \tag{5}$$

where $\Delta_s \in \{0, 1, \ldots, L_g - L_{\text{eff}}\}$ is a suitable *channel-shortening delay* and $\widetilde{\mathbf{G}}_{\text{win}} \in \mathbb{C}^{N_R \times N_T(L_{\text{eff}}+1)}$ contains the samples of the shortened channel impulse response.[1] By partitioning (3) as $\widetilde{\mathbf{H}} = [\widetilde{\mathbf{H}}_0, \widetilde{\mathbf{H}}_1, \ldots, \widetilde{\mathbf{H}}_{L_g}]$, with $\widetilde{\mathbf{H}}_d \in \mathbb{C}^{N_V(L_e+1) \times N_T}$ for $d \in \{0, 1, \ldots, L_g\}$, the channel-shortener output can be written as

$$\widetilde{\mathbf{y}}(n) = \sum_{d=0}^{L_g} \widetilde{\mathbf{F}}\, \widetilde{\mathbf{H}}_d\, \mathbf{u}(n-d) + \widetilde{\mathbf{d}}(n) \tag{6}$$

which shows that, to satisfy (5), matrix $\widetilde{\mathbf{F}}$ can be designed so as to preserve in (6) the $L_{\text{eff}} + 1$ *desired* signal contributions $\mathbf{u}(n-\Delta_s), \mathbf{u}(n-\Delta_s-1), \ldots, \mathbf{u}(n-\Delta_s-L_{\text{eff}})$ while, at the same time, canceling out all the remaining *undesired* ones. This task could be accomplished using different criteria, such as the ZF one, which imposes a set of linear equations on $\widetilde{\mathbf{F}}$, or the maximum shortening SNR [9]–[11] one. However, in the MIMO case, the latter amounts to solve a trace ratio maximization problem [32], which does not admit a closed-form solution; most important, both strategies are non-blind, since they would require knowledge of matrices $\widetilde{\mathbf{H}}_d$, i.e., of the MIMO channel to be shortened. Instead, we propose in Section III a blind design strategy based on the MMOE approach, which was originally introduced in [33] for code-division multiple-access (CDMA) multiuser detection, and is also known as the minimum variance distortionless response or Capon beamformer [34] in the array processing community.

III. MMOE MIMO Channel Shortening

Leveraging on the MMOE approach for SIMO channel shortening [27], we propose to design the matrix $\widetilde{\mathbf{F}}$ in (6) such that to minimize the cost function

$$\mathrm{E}[\|\widetilde{\mathbf{y}}(n)\|^2] = \mathrm{trace}(\widetilde{\mathbf{F}}\, \mathbf{R}_{\widetilde{\mathbf{r}}\widetilde{\mathbf{r}}}\, \widetilde{\mathbf{F}}^H) \tag{7}$$

where $\mathbf{R}_{\widetilde{\mathbf{r}}\widetilde{\mathbf{r}}} \triangleq \mathrm{E}[\widetilde{\mathbf{r}}(n)\, \widetilde{\mathbf{r}}^H(n)] \in \mathbb{C}^{N_V(L_e+1) \times N_V(L_e+1)}$ is the autocorrelation matrix of the received data, under suitable constraints aimed at preserving the desired signal contributions at the output of the channel shortener. Assuming that the zero-mean noise vector $\widetilde{\mathbf{v}}(n)$ is statistically independent of $\mathbf{u}(n)$ with $\mathbf{R}_{\widetilde{\mathbf{v}}\widetilde{\mathbf{v}}} \triangleq \mathrm{E}[\widetilde{\mathbf{v}}(n)\, \widetilde{\mathbf{v}}^H(n)] = \sigma_v^2 \mathbf{I}_{N_V}$, and accounting for (2), it results that $\mathbf{R}_{\widetilde{\mathbf{r}}\widetilde{\mathbf{r}}} = \widetilde{\mathbf{H}}\, \mathbf{R}_{\widetilde{\mathbf{u}}\widetilde{\mathbf{u}}}\, \widetilde{\mathbf{H}}^H + \sigma_v^2\, \mathbf{I}_{N_V(L_e+1)}$, with $\mathbf{R}_{\widetilde{\mathbf{u}}\widetilde{\mathbf{u}}} \triangleq \mathrm{E}[\widetilde{\mathbf{u}}(n)\, \widetilde{\mathbf{u}}^H(n)] \in \mathbb{C}^{N_T(L_g+1) \times N_T(L_g+1)}$ being the *nonsingular* autocorrelation matrix[2] of the transmitted signal. In the SIMO context [27], it has been proven that this blind constrained MMOE approach allows one to achieve perfect channel shortening in the limit of vanishingly small noise.

In our MIMO setting, the blind constraints are designed, similarly to [27], [28], by exploiting the block Toeplitz structure (3) of $\widetilde{\mathbf{H}}$, according to which each matrix $\widetilde{\mathbf{H}}_d$,

[1] We assume that the MIMO channel shortener introduces the same delay Δ_s for all the transmit antennas. Even though a better design could be developed by allowing different shortening delays, the computational complexity would be correspondingly higher.

[2] The matrix $\mathbf{R}_{\widetilde{\mathbf{u}}\widetilde{\mathbf{u}}}$ is time-invariant, provided that $L_g < M$ [21], [26], [27].

$$\widetilde{\mathbf{H}} \triangleq \begin{bmatrix} \mathbf{H}(0) & \mathbf{H}(1) & \dots & \mathbf{H}(L_h) & \mathbf{O}_{N_V\times N_T} & \dots & \mathbf{O}_{N_V\times N_T} \\ \mathbf{O}_{N_V\times N_T} & \mathbf{H}(0) & \mathbf{H}(1) & \dots & \mathbf{H}(L_h) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \mathbf{O}_{N_V\times N_T} \\ \mathbf{O}_{N_V\times N_T} & \dots & \mathbf{O}_{N_V\times N_T} & \mathbf{H}(0) & \mathbf{H}(1) & \dots & \mathbf{H}(L_h) \end{bmatrix} \in \mathbb{C}^{[N_V\,(L_e+1)]\times[N_T\,(L_g+1)]} \quad (3)$$

for $0 \le d \le \min(L_e, L_h)$, with $L_e > M_{\text{cp}}$, can be linearly parameterized as $\widetilde{\mathbf{H}}_d = \mathbf{\Theta}_d\,\mathbf{K}_d$, where $\mathbf{\Theta}_d = [\mathbf{I}_{N_V(d+1)}, \mathbf{O}^T_{N_V(L_e-d)\times N_V(d+1)}]^T \in \mathbb{R}^{N_V(L_e+1)\times N_V(d+1)}$ is a known full-column rank matrix satisfying $\mathbf{\Theta}_d^T\,\mathbf{\Theta}_d = \mathbf{I}_{N_V(d+1)}$, whereas $\mathbf{K}_d = [\mathbf{H}^T(d), \mathbf{H}^T(d-1), \dots, \mathbf{H}^T(0)]^T \in \mathbb{C}^{N_V(d+1)\times N_T}$ contains samples of the impulse responses of the channels to be shortened.[3] At this point, let us choose a particular *signal preserving delay* $\Delta_e \in \{0, 1, \dots, \min(L_e, L_h)\}$ and impose that the corresponding vector $\mathbf{u}(n-\Delta_e)$ is one of the desired signal contributions. Consequently, eq. (6) can be rewritten as

$$\widetilde{\mathbf{y}}(n) = \widetilde{\mathbf{F}}\,\mathbf{\Theta}_{\Delta_e}\,\mathbf{K}_{\Delta_e}\,\mathbf{u}(n-\Delta_e) + \sum_{\substack{d=0 \\ d\neq\Delta_e}}^{L_g} \widetilde{\mathbf{F}}\,\widetilde{\mathbf{H}}_d\,\mathbf{u}(n-d) + \widetilde{\mathbf{d}}(n)\,. \quad (8)$$

To preserve $\mathbf{u}(n-\Delta_e)$ at the channel shortener output, we impose the matrix linear constraint $\widetilde{\mathbf{F}}\,\mathbf{\Theta}_{\Delta_e} = \mathbf{\Gamma}$, where $\mathbf{\Gamma} \in \mathbb{C}^{N_R\times N_V(\Delta_e+1)}$ is a given matrix of constraint values, whose choice will be discussed in the next section; in this way, the signal vector $\mathbf{u}(n-\Delta_e)$ is passed to the output of the channel shortener with gain $\mathbf{\Gamma}\,\mathbf{K}_{\Delta_e}$. Therefore, the MMOE constrained optimization problem can be expressed as

$$\widetilde{\mathbf{F}}_{\text{mmoe}} = \arg\min_{\widetilde{\mathbf{F}}} \text{trace}\left(\widetilde{\mathbf{F}}\,\mathbf{R}_{\widetilde{\mathbf{r}}\widetilde{\mathbf{r}}}\,\widetilde{\mathbf{F}}^H\right) \;\text{ s.t. }\; \widetilde{\mathbf{F}}\,\mathbf{\Theta}_{\Delta_e} = \mathbf{\Gamma} \quad (9)$$

whose solution is given by (see [35])

$$\widetilde{\mathbf{F}}_{\text{mmoe}} \triangleq \mathbf{\Gamma}\,\mathbf{F}_{\text{mmoe}} \quad (10)$$

where $\mathbf{F}_{\text{mmoe}} \in \mathbb{C}^{N_V(\Delta_e+1)\times N_V(L_e+1)}$ is given by

$$\mathbf{F}_{\text{mmoe}} \triangleq \left(\mathbf{\Theta}^T_{\Delta_e}\,\mathbf{R}^{-1}_{\widetilde{\mathbf{r}}\widetilde{\mathbf{r}}}\,\mathbf{\Theta}_{\Delta_e}\right)^{-1}\,\mathbf{\Theta}^T_{\Delta_e}\,\mathbf{R}^{-1}_{\widetilde{\mathbf{r}}\widetilde{\mathbf{r}}}\,. \quad (11)$$

The matrix $\mathbf{F}_{\text{mmoe}}$ can also be rewritten as [27], [34]

$$\mathbf{F}_{\text{mmoe}} = \mathbf{\Theta}^T_{\Delta_e}\left[\mathbf{I}_{N_V(L_e+1)} - \mathbf{R}_{\widetilde{\mathbf{r}}\widetilde{\mathbf{r}}}\,\mathbf{\Pi}^T_{\Delta_e}\left(\mathbf{\Pi}_{\Delta_e}\,\mathbf{R}_{\widetilde{\mathbf{r}}\widetilde{\mathbf{r}}}\,\mathbf{\Pi}^T_{\Delta_e}\right)^{-1}\mathbf{\Pi}_{\Delta_e}\right] \quad (12)$$

with the rows of $\mathbf{\Pi}_{\Delta_e} \in \mathbb{R}^{N_V(L_e-\Delta_e)\times N_V(L_e+1)}$ forming an orthonormal basis for the null space of $\mathbf{\Theta}^T_{\Delta_e}$, i.e., $\mathbf{\Pi}_{\Delta_e}\,\mathbf{\Theta}_{\Delta_e} = \mathbf{O}_{N_V(L_e-\Delta_e)\times N_V(\Delta_e+1)}$, i.e., $\mathcal{R}(\mathbf{\Pi}^T_{\Delta_e}) \equiv \mathcal{R}^{\perp}(\mathbf{\Theta}_{\Delta_e})$.[4]

