Minimum-Mean-Output-Energy Blind Adaptive Channel Shortening for Multicarrier SIMO Transceivers

Donatella Darsena, *Member, IEEE*, and Francesco Verde

***Abstract*—In this paper, we propose a blind adaptive channel-shortening method for designing finite-impulse response time-domain equalizers (TEQs) in single-input multiple-output systems employing multicarrier modulations. The proposed algorithm, which relies on a constrained minimization of the mean-output-energy at the TEQ output, does not require *a priori* knowledge of the channel impulse response or transmission of training sequences, and admits an effective and computationally efficient adaptive implementation. Moreover, the proposed TEQ is narrowband-interference resistant and its synthesis only requires an upper bound (rather than the exact knowledge) of the channel order. Numerical simulations are provided to illustrate the advantages of the proposed technique over a recently developed blind channel shortener.**

***Index Terms*—Blind time-domain equalizer (TEQ), minimum-mean-output-energy (MMOE) criterion, multicarrier (MC) systems, single-input multiple-output (SIMO) transceivers.**

I. Introduction

MULTICARRIER (MC) modulations offer a practical and viable solution to cope with frequency-selectivity both in wireline and wireless communication channels [1], [2]. Within the family of MC technologies, the most commonly used schemes are the discrete multitone (DMT), which is employed in wireline applications such as several digital subscriber line (xDSL) standards [3] and power line communications standards (HomePlug) [4], and orthogonal-frequency-division multiplexing (OFDM), which is adopted in various wireless standards such as IEEE 802.11a/g [5] and HIPERLAN2 [6], digital audio and video broadcast (DAB/DVB) [7], [8]. The increasing popularity of MC modulation schemes is largely due to the fact that they rely on a low-cost and high-performance equalization method. Specifically, a cyclic prefix (CP) of length L_{cp} is inserted at the beginning of each inverse discrete Fourier transform (IDFT) transmitted block. If L_{cp} is larger than or equal to the order L_h of the discrete-time channel, i.e., $L_{cp} \geq L_h$, removing the CP at the receiver converts the convolutive channel matrix to a circulant one. Since circulant matrices are diagonalized by the discrete Fourier transform (DFT), the transmitted data can simply be recovered by one-tap frequency-domain equalization (FEQ). The only price to pay for CP insertion is an inherent reduction of the transmission data rate, which turns out to be negligible if the number of subcarriers is significantly greater than L_{cp}. However, for highly time-dispersive channels, for which L_h takes on huge values, fulfillment of the condition $L_{cp} \geq L_h$ leads to a significant reduction of the transmission data rate, since the number of subcarriers cannot excessively be increased in practice. If the CP is not long enough, i.e., $L_{cp} < L_h$, after CP removal, the received signal is corrupted by both intercarrier and interblock interference (ICI and IBI), which dramatically worsen the system performance. To avoid such a performance degradation, a channel-shortening finite-impulse response (FIR) filter, commonly referred to as time-domain equalizer (TEQ), is introduced before the DFT, whose goal is to force an overall system (channel-plus-equalizer) impulse response of order $L_{\text{eff}} \leq L_{cp} < L_h$.

Traditional TEQ designs for MC systems are targeted at single-input single-output (SISO) transmitter and receiver (the so-called "transceiver"). In this case, since perfect or zero-forcing (ZF) channel shortening cannot be achieved by using a FIR filter, even in the absence of noise, the TEQ parameters are synthesized in order to obtain an overall system impulse response with *almost* all of its energy concentrated in a window of length $L_{\text{eff}} + 1$. Along this line, a cornucopia of optimization criteria have been proposed, such as the maximum shortening signal-to-noise ratio (SNR) [9]–[11], the maximum geometric SNR [12], the minimum mean-square error (MMSE) [13], the maximum bit rate [14], and the minimum IBI [15]. However, all the aforementioned TEQ designs require transmission of training sequences or *a priori* channel knowledge. Besides reducing the channel throughput, the use of training sequences requires knowledge at the receiver of the position of training symbols within the incoming data stream and, moreover, turns out to be problematic in broadcasting or multicasting applications. To avoid the transmission of training symbols, one approach consists of first estimating the channel impulse response by resorting to existing *blind* channel identification methods, e.g., [16]–[18], and then performing channel shortening by using the estimated channel. However, doing so, several channel parameters have to be estimated for highly time-dispersive channels and, consequently, from an

Manuscript received December 8, 2006; revised April 8, 2007. This work was supported in part by the Italian National project Wireless 8O2.16 Multi-antenna mEsh Networks (WOMEN) under Grant 2005093248. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Erchin Serpedin.

D. Darsena is with the Department of Technology, University of Napoli Parthenope, I-80133 Napoli, Italy (e-mail: darsena@unina.it).

F. Verde is with the Department of Electronic and Telecommunication Engineering, University of Napoli Federico II, I-80125 Napoli, Italy (e-mail: f.verde@unina.it).

Digital Object Identifier 10.1109/TSP.2007.900087

estimation viewpoint, large data records have to be used in order to obtain reliable channel estimates. An alternative blind approach, which will be referred to as *direct* one, is extracting the TEQ parameters by the received data, without performing any explicit channel identification. With reference to MC-SISO transceivers, such an approach has been pursued in [19] and [20] to synthesize blind adaptive channel-shortening methods,[1] which do not require matrix inversions, eigendecompositions, or Cholesky decompositions. However, although the algorithm [19] is simple to implement and is globally convergent, it needs a large number of data records to converge. On the other hand, although the method [20] converges much faster than [19], its global convergence is not ensured and its implementation is computationally intensive. Furthermore, both methods [19], [20] poorly perform in the presence of narrowband interference (NBI), which is one of the major source of performance degradation in wireless MC systems operating in overlay mode or in nonlicensed band, and in wireline ones, where transmission cables are exposed to crosstalk or radio-frequency interference.

A great deal of attention has recently been focused on channel-shortening algorithms for MC single-input multiple-output (SIMO) transceivers (see Section II for the system model), which arise either by employing multiple receiving antennas (as may be the case in wireless systems) or by oversampling the received signal. MC-SIMO transceivers offer [22] the important advantage of allowing ZF channel shortening by means of FIR filters (see Section III), i.e., in the absence of noise, the TEQ can be designed such as to concentrate *all* the energy of the overall system impulse response within a window of length $L_{\rm eff}+1$. A blind direct equalization method for MC-SIMO systems has originally been proposed in [23] and, later on, in [24]. These methods exploit frequency redundancy arising from insertion of virtual carriers in the transmitted signal, in order to force the overall system impulse response to be a single spike (i.e., $L_{\rm eff}=0$), and, thus, they are basically suited to MC systems that do not use a CP. However, for CP-based MC systems, reducing the overall system impulse response to an impulse (rather than performing channel shortening) is not preferable, since it might lead to excessive noise enhancement at the TEQ output. More recently, relying on a ZF design, a subspace-based direct channel-shortening algorithm has been proposed in [25] for MC-SIMO transceivers, which performs significantly better than [19], [20]. Despite its performance advantage, the blind method [25] is not suited to real-time implementation, since it is based on batch processing. Even worse, due to the fact that it involves[2] two expensive singular value decompositions (SVD) and one eigenvalue decomposition (EVD), the algorithm of [25] is characterized by a high computational complexity. Additionally, the approach of [25] requires knowledge or estimation of the channel order and, as confirmed by our simulation results, it leads to a data-estimated TEQ whose NBI suppression capability is limited.

[1]Recently, the algorithm [19] has been generalized in [21] to multiple-input multiple-output (MIMO) systems.

[2]Two blind channel-shortening methods are presented in [25], which are referred to as methods A and B. Hereinafter, we exclusively consider method A since, in noisy channels, it performs more reliably than method B.

To overcome the aforementioned limitations, we propose in Section IV a blind direct channel shortener for MC-SIMO systems, whose synthesis is based on the minimum-mean-output-energy (MMOE) criterion. The MMOE criterion, originally proposed in the array processing literature [26], has also proven fruitful in the area of single-carrier channel equalization [27], [28], as well as in the context of multiuser detection for direct-sequence code-division multiple access communications systems [29]–[31]. Specifically, we analytically show that joint FIR-ZF channel shortening and perfect NBI suppression can blindly be achieved in the absence of noise, by minimizing the mean-output-energy (MOE) at the TEQ output, subject to a set of channel-irrespective linear constraints. Results of our analysis also provide useful insights into how the constraints can blindly be optimized. As in [19] and [20], the synthesis of the proposed TEQ can adaptively be carried out with a manageable computational burden and, moreover, differently from [25], it does not need knowledge or estimation of the channel order. Section V provides some numerical results which, besides assessing the performance advantages of the proposed method over the channel shortener developed in [25], shows that our TEQ is also able to satisfactorily operate in the presence of severe NBI. Finally, concluding remarks are given in Section VI.

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., $\boldsymbol{A}$ or $\boldsymbol{a}$); 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 $\mathbb{C}^{m\times 1}$ $[\mathbb{R}^{m\times 1}]$; $\{\boldsymbol{A}\}_{ij}$ or $\boldsymbol{A}_{ij}$ indicates the $(i+1, j+1)$th element of matrix $\boldsymbol{A} \in \mathbb{C}^{m\times n}$, with $i \in \{0, 1, \ldots, m-1\}$ and $j \in \{0, 1, \ldots, n-1\}$; the superscripts $*$, T, H, -1, $-$ and $\dagger$ denote the conjugate, the transpose, the Hermitian (conjugate transpose), the inverse, the generalized (1)-inverse [32] and the Moore–Penrose generalized inverse [32] of a matrix, respectively; $\mathbf{0}_m \in \mathbb{R}^m$, $\mathbf{O}_{m\times n} \in \mathbb{R}^{m\times n}$ and $\mathbf{I}_m \in \mathbb{R}^{m\times m}$ denote the null vector, the null matrix, and the identity matrix, respectively; for any $\mathbf{a} \in \mathbb{C}^n$, $\|\mathbf{a}\|$ denotes the Euclidean norm; $\mathcal{N}(\boldsymbol{A})$, $\mathcal{R}(\boldsymbol{A})$, and $\mathcal{R}^{\perp}(\boldsymbol{A})$ denote the null space, the range (column space), and the orthogonal complement of the column space of $\boldsymbol{A} \in \mathbb{C}^{m\times n}[\mathbb{R}^{m\times n}]$ in $\mathbb{C}^m[\mathbb{R}^m]$; when applied to a vector $\boldsymbol{A} = \mathrm{diag}(\boldsymbol{a})$ is the diagonal matrix with $\boldsymbol{A}_{ii} = \boldsymbol{a}_i$, whereas when applied to a matrix $\boldsymbol{a} = \mathrm{diag}(\boldsymbol{A})$ is the vector with $\boldsymbol{a}_i = \boldsymbol{A}_{ii}$; $\min[x, y]$ is the smallest of the two real number x and y; the subscript c stands for continuous-time signals, $E[\cdot]$ denotes ensemble averaging and, finally, $\lceil\cdot\rceil$ and $j \triangleq \sqrt{-1}$ denote the ceiling integer and the imaginary unit, respectively.

II. MC Transceiver Model With Time- and Frequency-Domain Equalization

Let us consider (see Fig. 1) the baseband equivalent of a MC system with M subcarriers. At the transmitter, the information data stream $s(k)$, with $k \in \mathbb{Z}$, is converted into M parallel substreams $s_m(n) \triangleq s(nM+m)$, where $n \in \mathbb{Z}$ and the index $m \in \{0, 1, \ldots, M-1\}$ refers to the subcarrier. Subsequently, the

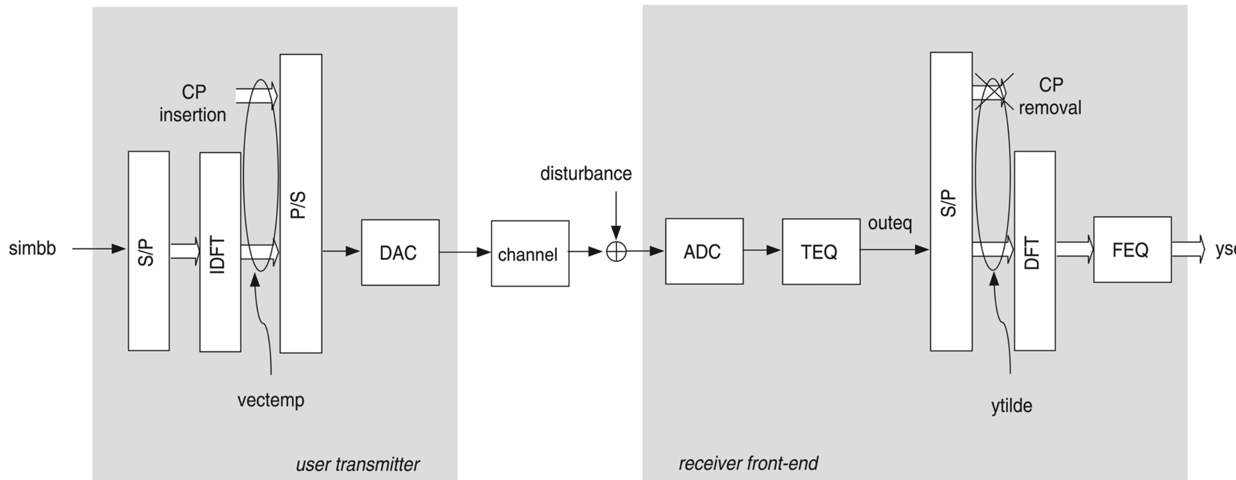


Fig. 1. MC transceiver with time- and frequency-domain equalization.