A key observation is that, besides $\mathbf{u}(n-\Delta_e)$, all the signal vectors $\mathbf{u}(n-d)$, for $d \in \{0, 1, \dots, \Delta_e-1\}$, are also preserved by the constraint in (9), since it can be theoretically proven that all the columns of the submatrices $\widetilde{\mathbf{H}}_0, \widetilde{\mathbf{H}}_1, \dots, \widetilde{\mathbf{H}}_{\Delta_e-1}$ belong to the column space of $\mathbf{\Theta}_{\Delta_e}$; hence, the corresponding signal contributions $\mathbf{u}(n), \mathbf{u}(n-1), \dots, \mathbf{u}(n-\Delta_e-1)$ are also passed to the output of the channel shortener with a fixed gain. This fact is formally stated by the following theorem, which enlightens the channel-shortening capabilities of the proposed MMOE approach:

Theorem 1: Given $\Delta_e \in \{0, 1, \dots, \min(L_e, L_h)\}$, let us decompose $\widetilde{\mathbf{H}} = [\widetilde{\mathbf{H}}_{\text{win}}, \widetilde{\mathbf{H}}_{\text{wall}}]$, with $\widetilde{\mathbf{H}}_{\text{win}} \triangleq [\widetilde{\mathbf{H}}_0, \widetilde{\mathbf{H}}_1, \dots, \widetilde{\mathbf{H}}_{\Delta_e}] \in \mathbb{C}^{N_V(L_e+1)\times N_T(\Delta_e+1)}$ and $\widetilde{\mathbf{H}}_{\text{wall}} \triangleq [\widetilde{\mathbf{H}}_{\Delta_e+1}, \widetilde{\mathbf{H}}_{\Delta_e+2}, \dots, \widetilde{\mathbf{H}}_{L_g}] \in \mathbb{C}^{N_V(L_e+1)\times N_T(L_g-\Delta_e)}$. Assume $N_V > N_T$ and the following conditions to hold:

(c1) $N_V(L_e-\Delta_e) \ge N_T(L_g-\Delta_e)$;

(c2) $\text{rank}(\mathbf{\Pi}\,\widetilde{\mathbf{H}}_{\text{wall}}) = N_T(L_g-\Delta_e)$.

Then, for $\sigma_v^2 \to 0$, the combined channel-prefilter matrix $\widetilde{\mathbf{G}}_{\text{mmoe}} \triangleq \widetilde{\mathbf{F}}_{\text{mmoe}}\widetilde{\mathbf{H}}$ assumes the target form (5), with $\Delta_s = 0$, $L_{\text{eff}} = \Delta_e$, and $\widetilde{\mathbf{G}}_{\text{win}} = \mathbf{\Gamma}\,\mathbf{\Theta}^T_{\Delta_e}\,\widetilde{\mathbf{H}}_{\text{win}}$, for any nonzero $\mathbf{\Gamma}$.

Proof: Substituting $\mathbf{R}_{\widetilde{\mathbf{r}}\widetilde{\mathbf{r}}} = \widetilde{\mathbf{H}}\,\mathbf{R}_{\widetilde{\mathbf{u}}\widetilde{\mathbf{u}}}\,\widetilde{\mathbf{H}}^H + \sigma_v^2\,\mathbf{I}_{N_V(L_e+1)}$ in (12), remembering that $\mathcal{R}(\mathbf{\Pi}^T_{\Delta_e}) \equiv \mathcal{R}^{\perp}(\mathbf{\Theta}_{\Delta_e})$ and $\mathbf{\Pi}_{\Delta_e}\,\mathbf{\Pi}^T_{\Delta_e} = \mathbf{I}_{N_V(L_e-\Delta_e)}$, and using the limit formula for the Moore-Penrose inverse [37], one obtains

$$\lim_{\sigma_v^2\to 0}\widetilde{\mathbf{G}}_{\text{mmoe}} = \mathbf{\Gamma}\,\mathbf{\Theta}^T_{\Delta_e}\,\widetilde{\mathbf{H}}\left[\mathbf{I}_{N_T(L_g+1)} - (\mathbf{\Pi}_{\Delta_e}\,\widetilde{\mathbf{H}})^{\dagger}\,(\mathbf{\Pi}_{\Delta_e}\,\widetilde{\mathbf{H}})\right] \quad (13)$$

which shows that the asymptotic expression of $\widetilde{\mathbf{G}}_{\text{mmoe}}$ does not depend on $\mathbf{R}_{\widetilde{\mathbf{u}}\widetilde{\mathbf{u}}}$, provided that such a matrix is nonsingular. It is seen by direct inspection that $\mathcal{R}^{\perp}(\mathbf{\Theta}_{\Delta_e-\delta}) \subseteq \mathcal{R}^{\perp}(\mathbf{\Theta}_{\Delta_e}) = \mathcal{R}(\mathbf{\Pi}^T_{\Delta_e})$, for $\delta \in \{0, 1, \dots, \Delta_e\}$, and, hence, it results that $\mathbf{\Pi}_{\Delta_e}\,\widetilde{\mathbf{H}}_{\Delta_e-\delta} = \mathbf{\Pi}_{\Delta_e}\,\mathbf{\Theta}_{\Delta_e-\delta}\,\mathbf{K}_{\Delta_e-\delta} = \mathbf{O}_{N_V(L_e-\Delta_e)\times N_T}$, which leads to $\mathbf{\Pi}_{\Delta_e}\,\widetilde{\mathbf{H}}_{\text{win}} = \mathbf{O}_{N_V(L_e-\Delta_e)\times N_T(\Delta_e+1)}$. Therefore, $\mathbf{\Pi}_{\Delta_e}\,\widetilde{\mathbf{H}} = [\mathbf{O}_{N_V(L_e-\Delta_e)\times N_T(\Delta_e+1)}, \mathbf{\Pi}_{\Delta_e}\,\widetilde{\mathbf{H}}_{\text{wall}}]$ and $(\mathbf{\Pi}_{\Delta_e}\,\widetilde{\mathbf{H}})^{\dagger} = [\mathbf{O}_{N_V(L_e-\Delta_e)\times N_T(\Delta_e+1)}, \{(\mathbf{\Pi}_{\Delta_e}\,\widetilde{\mathbf{H}}_{\text{wall}})^{\dagger}\}^T]^T$, which can be substituted in (13), thus obtaining

$$\lim_{\sigma_v^2\to 0}\widetilde{\mathbf{G}}_{\text{mmoe}} = \mathbf{\Gamma}\,\mathbf{\Theta}^T_{\Delta_e}\,[\widetilde{\mathbf{H}}_{\text{win}}, \widetilde{\mathbf{H}}_{\text{wall}}] \cdot \begin{bmatrix} \mathbf{I}_{N_T(\Delta_e+1)} & \mathbf{O}_{N_T(\Delta_e+1)\times N_T(L_g-\Delta_e)} \\ \mathbf{O}_{N_T(L_g-\Delta_e)\times N_T(\Delta_e+1)} & \mathbf{O}_{N_T(L_g-\Delta_e)\times N_T(L_g-\Delta_e)} \end{bmatrix} = \left[\mathbf{\Gamma}\,\mathbf{\Theta}^T_{\Delta_e}\,\widetilde{\mathbf{H}}_{\text{win}}, \mathbf{O}_{N_R\times N_T(L_g-\Delta_e)}\right] \quad (14)$$

where, since the matrix $\mathbf{\Pi}_{\Delta_e}\,\widetilde{\mathbf{H}}_{\text{wall}}$ is full-column rank by assumption, we have used the fact that $(\mathbf{\Pi}_{\Delta_e}\,\widetilde{\mathbf{H}}_{\text{wall}})^{\dagger}\,(\mathbf{\Pi}_{\Delta_e}\,\widetilde{\mathbf{H}}_{\text{wall}}) = \mathbf{I}_{N_T(L_g-\Delta_e)}$. The proof concludes by comparing (14) with (5) for $\Delta_s = 0$. ■

It is worth noticing that **(c2)** is not a restrictive assumption since, if **(c1)** is met, any matrix $\widetilde{\mathbf{H}}$ with nonzero statistically independent entries drawn from a continuous probability distribution will satisfy it with probability one. Some remarks regarding the implications of Theorem 1 are now in order.

Remark 1: A sufficient condition to assure perfect channel shortening, in the high SNR regime, is that the number of

[3] Contrary to the SIMO case [27], oversampling of the received signal is not indispensable in the MIMO context to obtain a composite channel matrix with a block Toeplitz structure and, thus, factorization $\widetilde{\mathbf{H}}_d = \mathbf{\Theta}_d\,\mathbf{K}_d$ holds even when $Q = 1$.

[4] Throughout the paper, we assume, without loss of generality, that $\mathbf{\Pi}_{\Delta_e}$ is semi-unitary, i.e., $\mathbf{\Pi}_{\Delta_e}\,\mathbf{\Pi}^T_{\Delta_e} = \mathbf{I}_{N_V(L_e-\Delta_e)}$.

virtual receive antennas $N_V = Q\,N_R$ is strictly greater than the number of transmit antennas N_T. This means that, for MIMO systems with $N_T = N_R$, oversampling at the receiver ($Q > 1$) is needed to shorten all the underlying channels. On the other hand, if $N_R > N_T$, perfect channel shortening may be achieved even without oversampling ($Q = 1$).

Remark 2: Remembering that $L_g = L_e + L_h$ and $N_V = Q\,N_R$, condition **(c1)** imposes the following upper bound on the order of the MIMO channel to be shortened:

$$L_h \leq (L_e - \Delta_e)\left(\frac{Q\,N_R}{N_T} - 1\right). \tag{15}$$

Remarkably, for a given pair (N_T, N_R), only an upper bound (rather than the exact knowledge) of L_h is required to select suitable values of Δ_e, Q, and L_e. In particular, for a fixed value of $L_e - \Delta_e$, inequality (15) shows that oversampling ($Q > 1$), although it is not strictly necessary when $N_R > N_T$, allows one to shorten longer channels.

Remark 3: As to the choice of Δ_e, it is apparent from (15) that, for a given value of Q, the upper bound on L_h linearly increases with $L_e - \Delta_e$. Therefore, one could conclude that the choice $\Delta_e = 0$ is preferable, since, in this case, the proposed MMOE channel shortener is asymptotically able to shorten longer channels. With this choice, one attains the minimum order $L_{\text{eff}} = 0$ of the shortened channel impulse response, which amounts to let the channel shortener completely suppress IBI. However, in noisy environments, the order of the shortened channel impulse response must be traded-off against noise enhancement at the equalizer output. Indeed, as shown in [27], for low-to-moderate values of the SNR, the best choice is $\Delta_e = M_{\text{cp}}$, and, hence, $L_{\text{eff}} = M_{\text{cp}}$, which amounts to shorten the channel up to the maximum length for which CP removal is still effective.

Remark 4: Our numerical simulations (see Section V) show that Theorem 1 accurately predicts the channel shortening capability of the MMOE channel shortener not only in the high-SNR regime, but also for moderate SNR values, for which the received MIMO-OFDM signal dominates the noise.

IV. Blind Optimization of the Constraint Matrix $\mathbf{\Gamma}$

In the previous section, we have proven that the optimal MMOE channel shortener (10) can be written as $\widetilde{\mathbf{F}}_{\text{mmoe}} = \mathbf{\Gamma}\,\mathbf{F}_{\text{mmoe}}$, i.e, it admits a two-stage implementation: the former stage, represented by $\mathbf{F}_{\text{mmoe}}$, performs actual shortening of the MIMO channel, whereas the latter one, represented by the constraint matrix $\mathbf{\Gamma}$, can be further optimized. In this section, we focus on blind (i.e., without requiring channel knowledge) design of $\mathbf{\Gamma}$, which is based on maximizing the energy of the useful signal. More specifically, we consider two criteria: the first one, similarly to [27] and [28], maximizes the SNR at the *output* of the channel shortener, whereas the second one performs SNR maximization at the *input* of the STFBC ML decoder, which in principle should assure a better performance. According to Theorem 1 we set $\Delta_s = 0$ and, moreover, accounting for Remark 3, we choose $\Delta_e = L_{\text{eff}} = M_{\text{cp}}$ in all the subsequent derivations.

Both criteria for determining $\mathbf{\Gamma}$ will be based on the assumption of perfect channel shortening, according to which it results that $\widetilde{\mathbf{G}}_{\text{mmoe}} = \widetilde{\mathbf{G}}_{\text{target}} = [\widetilde{\mathbf{G}}_{\text{win}}\,,\,\mathbf{O}_{N_R \times N_T(L_g - L_{\text{eff}})}]$, as given by (5) for $L_{\text{eff}} = M_{\text{cp}}$ and $\Delta_s = 0$, with $\widetilde{\mathbf{G}}_{\text{win}} = \widetilde{\mathbf{F}}_{\text{mmoe}}\,\widetilde{\mathbf{H}}_{\text{win}}$. Hence, by substituting in (4), we get

$$\widetilde{\mathbf{y}}(n) = \widetilde{\mathbf{F}}_{\text{mmoe}}\,\widetilde{\mathbf{r}}(n) = \widetilde{\mathbf{G}}_{\text{win}}\,\widetilde{\mathbf{u}}_{\text{win}}(n) + \widetilde{\mathbf{d}}(n) \tag{16}$$

where $\widetilde{\mathbf{u}}_{\text{win}}(n) \triangleq [\mathbf{u}^T(n), \mathbf{u}^T(n-1), \ldots, \mathbf{u}^T(n - L_{\text{eff}})]^T \in \mathbb{C}^{N_T(M_{\text{cp}}+1)}$. It can be observed that, in general, the noise $\widetilde{\mathbf{d}}(n) = \widetilde{\mathbf{F}}_{\text{mmoe}}\,\widetilde{\mathbf{v}}(n)$ at the output of the MMOE channel shortener is spatially colored, since its autocorrelation matrix is $\mathbf{R}_{\widetilde{\mathbf{d}}\widetilde{\mathbf{d}}} \triangleq \mathrm{E}[\widetilde{\mathbf{d}}(n)\,\widetilde{\mathbf{d}}^H(n)] = \sigma_v^2\,\widetilde{\mathbf{F}}_{\text{mmoe}}\,\widetilde{\mathbf{F}}_{\text{mmoe}}^H \in \mathbb{C}^{N_R \times N_R}$. To simplify the subsequent ML decoding process without incurring a significant performance loss, we impose on $\widetilde{\mathbf{F}}_{\text{mmoe}}$ the following design constraint: **(c3)** the matrix $\widetilde{\mathbf{F}}_{\text{mmoe}}$, which has more columns than rows by construction, is semi-unitary, i.e., $\widetilde{\mathbf{F}}_{\text{mmoe}}\,\widetilde{\mathbf{F}}_{\text{mmoe}}^H = \mathbf{I}_{N_R}$, which implies $\mathbf{\Gamma}\,\mathbf{F}_{\text{mmoe}}\,\mathbf{F}_{\text{mmoe}}^H\,\mathbf{\Gamma}^H = \mathbf{I}_{N_R}$. This condition ensures that $\mathbf{R}_{\widetilde{\mathbf{d}}\widetilde{\mathbf{d}}} = \sigma_v^2\,\mathbf{I}_{N_R}$ and, thus, accounting for the two-stage structure of $\widetilde{\mathbf{F}}_{\text{mmoe}}$, matrix $\mathbf{\Gamma}$ has to satisfy the constraint $\mathbf{\Gamma}\,\mathbf{F}_{\text{mmoe}}\,\mathbf{F}_{\text{mmoe}}^H\,\mathbf{\Gamma}^H = \mathbf{I}_{N_R}$.

A. Design 1: SNR maximization at the output of the channel shortener

A first strategy to blindly design $\mathbf{\Gamma}$ relies on maximization of the SNR at the channel-shortener output which, accounting for (16), can be expressed as

$$\text{SNR}_1 \triangleq \frac{\mathrm{E}\left[\left\|\widetilde{\mathbf{G}}_{\text{win}}\,\widetilde{\mathbf{u}}_{\text{win}}(n)\right\|^2\right]}{\mathrm{E}[\|\widetilde{\mathbf{d}}(n)\|^2]} = \frac{\text{trace}\left(\widetilde{\mathbf{G}}_{\text{win}}\,\mathbf{R}_{\widetilde{\mathbf{u}}_{\text{win}}\widetilde{\mathbf{u}}_{\text{win}}}\,\widetilde{\mathbf{G}}_{\text{win}}^H\right)}{\text{trace}\left(\mathbf{R}_{\widetilde{\mathbf{d}}\widetilde{\mathbf{d}}}\right)} \tag{17}$$

where $\mathbf{R}_{\widetilde{\mathbf{u}}_{\text{win}}\widetilde{\mathbf{u}}_{\text{win}}} \triangleq \mathrm{E}[\widetilde{\mathbf{u}}_{\text{win}}(n)\,\widetilde{\mathbf{u}}_{\text{win}}^H(n)] \in \mathbb{C}^{N_T(M_{\text{cp}}+1) \times N_T(M_{\text{cp}}+1)}$. Since, by (16), it results that $\widetilde{\mathbf{F}}_{\text{mmoe}}\,\mathbf{R}_{\widetilde{\mathbf{r}}\widetilde{\mathbf{r}}}\,\widetilde{\mathbf{F}}_{\text{mmoe}}^H = \widetilde{\mathbf{G}}_{\text{win}}\,\mathbf{R}_{\widetilde{\mathbf{u}}_{\text{win}}\widetilde{\mathbf{u}}_{\text{win}}}\,\widetilde{\mathbf{G}}_{\text{win}}^H + \mathbf{R}_{\widetilde{\mathbf{d}}\widetilde{\mathbf{d}}}$, by substituting in (17), one has

$$\text{SNR}_1 \triangleq \frac{\text{trace}\left(\widetilde{\mathbf{F}}_{\text{mmoe}}\,\mathbf{R}_{\widetilde{\mathbf{r}}\widetilde{\mathbf{r}}}\,\widetilde{\mathbf{F}}_{\text{mmoe}}^H\right)}{\text{trace}\left(\mathbf{R}_{\widetilde{\mathbf{d}}\widetilde{\mathbf{d}}}\right)} - 1\,. \tag{18}$$

Assuming that **(c3)** holds, and recalling that $\widetilde{\mathbf{F}}_{\text{mmoe}} = \mathbf{\Gamma}\,\mathbf{F}_{\text{mmoe}}$, maximizing (18) with respect to $\mathbf{\Gamma}$ reduces to solving the following constrained maximization problem:

$$\mathbf{\Gamma}_1 = \arg\max_{\mathbf{\Gamma}} \text{trace}\left(\mathbf{\Gamma}\,\mathbf{F}_{\text{mmoe}}\,\mathbf{R}_{\widetilde{\mathbf{r}}\widetilde{\mathbf{r}}}\,\mathbf{F}_{\text{mmoe}}^H\,\mathbf{\Gamma}^H\right) \quad \text{s.t. } \mathbf{\Gamma}\,\mathbf{F}_{\text{mmoe}}\,\mathbf{F}_{\text{mmoe}}^H\,\mathbf{\Gamma}^H = \mathbf{I}_{N_R}\,. \tag{19}$$

Since $\mathbf{F}_{\text{mmoe}} \in \mathbb{C}^{N_V(M_{\text{cp}}+1) \times N_V(L_e+1)}$ is full-row rank for all the SNR values of practical interest, i.e., $\text{rank}(\mathbf{F}_{\text{mmoe}}) = N_V(M_{\text{cp}}+1)$, the solution of such an optimization problem can be conveniently obtained by resorting [36], [37] to the QR decomposition $\mathbf{F}_{\text{mmoe}}^H = \widetilde{\mathbf{Q}}_{\text{mmoe}}\,\widetilde{\mathbf{R}}_{\text{mmoe}}$, where $\widetilde{\mathbf{R}}_{\text{mmoe}} \in \mathbb{C}^{N_V(M_{\text{cp}}+1) \times N_V(M_{\text{cp}}+1)}$ is a nonsingular upper-triangular matrix, and the columns of $\widetilde{\mathbf{Q}}_{\text{mmoe}} \in \mathbb{C}^{N_V(L_e+1) \times N_V(M_{\text{cp}}+1)}$ are orthonormal, i.e., $\widetilde{\mathbf{Q}}_{\text{mmoe}}^H\widetilde{\mathbf{Q}}_{\text{mmoe}} = \mathbf{I}_{N_V(M_{\text{cp}}+1)}$. Following this approach, the solution of (19) can be expressed as $\mathbf{\Gamma}_1 = \widetilde{\mathbf{\Gamma}}_1\,(\widetilde{\mathbf{R}}_{\text{mmoe}}^H)^{-1}$, where $\widetilde{\mathbf{\Gamma}}_1 \in \mathbb{C}^{N_R \times N_V(M_{\text{cp}}+1)}$ solves

the following optimization problem

$$\widetilde{\mathbf{\Gamma}}_1 = \arg\max_{\widetilde{\mathbf{\Gamma}}} \text{trace}\left(\widetilde{\mathbf{\Gamma}}\,\widetilde{\mathbf{Q}}_{\text{mmoe}}^H\,\mathbf{R}_{\widetilde{\mathbf{r}}\widetilde{\mathbf{r}}}\,\widetilde{\mathbf{Q}}_{\text{mmoe}}\,\widetilde{\mathbf{\Gamma}}^H\right) \quad \text{s.t. } \widetilde{\mathbf{\Gamma}}\,\widetilde{\mathbf{\Gamma}}^H = \mathbf{I}_{N_R} \quad (20)$$

and the adopted constraint is similar to the *orthonormality constraint* of [16]. It can be proven [37] that $\widetilde{\mathbf{\Gamma}}_1 = [\widetilde{\gamma}_1, \widetilde{\gamma}_2, \ldots, \widetilde{\gamma}_{N_R}]^H$, where $\widetilde{\gamma}_1, \widetilde{\gamma}_2, \ldots, \widetilde{\gamma}_{N_R}$ are the orthonormal eigenvectors corresponding to the N_R largest eigenvalues of the matrix $\widetilde{\mathbf{Q}}_{\text{mmoe}}^H\,\mathbf{R}_{\widetilde{\mathbf{r}}\widetilde{\mathbf{r}}}\,\widetilde{\mathbf{Q}}_{\text{mmoe}} \in \mathbb{C}^{N_V(M_{\text{cp}}+1)\times N_V(M_{\text{cp}}+1)}$. Accounting for the QR decomposition of $\mathbf{F}_{\text{mmoe}}^H$, the proposed channel-shortening matrix assumes the final expression $\widetilde{\mathbf{F}}_{\text{mmoe},1} \triangleq \mathbf{\Gamma}_1\,\mathbf{F}_{\text{mmoe}} = \mathbf{\Gamma}_1\,\widetilde{\mathbf{Q}}_{\text{mmoe}}^H$. The computational burden of synthesizing $\widetilde{\mathbf{F}}_{\text{mmoe},1}$ is evaluated in Table II, where a comparison with the computational complexity of [16] is also reported.

B. Design 2: SNR maximization at the input of the STFBC ML decoder

In this subsection, we consider a different blind design of the matrix $\mathbf{\Gamma}$, which is based on maximization of the SNR at the *input* of the STFBC ML decoder. The rationale behind such an alternative design is that, because of the non-reversible CP removal carried out between the channel-shortener output and the STFBC decoder input, maximizing the SNR at the channel-shortener output (design 1) leads to a waste of the available degrees of freedom, which are in part spent to perform optimization over data that are irrelevant to STFBC decoding.