nth data block $\boldsymbol{s}(n) \triangleq [s_0(n), s_1(n), \ldots, s_{M-1}(n)]^T \in \mathbb{C}^M$ is first subject to the IDFT and, then, a CP of length $L_{cp} \ll M$ is inserted at the beginning of the resulting vector, obtaining the time-domain block $\boldsymbol{u}(n) \triangleq [u_0(n), u_1(n), \ldots, u_{P-1}(n)]^T \in \mathbb{C}^P$, with $P \triangleq M + L_{cp}$. This block can compactly be expressed as $\boldsymbol{u}(n) \triangleq \boldsymbol{T}_{cp}\boldsymbol{W}_{\mathrm{IDFT}}\boldsymbol{s}(n)$, where $\{\boldsymbol{W}_{\mathrm{IDFT}}\}_{i_1 i_2} \triangleq \left(1/\sqrt{M}\right) e^{j(2\pi/M)i_1 i_2}$, with $i_1, i_2 \in \{0, 1, \ldots, M-1\}$, represents the unitary symmetric IDFT matrix,[3] $\boldsymbol{T}_{cp} \triangleq \left[\mathbf{I}_{cp}^T, \mathbf{I}_M\right]^T \in \mathbb{R}^{P \times M}$ and $\mathbf{I}_{cp} \in \mathbb{R}^{L_{cp} \times M}$ is obtained from $\mathbf{I}_M$ by picking its last L_{cp} rows. Vector $\boldsymbol{u}(n)$ undergoes parallel-to-serial (P/S) conversion, and the resulting sequence $u(nP+p) = u_p(n)$, for $p \in \{0, 1, \ldots, P-1\}$, feeds a digital-to-analog converter (DAC), operating at rate $1/T_c = P/T$, where T_c and T denote the sampling and the symbol period, respectively. After up-conversion, the signal at the DAC output is transmitted over a multipath channel, which is modeled as a linear time-invariant (LTI) system. Assuming perfect frequency offset-compensation, the received signal at the output of the analog-to-digital converter (ADC) is

$$r_c(t) = \sum_{n=-\infty}^{+\infty} \sum_{p=0}^{P-1} u_p(n) h_c(t - pT_c - nT) + w_c(t) \tag{1}$$

where $h_c(t)$ denotes the impulse response of the *composite* channel (encompassing the cascade of the DAC filter, the physical channel, and the ADC antialiasing filter), which spans L_h sampling periods, that is, $h_c(t) \equiv 0$ for $t \notin [0, L_h T_c]$, whereas $w_c(t) = w_{c,\mathrm{nbi}}(t) + w_{c,\mathrm{noise}}(t)$ accounts for the disturbance (NBI plus thermal noise) at the output of the ADC filter. At the receiver, the signal $r_c(t)$ is first fractionally sampled at rate N/T_c, with $N > 1$ denoting the *oversampling* factor, and, then, the resulting sequence is subject to a polyphase decomposition [33] with respect to N, thus yielding the *phases*

$$r_q(k) \triangleq r_c(t_{k,q}) = \sum_{i=0}^{L_h} h_q(i) u(k-i) + w_q(k) \tag{2}$$

[3]Its inverse $\boldsymbol{W}_{\mathrm{DFT}} \triangleq \boldsymbol{W}_{\mathrm{IDFT}}^{-1} = \boldsymbol{W}_{\mathrm{IDFT}}^{*}$ defines the DFT matrix.

for $q \in \{0, 1, \ldots, N-1\}$, with $t_{k,q} \triangleq kT_c + q(T_c/N)$, where $h_q(k) \triangleq h_c(t_{k,q})$ and $w_q(k) \triangleq w_c(t_{k,q})$ are the channel and disturbance phases, respectively. By collecting the different phases $\{r_q(k)\}_{q=0}^{N-1}$, $\{h_q(k)\}_{q=0}^{N-1}$, and $\{w_q(k)\}_{q=0}^{N-1}$ in the vectors $\boldsymbol{r}(k) \triangleq [r_0(k), r_1(k), \ldots, r_{N-1}(k)]^T \in \mathbb{C}^N$, $\boldsymbol{h}(k) \triangleq [h_0(k), h_1(k), \ldots, h_{N-1}(k)]^T \in \mathbb{C}^N$, and $\boldsymbol{w}(k) \triangleq [w_0(k), w_1(k), \ldots, w_{N-1}(k)]^T \in \mathbb{C}^N$, one obtains the SIMO representation

$$\boldsymbol{r}(k) = \sum_{i=0}^{L_h} \boldsymbol{h}(i) u(k-i) + \boldsymbol{w}(k). \tag{3}$$

It is noteworthy that, in wireless applications, a SIMO model equivalent to (3) can also be obtained by sampling, with rate $1/T_c$, the signal received by N multiple antennas.

When the CP length is insufficient, i.e., $L_{cp} < L_h$, a TEQ can be employed to suitably shorten the channel impulse response (see Fig. 1). Denoting with L_e the equalizer order (expressed in sampling intervals), the input-output relationship of a FIR TEQ can be expressed as $\widetilde{y}(k) = \boldsymbol{f}^H \boldsymbol{z}(k)$, where $\boldsymbol{f} \in \mathbb{C}^{N(L_e+1)}$ is the weight vector to be designed according to a specified optimization criterion and, accounting for (3), the input vector $\boldsymbol{z}(k) \triangleq \left[\boldsymbol{r}^T(k), \boldsymbol{r}^T(k-1), \ldots, \boldsymbol{r}^T(k-L_e)\right]^T \in \mathbb{C}^{N(L_e+1)}$ assumes the form

$$\boldsymbol{z}(k) = \boldsymbol{H}\, \widetilde{\boldsymbol{u}}(k) + \boldsymbol{v}(k) \tag{4}$$

where, by setting $L_g \triangleq L_e + L_h$

$$\boldsymbol{H} \triangleq \begin{bmatrix} \boldsymbol{h}(0) & \boldsymbol{h}(1) & \ldots & \boldsymbol{h}(L_h) & \mathbf{0}_N & \ldots & \mathbf{0}_N \\ \mathbf{0}_N & \boldsymbol{h}(0) & \boldsymbol{h}(1) & \ldots & \boldsymbol{h}(L_h) & \ddots & \mathbf{0}_N \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \mathbf{0}_N \\ \mathbf{0}_N & \ldots & \ddots & \boldsymbol{h}(0) & \boldsymbol{h}(1) & \ldots & \boldsymbol{h}(L_h) \end{bmatrix} \tag{5}$$

is the $N(L_e+1) \times (L_g+1)$ block Toeplitz channel matrix, whereas

$$\widetilde{\boldsymbol{u}}(k) \triangleq [u(k), u(k-1), \ldots, u(k-L_g)]^T \in \mathbb{C}^{L_g+1} \tag{6}$$

$$\boldsymbol{v}(k) \triangleq \left[\boldsymbol{w}^T(k), \boldsymbol{w}^T(k-1), \ldots, \boldsymbol{w}^T(k-L_e)\right]^T \in \mathbb{C}^{N(L_e+1)}. \tag{7}$$

By virtue of (4), the TEQ output can be written as

$$\widetilde{y}(k) = \boldsymbol{g}^H \widetilde{\boldsymbol{u}}(k) + \widetilde{d}(k) \tag{8}$$

where

$$\boldsymbol{g} = [g(0), g(1), \ldots, g(L_g)]^T \triangleq \boldsymbol{H}^H \boldsymbol{f} \in \mathbb{C}^{L_g+1} \tag{9}$$

collects the samples of the *combined channel-TEQ impulse response*, whose order is L_g, and $\widetilde{d}(k) \triangleq \boldsymbol{f}^H \boldsymbol{v}(k)$ is the disturbance contribution. In principle, the weight vector $\boldsymbol{f}$ has to be chosen so as to nullify all the entries of $\boldsymbol{g}$, except for a preselected part of $L_{\text{eff}}+1$ consecutive samples, thus obtaining a *target* combined channel-TEQ response vector having the form

$$\boldsymbol{g}_{\text{target}} \triangleq [0, \ldots, 0, g_{\text{target}}(\Delta), g_{\text{target}}(\Delta+1), \ldots \\ g_{\text{target}}(\Delta + L_{\text{eff}}), 0, \ldots, 0]^T \tag{10}$$

where $0 \leq \Delta \leq L_g - L_{\text{eff}}$ is a suitable *shortening delay* at the designer's disposal, and L_{eff}, which represents the effective order of the combined channel-TEQ impulse response, obeys $L_{\text{eff}} \leq L_{cp} < L_h$ in order to achieve perfect IBI cancellation through CP removal. For the problem at hand, a reasonable performance measure is the shortening signal-to-noise-plus-interference ratio (SSINR) at the TEQ output $\widetilde{y}(k)$, which is defined as

$$\text{SSINR}(\boldsymbol{f}) \triangleq \frac{E\left[\left|\boldsymbol{g}_{\text{win}}^H \widetilde{\boldsymbol{u}}_{\text{win}}(k)\right|^2\right]}{E\left[\left|\boldsymbol{g}_{\text{wall}}^H \widetilde{\boldsymbol{u}}_{\text{wall}}(k) + \widetilde{d}(k)\right|^2\right]} \tag{11}$$

where[4]

$$\boldsymbol{g}_{\text{win}} \triangleq [g(\Delta), g(\Delta+1), \ldots, g(\Delta+L_{\text{eff}})]^T \in \mathbb{C}^{L_{\text{eff}}+1} \tag{12}$$

$$\boldsymbol{g}_{\text{wall}} \triangleq [g(0), \ldots, g(\Delta-1), g(\Delta+L_{\text{eff}}+1) \\ \ldots, g(L_g)]^T \in \mathbb{C}^{L_g - L_{\text{eff}}} \tag{13}$$

$$\widetilde{\boldsymbol{u}}_{\text{win}}(k) \triangleq [u(k-\Delta), u(k-\Delta-1), \ldots \\ u(k-\Delta-L_{\text{eff}})]^T \in \mathbb{C}^{L_{\text{eff}}+1}, \tag{14}$$

$$\widetilde{\boldsymbol{u}}_{\text{wall}}(k) \triangleq [u(k), \ldots, u(k-\Delta+1) \\ u(k-\Delta-L_{\text{eff}}-1), \ldots, u(k-L_g)]^T \\ \in \mathbb{C}^{L_g - L_{\text{eff}}}. \tag{15}$$

After performing channel shortening in the time domain, the resulting signal is equalized in the frequency domain. Specifically, a time-advanced version of $\widetilde{y}(k)$ at the TEQ output is converted into P parallel substreams $\widetilde{y}_p(n) \triangleq \widetilde{y}(nP+p+\Delta)$, for $p \in \{0, 1, \ldots, P-1\}$. Provided that $L_{\text{eff}} \leq L_{cp} < L_h$, after gathering the samples of $\{\widetilde{y}_p(n)\}_{p=0}^{P-1}$ into the vector $\widetilde{\boldsymbol{y}}(n) \triangleq [\widetilde{y}_0(n), \widetilde{y}_1(n), \ldots, \widetilde{y}_{P-1}(n)]^T \in \mathbb{C}^P$ and removing the CP, one obtains [2] the IBI-free model

$$\boldsymbol{y}(n) \triangleq \boldsymbol{R}_{cp} \widetilde{\boldsymbol{y}}(n) = \boldsymbol{G}\boldsymbol{T}_{cp} \boldsymbol{W}_{\text{IDFT}} \boldsymbol{s}(n) + \boldsymbol{d}(n) \\ = \boldsymbol{W}_{\text{IDFT}} \boldsymbol{\mathcal{G}} \boldsymbol{s}(n) + \boldsymbol{d}(n) \tag{16}$$

where the matrix $\boldsymbol{R}_{cp} \triangleq [\mathbf{O}_{M \times L_{cp}}, \mathbf{I}_M] \in \mathbb{R}^{M \times P}$ discards the first L_{cp} entries of $\widetilde{\boldsymbol{y}}(n)$, $\boldsymbol{G} \in \mathbb{C}^{M \times P}$ is the Toeplitz channel matrix, whose first column and row are given by $[g_{\text{target}}^*(\Delta + L_{\text{eff}}), 0, \ldots, 0]^T$ and $[g_{\text{target}}^*(\Delta + L_{\text{eff}}), \ldots, g_{\text{target}}^*(\Delta), 0, \ldots, 0]$, respectively, and $\boldsymbol{d}(n) \triangleq \boldsymbol{R}_{cp} \widetilde{\boldsymbol{d}}(n)$, with the disturbance vector $\widetilde{\boldsymbol{d}}(n) \triangleq [\widetilde{d}_0(n), \widetilde{d}_1(n), \ldots, \widetilde{d}_{P-1}(n)]^T \in \mathbb{C}^P$ collecting $\widetilde{d}_p(n) \triangleq \widetilde{d}(nP+p+\Delta)$. Moreover, the last equality in (16) comes from the fact that, since $\boldsymbol{G}\boldsymbol{T}_{cp}$ is a circulant matrix [34], by using standard eigenstructure concepts, it can equivalently be expressed [2] as $\boldsymbol{G}\boldsymbol{T}_{cp} = \boldsymbol{W}_{\text{IDFT}} \boldsymbol{\mathcal{G}} \boldsymbol{W}_{\text{DFT}}$, where the diagonal entries of $\boldsymbol{\mathcal{G}} \triangleq \text{diag}[G(z_0), G(z_1), \ldots, G(z_{M-1})] \in \mathbb{C}^{M \times M}$ are the values of the (conjugate) combined channel-TEQ transfer function $G(z) \triangleq \sum_{\ell=0}^{L_{\text{eff}}} g_{\text{target}}^*(\Delta+\ell) z^{-\ell}$ evaluated at the subcarriers $z_m \triangleq e^{j(2\pi/M)m}$. If $G(z)$ has no zero located at $\{z_m\}_{m=0}^{M-1}$, the matrix $\boldsymbol{\mathcal{G}}$ is nonsingular, and, thus, perfect symbol recovery in the absence of disturbance can be obtained by resorting to the conventional receiver, which performs (in the given order) DFT and FEQ, thus yielding (see Fig. 1) the ultimate output $\boldsymbol{x}(n) = \boldsymbol{\mathcal{G}}^{-1} \boldsymbol{W}_{\text{DFT}} \boldsymbol{y}(n)$. After FEQ, the mth entry of $\boldsymbol{x}(n)$ is quantized to the nearest (in Euclidean distance) information symbol to form the estimate of the symbol $s_m(n)$ belonging to the mth data substream. In the sequel, the following customary assumptions are considered: **(a1)** the information symbols $s(k)$ are modeled as a sequence of zero-mean independent and identically distributed (i.i.d.) circular random variables, with variance $\sigma_s^2 \triangleq E[|s(k)|^2]$; **(a2)** the disturbance vector $\boldsymbol{v}(k)$ in (4) is modeled as a zero-mean complex circular wide-sense stationary (WSS) random vector, statistically independent of $s(k)$, with autocorrelation matrix $\boldsymbol{R}_{\boldsymbol{vv}} \triangleq E[\boldsymbol{v}(k)\boldsymbol{v}^H(k)] \in \mathbb{C}^{N(L_e+1) \times N(L_e+1)}$.