In order to develop the design criterion, with reference to (16), we partition $\widetilde{\mathbf{F}}_{\text{mmoe}} = [\widetilde{\mathbf{f}}_1, \widetilde{\mathbf{f}}_2, \ldots, \widetilde{\mathbf{f}}_{N_R}]^H$, with $\widetilde{\mathbf{f}}_\beta \in \mathbb{C}^{N_V(L_e+1)}$, and $\widetilde{\mathbf{G}}_{\text{win}} = [\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_{N_R}]^H$, where $\mathbf{g}_\beta^H = [\mathbf{g}_{\beta,0}^H, \mathbf{g}_{\beta,1}^H, \ldots, \mathbf{g}_{\beta,L_{\text{eff}}}^H]$, with $\mathbf{g}_{\beta,\ell} = [g_{\beta,\ell,1}, g_{\beta,\ell,2}, \ldots, g_{\beta,\ell,N_T}]^T \in \mathbb{C}^{N_T}$, for $\beta \in \{1,2,\ldots,N_R\}$ and $\ell \in \{0,1,\ldots,L_{\text{eff}}\}$. Thus, the channel-shortener output (16) at the βth receive antenna can be written as

$$\begin{aligned}\widetilde{y}_\beta(n) &= \widetilde{\mathbf{f}}_\beta^H\,\widetilde{\mathbf{r}}(n) = \widetilde{\mathbf{r}}^T(n)\,\widetilde{\mathbf{f}}_\beta^* = \mathbf{g}_\beta^H\,\widetilde{\mathbf{u}}_{\text{win}}(n) + \widetilde{d}_\beta(n)\\ &= \sum_{\ell=0}^{L_{\text{eff}}} \mathbf{g}_{\beta,\ell}^H\,\mathbf{u}(n-\ell) + \widetilde{d}_\beta(n)\,. \end{aligned} \quad (21)$$

To recover the symbols belonging to the kth OFDM symbol of the ith codeword, the signal $\widetilde{y}_\beta(n)$ is converted into P parallel substreams $\widetilde{y}_\beta^{(k,p)}(i) \triangleq \widetilde{y}_\beta[(iK+k)P+p]$, for $p \in \{0,1,\ldots,P-1\}$, which are subsequently used to build $\widetilde{\mathbf{y}}_\beta^{(k)}(i) \triangleq [\widetilde{y}_\beta^{(k,0)}(i), \widetilde{y}_\beta^{(k,1)}(i), \ldots, \widetilde{y}_\beta^{(k,P-1)}(i)]^T \in \mathbb{C}^P$. Since $L_{\text{eff}} \leq M_{\text{cp}}$, after removing the CP and carrying out the DFT, one obtains [1] the IBI-free model

$$\begin{aligned}\mathbf{y}_\beta^{(k)}(i) &\triangleq \mathbf{W}_{\text{DFT}}\,\mathbf{R}_{\text{cp}}\,\widetilde{\mathbf{y}}_\beta^{(k)}(i) = \boldsymbol{\mathcal{R}}^{(k)}(i)\,\widetilde{\mathbf{f}}_\beta^*\\ &= \sum_{\alpha=1}^{N_T} \text{diag}[\mathbf{c}_\alpha^{(k)}(i)]\,\boldsymbol{\psi}_{\beta,\alpha} + \mathbf{d}_\beta^{(k)}(i)\\ &= \overline{\mathbf{C}}^{(k)}(i)\,\boldsymbol{\psi}_\beta + \mathbf{d}_\beta^{(k)}(i) \end{aligned} \quad (22)$$

where $\{\mathbf{W}_{\text{DFT}}\}_{mp} \triangleq \frac{1}{\sqrt{M}}\,e^{-j\frac{2\pi}{M}mp}$, with $m,p \in \{0,1,\ldots,M-1\}$, represents the unitary symmetric DFT matrix, the matrix $\mathbf{R}_{\text{cp}} \triangleq [\mathbf{O}_{M\times M_{\text{cp}}}, \mathbf{I}_M] \in \mathbb{R}^{M\times P}$ discards the first M_{cp} entries of $\widetilde{\mathbf{y}}_\beta^{(k)}(i)$, and we have defined the data matrix $\boldsymbol{\mathcal{R}}^{(k)}(i) \triangleq \mathbf{W}_{\text{DFT}}\,\mathbf{R}_{\text{cp}}\,\widetilde{\boldsymbol{\mathcal{R}}}^{(k)}(i) \in \mathbb{C}^{M\times N_V(L_e+1)}$, with $\widetilde{\boldsymbol{\mathcal{R}}}^{(k)}(i) \triangleq [\widetilde{\mathbf{r}}^{(k,0)}(i), \widetilde{\mathbf{r}}^{(k,1)}(i), \ldots, \widetilde{\mathbf{r}}^{(k,P-1)}(i)]^T \in \mathbb{C}^{P\times N_V(L_e+1)}$ gathering the received phases $\widetilde{\mathbf{r}}^{(k,p)}(i) \triangleq \widetilde{\mathbf{r}}[(iK+k)P+p]$, for $i \in \mathbb{Z}$, $k \in \{0,1,\ldots,K-1\}$, and $p \in \{0,1,\ldots,P-1\}$. In addition, the entries of the vector $\boldsymbol{\psi}_{\beta,\alpha} \triangleq [\psi_{\beta,\alpha}(z_0), \psi_{\beta,\alpha}(z_1), \ldots, \psi_{\beta,\alpha}(z_{M-1})] \in \mathbb{C}^M$ represent the values of the conjugate combined channel-prefilter transfer functions $\psi_{\beta,\alpha}(z) \triangleq \sum_{\ell=0}^{L_{\text{eff}}} g_{\beta,\ell,\alpha}^*\,z^{-\ell}$ evaluated at $z_m \triangleq e^{j\frac{2\pi}{M}m}$, and we also defined $\mathbf{d}_\beta^{(k)}(i) \triangleq \mathbf{W}_{\text{DFT}}\,\mathbf{R}_{\text{cp}}\,\widetilde{\mathbf{d}}_\beta^{(k)}(i)$, with $\widetilde{\mathbf{d}}_\beta^{(k)}(i) \triangleq [\widetilde{d}_\beta^{(k,0)}(i), \widetilde{d}_\beta^{(k,1)}(i), \ldots, \widetilde{d}_\beta^{(k,P-1)}(i)]^T \in \mathbb{C}^P$ collecting the noise phases $\widetilde{d}_\beta^{(k,p)}(i) \triangleq \widetilde{d}_\beta[(iK+k)P+p]$, for $i \in \mathbb{Z}$, $k \in \{0,1,\ldots,K-1\}$, and $p \in \{0,1,\ldots,P-1\}$. In the last hand of (22), finally, we defined $\overline{\mathbf{C}}^{(k)}(i) \triangleq [\text{diag}[\mathbf{c}_1^{(k)}(i)], \text{diag}[\mathbf{c}_2^{(k)}(i)], \ldots, \text{diag}[\mathbf{c}_{N_T}^{(k)}(i)]] \in \mathbb{C}^{M\times MN_T}$ and $\boldsymbol{\psi}_\beta \triangleq [\boldsymbol{\psi}_{\beta,1}^T, \boldsymbol{\psi}_{\beta,2}^T, \ldots, \boldsymbol{\psi}_{\beta,N_T}^T]^T \in \mathbb{C}^{MN_T}$.

The K blocks received at the βth antenna within the ith codeword are gathered in $\mathbf{y}_\beta(i) \triangleq [\mathbf{y}_\beta^{(0)}(i)^T, \mathbf{y}_\beta^{(1)}(i)^T, \ldots, \mathbf{y}_\beta^{(K-1)}(i)^T]^T \in \mathbb{C}^{KM}$, thus yielding

$$\mathbf{y}_\beta(i) = \boldsymbol{\mathcal{R}}(i)\,\widetilde{\mathbf{f}}_\beta^* = \overline{\mathbf{C}}(i)\,\boldsymbol{\psi}_\beta + \mathbf{d}_\beta(i) \quad (23)$$

where $\boldsymbol{\mathcal{R}}(i) \triangleq [\boldsymbol{\mathcal{R}}^{(0)}(i)^T, \boldsymbol{\mathcal{R}}^{(1)}(i)^T, \ldots, \boldsymbol{\mathcal{R}}^{(K-1)}(i)^T]^T \in \mathbb{C}^{KM\times N_V(L_e+1)}$, $\mathbf{d}_\beta(i) \triangleq [\mathbf{d}_\beta^{(0)}(i)^T, \mathbf{d}_\beta^{(1)}(i)^T, \ldots, \mathbf{d}_\beta^{(K-1)}(i)^T]^T \in \mathbb{C}^{KM}$, and $\overline{\mathbf{C}}(i) \triangleq [\overline{\mathbf{C}}^{(0)}(i)^T, \overline{\mathbf{C}}^{(1)}(i)^T, \ldots, \overline{\mathbf{C}}^{(K-1)}(i)^T]^T \in \mathbb{C}^{KM\times N_TM}$ conveys the same information[5] of the codeword matrix $\mathbf{C}(i)$.

By defining the matrix $\mathbf{Y}(i) \triangleq [\mathbf{y}_1(i), \mathbf{y}_2(i), \ldots, \mathbf{y}_{N_R}(i)] \in \mathbb{C}^{KM\times N_R}$, the data at the input of the STFBC decoder are given by

$$\mathbf{Y}(i) = \boldsymbol{\mathcal{R}}(i)\,\widetilde{\mathbf{F}}_{\text{mmoe}}^T = \overline{\mathbf{C}}(i)\,\mathbf{\Psi} + \mathbf{D}(i) \quad (24)$$

where $\mathbf{D}(i) \triangleq [\mathbf{d}_1(i), \mathbf{d}_2(i), \ldots, \mathbf{d}_{N_R}(i)] \in \mathbb{C}^{KM\times N_R}$ and $\mathbf{\Psi} \triangleq [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_{N_R}] \in \mathbb{C}^{N_TM\times N_R}$ collects all the frequency responses of the shortened channels, which can be estimated for instance by resorting to conventional training-based channel-estimation techniques [19]. Based on (24), the SNR at the input of the STFBC decoder can be defined as

$$\text{SNR}_2 \triangleq \frac{\text{E}\left[\left\|\overline{\mathbf{C}}(i)\,\mathbf{\Psi}\right\|^2\right]}{\text{E}\left[\left\|\mathbf{D}(i)\right\|^2\right]} = \frac{\text{trace}\left(\mathbf{\Psi}^H\mathbf{R}_{\overline{\mathbf{C}}\overline{\mathbf{C}}}\,\mathbf{\Psi}\right)}{\text{trace}\left(\mathbf{R}_{\mathbf{DD}}\right)} \quad (25)$$

where $\mathbf{R}_{\mathbf{DD}} \triangleq \text{E}[\mathbf{D}^H(i)\,\mathbf{D}(i)] \in \mathbb{C}^{N_R\times N_R}$ and $\mathbf{R}_{\overline{\mathbf{C}}\overline{\mathbf{C}}} \triangleq \text{E}[\overline{\mathbf{C}}^H(i)\,\overline{\mathbf{C}}(i)] \in \mathbb{C}^{MN_T\times MN_T}$ is the autocorrelation matrix of the STFBC codeword. By (24), it results that $\widetilde{\mathbf{F}}_{\text{mmoe}}^*\,\mathbf{R}_{\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{R}}}\,\widetilde{\mathbf{F}}_{\text{mmoe}}^T = \mathbf{\Psi}^H\mathbf{R}_{\overline{\mathbf{C}}\overline{\mathbf{C}}}\,\mathbf{\Psi} + \mathbf{R}_{\mathbf{DD}}$, where $\mathbf{R}_{\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{R}}} \triangleq \text{E}[\boldsymbol{\mathcal{R}}^H(i)\,\boldsymbol{\mathcal{R}}(i)] \in \mathbb{C}^{N_V(L_e+1)\times N_V(L_e+1)}$, hence, by substituting this last relation in (25), one has

$$\text{SNR}_2 \triangleq \frac{\text{trace}\left(\widetilde{\mathbf{F}}_{\text{mmoe}}\,\mathbf{R}_{\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{R}}}^*\,\widetilde{\mathbf{F}}_{\text{mmoe}}^H\right)}{\text{trace}\left(\mathbf{R}_{\mathbf{DD}}\right)} - 1\,. \quad (26)$$

Observe that, by virtue of **(c3)**, it turns out that $\text{trace}(\mathbf{R}_{\mathbf{DD}}) = \sigma_v^2\,K\,M\,N_R$. Reasoning as in the previous subsection, and

[5]The matrix $\overline{\mathbf{C}}(i)$ has the same form as $\mathbf{C}(i)$, but with the diagonal matrices $\text{diag}[\mathbf{c}_\alpha^{(k)}(i)]$ *in lieu* of the vectors $\mathbf{c}_\alpha^{(k)}(i)$.

- Computational burden of evaluating $\mathbf{F}_{\text{mmoe}}$ given by (12).
  Since both the matrices $\mathbf{\Theta}_{\Delta_e}$ and $\mathbf{\Pi}_{\Delta_e}$ can be pre-computed off-line, the synthesis of $\mathbf{F}_{\text{mmoe}}$ requires only the real-time computation of the inverse of $\mathbf{\Pi}_{\Delta_e}\,\mathbf{R}_{\widetilde{\mathbf{r}}\widetilde{\mathbf{r}}}\,\mathbf{\Pi}_{\Delta_e}^T$, which can consistently be estimated from the received data either in batch mode or adaptively (e.g., a fast recursive least-square (RLS) algorithm for direct estimation of $\mathbf{F}_{\text{mmoe}}$ from the received data can be found in [27]). The computational burden of the matrix inversion $(\mathbf{\Pi}_{\Delta_e}\,\mathbf{R}_{\widetilde{\mathbf{r}}\widetilde{\mathbf{r}}}\,\mathbf{\Pi}_{\Delta_e}^T)^{-1}$ involves:
  - $\mathcal{O}[N_V^3\,(L_e-\Delta_e)^3]$ floating point operations (flops) [36] if one resorts to batch algorithms;
  - $\mathcal{O}[N_V^2\,(L_e-\Delta_e)^2]$ flops per iteration if the RLS-like adaptive algorithm of [27] is employed.

  As a comparison, the synthesis of the non-blind MMSE channel shortener proposed in [16] requires inversion of the matrix $\mathbf{R}_{\widetilde{\mathbf{r}}\widetilde{\mathbf{r}}}$, with a higher computational complexity of:
  - $\mathcal{O}[N_V^3\,(L_e+1)^3]$ flops in batch mode;
  - $\mathcal{O}[N_V^2\,(L_e+1)^2]$ flops per iteration in RLS-based adaptive mode.
- Computational burden of evaluating $\mathbf{\Gamma}_1$ given by (20).
  The matrix $\widetilde{\mathbf{Q}}_{\text{mmoe}}$ can be efficiently obtained from $\mathbf{F}_{\text{mmoe}}^H$ by performing $N_V(M_{\text{cp}}+1)$ successive Givens or Householder rotations (see [27] for details), which involves:
  - $\mathcal{O}[N_V^3\,(M_{\text{cp}}+1)^2(L_e+1)]$ flops if computed from scratch;
  - $\mathcal{O}[N_V^2\,(L_e+1)^2]$ flops per iteration if evaluated recursively [36].

  After computing $\widetilde{\mathbf{Q}}_{\text{mmoe}}$, the N_R dominant eigenvectors of $\widetilde{\mathbf{Q}}_{\text{mmoe}}^H\,\mathbf{R}_{\widetilde{\mathbf{r}}\widetilde{\mathbf{r}}}\,\widetilde{\mathbf{Q}}_{\text{mmoe}}$ can be blindly estimated from the received data, either in batch mode [36] or adaptively by adopting the projection approximation subspace tracking (PAST) algorithm [38] operating on the transformed vector $\widetilde{\boldsymbol{z}}(n) \triangleq \widetilde{\mathbf{Q}}_{\text{mmoe}}^H\,\widetilde{\mathbf{r}}(n) \in \mathbb{C}^{N_V(M_{\text{cp}}+1)}$. Such operation requires:
  - $\mathcal{O}[N_V^3\,(M_{\text{cp}}+1)^3]$ flops with batch algorithms;
  - $\mathcal{O}[N_R\,N_V\,(M_{\text{cp}}+1)]$ flops per iteration with the PAST algorithm.

  As a comparison, the synthesis of the constraint for the non-blind MMSE channel shortener of [16] involves:
  - $\mathcal{O}[N_T^3\,(M_{\text{cp}}+1)^3]$ flops in batch mode;
  - $\mathcal{O}[N_T^2\,(M_{\text{cp}}+1)^2]$ flops per iteration with the PAST algorithm.

  Hence, in general, the computational complexity of evaluating $\mathbf{\Gamma}_1$ is slightly higher in batch mode, whereas it tends to be smaller when the PAST algorithm is employed.

TABLE II
COMPUTATIONAL COMPLEXITY OF THE MMOE CHANNEL SHORTENER (DESIGN 1) IN COMPARISON WITH THAT OF [16]. A SIMILAR EVALUATION CAN BE CARRIED OUT FOR DESIGN 2.

recalling that $\widetilde{\mathbf{F}}_{\text{mmoe}} = \mathbf{\Gamma}\,\mathbf{F}_{\text{mmoe}}$, the value of $\mathbf{\Gamma}$ maximizing (26) can be expressed as $\mathbf{\Gamma}_2 = \widetilde{\mathbf{\Gamma}}_2\,(\widetilde{\mathbf{R}}_{\text{mmoe}}^H)^{-1}$, where $\widetilde{\mathbf{\Gamma}}_2 \triangleq [\widetilde{\boldsymbol{\gamma}}_1, \widetilde{\boldsymbol{\gamma}}_2, \ldots, \widetilde{\boldsymbol{\gamma}}_{N_R}]^H \in \mathbb{C}^{N_R \times N_V(M_{\text{cp}}+1)}$ is the solution of the following constrained optimization problem

$$\widetilde{\mathbf{\Gamma}}_2 = \arg\max_{\widetilde{\mathbf{\Gamma}}} \text{trace}\left(\widetilde{\mathbf{\Gamma}}\,\widetilde{\mathbf{Q}}_{\text{mmoe}}^H\,\mathbf{R}_{\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{R}}}^*\,\widetilde{\mathbf{Q}}_{\text{mmoe}}\,\widetilde{\mathbf{\Gamma}}^H\right), \quad \text{s.t. } \widetilde{\mathbf{\Gamma}}\,\widetilde{\mathbf{\Gamma}}^H = \mathbf{I}_{N_R} \quad (27)$$

with $\widetilde{\boldsymbol{\gamma}}_1, \widetilde{\boldsymbol{\gamma}}_2, \ldots, \widetilde{\boldsymbol{\gamma}}_{N_R}$ representing the orthonormal eigenvectors [37] corresponding to the N_R largest eigenvalues of the Hermitian matrix $\widetilde{\mathbf{Q}}_{\text{mmoe}}^H\,\mathbf{R}_{\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{R}}}^*\,\widetilde{\mathbf{Q}}_{\text{mmoe}} \in \mathbb{C}^{N_V(M_{\text{cp}}+1)\times N_V(M_{\text{cp}}+1)}$. Accounting for the QR decomposition of $\mathbf{F}_{\text{mmoe}}^H$, the proposed channel-shortening matrix assumes the final expression $\widetilde{\mathbf{F}}_{\text{mmoe},2} \triangleq \mathbf{\Gamma}_2\,\mathbf{F}_{\text{mmoe}} = \widetilde{\mathbf{\Gamma}}_2\,\widetilde{\mathbf{Q}}_{\text{mmoe}}^H$. Interestingly, for a given $\widetilde{\mathbf{Q}}_{\text{mmoe}}$, the N_R dominant eigenvalues of $\widetilde{\mathbf{Q}}_{\text{mmoe}}^H\,\mathbf{R}_{\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{R}}}^*\,\widetilde{\mathbf{Q}}_{\text{mmoe}}$ can be blindly estimated from the received data either in batch mode [36], or adaptively by adopting the PAST algorithm [38] operating on the transformed vector $\widetilde{\boldsymbol{z}}(n) \triangleq \widetilde{\mathbf{Q}}_{\text{mmoe}}^H\,\widetilde{\mathbf{r}}(n)$, as in design 1. Indeed, evaluation of $\mathbf{\Gamma}_2$ essentially involves the same computational complexity of $\mathbf{\Gamma}_1$ (see Table II).

V. NUMERICAL PERFORMANCE ANALYSIS

Theoretical performance analysis (e.g., BER or diversity order evaluation) of the STFBC MIMO-OFDM ML transceivers equipped with the proposed channel shorteners is a formidable task, since the shortening matrices, for both designs, are complicated functions of the physical channel coefficients. Thus, in order to assess the system performances, we present in this section the results of extensive Monte Carlo computer simulations, aimed at evaluating the average BER (ABER) at the output of the STFBC ML decoder.

The considered MIMO-OFDM system works at sampling period $T = 50$ ns and employs $N_T = 2$ antennas, $M = 32$ subcarriers, and a CP of length $M_{\text{cp}} = 8$ samples. As STFBC rule, we implemented the standard Alamouti code [29] (see Section II), where the blocks $\mathbf{c}_1^{(0)}(i)$ and $\mathbf{c}_2^{(0)}(i)$ are chosen as statistically independent of each other, with i.i.d. entries drawn from a QPSK alphabet $\left\{\pm\frac{1}{\sqrt{2}} \pm j\frac{1}{\sqrt{2}}\right\}$. For each $\ell \in \{0, 1, \ldots, L_h\}$, the entries of the MIMO channel matrix $\mathbf{H}(\ell)$ in (1) are generated as i.i.d. circularly symmetric complex Gaussian random variables with zero mean and variance $\sigma_h^2(\ell)$, with $\mathbf{H}(\ell_1)$ and $\mathbf{H}(\ell_2)$ statistically independent of each other for $\ell_1 \neq \ell_2$. This channel model complies with the widely accepted Rayleigh fading scenario and, moreover, does not take into account both transmit and receive correlation, as it is reasonable in rich scattering environments [39]. Two channel profiles are considered: an uniform power delay profile (PDP) model, with $\sigma_h^2(\ell) = \frac{1}{L_h+1}$, for $\ell = 0, 1, \ldots, L_h$, as well as an exponential one, with $\sigma_h^2(\ell) = \frac{1/\overline{T}_m}{\sum_{k=0}^{L_h}\sigma_h^2(k)}\,e^{-\ell\,(T_c/\overline{T}_m)}$, for $\ell = 0, 1, \ldots, L_h$, with delay spread set to $\overline{T}_m = 300$ ns. In addition to the assumption already made in Section III, the noise $\widetilde{\mathbf{v}}(n)$ is modeled as a circularly symmetric complex Gaussian random vector, with $\widetilde{\mathbf{v}}(n_1)$ and $\widetilde{\mathbf{v}}(n_2)$ statistically independent of each other for $n_1 \neq n_2$. For the purpose of channel estimation, the OFDM signal piggybacks one training symbol per antenna.