III. ZF and Maximum-SSINR Channel Shortening

Before getting deeply into the description of the proposed channel-shortening approach, we briefly review two conventional *nonblind* TEQ designs, which serve as useful clairvoyant benchmarks to evaluate the performance of any blind channel-shortening method.

In the absence of disturbance, i.e., when $\boldsymbol{v}(k) = \mathbf{0}_{N(L_e+1)}$, the optimum channel-shortening approach consists of solving the system of linear equations $\boldsymbol{H}^H \boldsymbol{f} = \boldsymbol{g}_{\text{target}}$, which allows one to perform perfect or ZF channel shortening [22]. Mathematically, this system is consistent [32] if and only if (iff) $\boldsymbol{H}^H (\boldsymbol{H}^H)^- \boldsymbol{g}_{\text{target}} = \boldsymbol{g}_{\text{target}}$. However, when $\boldsymbol{v}(k) \neq \mathbf{0}_{N(L_e+1)}$, to avoid excessive disturbance enhancement at the TEQ output, one has to impose the additional constraint that $\|\boldsymbol{f}\|$ be minimal. The minimal-norm least-squares solution of $\boldsymbol{H}^H \boldsymbol{f} = \boldsymbol{g}_{\text{target}}$ is given [32] by $(\boldsymbol{H}^H)^\dagger \boldsymbol{g}_{\text{target}}$, which leads to perfect channel shortening in the absence of disturbance if the channel matrix (5) is full-column rank, i.e., $\text{rank}(\boldsymbol{H}) = L_g + 1$, in which case, one obtains $\boldsymbol{f}_{zf} = \boldsymbol{H}(\boldsymbol{H}^H \boldsymbol{H})^{-1} \boldsymbol{g}_{\text{target}}$. It is interesting to observe that, if $\boldsymbol{H}$ is full-column rank, one has $\boldsymbol{H}^H (\boldsymbol{H}^H)^- = \mathbf{I}_{L_g+1}$ and, then, the system $\boldsymbol{H}^H \boldsymbol{f} = \boldsymbol{g}_{\text{target}}$ turns out to be consistent independently of $\boldsymbol{g}_{\text{target}}$. Apparently, the synthesis of $\boldsymbol{f}_{zf}$ requires *a priori* knowledge or estimation of the channel matrix $\boldsymbol{H}$.

[4] For the sake of notation simplicity, we do not explicitly indicate the dependence of $\boldsymbol{g}_{\text{target}}$, $\boldsymbol{g}_{\text{win}}$, and $\boldsymbol{g}_{\text{wall}}$ on $\boldsymbol{f}$.

In the presence of disturbance, the ZF-based TEQ perfectly shortens the channel at the price of disturbance enhancement. To better counteract the disturbance, one can choose $\boldsymbol{f}$ in order to maximize the output SSINR which, accounting for (11) and assumptions (a1) and (a2), can be written as

$$\mathrm{SSINR}(\boldsymbol{f}) = \frac{\boldsymbol{f}^H \left(\boldsymbol{H}_{\mathrm{win}} \widetilde{\boldsymbol{R}}_{\mathrm{win}} \boldsymbol{H}_{\mathrm{win}}^H\right) \boldsymbol{f}}{\boldsymbol{f}^H \left(\boldsymbol{H}_{\mathrm{wall}} \widetilde{\boldsymbol{R}}_{\mathrm{wall}} \boldsymbol{H}_{\mathrm{wall}}^H + \boldsymbol{R}_{\boldsymbol{vv}}\right) \boldsymbol{f}} \tag{17}$$

where

$$\boldsymbol{H}_{\mathrm{win}} \triangleq [\boldsymbol{h}_\Delta, \boldsymbol{h}_{\Delta+1}, \ldots, \boldsymbol{h}_{\Delta+L_{\mathrm{eff}}}] \in \mathbb{C}^{N(L_e+1)\times(L_{\mathrm{eff}}+1)} \tag{18}$$

$$\widetilde{\boldsymbol{R}}_{\mathrm{win}} \triangleq E[\widetilde{\boldsymbol{u}}_{\mathrm{win}} \widetilde{\boldsymbol{u}}_{\mathrm{win}}^H] \in \mathbb{C}^{(L_{\mathrm{eff}}+1)\times(L_{\mathrm{eff}}+1)} \tag{19}$$

$$\boldsymbol{H}_{\mathrm{wall}} \triangleq [\boldsymbol{h}_0, \ldots, \boldsymbol{h}_{\Delta-1}, \boldsymbol{h}_{\Delta+L_{\mathrm{eff}}+1}, \ldots, \boldsymbol{h}_{L_g}] \in \mathbb{C}^{N(L_e+1)\times(L_g-L_{\mathrm{eff}})} \tag{20}$$

$$\widetilde{\boldsymbol{R}}_{\mathrm{wall}} \triangleq E[\widetilde{\boldsymbol{u}}_{\mathrm{wall}} \widetilde{\boldsymbol{u}}_{\mathrm{wall}}^H] \in \mathbb{C}^{(L_g-L_{\mathrm{eff}})\times(L_g-L_{\mathrm{eff}})}. \tag{21}$$

It should be observed that (17) is different from the shortening SNR (SSNR) considered in [9]. Specifically, the SSINR given by (17) boils down to the SSNR defined in [9] if: 1) there is no disturbance, i.e., $\boldsymbol{v}(k) = \boldsymbol{0}_{N(L_e+1)}$; 2) the vector $\widetilde{\boldsymbol{u}}(k)$ is white,[5] which implies $\widetilde{\boldsymbol{R}}_{\mathrm{win}} = \sigma_s^2 \mathbf{I}_{L_{\mathrm{eff}}+1}$ and $\widetilde{\boldsymbol{R}}_{\mathrm{wall}} = \sigma_s^2 \mathbf{I}_{L_g-L_{\mathrm{eff}}}$. The vector $\boldsymbol{f}_{\max}$ maximizing (17) is given [35] by $\boldsymbol{f}_{\max} = \varrho \boldsymbol{a}_{\max}$, where $\varrho \in \mathbb{C}$ and $\boldsymbol{a}_{\max}$ is the generalized eigenvector corresponding to the largest generalized eigenvalue $\alpha_{\max}$ of $\boldsymbol{H}_{\mathrm{win}} \widetilde{\boldsymbol{R}}_{\mathrm{win}} \boldsymbol{H}_{\mathrm{win}}^H$ and $\boldsymbol{H}_{\mathrm{wall}} \widetilde{\boldsymbol{R}}_{\mathrm{wall}} \boldsymbol{H}_{\mathrm{wall}}^H + \boldsymbol{R}_{\boldsymbol{vv}}$, i.e.,

$$(\boldsymbol{H}_{\mathrm{win}} \widetilde{\boldsymbol{R}}_{\mathrm{win}} \boldsymbol{H}_{\mathrm{win}}^H) \boldsymbol{a}_{\max} = \alpha_{\max} (\boldsymbol{H}_{\mathrm{wall}} \widetilde{\boldsymbol{R}}_{\mathrm{wall}} \boldsymbol{H}_{\mathrm{wall}}^H + \boldsymbol{R}_{\boldsymbol{vv}}) \boldsymbol{a}_{\max}. \tag{22}$$

Let [see (10)]

$$\boldsymbol{g}_{\mathrm{target,win}} \triangleq [g_{\mathrm{target}}(\Delta), g_{\mathrm{target}}(\Delta+1), \ldots g_{\mathrm{target}}(\Delta+L_{\mathrm{eff}})]^T \in \mathbb{C}^{(L_{\mathrm{eff}}+1)} \tag{23}$$

it can be shown [13] that maximizing (17) is equivalent to minimizing the mean-square error

$$\mathrm{MSE}(\boldsymbol{f}, \boldsymbol{g}_{\mathrm{target,win}}) \triangleq E\left[\left|\widetilde{y}(k) - \boldsymbol{g}_{\mathrm{target,win}}^H \widetilde{\boldsymbol{u}}_{\mathrm{win}}(k)\right|^2\right] \tag{24}$$

with respect to both $\boldsymbol{f}$ and $\boldsymbol{g}_{\mathrm{target,win}}$, subject to $\boldsymbol{g}_{\mathrm{target,win}}^H \widetilde{\boldsymbol{R}}_{\mathrm{win}} \boldsymbol{g}_{\mathrm{target,win}} = 1$. It is worth observing that $\boldsymbol{f}_{\max}$ ends up to the ZF solution $\boldsymbol{f}_{zf}$, as the mean power $E[\|\boldsymbol{v}(k)\|^2]$ of the disturbance approaches to zero.

Some comments are now in order about the assumption $\mathrm{rank}(\boldsymbol{H}) = L_g + 1$. The matrix $\boldsymbol{H}$ turns out to be full-column rank iff the following three conditions [36] hold: **(c1)** $N > 1$; **(c2)** $L_e \geq L_h - 1$; **(c3)** the $\mathcal{Z}$-transforms of the channel phases $h_q(k)$, for $q \in \{0, 1, \ldots, N-1\}$, have no common zeros. As already pointed out in [22], condition (c1) states that, even in the absence of disturbance, FIR TEQs cannot guarantee perfect channel shortening for MC-SISO transceivers (i.e., when $N = 1$). Condition (c2) infers the well-known fact that the equalizer length $L_e + 1$ must be greater than the channel order L_h. Finally, condition (c3) assures the existence of enough temporal (or spatial) diversity produced by fractionally oversampling (or by using multiple antennas) at the receiver. It should be observed that, with reference to diversity arising from oversampling, condition (c3) cannot be fulfilled in all those situations where the excess bandwidth is limited [37].

[5]Under assumption (a1), the vector $\widetilde{\boldsymbol{u}}(k)$ is white iff the TEQ order L_e satisfies [25] the inequality $L_e \leq M - L_h - 1$.

IV. Proposed MMOE-Based Blind Channel-Shortening Approach

A blind ZF channel shortener was developed in [25], by exploiting the eigenstructure of the correlation matrices $E[\boldsymbol{z}(k)\boldsymbol{z}(k+i)]$ of the received vector $\boldsymbol{z}(k)$, with $i \in \{0, \Delta, \Delta + L_{\mathrm{eff}} + 1\}$. Besides exhibiting a high computational complexity cubic in $N(L_e+1)$ and requiring that conditions (c1)–(c3) be fulfilled, this method is essentially designed to work in the absence of disturbance and requires estimation of the channel order. Herein, we show that, by pursuing a completely different approach, blind NBI-resistant channel shortening can adaptively be carried out with a computational complexity per iteration quadratic in $N(L_e - L_{cp})$, without requiring channel-order estimation. To obtain a combined channel-TEQ response vector as in (10), we rely on minimization of the TEQ mean-output-energy $\mathrm{MOE}(\boldsymbol{f}) \triangleq E\left[|\widetilde{y}(k)|^2\right] = \boldsymbol{f}^H \boldsymbol{R}_{\boldsymbol{zz}} \boldsymbol{f}$, where $\boldsymbol{R}_{\boldsymbol{zz}} \triangleq E\left[\boldsymbol{z}(k)\boldsymbol{z}^H(k)\right] \in \mathbb{C}^{N(L_e+1)\times N(L_e+1)}$ denotes the autocorrelation matrix of $\boldsymbol{z}(k)$, subject to appropriate linear blind constraints. These constraints have the effect of preserving (a delayed version of) the transmitted information-bearing sequence $u(k)$, without requiring knowledge or estimation of the channel impulse response to be shortened. To this aim, we observe that the TEQ output can be explicitly written as

$$\widetilde{y}(k) = \sum_{d=0}^{L_g} \underbrace{\boldsymbol{f}^H \boldsymbol{h}_d}_{g^*(d)} u(k-d) + \widetilde{d}(k) = \sum_{d=0}^{L_g} g^*(d) u(k-d) + \widetilde{d}(k) \tag{25}$$

where $\boldsymbol{h}_d$ denotes the $(d+1)$th column of the channel matrix $\boldsymbol{H}$. Interestingly, it has been shown in [28] that, by exploiting the block Toeplitz structure of $\boldsymbol{H}$ arising from oversampling of the received signal, each column vector $\boldsymbol{h}_d$ can be linearly parameterized, $\forall d \in \{0, 1, \ldots, L_g\}$, as follows:

$$\boldsymbol{h}_d = \boldsymbol{\Theta}_d \boldsymbol{\xi}_d, \tag{26}$$

with $\boldsymbol{\Theta}_d \in \mathbb{R}^{N(L_e+1)\times N(L_d+1)}$ and $\boldsymbol{\xi}_d \in \mathbb{C}^{N(L_d+1)}$, where, depending on the relative values of d, L_e and L_h, the following cases hold.

Case $P_1 (0 \leq d \leq \min[L_e, L_h])$: In this case

$$\boldsymbol{\Theta}_d = \left[\mathbf{I}_{N(d+1)\times N(d+1)}, \mathbf{O}_{N(L_e-d)\times N(d+1)}^T\right]^T \tag{27}$$

and

$$\boldsymbol{\xi}_d = \left[\boldsymbol{h}^T(d), \boldsymbol{h}^T(d-1), \ldots, \boldsymbol{h}^T(0)\right]^T \tag{28}$$

with $L_d = d$.

Case $P_2 (\min[L_e, L_h] < d < \max[L_e, L_h])$: In this case, it can be verified that there is no parameterization matrix of full-column rank when $L_h > L_e$; therefore, we focus attention only on the case $L_h \leq L_e$, wherein

$$\boldsymbol{\Theta}_d = \Big[\mathbf{O}^T_{N(d-L_h)\times N(L_h+1)}, \mathbf{I}_{N(L_h+1)\times N(L_h+1)}, \mathbf{O}^T_{N(L_e-d)\times N(L_h+1)}\Big]^T \quad (29)$$

and

$$\boldsymbol{\xi}_d = \Big[\boldsymbol{h}^T(L_h), \boldsymbol{h}^T(L_h-1), \ldots, \boldsymbol{h}^T(0)\Big]^T \quad (30)$$

with $L_d = L_h$.