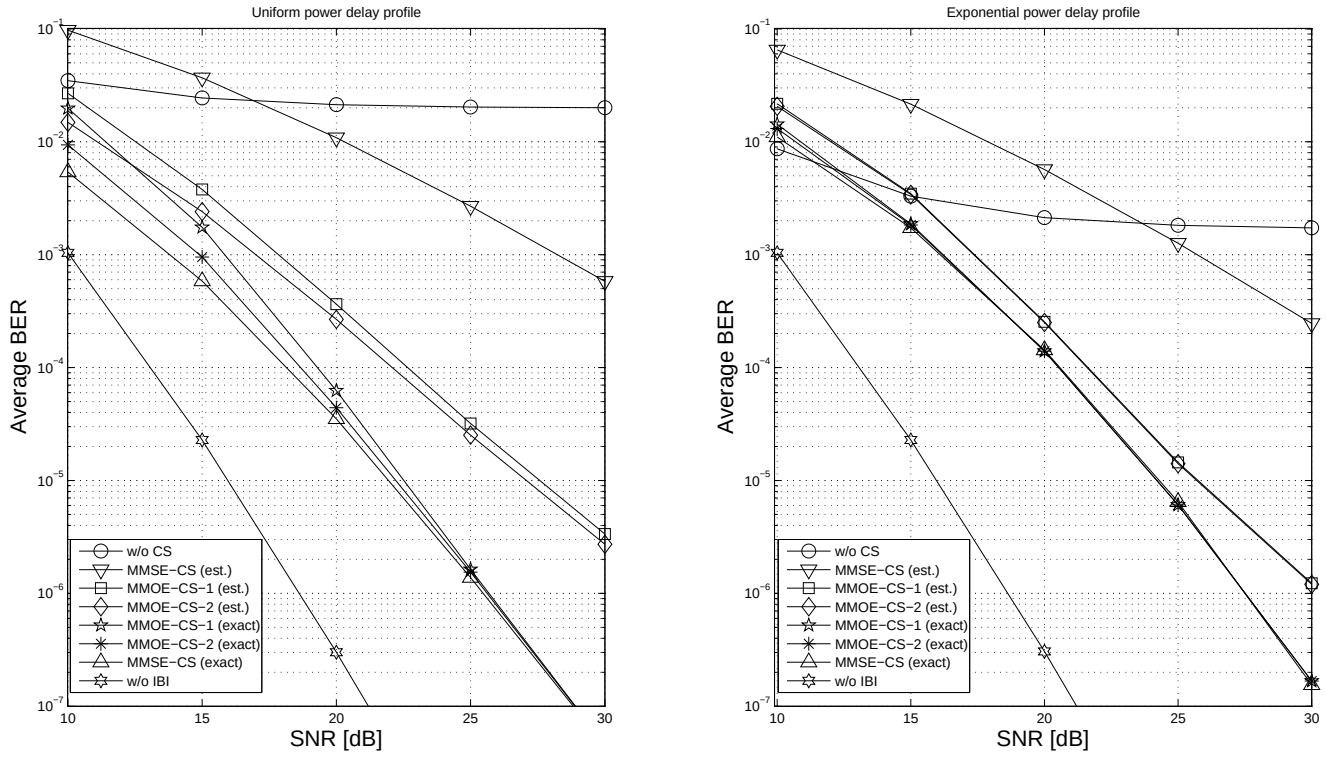

Fig. 1. Average BER versus SNR for the scenario with oversampling ($N_R = 2$, $Q = 2$).

Finally, according to (1), the received SNR is defined as

$$\text{SNR} \triangleq \frac{\text{E}\left[\left\|\sum_{\ell=0}^{L_h} \mathbf{H}(\ell)\,\mathbf{u}(n-\ell)\right\|^2\right]}{\text{E}\left[\|\mathbf{v}(n)\|^2\right]} = \frac{\sum_{\ell=0}^{L_h} \sigma_h^2(\ell)\; N_T}{\sigma_v^2}\,. \tag{28}$$

We evaluated the performance of the system when the two proposed blind channel shorteners referred to as "MMOE-CS-1" (design 1) and "MMOE-CS-2" (design 2), respectively, are used. In each of the 10^5 Monte Carlo runs carried out (wherein, besides all the channel coefficients, independent sets of noise and data sequences are randomly generated), an independent record of 10^2 OFDM symbols was generated to accurately evaluate the BER. We considered two scenarios: in the first scenario (with oversampling), $N_R = 2$ antennas are employed at the receiver and $Q = 2$; in the second one (without oversampling), $N_R = 4$ and $Q = 1$. In both scenarios, the performances of our blind channel shorteners are compared with those of the non-blind MMSE channel-shortener of [16] (referred to as "MMSE-CS"), which requires a priori knowledge (or estimation) of the channel matrix $\widetilde{\mathbf{H}}$ in (3) and is implemented with $\Delta_s = 0$. Indeed, we found through extensive computer simulations that $\Delta_s = 0$ is the optimal value of the shortening delay for the MMSE channel shortener of [16] with the orthonormality constraint when it operates in the considered scenarios. This value is different from that originally recommended in [16], due to the fact that a different MIMO system is considered in [16]. Moreover, the MMSE shortening matrix of [16] does not fulfill a design constraint similar to **(c3)** and, consequently, the noise at the output of the MMSE channel shortener is colored. Therefore, only for the MMSE channel shortener, an invertible whitening transformation has been implemented before STFBC decoding.

All the considered receivers are implemented in both their exact and data-dependent versions. In the exact versions (labeled with "exact"), we assumed perfect knowledge of the autocorrelation matrix $\mathbf{R}_{\widetilde{\mathbf{r}}\widetilde{\mathbf{r}}}$ needed for the synthesis of the channel shorteners,[6] and of the frequency response of the "shortened" MIMO channel required for coherent ML detection. In the data-dependent versions (labeled with "est."), the matrix $\mathbf{R}_{\widetilde{\mathbf{r}}\widetilde{\mathbf{r}}}$ is estimated in batch mode on the basis of a sample-size of $K_s = 50$ OFDM symbols, and the shortened MIMO channel is estimated by using a least-squares (LS) approach [19]. Additionally, since the MMSE channel shortener of [16] is not blind, the exact version of the "MMSE-CS" receiver is obtained by assuming also the perfect knowledge of $\widetilde{\mathbf{H}}$; accordingly, the data-dependent version of "MMSE-CS" is implemented by also performing LS estimation of the "long" MIMO channel impulse response, i.e., of the matrices $\{\mathbf{H}(\ell)\}_{\ell=0}^{L_h}$. As a reference, to show the degradation due to IBI, we also reported the performance of the exact MIMO-OFDM ML transceiver (referred to as "w/o CS") that operates without channel shortening, without oversampling ($Q = 1$), and assuming perfect knowledge of $\widetilde{\mathbf{H}}$, as well as

[6]For the setting at hand, the autocorrelation matrix $\mathbf{R}_{\widetilde{\mathbf{u}}\widetilde{\mathbf{u}}}$ of the transmitted OFDM signal reduces to $\mathbf{R}_{\widetilde{\mathbf{u}}\widetilde{\mathbf{u}}} = \mathbf{I}_{N_T(L_e+L_h+1)}$ when $L_g = L_e + L_h < M$ [26]; when this condition is violated, instead, $\mathbf{R}_{\widetilde{\mathbf{u}}\widetilde{\mathbf{u}}}$ is no longer white and possibly time-variant [40]. Nevertheless, we consistently used in the simulations the simplified $\mathbf{R}_{\widetilde{\mathbf{u}}\widetilde{\mathbf{u}}} = \mathbf{I}_{N_T(L_e+L_h+1)}$ in the synthesis of all exact versions of the receivers.

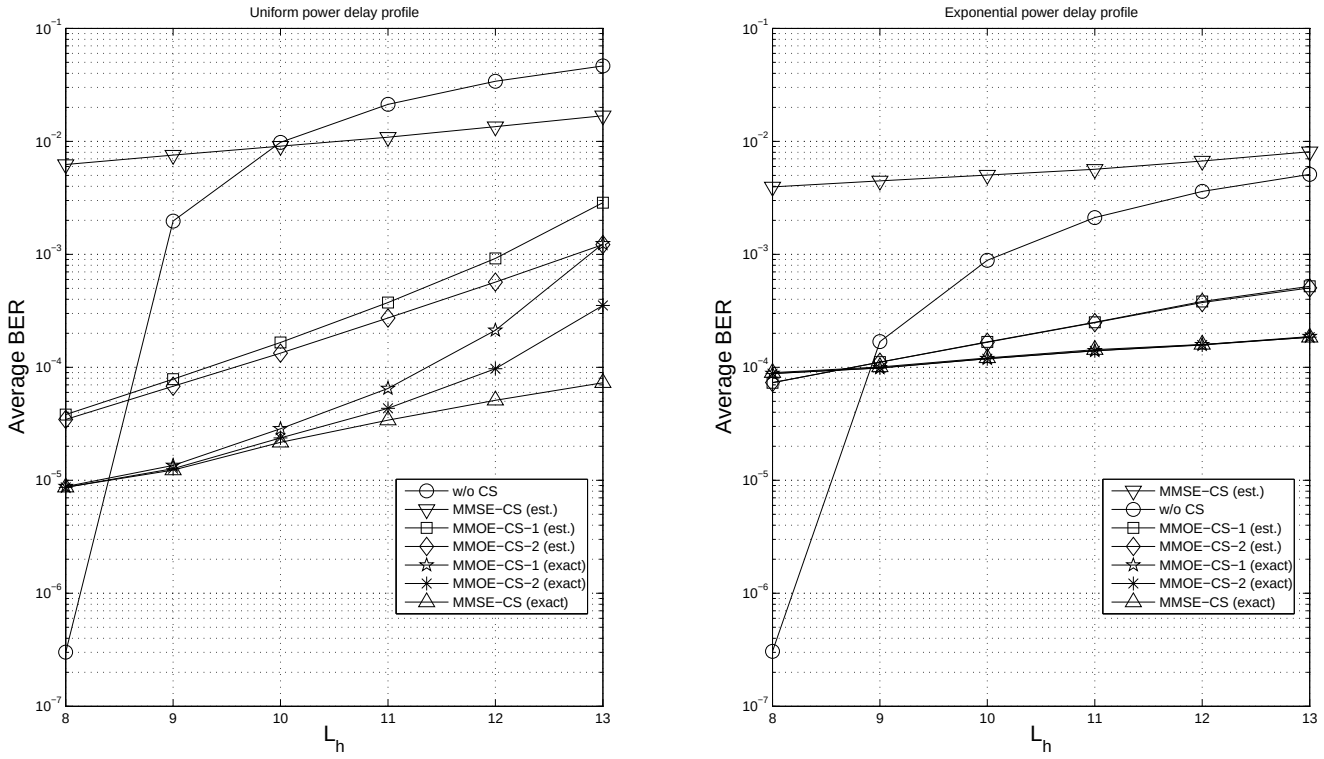

Fig. 2. Average BER versus channel order L_h for the scenario with oversampling ($N_R = 2$, $Q = 2$).

the performance of the same transceiver (referred to as "w/o IBI") in the case of no IBI ($L_h = M_{\text{cp}}$).

A. Scenario with oversampling ($N_R = 2$, $Q = 2$)

In the first experiment we evaluated (Fig. 1) the ABER as a function of SNR, for both the uniform (left plot) and exponential (right plot) channel PDPs, with $L_h = 11$. According to (15), the minimum allowable value of L_e in this scenario is $L_h + M_{\text{cp}} = 19$, hence we choose $L_e = 20$. Results show that, since $M_{\text{cp}} < L_h$, the performance of the conventional receiver "w/o CS" is completely unsatisfactory for both channel profiles, due to the presence of a BER floor, which is lower for the exponential PDP. In contrast, the exact versions of the proposed blind "MMOE-CS" receivers (both designs) exhibit satisfactory performances for both channel profiles. Moreover, their performances are comparable to those of the non-blind "MMSE-CS" receiver, except for very low values of the SNR in the uniform PDP case. In particular, the exact versions of the "MMOE-CS" (both designs) and "MMSE-CS" receivers exhibit the same diversity order,[7] which turns out to be smaller than that of the "w/o IBI" transceiver. Indeed, in the case of the Alamouti code and in the absence of IBI (i.e., without channel shortening), the diversity order of the MIMO-OFDM system with ML decoding can be also evaluated analytically [3] and is given by $2\,N_R = 4$. A way to reduce this loss of diversity is to gain additional degrees of freedom for getting nearer to the ideal condition of perfect channel shortening (5). For instance, this can be obtained by increasing the oversampling factor Q and, hence, the number of virtual receiving antennas N_V, at the expense however of a significant increase in computational complexity (see Table II). Turning to data-estimated versions, the "MMOE-CS" receivers (both designs) pay a moderate performance loss to their exact versions,[8] while significantly outperforming the data-estimated "MMSE-CS". Differently from the "MMOE-CS" receivers, the performance of the data-estimated "MMSE-CS" receiver departs significantly from that of its exact counterpart, mainly because the data-estimated "MMSE-CS" receiver requires estimation of the long (i.e., before shortening) MIMO channel impulse response, which tends to be inaccurate unlike a large number of OFDM training symbols is employed.

In the second experiment, we evaluated (Fig. 2) the ABER as a function of the MIMO channel order L_h, for both channel PDPs, when SNR $= 20$ dB and $L_e = 20$. When $L_h = 13$, condition (15) is violated, i.e., perfect channel shortening is not ensured for the proposed blind methods even in the absence of noise; whereas, when $L_h = M_{\text{cp}} = 8$, CP removal is capable to completely suppress IBI: in this case, the conventional receiver "w/o CS" is expected to exhibit the best performance. Results show that the performances of the blind exact "MMOE-CS" receivers (both designs) are similar to those of the non-blind exact "MMSE-CS" one, paying a marked performance penalty only for $L_h \in \{12, 13\}$ in the more pessimistic uniform channel PDP scenario. With

[7]Since the curve of the error probability at the output of the decoder versus SNR in log-log scale becomes a straight line in the high SNR region, the diversity order is the slope of such a curve [4].

[8]This performance loss can be reduced only by increasing the sample-size K_s, whereas data-estimated receivers exhibit an irreducible floor when only the SNR is increased.

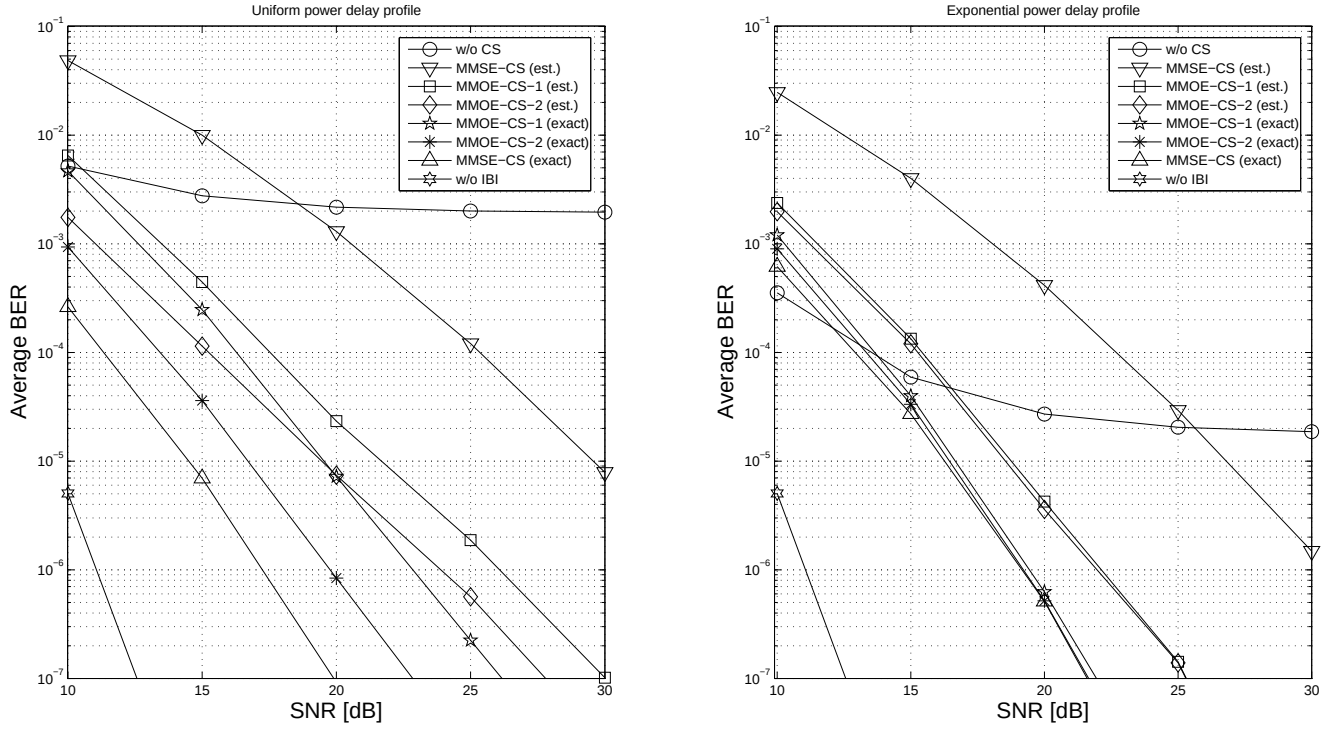

Fig. 3. Average BER versus SNR for the scenario without oversampling ($N_R = 4$, $Q = 1$).

reference to the data-estimated versions, similarly to the first experiment, the "MMSE-CS" receiver exhibits a larger performance degradation compared to "MMOE-CS" receivers (both designs), for all values of L_h, which results in a rather poor behavior for both channel PDPs.

B. Scenario without oversampling ($N_R = 4$, $Q = 1$)

In the third experiment we evaluated (Fig. 3) the ABER as a function of SNR, for both channel PDPs, when $L_h = 11$ and $L_e = 20$. For the exponential channel PDP, similarly to results of the first experiment (compare with right plot of Fig. 1), the exact versions of "MMOE-CS" receivers (both designs) attain the same performance of non-blind exact "MMSE-CS". When an uniform channel PDP is considered, however, differently from the first experiment (compare with left plot of Fig. 1), the exact version of "MMOE-CS-2" performs better than the exact version of "MMOE-CS-1" for all the considered SNR values, exhibiting a SNR performance gain of about 2 dB. In both PDP scenarios, the ABER curves of the "MMOE-CS" (both designs) and "MMSE-CS" receivers exhibit the same diversity order, which is smaller than that of the "w/o IBI" transceiver (expected to be equal to $2\,N_R = 8$). Turning to data-estimated receivers, similarly to the results of the first experiment, the "MMOE-CS" receivers (both designs) exhibit a moderate performance degradation with respect to their exact counterparts, for both channel profiles, whereas the data-estimated "MMSE-CS" receiver pays a severe performance penalty to its exact version.

Finally, in the fourth experiment we evaluated the ABER (Fig. 4) as a function of the MIMO channel order L_h, for both channel PDPs, when SNR $= 20$ dB and $L_e = 20$. The remark regarding L_h stated in the third experiment is also valid for the experiment at hand. Results confirm, in this scenario without oversampling, the superiority of the "MMOE-CS-2" receiver over the "MMOE-CS-1" one for the uniform channel PDP, whereas in the exponential PDP case the two designs perform comparably.

VI. Conclusions

Two channel-shortening blind designs for MIMO-OFDM systems employing STFBC have been proposed and analyzed, based on the constrained MMOE criterion followed by SNR maximization, which admit a closed-form solution. The first design is based on SNR maximization at the channel shortener output, whereas the second one directly relies on SNR maximization at the input of the STFBC ML decoder. Simulation results show that receivers equipped with the proposed blind channel-shortening prefilters pay only a slight performance loss to the non-blind MMSE channel shortener, when all the receivers are implemented exactly. Interestingly, the data-estimated versions of the proposed receivers are significantly more robust than the MMSE one, exhibiting a distinctive superior performance for both designs. In particular, the second design turns out to be especially advantageous in the absence of oversampling and for a channel with uniform power delay profile. A future research issue consists of analytically evaluating the diversity order of a MIMO-OFDM transceiver employing the considered channel shortening strategies.

References

[1] Z. Wang and G.B. Giannakis, "Wireless multicarrier communications – Where Fourier meets Shannon," *IEEE Signal Process. Mag.*, vol. 17, pp. 29–48, May 2000.

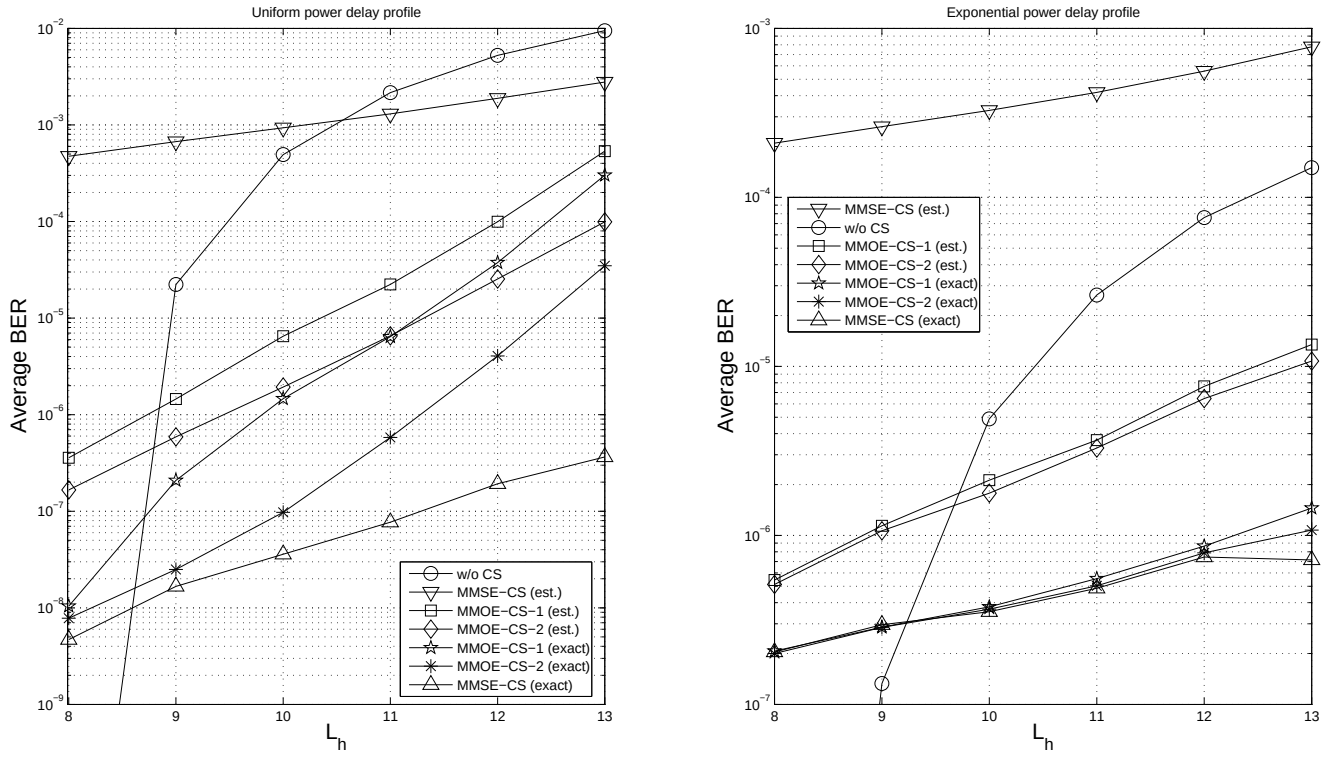

Fig. 4. Average BER versus channel order L_h for the scenario without oversampling ($N_R = 4$, $Q = 1$).