Case $P_3(\max[L_e, L_h] \leq d \leq L_g)$: One has

$$\boldsymbol{\Theta}_d = \Big[\mathbf{O}^T_{N(d-L_h)\times N(L_g-d+1)}, \boldsymbol{I}_{N(L_g-d+1)\times N(L_g-d+1)}\Big]^T \quad (31)$$

and

$$\boldsymbol{\xi}_d = \Big[\boldsymbol{h}^T(L_h), \boldsymbol{h}^T(L_h-1), \ldots, \boldsymbol{h}^T(d-L_e)\Big]^T \quad (32)$$

with $L_d = L_g - d$.

Observe that, in all of the three aforementioned cases, $\boldsymbol{\Theta}_d$ is a *known* full-column rank parameterization matrix satisfying $\boldsymbol{\Theta}_d^T \boldsymbol{\Theta}_d = \mathbf{I}_{N(L_d+1)}$, whose column dimension $N(L_d+1)$ depends on the particular parameterization used, whereas $\boldsymbol{\xi}_d$ is an *unknown* vector collecting some samples of the channel impulse response. By virtue of (26), the TEQ output can be re-expressed as

$$\widetilde{y}(k) = \boldsymbol{f}^H \boldsymbol{\Theta}_\delta \boldsymbol{\xi}_\delta u(k-\delta) + \sum_{\substack{d=0 \\ d\neq\delta}}^{L_g} \boldsymbol{f}^H \boldsymbol{\Theta}_d \boldsymbol{\xi}_d u(k-d) + \widetilde{d}(k) \quad (33)$$

where $\delta \in \{0, \ldots, L_g\}$ represents a design parameter. As will be clear in the sequel (see, in particular, Theorem 1), the value of the shortening delay Δ strictly depends on the choice of δ. Equation (33) shows that, to blindly preserve the signal contribution $u(k-\delta)$ while minimizing the object function $\mathrm{MOE}(\boldsymbol{f})$, one can impose the set of $N(L_\delta+1)$ linear constraints $\boldsymbol{\Theta}_\delta^T \boldsymbol{f} = \boldsymbol{\gamma}_\delta$, where $\boldsymbol{\gamma}_\delta \in \mathbb{C}^{N(L_\delta+1)}$ is *the vector of constraint values*, whose choice will be discussed in Section IV-B. Doing so, the sequence $u(k-\delta)$ is passed to the output of the TEQ with gain $\boldsymbol{\gamma}_\delta^H \boldsymbol{\xi}_\delta$. As a first remark, it is worth noting that unit-norm or unit-tap constraints on $\boldsymbol{f}$, which are commonly employed [14] in channel-shortening optimization criteria to avoid the trivial all-zero solution, cannot be used in this case, since they do not ensure that $u(k-\delta)$ is passed to the TEQ output with a specified response. Moreover, observe that, although the constraint $\boldsymbol{\Theta}_\delta^T \boldsymbol{f} = \boldsymbol{\gamma}_\delta$ is designed to preserve the signal contribution associated with the $(\delta+1)$th column of $\boldsymbol{H}$, more generally, it has the effect of preserving the signal contributions associated with *all* those columns of $\boldsymbol{H}$ belonging to the column space of $\boldsymbol{\Theta}_\delta$. Indeed, if $\boldsymbol{h}_d \in \mathcal{R}(\boldsymbol{\Theta}_\delta)$, for some $d \in \{0, 1, \ldots, L_g\} - \{\delta\}$, there exists a nonzero vector $\boldsymbol{\alpha}_d \in \mathbb{C}^{N(L_\delta+1)}$ such that $\boldsymbol{h}_d = \boldsymbol{\Theta}_\delta \boldsymbol{\alpha}_d$ and, consequently, the signal contribution $u(k-d)$ is passed to the TEQ output with gain $\boldsymbol{f}^H \boldsymbol{h}_d = \boldsymbol{f}^H \boldsymbol{\Theta}_\delta \boldsymbol{\alpha}_d = \boldsymbol{\gamma}_\delta^H \boldsymbol{\alpha}_d$. On the other hand, if none of the columns $\boldsymbol{h}_d$ of $\boldsymbol{H}$ belongs to $\mathcal{R}(\boldsymbol{\Theta}_\delta)$, for $d \in \{0, 1, \ldots, L_g\} - \{\delta\}$, then the imposed constraint preserves only the sequence $u(k-\delta)$ and attempts to suppress all the remaining signal contributions: in this case, the TEQ forces the overall system impulse response to be a single spike (i.e., $L_{\text{eff}} = 0$). However, it can be seen by direct inspection (see also Appendix I) that, for certain values of δ, besides the vector $\boldsymbol{h}_\delta$, some other columns of $\boldsymbol{H}$ also belong to $\mathcal{R}(\boldsymbol{\Theta}_\delta)$, and, thus, the constraint $\boldsymbol{\Theta}_\delta^T \boldsymbol{f} = \boldsymbol{\gamma}_\delta$ preserves more than one signal contribution: in this case, the TEQ forces the overall system impulse response to have a nonzero order L_{eff}, where L_{eff} is equal to the number of preserved signal contributions. These intuitive findings are analytically corroborated by Theorem 1. Therefore, the TEQ can be synthesized accordingly to the following blind optimization criterion

$$\boldsymbol{f}_{\delta,\text{mmoe}} = \arg \min_{\boldsymbol{f} \in \mathbb{C}^{N(L_e+1)}} \boldsymbol{f}^H \boldsymbol{R}_{\boldsymbol{zz}} \boldsymbol{f} \quad \text{subject to } \boldsymbol{\Theta}_\delta^T \boldsymbol{f} = \boldsymbol{\gamma}_\delta \quad (34)$$

whose solution is given (see, e.g., [26]) by

$$\boldsymbol{f}_{\delta,\text{mmoe}} = \boldsymbol{F}_{\delta,\text{mmoe}} \boldsymbol{\gamma}_\delta \quad (35)$$

with

$$\boldsymbol{F}_{\delta,\text{mmoe}} \triangleq \boldsymbol{R}_{\boldsymbol{zz}}^{-1} \boldsymbol{\Theta}_\delta \left(\boldsymbol{\Theta}_\delta^T \boldsymbol{R}_{\boldsymbol{zz}}^{-1} \boldsymbol{\Theta}_\delta\right)^{-1} \in \mathbb{C}^{N(L_e+1)\times N(L_\delta+1)}. \quad (36)$$

It is interesting to point out that solution $\boldsymbol{f}_{\delta,\text{mmoe}}$ has the same mathematical form as the linearly constrained minimum power (LCMP) beamformer [26], which is widely considered in the framework of array processing. Such an equivalence allows one to decompose the weight vector $\boldsymbol{f}_{\delta,\text{mmoe}}$, according to the generalized sidelobe canceler (GSC) scheme [26], into two orthogonal components as

$$\boldsymbol{f}_{\delta,\text{mmoe}} = \boldsymbol{f}^{(0)}_{\delta,\text{mmoe}} - \boldsymbol{\Pi}_\delta \boldsymbol{f}^{(a)}_{\delta,\text{mmoe}} \quad (37)$$

where the fixed *quiescent* vector $\boldsymbol{f}^{(0)}_{\delta,\text{mmoe}} = \boldsymbol{\Theta}_\delta \left(\boldsymbol{\Theta}_\delta^T \boldsymbol{\Theta}_\delta\right)^{-1} \boldsymbol{\gamma}_\delta = \boldsymbol{\Theta}_\delta \boldsymbol{\gamma}_\delta$ belongs to the constraint subspace $\mathcal{R}(\boldsymbol{\Theta}_\delta)$, the *signal blocking* matrix[6] $\boldsymbol{\Pi}_\delta \in \mathbb{R}^{N(L_e+1)\times N(L_e-L_\delta)}$ is chosen so that its columns form a basis for the null space of $\boldsymbol{\Theta}_\delta^T$, i.e., $\mathcal{R}(\boldsymbol{\Pi}_\delta) \equiv \mathcal{R}^\perp(\boldsymbol{\Theta}_\delta)$, and the optimal solution of the data-dependent component $\boldsymbol{f}^{(a)}_{\delta,\text{mmoe}}$ is given by

$$\boldsymbol{f}^{(a)}_{\delta,\text{mmoe}} = \left(\boldsymbol{\Pi}_\delta^T \boldsymbol{R}_{\boldsymbol{zz}} \boldsymbol{\Pi}_\delta\right)^{-1} \boldsymbol{\Pi}_\delta^T \boldsymbol{R}_{\boldsymbol{zz}} \boldsymbol{f}^{(0)}_{\delta,\text{mmoe}}. \quad (38)$$

As a direct consequence of the GSC decomposition given by (37), it can easily be verified by direct inspection that $\boldsymbol{F}_{\delta,\text{mmoe}}$ defined in (35) can be equivalently decomposed as

$$\begin{aligned}\boldsymbol{F}_{\delta,\text{mmoe}} &= \boldsymbol{\Theta}_\delta - \boldsymbol{\Pi}_\delta \underbrace{\left(\boldsymbol{\Pi}_\delta^T \boldsymbol{R}_{\boldsymbol{zz}} \boldsymbol{\Pi}_\delta\right)^{-1} \boldsymbol{\Pi}_\delta^T \boldsymbol{R}_{\boldsymbol{zz}} \boldsymbol{\Theta}_\delta}_{\boldsymbol{F}^{(a)}_{\delta,\text{mmoe}} \in \mathbb{C}^{N(L_e-L_\delta)\times N(L_\delta+1)}} \\ &= \boldsymbol{\Theta}_\delta - \boldsymbol{\Pi}_\delta \boldsymbol{F}^{(a)}_{\delta,\text{mmoe}}. \end{aligned} \quad (39)$$

It is worth noting that, for a given vector $\boldsymbol{\gamma}_\delta$, synthesis of $\boldsymbol{f}_{\delta,\text{mmoe}}$ requires only the computation in real-time of $\boldsymbol{F}^{(a)}_{\delta,\text{mmoe}}$, since both the matrices $\boldsymbol{\Theta}_\delta$ and $\boldsymbol{\Pi}_\delta$ can be precomputed off-line.

[6]In the sequel, we assume, without loss of generality, that $\boldsymbol{\Pi}_\delta$ is unitary, i.e., $\boldsymbol{\Pi}_\delta^T \boldsymbol{\Pi}_\delta = \mathbf{I}_{N(L_e-L_\delta)}$.

In its turn, the synthesis of the matrix $\boldsymbol{F}^{(a)}_{\delta,\mathrm{mmoe}}$ does not involve EVD/SVD or Cholesky decompositions and only requires the inversion of $\boldsymbol{\Pi}_\delta^T \boldsymbol{R_{zz}} \boldsymbol{\Pi}_\delta \in \mathbb{C}^{N(L_e-L_\delta)\times N(L_e-L_\delta)}$, which can consistently be estimated from the received data. Henceforth, if one resorts to batch algorithms, the computational burden of the proposed method is dominated by the matrix inversion in (39), which involves $\mathcal{O}\left[N^3(L_e-L_\delta)^3\right]$ floating point operations (flops) [35]. However, the matrix $\boldsymbol{F}^{(a)}_{\delta,\mathrm{mmoe}}$ can also be estimated adaptively by means of a simple and effective recursion (see Section IV-C), similar to the well-known recursive least square (RLS) algorithm [38], with a complexity of only $\mathcal{O}[N^2(L_e-L_\delta)^2]$ flops per iteration.

In Section IV-A, we analyze the expression of the combined channel-TEQ response vector $\boldsymbol{g}_{\delta,\mathrm{mmoe}} \triangleq \boldsymbol{H}^H \boldsymbol{f}_{\delta,\mathrm{mmoe}} \in \mathbb{C}^{L_g+1}$ and evaluate the mean power $\boldsymbol{f}^H_{\delta,\mathrm{mmoe}} \boldsymbol{R_{vv}} \boldsymbol{f}_{\delta,\mathrm{mmoe}}$ of the disturbance at the TEQ output, when the OFDM signal dominates the background noise (a common occurrence in many practical situations). This analysis allows us to enlighten the channel shortening process, as well as the NBI suppression capability of the proposed TEQ, and, moreover, provides a basis for further optimizations in Section IV-B.