[2] A.F. Molish, M.Z. Win, and J.H. Winters, "Space-time-frequency (STF) coding for MIMO-OFDM systems," *IEEE Commun. Lett.*, vol. 6, pp. 370–372, Sept. 2002.

[3] W. Su, Z. Safar, and K.J.R Liu, "Towards maximum achievable diversity in space, time, and frequency: Performance analysis and code design," *IEEE Trans. Signal Process.*, vol. 4, pp. 1847–1857, July 2005.

[4] V. Tarokh, N. Seshadri, and A.R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, pp. 744–765, Mar. 1998.

[5] H. ̈Bolcskei and A.J. Paulraj, "Space-frequency coded broadband OFDM systems," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Chicago, IL, USA, Sept. 2000, pp. 1–6.

[6] W. Su, Z. Safar, M. Olfat, and K.J.R. Liu, "Obtaining full-diversity space-frequency codes from space-time codes via mapping," *IEEE Trans. Signal Process.*, vol. 51, pp. 2905–2916, Nov. 2003.

[7] H. ̈Bolcskei, M. Borgmann, and A.J. Paulraj, "Impact of the propagation environment on the performance of space-frequency coded MIMO-OFDM," *IEEE J. Sel. Areas Commun.*, vol. 21, pp. 427–439, Apr. 2003.

[8] W. Su, Z. Safar, and K.J.R. Liu, "Full-rate full-diversity space-frequency codes with optimum coding advantage," *IEEE Trans. Inf. Theory*, vol. 51, pp. 229–249, Jan. 2005.

[9] P.J.W. Melsa, R.C. Younce, and C.E. Rohrs, "Impulse response shortening for discrete multitone transceivers," *IEEE Trans. Commun.*, vol. 44, pp. 1662–1672, Dec. 1996.

[10] C. Yin and G. Yue, "Optimal impulse response shortening for discrete multitone transceivers," *Electron. Lett.*, vol. 34, pp. 35–36, Jan. 1998.

[11] A. Tkacenko and P.P. Vaidyanathan, "A low-complexity eigenfilter design method for channel shortening equalizers for DMT systems," *IEEE Trans. Commun.*, vol. 51, pp. 1069–1072, July 2003.

[12] N. Al-Dhahir and J.M. Cioffi, "Optimum finite-length equalization for multicarrier transceivers," *IEEE Trans. Commun.*, vol. 44, pp. 56–64, Jan. 1996.

[13] D. Daly, C. Heneghan, and A.D. Fagan, "Minimum mean-squared error impulse response shortening for discrete multitone transceivers," *IEEE Trans. Signal Process.*, vol. 52, pp. 301-306, Jan. 2004.

[14] G. Arslan, B. Evans, and S. Kiaei, "Equalization for discrete multitone transceivers to maximize bit rate," *IEEE Trans. Signal Process.*, vol. 49, pp. 3123-3135, Dec. 2001.

[15] S. Celebi, "Interblock interference (IBI) minimizing time-domain equalizer (TEQ) for OFDM", *IEEE Signal Process. Lett.*, vol. 10, pp. 232-234, Aug. 2003.

[16] N. Al-Dhahir, "FIR channel-shortening equalizers for MIMO ISI channels," *IEEE Trans. Commun.*, vol. 49, pp. 213–218, Feb. 2001.

[17] C. Toker, S. Lambotharan, J.A. Chambers, and B. Baykal, "Joint spatial and temporal channel-shortening techniques for frequency selective fading MIMO channels," *IEE Proc. Commun.*, vol. 152, pp. 89–94, Feb. 2005.

[18] R.K. Martin, G. Ysebaert, and K. Vanbleu, "Bit error rate minimizing channel shortening equalizers for cyclic prefixed systems," *IEEE Trans. Signal Process.*, vol. 55, pp. 2605-2616, June 2007.

[19] H. Minn and N. Al-Dhahir, "Optimal training signals for MIMO OFDM channel estimation," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 1158–1168, May 2006.

[20] R.K. Martin, J. Balakrishnan, W.A. Sethares, and C.R. Johnson, "A blind adaptive TEQ for multicarrier systems," *IEEE Signal Process. Lett.*, vol. 9, pp. 341-343, Nov. 2002.

[21] J. Balakrishnan, R.K. Martin, and C.R. Johnson, "Blind adaptive channel shortening by sum-squared auto-correlation minimization (SAM)," *IEEE Trans. Signal Process.*, vol. 51, pp. 3086-3093, Dec. 2003.

[22] R.K. Martin, "Joint blind adaptive carrier frequency offset estimation and channel shortening," *IEEE Trans. Signal Process.*, vol. 54, pp. 4194-4203, Nov. 2006.

[23] R.K. Martin, "Fast-converging blind adaptive channel-shortening and frequency-domain equalization," *IEEE Trans. Signal Process.*, vol. 55, pp. 102-110, Jan. 2007.

[24] R.K. Martin, J.M. Walsh, and C.R. Johnson, "Low-complexity MIMO blind adaptive channel shortening," *IEEE Trans. Signal Process.*, vol. 53, pp. 1324-1334, Apr. 2005.

[25] G. Altin and R.K. Martin, "Adaptive MIMO channel shortening with post-FEQ diversity combining," in *Proc. IEEE Asilomar Conf. Signals, Syst., Comp.*, Pacific Grove, CA, USA, Nov. 2008, pp. 2183–2187.

[26] T. Miyajima and Z. Ding, "Second-order statistical approaches to channel shortening in multicarrier systems," *IEEE Trans. Signal Process.*, vol. 52, pp. 3253-3264, Nov. 2004.

[27] D. Darsena and F. Verde, "Minimum-mean-output-energy blind adaptive channel shortening for multicarrier SIMO transceivers," *IEEE Trans. Signal Process.*, vol. 55, pp. 5755–5771, Jan. 2007.

[28] D. Darsena, G. Gelli, L. Paura, and F. Verde, "Blind MMOE channel shortening for MIMO-OFDM systems operating over highly frequency-

selective chennels," in *Proc. IEEE International Conf. Acoustics, Speech Signal Process.*, Dallas, TX, USA, Mar. 2010, pp. 3198–3201.
[29] S.M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.
[30] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses*. NY: Springer-Verlag, 2002.
[31] F. Romano and S. Barbarossa, "Non-data aided adaptive channel shortening for efficient multicarrier systems," in *Proc. IEEE Int. Conf. Acoustic, Speech, Signal Process.*, Hong Kong, China, Apr. 2003, pp. 233–236.
[32] Y. Jia, U. Nie, and C. Zhang, "Trace ratio problem revisited," *IEEE Trans. Neural Netw.*, vol. 20, pp. 729–735, Apr. 2009.
[33] M. Honig, U. Madhow, and S. Verdù, "Blind adaptive multiuser detection," *IEEE Trans. Inf. Theory*, vol. 41, pp. 944–960, July 1995.
[34] H.L. Van Trees, *Optmimum Array Processing*. NY: John Wiley & Sons, 2002.
[35] G. Gelli and F. Verde, "Two-stage interference-resistant adaptive periodically time-varying CMA blind equalization," *IEEE Trans. Signal Process.*, vol. 50, pp. 662–672, Mar. 2002.
[36] G.H. Golub and C.F. Van Loan, *Matrix Computations*. Baltimore: John Hopkins Univ. Press, 1996.
[37] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge: Cambridge University Press, 1990.
[38] B. Yang, "Projection approximation subspace tracking," *IEEE Trans. Signal Process.*, vol. 43, pp. 95–107, Jan. 1995.
[39] D. Tse and P. Wiswanath, *Fundamentals of Wireless Communications*. Cambridge: Cambridge University Press, 2005.
[40] D. Darsena, G. Gelli, L. Paura, and F. Verde, "Blind periodically time-varying MMOE channel shortening for OFDM systems," in *Proc. IEEE International Conf. Acoustics, Speech Signal Process.*, Prague, Czech Republic, May 2011.

Donatella Darsena (M'06) received the Dr. Eng. degree *summa cum laude* in telecommunications engineering in 2001, and the Ph.D. degree in electronic and telecommunications engineering in 2005, both from the University of Napoli Federico II, Italy.

From 2001 to 2002, she was an engineer in the Telecommunications, Peripherals and Automotive Group, STMicroelectronics, Milano, Italy. Since 2005, she has been an Assistant Professor with the Department for Technologies, University of Napoli Parthenope, Italy. Her research activities lie in the area of statistical signal processing, digital communications, and communication systems. In particular, her current interests are focused on equalization, channel identification, narrowband-interference suppression for multicarrier systems, and space-time processing for cooperative communications systems.

Giacinto Gelli was born in Napoli, Italy, on July 29, 1964. He received the Dr. Eng. degree *summa cum laude* in electronic engineering in 1990, and the Ph.D. degree in computer science and electronic engineering in 1994, both from the University of Napoli Federico II.

From 1994 to 1998, he was an Assistant Professor with the Department of Information Engineering, Second University of Napoli. Since 1998 he has been with the Department of Biomedical, Electronic and Telecommunication Engineering, University of Napoli Federico II, first as an Associate Professor, and since November 2006 as a Full Professor of Telecommunications. He also held teaching positions at the University Parthenope of Napoli. His research interests are in the fields of statistical signal processing, array processing, image processing, and wireless communications, with current emphasis on code-division multiple-access systems and multicarrier modulation.

Luigi Paura (M'11) was born in Napoli, Italy, on February 20, 1950. He received the Dr. Eng. degree *summa cum laude* in electronic engineering in 1974 from the University of Napoli Federico II.

From 1979 to 1984 he was with the Department of Biomedical, Electronic and Telecommunication Engineering, University of Napoli, Italy, first as an Assistant Professor and then as an Associate Professor. Since 1994, he has been a Full Professor of Telecommunications: first, with the Department of Mathematics, University of Lecce, Italy; then, with the Department of Information Engineering, Second University of Napoli; and, finally, from 1998 he has been with the Department of Biomedical, Electronic and Telecommunication Engineering, University of Napoli Federico II. He also held teaching positions at the University of Salerno, at the University of Sannio, and at the University Parthenope of Napoli, In 1985-86 and 1991 he was a Visiting Researcher at the Signal and Image Processing Laboratory, University of California, Davis. At the present time, his research activities are mainly concerned with statistical signal processing, digital communication systems, and medium access control in wireless networks.

Francesco Verde (M'10) was born in Santa Maria Capua Vetere, Italy, on June 12, 1974. He received the Dr. Eng. degree *summa cum laude* in electronic engineering in 1998 from the Second University of Napoli, Italy, and the Ph.D. degree in information engineering in 2002, from the University of Napoli Federico II, Italy.

Since December 2002, he has been with the Department of Biomedical, Electronic and Telecommunication Engineering, University of Napoli Federico II, first as an Assistant Professor of Signal Theory and Mobile Communications and, since December 2011, as an Associate Professor of Telecommunications. He also held teaching positions at the Second University of Napoli. Since September 2010, he has served as Associate Editor for the IEEE Transactions on Signal Processing. His research activities lie in the broad area of statistical signal processing, digital communications, and communication systems. In particular, his current interests include cyclostationarity-based techniques for blind identification, equalization and interference suppression for narrowband modulation systems, code-division multiple-access systems, multicarrier modulation systems, and space-time processing for cooperative communications systems.