A. Channel Shortening and NBI Suppression Analysis

Our aim is to establish whether the proposed MMOE-based TEQ allows one to truly shorten the channel, by leading to a combined channel-TEQ response vector $\boldsymbol{g}_{\delta,\mathrm{mmoe}}$ having a form similar to that reported in (10), and, at the same time, reject the NBI. To evaluate the NBI suppression capability of the proposed TEQ, we maintain that the disturbance vector in (4) is composed of two terms $\boldsymbol{v}(k) = \boldsymbol{v}_{\mathrm{nbi}}(k) + \boldsymbol{v}_{\mathrm{noise}}(k)$, where $\boldsymbol{v}_{\mathrm{nbi}}(k)$ and $\boldsymbol{v}_{\mathrm{noise}}(k)$ account for NBI and noise, respectively. In addition to assumption (a2), we assume that: **(a3)** the first $R_{\mathrm{nbi}} \ll N(L_e+1)$ eigenvalues of the NBI autocorrelation matrix $\boldsymbol{R}_{\mathrm{nbi}} \triangleq E\left[\boldsymbol{v}_{\mathrm{nbi}}(k)\boldsymbol{v}^H_{\mathrm{nbi}}(k)\right]$ are significantly different from zero, whereas the remaining ones are vanishingly small; **(a4)** the vector $\boldsymbol{v}_{\mathrm{noise}}(k)$ is a white random process, statistically independent of $\boldsymbol{v}_{\mathrm{nbi}}(k)$, with autocorrelation matrix $\boldsymbol{R}_{\mathrm{noise}} \triangleq E[\boldsymbol{v}_{\mathrm{noise}}(k)\boldsymbol{v}^H_{\mathrm{noise}}(k)] = \sigma_v^2 \mathbf{I}_{N(L_e+1)}$. It is worth noting that, by invoking some results [39] regarding the approximate dimensionality of exactly time-limited and nominally band-limited signals, assumption (a3) is well verified for reasonably large values of $N(L_e+1)$, with $R_{\mathrm{nbi}} = \lceil N(L_e+1)T_c W_{\mathrm{nbi}}\rceil + 1$, where W_{nbi} is the (nominal) bandwidth of the continuous-time NBI process. In the case of NBI, it happens in practice that, compared with the bandwidth of the multicarrier system, the bandwidth W_{nbi} is significantly smaller, i.e., $T_c W_{\mathrm{nbi}} \ll 1$, and, thus, it turns out that $R_{\mathrm{nbi}} \ll N(L_e+1)$. Under assumption (a3), the NBI autocorrelation matrix can be well modeled by the following full-rank factorization (see [32]) $\boldsymbol{R}_{\mathrm{nbi}} = \boldsymbol{J}\boldsymbol{J}^H$, where the matrix $\boldsymbol{J} \in \mathbb{C}^{N(L_e+1)\times R_{\mathrm{nbi}}}$ is full-column rank, i.e., $\mathrm{rank}(\boldsymbol{J}) = R_{\mathrm{nbi}}$. By virtue of this fact, the NBI suppression capability of the MMOE channel shortener can be quantified by evaluating the mean power $\mathcal{E}_{\mathrm{nbi}} \triangleq E\left[|\boldsymbol{f}^H_{\delta,\mathrm{mmoe}}\boldsymbol{v}_{\mathrm{nbi}}(k)|^2\right] = \boldsymbol{f}^H_{\delta,\mathrm{mmoe}}\boldsymbol{J}\boldsymbol{J}^H\boldsymbol{f}_{\delta,\mathrm{mmoe}} = \|\boldsymbol{J}^H\boldsymbol{f}_{\delta,\mathrm{mmoe}}\|^2$ of the NBI at the TEQ output, which depends on the expression of the vector $\boldsymbol{\ell}_{\delta,\mathrm{mmoe}} \triangleq \boldsymbol{J}^H\boldsymbol{f}_{\delta,\mathrm{mmoe}} \in \mathbb{C}^{R_{\mathrm{nbi}}}$. Therefore, the channel shortening and the NBI suppression capabilities of the proposed TEQ can jointly be characterized by the vector $\boldsymbol{b}_{\delta,\mathrm{mmoe}} \triangleq \boldsymbol{\mathcal{H}}^H\boldsymbol{f}_{\delta,\mathrm{mmoe}}$, with $\boldsymbol{\mathcal{H}} \triangleq [\boldsymbol{H},\boldsymbol{J}] \in \mathbb{C}^{N(L_e+1)\times(L_g+R_{\mathrm{nbi}}+1)}$. On the basis of (37) and (38), the vector $\boldsymbol{b}_{\delta,\mathrm{mmoe}}$ can be explicitly written as follows:

$$
\begin{aligned}
\boldsymbol{b}_{\delta,\mathrm{mmoe}} &= \begin{bmatrix} \boldsymbol{g}_{\delta,\mathrm{mmoe}} \\ \boldsymbol{\ell}_{\delta,\mathrm{mmoe}} \end{bmatrix} \\
&= \boldsymbol{\mathcal{H}}^H\left[\mathbf{I}_{N(L_e+1)} - \boldsymbol{\Pi}_\delta\left(\boldsymbol{\Pi}_\delta^T\boldsymbol{R_{zz}}\boldsymbol{\Pi}_\delta\right)^{-1}\boldsymbol{\Pi}_\delta^T\boldsymbol{R_{zz}}\right] \\
&\quad \times \boldsymbol{\Theta}_\delta\boldsymbol{\gamma}_\delta
\end{aligned} \tag{40}
$$

where, accounting for (4) and invoking assumptions (a1)–(a4), the autocorrelation matrix of the vector $\boldsymbol{z}(k)$ is given by

$$
\begin{aligned}
\boldsymbol{R_{zz}} &= \boldsymbol{H}\widetilde{\boldsymbol{R}}_{\boldsymbol{uu}}\boldsymbol{H}^H + \boldsymbol{J}\boldsymbol{J}^H + \sigma_v^2\mathbf{I}_{N(L_e+1)} \\
&= \boldsymbol{\mathcal{H}}\underbrace{\begin{bmatrix} \widetilde{\boldsymbol{R}}_{\boldsymbol{uu}} & \mathbf{O}_{(L_g+1)\times R_{\mathrm{nbi}}} \\ \mathbf{O}_{R_{\mathrm{nbi}}\times(L_g+1)} & \mathbf{I}_{R_{\mathrm{nbi}}} \end{bmatrix}}_{\boldsymbol{\Delta}\in\mathbb{C}^{(L_g+R_{\mathrm{nbi}}+1)\times(L_g+R_{\mathrm{nbi}}+1)}}\boldsymbol{\mathcal{H}}^H \\
&\quad + \sigma_v^2\mathbf{I}_{N(L_e+1)} \\
&= \boldsymbol{\mathcal{H}}\boldsymbol{\Delta}\boldsymbol{\mathcal{H}}^H + \sigma_v^2\mathbf{I}_{N(L_e+1)}
\end{aligned} \tag{41}
$$

with $\widetilde{\boldsymbol{R}}_{\boldsymbol{uu}} \triangleq E[\widetilde{\boldsymbol{u}}(k)\widetilde{\boldsymbol{u}}^H(k)] \in \mathbb{C}^{(L_g+1)\times(L_g+1)}$ representing the *nonsingular* autocorrelation matrix of the vector $\widetilde{\boldsymbol{u}}(k)$ defined in (4). As a consequence of the nonsingularity of $\widetilde{\boldsymbol{R}}_{\boldsymbol{uu}}$, the block diagonal matrix $\boldsymbol{\Delta}$ is nonsingular, too. Substituting (41) in (40) and recalling that $\boldsymbol{\Pi}_\delta^T\boldsymbol{\Pi}_\delta = \mathbf{I}_{N(L_e-L_\delta)}$ and $\boldsymbol{\Pi}_\delta^T\boldsymbol{\Theta}_\delta = \mathbf{O}_{N(L_e-L_\delta)\times N(L_\delta+1)}$, one obtains

$$
\begin{aligned}
\boldsymbol{b}_{\delta,\mathrm{mmoe}} = \boldsymbol{\mathcal{H}}^H\Big\{&\mathbf{I}_{N(L_e+1)} - \boldsymbol{\Pi}_\delta \\
&\cdot\Big[(\boldsymbol{\Pi}_\delta^T\boldsymbol{\mathcal{H}}\boldsymbol{\Delta}^{1/2})\cdot(\boldsymbol{\Pi}_\delta^T\boldsymbol{\mathcal{H}}\boldsymbol{\Delta}^{1/2})^H \\
&+\sigma_v^2\mathbf{I}_{N(L_e+1)}\Big]^{-1} \\
&\cdot(\boldsymbol{\Pi}_\delta^T\boldsymbol{\mathcal{H}}\boldsymbol{\Delta}^{1/2})\boldsymbol{\Delta}^{1/2}\boldsymbol{\mathcal{H}}^H\Big\}\boldsymbol{\Theta}_\delta\boldsymbol{\gamma}_\delta.
\end{aligned} \tag{42}
$$

Since our aim is to understand both the channel shortening and the NBI suppression features, we derive the analytical expression of $\boldsymbol{b}_{\delta,\mathrm{mmoe}}$ in the high SNR regime, i.e., as σ_v^2 approaches to zero. In this regard, let us define the asymptotic counterparts of $\boldsymbol{b}_{\delta,\mathrm{mmoe}}$ and $\boldsymbol{\ell}_{\delta,\mathrm{mmoe}}$ as $\bar{\boldsymbol{b}}_{\delta,\mathrm{mmoe}} \triangleq \lim_{\sigma_v^2\to 0}\boldsymbol{b}_{\delta,\mathrm{mmoe}}$ and $\bar{\boldsymbol{\ell}}_{\delta,\mathrm{mmoe}} \triangleq \lim_{\sigma_v^2\to 0}\boldsymbol{\ell}_{\delta,\mathrm{mmoe}}$; as a consequence of the limit formula for the Moore–Penrose inverse [32], it can be verified that

$$
\begin{aligned}
\bar{\boldsymbol{b}}_{\delta,\mathrm{mmoe}} &\triangleq \lim_{\sigma_v^2\to 0}\boldsymbol{b}_{\delta,\mathrm{mmoe}} = \begin{bmatrix} \bar{\boldsymbol{g}}_{\delta,\mathrm{mmoe}} \\ \bar{\boldsymbol{\ell}}_{\delta,\mathrm{mmoe}} \end{bmatrix} \\
&= \boldsymbol{\mathcal{H}}^H\left[\mathbf{I}_{N(L_e+1)} - \boldsymbol{\Pi}_\delta\left(\boldsymbol{\Delta}^{1/2}\boldsymbol{\mathcal{H}}^H\boldsymbol{\Pi}_\delta\right)^\dagger\boldsymbol{\Delta}^{1/2}\boldsymbol{\mathcal{H}}^H\right] \\
&\quad \times\boldsymbol{\Theta}_\delta\boldsymbol{\gamma}_\delta \\
&= \underbrace{\left[\mathbf{I}_{L_g+R_{\mathrm{nbi}}+1} - \boldsymbol{\mathcal{H}}^H\boldsymbol{\Pi}_\delta\left(\boldsymbol{\mathcal{H}}^H\boldsymbol{\Pi}_\delta\right)^\dagger\right]}_{\boldsymbol{P}_\delta\in\mathbb{C}^{(L_g+R_{\mathrm{nbi}}+1)\times(L_g+R_{\mathrm{nbi}}+1)}}\boldsymbol{\mathcal{H}}^H\boldsymbol{\Theta}_\delta\boldsymbol{\gamma}_\delta \\
&= \boldsymbol{P}_\delta\boldsymbol{\mathcal{H}}^H\boldsymbol{\Theta}_\delta\boldsymbol{\gamma}_\delta
\end{aligned} \tag{43}
$$

where, because of the nonsingularity of $\boldsymbol{\Delta}$, we have used in the second equality the fact that $\left(\boldsymbol{\Delta}^{1/2}\boldsymbol{\mathcal{H}}^H\boldsymbol{\Pi}_\delta\right)^\dagger = \left(\boldsymbol{\mathcal{H}}^H\boldsymbol{\Pi}_\delta\right)^\dagger \boldsymbol{\Delta}^{-1/2}$. As a first remark, it is important to note that the asymptotic expression of $\boldsymbol{b}_{\delta,\text{mmoe}}$ does not depend[7] on $\boldsymbol{\Delta}$ and, hence, on the structure of $\widetilde{\boldsymbol{R}}_{\boldsymbol{uu}}$. Furthermore, since $\boldsymbol{P}_\delta$ represents the orthogonal projector [32] onto the null space of $\boldsymbol{\Pi}_\delta^T\boldsymbol{\mathcal{H}}$, it is apparent from (43) that $\bar{\boldsymbol{b}}_{\delta,\text{mmoe}}$ belongs to the subspace $\mathcal{N}\left(\boldsymbol{\Pi}_\delta^T\boldsymbol{\mathcal{H}}\right) \equiv \mathcal{R}^\perp\left(\boldsymbol{\mathcal{H}}^H\boldsymbol{\Pi}_\delta\right)$. Relying on (43), the following Theorem provides a complete characterization of both the channel shortening and NBI suppression capabilities of the proposed approach.

Theorem 1: Assume that $N > 1$ and let $\boldsymbol{\mathcal{H}}_\delta \in \mathbb{C}^{N(L_e+1)\times R_\delta}$ define the matrix whose expression depends on the particular used parameterization (P1, P2, or P3) as shown in (44), at the bottom of the page.

If the subspaces $\mathcal{R}(\boldsymbol{\Theta}_\delta)$ and $\mathcal{R}(\boldsymbol{\mathcal{H}}_\delta)$ are nonoverlapping or disjoint, that is

$$\mathcal{R}(\boldsymbol{\Theta}_\delta) \cap \mathcal{R}(\boldsymbol{\mathcal{H}}_\delta) = \{\mathbf{0}_{N(L_e+1)}\} \tag{45}$$

then, for $\sigma_v^2 \to 0$, the combined channel-TEQ vector $\boldsymbol{g}_{\delta,\text{mmoe}}$ assumes the target form (10) where: for $0 \le \delta \le \min[L_e, L_h]$ (case P1), $\Delta = 0$, $L_{\text{eff}} = \delta$ and $g_{\text{target}}(\ell) = \boldsymbol{h}_\ell^H\boldsymbol{\Theta}_\delta\boldsymbol{\gamma}_\delta$, with $\ell \in \{0, 1, \ldots, \delta\}$; for $L_h < \delta < L_e$ (case P2), $\Delta = \delta$, $L_{\text{eff}} = 0$ and $g_{\text{target}}(\delta) = \boldsymbol{\xi}_\delta^H\boldsymbol{\gamma}_\delta$; for $\max[L_e, L_h] \le \delta \le L_g$ (case P3), $\Delta = \delta + 1$, $L_{\text{eff}} = L_g - \delta$ and $g_{\text{target}}(\ell) = \boldsymbol{h}_\ell^H\boldsymbol{\Theta}_\delta\boldsymbol{\gamma}_\delta$, with $\ell \in \{\delta+1, \delta+2, \ldots, L_g\}$. Furthermore, if condition (45) is fulfilled, perfect NBI suppression is achieved in the limiting case of vanishingly small noise, i.e., $\boldsymbol{\ell}_{\delta,\text{mmoe}} \to \mathbf{0}_{R_{\text{nbi}}}$ (and, thus, $\mathcal{E}_{\text{nbi}} \to 0$) as $\sigma_v^2 \to 0$. ∎

Proof: See Appendix I.

At this point, several remarks are in order concerning the implications of Theorem 1. First of all, it is now apparent that the choice of δ strongly influences both the actual values of the shortening delay Δ and the effective order L_{eff} of $\boldsymbol{g}_{\text{target}}$, as well as the values of its nonzero entries. It is important to observe that, unlike previously proposed blind channel-shortening approaches [19], [20], [25], the proposed MMOE-based TEQ is able to completely reject the NBI in the high SNR region. In particular, fulfillment of (45) necessarily imposes that condition (c1) holds, but it does not require that condition (c3) be satisfied. In other words, unlike [25], the proposed channel-shortening approach is able to work even when the channel phases $\{h_q(k)\}_{q=0}^{N-1}$ exhibit common zeros in their $\mathcal{Z}$-transforms, a situation that happens very often in wireless channels [37]. It is shown in Appendix I that condition (45) is equivalent to require that the matrix $\boldsymbol{\Pi}_\delta^T\boldsymbol{H}_\delta \in \mathbb{C}^{N(L_e-L_\delta)\times R_\delta}$ be full-column rank, i.e., $\text{rank}(\boldsymbol{\Pi}_\delta^T\boldsymbol{H}_\delta) = R_\delta$. A necessary condition for satisfying this rank requirement is that $N(L_e - L_\delta) \ge R_\delta$ which, depending on the employed parameterization, imposes the following upper bound on the order of the channel to be shortened [see (46), shown at the bottom of the page]. If inequality (46) holds, as confirmed by our simulation results, it is very unlikely that condition (45) is violated in practice. Let us first focus attention on parameterization P1. In this case, if the channel order L_h does not exceed the maximum value $L_{1,\max}(\delta) \triangleq (N-1)(L_e-\delta) - R_{\text{nbi}}$, for $0 \le \delta \le \min[L_e, L_h]$, the asymptotic combined channel-TEQ impulse response has effective order $L_{\text{eff}} = \delta$, and, thus, to allow perfect IBI cancellation by means of CP removal, it is required that $L_{cp} \ge \delta$. At first sight, one might conclude that the best choice of the equalization delay is $\delta = 0$ since, in this case, for a given value of N, L_e, and R_{nbi}, the upper bound $L_{1,\max}(\delta)$ achieves its maximum value $(N-1)L_e - R_{\text{nbi}}$, i.e., a longer channel can be shortened, and, similarly to [24], the proposed TEQ performs perfect equalization, i.e., $L_{\text{eff}} = 0$, thus avoiding CP insertion at the transmitter. This is surely true when $\sigma_v^2 \to 0$, but it is not always the best choice in noisy environments: indeed, using too small values of δ does not allow one to exploit the entire channel energy, since the vectors $\boldsymbol{h}_0, \boldsymbol{h}_1, \ldots, \boldsymbol{h}_\delta$, which determine the nonzero entries of $\boldsymbol{g}_{\delta,\text{mmoe}}$, only collect a limited number of channel samples and might, thus, lead to a significant SSINR degradation at the TEQ output. Furthermore, it is well-known that perfect equalization may lead to excessive noise enhancement at the equalizer output. Therefore, for low-to-moderate values of the SNR, it is preferable to perform channel shortening (rather than perfect equalization) by choosing $\delta = L_{\text{eff}} = L_{cp}$ and, thus, leaving to CP removal the duty of suppressing the residual IBI. In this case,

[7]In contrast, the channel-shortening methods proposed in [25] can work only when the vector $\widetilde{\boldsymbol{u}}(k)$ is white.

$$\boldsymbol{\mathcal{H}}_\delta = \begin{cases} [\boldsymbol{h}_{\delta+1}, \boldsymbol{h}_{\delta+2}, \ldots, \boldsymbol{h}_{L_g}, \boldsymbol{J}], & \text{with } R_\delta = L_g + R_{\text{nbi}} - \delta, \text{ for } 0 \le \delta \le \min[L_e, L_h] \\ [\boldsymbol{h}_0, \ldots, \boldsymbol{h}_{\delta-1}, \boldsymbol{h}_{\delta+1}, \ldots, \boldsymbol{h}_{L_g}, \boldsymbol{J}], & \text{with } R_\delta = L_g + R_{\text{nbi}}, \text{ for } L_h < \delta < L_e \\ [\boldsymbol{h}_0, \boldsymbol{h}_1, \ldots, \boldsymbol{h}_\delta, \boldsymbol{J}], & \text{with } R_\delta = R_{\text{nbi}} + \delta + 1, \text{ for } \max[L_e, L_h] \le \delta \le L_g. \end{cases} \tag{44}$$

$$L_h \le \begin{cases} (N-1)(L_e-\delta) - R_{\text{nbi}}, & \text{for } 0 \le \delta \le \min[L_e, L_h] \\ \left\lceil \frac{(N-1)L_e - R_{\text{nbi}}}{N+1} \right\rceil, & \text{for } L_h < \delta < L_e \\ \left\lceil \frac{(N-1)\delta - R_{\text{nbi}} - 1}{N} \right\rceil, & \text{for } \max[L_e, L_h] \le \delta \le L_g \end{cases} \tag{46}$$

the channel order must obey $L_h \leq (N-1)(L_e - L_{cp}) - R_{\text{nbi}}$, which can be equivalently expressed as

$$L_e \geq L_{cp} + \left\lceil \frac{L_h + R_{\text{nbi}}}{N-1} \right\rceil . \tag{47}$$

In the absence of NBI (i.e., $R_{\text{nbi}} = 0$), when the received signal is sampled at rate $2/T_c$, inequality (47) becomes $L_e \geq L_{cp} + L_h$. This lower bound on L_e is more restrictive than condition (c2), in the sense that, for a given channel order, a larger TEQ order L_e is required to achieve perfect channel shortening. However, for NBI-free systems with $N > 2$, inequality (47) is less restrictive than condition (c2) if the channel order satisfies the relation

$$L_h \geq \left\lceil \frac{N-1}{N-2}(L_{cp}+1) \right\rceil . \tag{48}$$

In this latter case, perfect channel shortening can be obtained even when the channel order is greater than the TEQ length, i.e., $L_h > L_e + 1$, which leads to a significant computational saving for highly time-dispersive channels. A remarkable property of parameterization P1 is that the structure of $\boldsymbol{\Theta}_\delta$ is *independent* of the channel order L_h. Henceforth, *only* upper bounds (rather than the exact knowledge) of the channel order and the NBI rank R_{nbi} (i.e., the NBI bandwidth W_{nbi}) are required to select a suitable value of L_e. This is a reasonable assumption in practice since, depending on the transmitted signal parameters (carrier frequency and bandwidth) and application (wireline or wireless), the maximum channel order and NBI bandwidth are known.

Contrary to the case P1, when parameterization P2 is employed, the structure of $\boldsymbol{\Theta}_\delta$ explicitly depends on the channel order L_h, which must be smaller than L_e, i.e., $L_h \leq L_e$. This implies that, similarly to [25], exact knowledge or estimation of L_h is a crucial issue. Furthermore, in this case, at the expense of increased noise enhancement at the TEQ output, it is only possible to perfectly equalize the channel, i.e., $L_{\text{eff}} = 0$, provided that the channel order L_h does not exceed the maximum value

$$L_{2,\max} \triangleq \left\lceil \frac{(N-1)L_e - R_{\text{nbi}}}{N+1} \right\rceil . \tag{49}$$

The advantage of using the parameterization P2 only lies in the fact that it allows full exploitation of the channel energy, since the vector $\boldsymbol{\xi}_\delta$, which determines the nonzero entry $g(\delta)$ of $\boldsymbol{g}_{\delta,\text{mmoe}}$, collects all the channel samples, for $\delta \in \{L_h + 1, L_h + 2, \ldots, L_e - 1\}$. Finally, regarding parameterization P3, observe that, since $\Delta + L_{\text{eff}} > L_g$, the $L_{\text{eff}} + 1$ nonzero entries of $\boldsymbol{g}_{\delta,\text{mmoe}}$ are not consecutive. Moreover, similarly to the P2 case, the structure of $\boldsymbol{\Theta}_\delta$ explicitly depends on the channel order L_h. In summary, we can state that parameterization P1, with $\delta = L_{cp}$, turns out to be the best choice in practice.

B. Optimization of the Vector of Constraint Values

Theorem 1 shows that there is ample freedom to select $\boldsymbol{\gamma}_\delta$ in the optimization problem (34). At this point, it is, thus, paramount to derive a blind criterion that will guide us in selecting optimal constraining values in the presence of disturbance. Towards this aim, the vector $\boldsymbol{\gamma}_\delta$ is chosen so that it approximatively maximizes the SSINR at the output of the proposed TEQ, which is defined in (11) and whose explicit expression is given by (17). As a first step, observe that, by virtue of Theorem 1, it results that $\text{SSINR}(\boldsymbol{f}_{\delta,\text{mmoe}}) \to +\infty$ as $\sigma_w^2 \to 0$, regardless of the choice of $\boldsymbol{\gamma}_\delta$. On the other hand, as a consequence of Theorem 1, if the OFDM signal dominates the background noise, we can assume that the proposed TEQ is able to perform almost perfect channel shortening, i.e., $\|\boldsymbol{g}_{\text{wall}}\|^2 \approx 0$, and almost perfect NBI suppression, i.e., $\mathcal{E}_{\text{nbi}} \approx 0$. Consequently, accounting for (35) and assumptions (a3)–(a4), the SSINR at the output of the MMOE-based TEQ can approximatively be written as

$$\begin{aligned} &\text{SSINR}(\boldsymbol{f}_{\delta,\text{mmoe}}) \\ &\quad = \frac{\boldsymbol{f}_{\delta,\text{mmoe}}^H \left(\boldsymbol{H}_{\text{win}} \widetilde{\boldsymbol{R}}_{\text{win}} \boldsymbol{H}_{\text{win}}^H\right) \boldsymbol{f}_{\delta,\text{mmoe}}}{\sigma_v^2 \|\boldsymbol{f}_{\delta,\text{mmoe}}\|^2} \\ &\quad = \frac{\boldsymbol{\gamma}_\delta^H \left(\boldsymbol{F}_{\delta,\text{mmoe}}^H \boldsymbol{H}_{\text{win}} \widetilde{\boldsymbol{R}}_{\text{win}} \boldsymbol{H}_{\text{win}}^H \boldsymbol{F}_{\delta,\text{mmoe}}\right) \boldsymbol{\gamma}_\delta}{\sigma_v^2 \boldsymbol{\gamma}_\delta^H \boldsymbol{F}_{\delta,\text{mmoe}}^H \boldsymbol{F}_{\delta,\text{mmoe}} \boldsymbol{\gamma}_\delta}. \end{aligned} \tag{50}$$

Therefore, on the basis of (50), maximizing $\text{SSINR}(\boldsymbol{f}_{\delta,\text{mmoe}})$ amounts to choose $\boldsymbol{\gamma}_\delta$ as the solution $\boldsymbol{\gamma}_{\delta,\text{opt}}$ of the following maximization problem:

$$\arg \max_{\boldsymbol{\gamma}_\delta \in \mathbb{C}^{N(L_\delta+1)}} \frac{\boldsymbol{\gamma}_\delta^H \boldsymbol{F}_{\delta,\text{mmoe}}^H \boldsymbol{H}_{\text{win}} \widetilde{\boldsymbol{R}}_{\text{win}} \boldsymbol{H}_{\text{win}}^H \boldsymbol{F}_{\delta,\text{mmoe}} \boldsymbol{\gamma}_\delta}{\boldsymbol{\gamma}_\delta^H \boldsymbol{F}_{\delta,\text{mmoe}}^H \boldsymbol{F}_{\delta,\text{mmoe}} \boldsymbol{\gamma}_\delta}. \tag{51}$$

The solution of (51) is given [35] by $\boldsymbol{\gamma}_{\delta,\text{opt}} = \varrho \widetilde{\boldsymbol{a}}_{\max}$, where $\varrho \in \mathbb{C}$ and $\widetilde{\boldsymbol{a}}_{\max}$ is the generalized eigenvector associated with the largest generalized eigenvalue of the two matrices $\boldsymbol{F}_{\delta,\text{mmoe}}^H \boldsymbol{H}_{\text{win}} \widetilde{\boldsymbol{R}}_{\text{win}} \boldsymbol{H}_{\text{win}}^H \boldsymbol{F}_{\delta,\text{mmoe}}$ and $\boldsymbol{F}_{\delta,\text{mmoe}}^H \boldsymbol{F}_{\delta,\text{mmoe}}$. Since $\boldsymbol{F}_{\delta,\text{mmoe}} \in \mathbb{C}^{N(L_e+1)\times N(L_\delta+1)}$ is full-column rank for all the SNR values of practical interest, i.e., $\text{rank}(\boldsymbol{F}_{\delta,\text{mmoe}}) = N(L_\delta + 1)$, the solution of the optimization problem (51) can be more conveniently obtained by resorting to any one of several orthogonalization procedures, such as the SVD, Cholesky or QR decomposition [34], [35]. In what follows, we use the QR decomposition (QRD) since it is not only mathematically elegant, but also computationally powerful and highly versatile. Following this approach, the matrix $\boldsymbol{F}_{\delta,\text{mmoe}}$ can be factorized as follows:

$$\boldsymbol{F}_{\delta,\text{mmoe}} = \boldsymbol{Q}_{\delta,\text{mmoe}} \boldsymbol{R}_{\delta,\text{mmoe}} \tag{52}$$

where $\boldsymbol{R}_{\delta,\text{mmoe}} \in \mathbb{C}^{N(L_\delta+1)\times N(L_\delta+1)}$ is a nonsingular upper triangular matrix, whereas the columns of the matrix $\boldsymbol{Q}_{\delta,\text{mmoe}} \in \mathbb{C}^{N(L_e+1)\times N(L_\delta+1)}$ are orthonormal, i.e., $\boldsymbol{Q}_{\delta,\text{mmoe}}^H \boldsymbol{Q}_{\delta,\text{mmoe}} = \mathbf{I}_{N(L_\delta+1)}$. Relying on this decomposition, it is readily seen that the solution of (51) can equivalently be expressed as

$$\boldsymbol{\gamma}_{\delta,\text{opt}} = \boldsymbol{R}_{\delta,\text{mmoe}}^{-1} \boldsymbol{\zeta}_{\delta,\text{opt}} \tag{53}$$

where, in its turn, $\boldsymbol{\zeta}_{\delta,\text{opt}} \in \mathbb{C}^{N(L_\delta+1)}$ is determined by the maximization problem

$$\boldsymbol{\zeta}_{\delta,\text{opt}} = \arg \max_{\boldsymbol{\zeta}_\delta \in \mathbb{C}^{N(L_\delta+1)}} \frac{\boldsymbol{\zeta}_\delta^H \boldsymbol{M}_{\delta,\text{mmoe}} \boldsymbol{\zeta}_\delta}{\boldsymbol{\zeta}_\delta^H \boldsymbol{\zeta}_\delta} \tag{54}$$

with $\boldsymbol{M}_{\delta,\text{mmoe}} \triangleq \boldsymbol{Q}_{\delta,\text{mmoe}}^H \boldsymbol{H}_{\text{win}} \widetilde{\boldsymbol{R}}_{\text{win}} \boldsymbol{H}_{\text{win}}^H \boldsymbol{Q}_{\delta,\text{mmoe}}$, whose solution [34] is the eigenvector associated with the largest eigenvalue of $\boldsymbol{M}_{\delta,\text{mmoe}}$. Accounting for the QRD of $\boldsymbol{F}_{\delta,\text{mmoe}}$ and (53), the weight vector (35) reduces to

$$\boldsymbol{f}_{\delta,\text{opt}} \triangleq \boldsymbol{F}_{\delta,\text{mmoe}}\boldsymbol{\gamma}_{\delta,\text{opt}} = \boldsymbol{Q}_{\delta,\text{mmoe}}\boldsymbol{\zeta}_{\delta,\text{opt}} \tag{55}$$

which shows that the synthesis of $\boldsymbol{f}_{\delta,\text{opt}}$ does not need neither evaluation nor inversion of $\boldsymbol{R}_{\delta,\text{mmoe}}$, but only computation of $\boldsymbol{Q}_{\delta,\text{mmoe}}$. On the other hand, it is apparent that evaluation of $\boldsymbol{\zeta}_{\delta,\text{opt}}$ requires *a priori* knowledge of the channel-dependent matrix $\boldsymbol{H}_{\text{win}}$, which is unknown at the receiver. Interestingly, the computation of $\boldsymbol{\zeta}_{\delta,\text{opt}}$ can blindly be carried out, without requiring knowledge or estimation of $\boldsymbol{H}_{\text{win}}$. To show this, let us consider the linear transformation $\boldsymbol{\jmath}(k) \triangleq \boldsymbol{Q}^H_{\delta,\text{mmoe}}\boldsymbol{z}(k) \in \mathbb{C}^{N(L_\delta+1)}$ of the received vector $\boldsymbol{z}(k)$ which, accounting for (4) and invoking Theorem 1, for moderate-to-high values of the SNR, can be approximatively decomposed as

$$\boldsymbol{\jmath}(k) = \boldsymbol{Q}^H_{\delta,\text{mmoe}}\boldsymbol{H}_{\text{win}}\widetilde{\boldsymbol{u}}_{\text{win}}(k) + \boldsymbol{Q}^H_{\delta,\text{mmoe}}\boldsymbol{v}_{\text{noise}}(k) \tag{56}$$

whose autocorrelation matrix $\boldsymbol{R}_{\boldsymbol{\jmath}\boldsymbol{\jmath}} \triangleq E[\boldsymbol{\jmath}(k)\boldsymbol{\jmath}^H(k)] \in \mathbb{C}^{N(L_\delta+1)\times N(L_\delta+1)}$ is given by

$$\boldsymbol{R}_{\boldsymbol{\jmath}\boldsymbol{\jmath}} = \boldsymbol{M}_{\delta,\text{mmoe}} + \sigma_v^2\mathbf{I}_{N(L_\delta+1)} \tag{57}$$

which shows that $\boldsymbol{\zeta}_{\delta,\text{opt}}$ corresponding to the largest eigenvalue of $\boldsymbol{M}_{\delta,\text{mmoe}}$ can blindly be evaluated as the eigenvector corresponding to the largest eigenvalue of $\boldsymbol{R}_{\boldsymbol{\jmath}\boldsymbol{\jmath}}$, which can consistently be estimated from the received data. From a computational point of view, it turns out that the evaluation of $\boldsymbol{\zeta}_{\delta,\text{opt}}$ is dominated by the QRD of $\boldsymbol{F}_{\delta,\text{mmoe}}$, which involves $\mathcal{O}\left[N^3(L_\delta+1)^2(L_e+1)\right]$ flops if computed from scratch, and, additionally, by the calculus of the dominant eigenvector of $\boldsymbol{R}_{\boldsymbol{\jmath}\boldsymbol{\jmath}}$, which requires $\mathcal{O}\left[N^3(L_\delta+1)^3\right]$ flops if one resorts to batch algorithms.

In Section IV-C, we will show that $\boldsymbol{\zeta}_{\delta,\text{opt}}$ can adaptively be estimated by means of a fast recursive rule, which only takes $\mathcal{O}\{N^2(L_e-L_\delta)\cdot\min\left[L_e-L_\delta,(L_\delta+1)^2\right]\}$ flops per iteration.

C. Adaptive Implementation

Overall, from the previous sections, we can maintain that, in practice, the synthesis of the proposed MMOE-based TEQ can be subdivided into the following four steps.

- Step 1) Given K_c consecutive samples of the received data vectors $\{\boldsymbol{z}(k)\}_{k=0}^{K_c-1}$, produce a reliable estimate $\widehat{\boldsymbol{F}}^{(a)}_{\delta,\text{mmoe}}$ of the matrix $\boldsymbol{F}^{(a)}_{\delta,\text{mmoe}} = \left(\boldsymbol{\Pi}_\delta^T\boldsymbol{R}_{\boldsymbol{z}\boldsymbol{z}}\boldsymbol{\Pi}_\delta\right)^{-1}\boldsymbol{\Pi}_\delta^T\boldsymbol{R}_{\boldsymbol{z}\boldsymbol{z}}\boldsymbol{\Theta}_\delta$, and build the matrix $\widehat{\boldsymbol{F}}_{\delta,\text{mmoe}} = \boldsymbol{\Theta}_\delta - \boldsymbol{\Pi}_\delta\widehat{\boldsymbol{F}}^{(a)}_{\delta,\text{mmoe}}$ [see (39)].
- Step 2) Calculate the QRD of $\widehat{\boldsymbol{F}}_{\delta,\text{mmoe}}$, thus obtaining an estimate $\widehat{\boldsymbol{Q}}_{\delta,\text{mmoe}}$ of $\boldsymbol{Q}_{\delta,\text{mmoe}}$ [see (52)].
- Step 3) Form an estimate $\widehat{\boldsymbol{\zeta}}_{\delta,\text{opt}}$ of $\boldsymbol{\zeta}_{\delta,\text{opt}}$ by evaluating the dominant eigenvector of the autocorrelation matrix of vector $\widehat{\boldsymbol{Q}}^H_{\delta,\text{mmoe}}\boldsymbol{z}(k)$ [see (56) and (57)].
- Step 4) Build the weight vector of the proposed TEQ as $\widehat{\boldsymbol{f}}_{\delta,\text{opt}} = \widehat{\boldsymbol{Q}}_{\delta,\text{mmoe}}\widehat{\boldsymbol{\zeta}}_{\delta,\text{opt}}$ [see (55)].

For the time being, let us suppose that $\delta = L_{cp}$ (and, thus, $L_\delta = L_{cp}$) which, as pointed out in Section IV-A, turns out to be the most suitable choice in practice. In this case, if steps 1) and 3) are carried out in batch mode, and $\widehat{\boldsymbol{Q}}_{\delta,\text{mmoe}}$ in step 2) is obtained by computing the QRD of $\widehat{\boldsymbol{F}}_{\delta,\text{mmoe}}$ from scratch, the synthesis of the weight vector $\widehat{\boldsymbol{f}}_{\delta,\text{opt}}$ requires $\mathcal{O}\left\{N^3\left[(L_e-L_{cp})^3+(L_{cp}+1)^3+(L_{cp}+1)^2(L_e+1)\right]\right\}$ flops. Since L_e is significantly larger than L_{cp} in practical settings, this computational load is lower than that of both the methods (A and B) proposed in [25], whose total computational complexity increases approximatively by a factor of $N^3(L_e+1)^3$. For instance, for the representative parameters used in Section V, wherein $N=2$, $L_{cp}=4$ and $L_e=18$, the methods of [25] take $\mathcal{O}\,(54872)$ flops, whereas the proposed technique requires only $\mathcal{O}\,(26752)$ flops. Nevertheless, the computational burden of the proposed method might not be affordable in real-time applications, especially for large values of L_e. However, unlike [25], the weight vector $\widehat{\boldsymbol{f}}_{\delta,\text{opt}}$ can also be estimated by means of a simple and efficient recursive algorithm, which additionally makes it possible for the proposed method to track time variations in the second-order statistics (SOS) of the received data, provided that the variations are sufficiently slow.

In the first step, relying on the RLS algorithm [38], the matrix $\boldsymbol{F}^{(a)}_{\delta,\text{mmoe}}$ can be estimated from the incoming data through sample-by-sample updating. Specifically, after some lengthy calculations, it can be shown [31] that, for $k \in \{0,1,\ldots,K_c-1\}$, the recursion for updating $\widehat{\boldsymbol{F}}^{(a)}_{\delta,\text{mmoe}}$ is

$$\begin{aligned}\widehat{\boldsymbol{F}}^{(a)}_{\delta,\text{mmoe}}(k) &= \widehat{\boldsymbol{F}}^{(a)}_{\delta,\text{mmoe}}(k-1)\\ &\quad + \boldsymbol{p}(k)\Bigg[\underbrace{\widehat{\boldsymbol{F}}^H_{\delta,\text{mmoe}}(k-1)\boldsymbol{z}(k)}_{\boldsymbol{q}(k)\in\mathbb{C}^{N(L_\delta+1)}}\Bigg]^H\\ &= \widehat{\boldsymbol{F}}^{(a)}_{\delta,\text{mmoe}}(k-1) + \boldsymbol{p}(k)\boldsymbol{q}^H(k)\end{aligned} \tag{58}$$

where $\widehat{\boldsymbol{F}}_{\delta,\text{mmoe}}(k-1) = \boldsymbol{\Theta}_\delta - \boldsymbol{\Pi}_\delta\widehat{\boldsymbol{F}}^{(a)}_{\delta,\text{mmoe}}(k-1)$ and

$$\boldsymbol{p}(k) \triangleq \frac{\boldsymbol{\Phi}_{\boldsymbol{z}\boldsymbol{z}}(k-1)\boldsymbol{\Pi}_\delta^T\boldsymbol{z}(k)}{\lambda_1 + \boldsymbol{z}^H(k)\boldsymbol{\Pi}_\delta\boldsymbol{\Phi}_{\boldsymbol{z}\boldsymbol{z}}(k-1)\boldsymbol{\Pi}_\delta^T\boldsymbol{z}(k)} \in \mathbb{C}^{N(L_e-L_\delta)} \tag{59}$$

is the *overall gain* vector, with $\boldsymbol{\Phi}_{\boldsymbol{z}\boldsymbol{z}}(k-1)$ and $\lambda_1 \in (0,1]$ denoting the estimate of $\left(\boldsymbol{\Pi}_\delta^T\boldsymbol{R}_{\boldsymbol{z}\boldsymbol{z}}\boldsymbol{\Pi}_\delta\right)^{-1}$ at iteration $k-1$ and the forgetting factor of the recursive algorithm, respectively. According to the standard initialization strategy for the RLS algorithm, we set $\boldsymbol{\Phi}_{\boldsymbol{z}\boldsymbol{z}}(-1) = \delta^{-1}\boldsymbol{I}_{N(L_e-L_\delta)}$ and $\widehat{\boldsymbol{F}}^{(a)}_{\delta,\text{mmoe}}(-1) = \left[\mathbf{I}_{N(L_\delta+1)}, \mathbf{O}_{N(L_\delta+1)\times N(L_e-2L_\delta-1)}\right]^T$, where $\delta \in \mathbb{R}$ is a positive constant. By resorting to standard analysis tools [38], it can be proven that, as k grows, matrix $\widehat{\boldsymbol{F}}^{(a)}_{\delta,\text{mmoe}}(k)$ converges in mean square to the optimal matrix $\boldsymbol{F}^{(a)}_{\delta,\text{mmoe}}$, regardless of the eigenstructure of $\boldsymbol{R}_{\boldsymbol{z}\boldsymbol{z}}$. It is worth noting that the recursive (58) requires $\mathcal{O}\left[N^2(L_e-L_\delta)^2\right]$ flops per iteration.

The second step involves the QRD of $\widehat{\boldsymbol{F}}_{\delta,\text{mmoe}}(k)$ in order to obtain the following factorization:

$$\widehat{\boldsymbol{F}}_{\delta,\text{mmoe}}(k) = \widehat{\boldsymbol{Q}}_{\delta,\text{mmoe}}(k)\widehat{\boldsymbol{R}}_{\delta,\text{mmoe}}(k) \tag{60}$$

where $\widehat{\boldsymbol{R}}_{\delta,\text{mmoe}}(k) \in \mathbb{C}^{N(L_\delta+1)\times N(L_\delta+1)}$ is a nonsingular upper triangular matrix, whereas the columns of the matrix $\widehat{\boldsymbol{Q}}_{\delta,\text{mmoe}}(k) \in \mathbb{C}^{N(L_e+1)\times N(L_\delta+1)}$ are orthonormal, i.e., $\widehat{\boldsymbol{Q}}^{H}_{\delta,\text{mmoe}}(k)\widehat{\boldsymbol{Q}}_{\delta,\text{mmoe}}(k) = \mathbf{I}_{N(L_\delta+1)}$. By observing that $\widehat{\boldsymbol{F}}_{\delta,\text{mmoe}}(k) = \boldsymbol{\Theta}_\delta - \boldsymbol{\Pi}_\delta \widehat{\boldsymbol{F}}^{(a)}_{\delta,\text{mmoe}}(k)$, the matrix $\widehat{\boldsymbol{Q}}_{\delta,\text{mmoe}}(k)$ can be conveniently obtained from the QRD of $\widehat{\boldsymbol{F}}^{(a)}_{\delta,\text{mmoe}}(k)$ rather than from the QRD of $\widehat{\boldsymbol{F}}_{\delta,\text{mmoe}}(k)$. Indeed, the matrix $\widehat{\boldsymbol{F}}_{\delta,\text{mmoe}}(k)$ can be rewritten as shown by (61), shown at the bottom of the page, where $\widehat{\boldsymbol{Q}}^{(a)}_{\delta,\text{mmoe}}(k)\widehat{\boldsymbol{R}}^{(a)}_{\delta,\text{mmoe}}(k)$ is the QRD of $\widehat{\boldsymbol{F}}^{(a)}_{\delta,\text{mmoe}}(k)$, with $\widehat{\boldsymbol{R}}^{(a)}_{\delta,\text{mmoe}}(k) \in \mathbb{C}^{N(L_\delta+1)\times N(L_\delta+1)}$ being a nonsingular upper triangular matrix and $\widehat{\boldsymbol{Q}}_{\delta,\text{mmoe}}(k) \in \mathbb{C}^{N(L_e-L_\delta)\times N(L_\delta+1)}$ satisfying $\left[\widehat{\boldsymbol{Q}}^{(a)}_{\delta,\text{mmoe}}(k)\right]^{H} \widehat{\boldsymbol{Q}}^{(a)}_{\delta,\text{mmoe}}(k) = \mathbf{I}_{N(L_\delta+1)}$. It is important to observe that, by construction, the columns of $\widehat{\boldsymbol{\Xi}}^{(a)}_{\delta,\text{mmoe}}(k)$ are orthonormal, i.e., $\left[\widehat{\boldsymbol{\Xi}}^{(a)}_{\delta,\text{mmoe}}(k)\right]^{H} \widehat{\boldsymbol{\Xi}}^{(a)}_{\delta,\text{mmoe}}(k) = \mathbf{I}_{2N(L_\delta+1)}$. See equation (61).

Capitalizing on (61), the QRD of $\widehat{\boldsymbol{F}}_{\delta,\text{mmoe}}(k)$ is obtained by performing $N(L_\delta+1)$ successive Givens or Householder rotations [35] $\widehat{\boldsymbol{\Omega}}^{(a)}_{\ell}(k) \in \mathbb{C}^{2N(L_\delta+1)\times 2N(L_\delta+1)}$, for $\ell \in \{1, 2, \ldots, N(L_\delta+1)\}$, thus getting

$$\widehat{\boldsymbol{\Omega}}^{(a)}_{N(L_\delta+1)}(k)\ldots\widehat{\boldsymbol{\Omega}}^{(a)}_{1}(k)\begin{bmatrix}\widehat{\boldsymbol{R}}^{(a)}_{\delta,\text{mmoe}}(k)\\ \mathbf{I}_{N(L_\delta+1)}\end{bmatrix} = \begin{bmatrix}\widehat{\boldsymbol{R}}_{\delta,\text{mmoe}}(k)\\ \mathbf{O}_{N(L_\delta+1)\times N(L_\delta+1)}\end{bmatrix}. \quad (62)$$

At this point, accounting for (61) and (62), and observing that Givens/Householder matrices are unitary by definition, the matrix $\widehat{\boldsymbol{Q}}_{\delta,\text{mmoe}}(k)$ can be obtained by picking up the first $N(L_\delta+1)$ columns of the matrix $\widehat{\boldsymbol{\Xi}}^{(a)}_{\delta,\text{mmoe}}(k)\left[\widehat{\boldsymbol{\Omega}}^{(a)}_{N(L_\delta+1)}(k)\right]^{H}\ldots\left[\widehat{\boldsymbol{\Omega}}^{(a)}_{1}(k)\right]^{H}$. Therefore, given the QRD of $\widehat{\boldsymbol{F}}^{(a)}_{\delta,\text{mmoe}}(k)$, the calculus of $\widehat{\boldsymbol{Q}}_{\delta,\text{mmoe}}(k)$ requires $\mathcal{O}\left[N^2(L_\delta+1)^2\right]$ flops if standard Givens/Householder transformations are used [35]. Let us now focus attention on evaluation of the QRD of $\widehat{\boldsymbol{F}}^{(a)}_{\delta,\text{mmoe}}(k)$, whose computation from scratch takes $\mathcal{O}\left[N^2(L_e-L_\delta)(L_\delta+1)^2\right]$ flops. To develop an alternative procedure for computing the QRD of $\widehat{\boldsymbol{F}}^{(a)}_{\delta,\text{mmoe}}(k)$, we observe from (58) that $\widehat{\boldsymbol{F}}^{(a)}_{\delta,\text{mmoe}}(k)$ comes from a rank-one change of $\widehat{\boldsymbol{F}}^{(a)}_{\delta,\text{mmoe}}(k-1)$. If the complete QRD of $\widehat{\boldsymbol{F}}^{(a)}_{\delta,\text{mmoe}}(k-1)$ is known and $\text{rank}\left[\widehat{\boldsymbol{F}}^{(a)}_{\delta,\text{mmoe}}(k-1)\right] = \text{rank}\left[\widehat{\boldsymbol{F}}^{(a)}_{\delta,\text{mmoe}}(k)\right] = N(L_\delta+1)$, there exists [35] a simple procedure to compute the QRD of the rank-one update $\widehat{\boldsymbol{F}}^{(a)}_{\delta,\text{mmoe}}(k)$ of $\widehat{\boldsymbol{F}}^{(a)}_{\delta,\text{mmoe}}(k-1)$, which involves a computational burden of $\mathcal{O}\left[N^2(L_e-L_\delta)^2\right]$ flops. For the sake of conciseness, we defer directly to [35] for details regarding the structure of the rotations matrices $\{\widehat{\boldsymbol{\Omega}}^{(a)}_{\ell}(k)\}_{\ell=1}^{N(L_\delta+1)}$ and the rank-one updating of the QRD. However, it is worth noting that, if $L_e - L_\delta$ is larger than $(L_\delta+1)^2$, updating the QRD of $\widehat{\boldsymbol{F}}^{(a)}_{\delta,\text{mmoe}}(k)$ from $\widehat{\boldsymbol{F}}^{(a)}_{\delta,\text{mmoe}}(k-1)$ turns out to be more expensive than computing the QRD of $\widehat{\boldsymbol{F}}^{(a)}_{\delta,\text{mmoe}}(k)$ from scratch. Hence, we can conclude that computation of $\widehat{\boldsymbol{Q}}_{\delta,\text{mmoe}}(k)$ requires an overall complexity of $\mathcal{O}\left\{N^2(L_e-L_\delta)\cdot\min\left[L_e-L_\delta,(L_\delta+1)^2\right]\right\}$ flops.

In the third step, starting from the samples of $\widehat{\boldsymbol{\jmath}}(k) \triangleq \widehat{\boldsymbol{Q}}^{H}_{\delta,\text{mmoe}}(k)\boldsymbol{z}(k)$, for $k \in \{0, 1, \ldots, K_c-1\}$, tracking of the most dominant eigenvector of the autocorrelation matrix of $\widehat{\boldsymbol{\jmath}}(k)$ has to be performed. To this aim, we adopt the projection approximation subspace tracking (PAST) algorithm [40], which admits a RLS implementation and, moreover, assures almost sure global convergence to the most dominant eigenvector $\boldsymbol{\zeta}_{\delta,\text{opt}}$ of $\boldsymbol{R}_{\boldsymbol{\jmath}\boldsymbol{\jmath}}$. The PAST algorithm for updating $\widehat{\boldsymbol{\zeta}}_{\delta,\text{opt}}$ is given [40] by

$$\begin{aligned}\widehat{\boldsymbol{\zeta}}_{\delta,\text{opt}}(k) = {} & \widehat{\boldsymbol{\zeta}}_{\delta,\text{opt}}(k-1)\\ & +\mu^{-1}(k)\left[\widehat{\boldsymbol{\jmath}}(k) - \widehat{\boldsymbol{\zeta}}_{\delta,\text{opt}}(k-1)\psi(k)\right]\psi^*(k) \quad (63)\end{aligned}$$

where $\psi(k) \triangleq \widehat{\boldsymbol{\zeta}}^{H}_{\delta,\text{opt}}(k-1)\widehat{\boldsymbol{\jmath}}(k)$ and $\mu(k) \triangleq \lambda_2\mu(k-1) + |\psi(k)|^2$, with $\lambda_2 \in (0,1]$ denoting the forgetting factor of the recursive algorithm. As suggested in [40], the simplest way of choosing the initial values $\mu(-1)$ and $\widehat{\boldsymbol{\zeta}}_{\delta,\text{opt}}(-1)$ is to set $\mu(-1) = 1$ and $\widehat{\boldsymbol{\zeta}}_{\delta,\text{opt}}(-1) = [1, 0, \ldots, 0]^T$. Remarkably, the recursive rule (63) requires only $\mathcal{O}[N(L_\delta+1)]$ flops per iteration.

In conclusion, we can state that, in practical settings wherein $\delta = L_{cp}$ (which implies that $L_\delta = L_{cp}$) and L_e is significantly larger then L_{cp}, the overall adaptive implementation of the proposed method involves a computational burden of only $\mathcal{O}\left[N^2(L_e-L_{cp})^2\right]$ flops per iteration. For instance, for the representative parameters used in Section V, wherein $N=2$, $L_{cp}=4$ and $L_e=18$, the proposed adaptive technique requires $\mathcal{O}(784)$ flops per iteration.

V. Numerical Performance Assessment

In the following, we present the results of Monte Carlo computer simulations. Specifically, in all the examples, we consider an MC-SIMO transceiver employing $M = 64$ subcarriers with QPSK signaling and a CP of length $L_{cp} = 4$,

$$\widehat{\boldsymbol{F}}_{\delta,\text{mmoe}}(k) = \underbrace{\left[\boldsymbol{\Theta}_\delta \quad -\boldsymbol{\Pi}_\delta\right]\begin{bmatrix}\mathbf{O}_{N(L_\delta+1)\times N(L_\delta+1)} & \mathbf{I}_{N(L_\delta+1)}\\ \widehat{\boldsymbol{Q}}^{(a)}_{\delta,\text{mmoe}}(k) & \mathbf{O}_{N(L_e-L_\delta)\times N(L_\delta+1)}\end{bmatrix}}_{\widehat{\boldsymbol{\Xi}}^{(a)}_{\delta,\text{mmoe}}(k)\in\mathbb{C}^{N(L_e+1)\times 2N(L_\delta+1)}}\begin{bmatrix}\widehat{\boldsymbol{R}}^{(a)}_{\delta,\text{mmoe}}(k)\\ \mathbf{I}_{N(L_\delta+1)}\end{bmatrix} \quad (61)$$

which works at half-sampling spacing $T_c/2$ (i.e., $N = 2$). The system operates over a random FIR channel of order $L_h = 14$, whose taps $h_0(0), h_1(0), h_0(1), h_1(1), \ldots, h_0(14), h_1(14)$ are modeled as i.i.d. complex proper zero-mean Gaussian random variables, with variance $\sigma_h^2 = 2$. We consider either an NBI-free scenario, wherein $\boldsymbol{v}(k) = \boldsymbol{v}_{\text{noise}}(k)$, or an NBI-contaminated environment, wherein $\boldsymbol{v}(k) = \boldsymbol{v}_{\text{nbi}}(k) + \boldsymbol{v}_{\text{noise}}(k)$. In the latter case, the continuous-time NBI signal is modeled as $w_{c,\text{nbi}}(t) = i(t) e^{j 2\pi f_I t}$, where $i(t)$ is a zero-mean WSS circular Gaussian process, with autocorrelation function $r_{ii}(\tau) \triangleq E\left[i(t)\, i^*(t-\tau)\right] = \sigma_I^2\, e^{-a|\tau|}$, where σ_I^2 is the NBI power, $a > 0$ can be related to the 3-dB NBI one-sided bandwidth W_{nbi} by the relationship $W_{\text{nbi}} \approx 0.1a$, and f_I is the NBI carrier frequency-offset. According to assumption (a4), the additive noise $\boldsymbol{v}_{\text{noise}}(k)$ is modeled as a zero-mean WSS circular white Gaussian random vector, with autocorrelation matrix $\boldsymbol{R}_{\text{noise}} = \sigma_v^2\, \mathbf{I}_{N(L_e+1)}$. The SNR is defined[8] as σ_s^2/σ_v^2, whereas, for the NBI-contaminated scenario, the SIR is defined as σ_s^2/σ_I^2, and the parameters a and f_I are set equal to $0.05(N/T_c)$ (corresponding to $W_{\text{nbi}} \approx 0.01/T_c$) and 31.5 (N/T_c), respectively.

Since the approach of [25] is implemented[9] through batch processing [the resulting TEQ is referred to as "B-ZF (batch)"], we initially implement the proposed TEQ in its batch version [referred to as "Proposed (batch)"]. Soon thereafter, we also provide a detailed performance analysis of the adaptive implementation of our channel-shortening method. In addition to the data-estimated versions of the two blind TEQs under comparison, we report the performances of their exact counterparts [referred to as "B-ZF (exact)" and "Proposed (exact)"], which have exact knowledge of the correlation matrices involved in the algorithms. This choice allows us not only to carry out a meaningful comparison with the method of [25], but also to accurately determine the ultimate performance penalty with respect to the ideal nonblind maximum-SSINR channel shortener [referred to as "Max-SSINR (ideal)"], which has exact knowledge of the channel matrix $\boldsymbol{H}$ and the disturbance autocorrelation matrix $\boldsymbol{R_{vv}}$. Moreover, for the sake of comparison, we additionally report in some plots the performance of the FEQ receiver without a channel-shortening equalizer [referred to as "w/o TEQ"], which is obtained by setting $\boldsymbol{f} = [1, 0, \ldots, 0]^T$. For both the maximum-SSINR equalizer and the method of [25], we set $\Delta = 0$ and $L_{\text{eff}} = L_{cp}$. Regarding the proposed method, we consider parameterization P1 and we set $\delta = L_{cp}$, which, according to Theorem 1, leads to $\Delta = 0$ and $L_{\text{eff}} = L_{cp}$.

As performance measure, in addition to the SSINR at the output of the considered TEQs, which is defined in (11), we resort to the bit-error-rate (BER) at the output of their corresponding FEQs, which is defined as $\text{BER} \triangleq \sum_{m=0}^{M-1} \text{BER}_m / M$, where BER_m is the output BER at the mth subcarrier. All the FEQs under comparison are synthesized by assuming exact knowledge of the channel impulse response shortened by the corresponding blind TEQ, i.e., perfect knowledge of matrix $\boldsymbol{\mathcal{G}}$ in (16) is assumed.[10] For each Monte Carlo trial, after estimating the TEQ weights on the basis of the given data record of length K (expressed in OFDM symbols), an independent record of $K_{\text{aber}} = 200$ OFDM symbols is considered to evaluate the BER at the FEQ output of the considered receivers. All the results are obtained by carrying out 10^4 independent trials, with each run using a different set of symbols, disturbance, and channel realizations.

Example 1—Average SSINR and BER Versus SNR (NBI-Free Environment): Preliminarily, we study the performances of the considered receivers in the absence of NBI, as a function of the SNR. For all the TEQs under comparison, the equalizer order L_e is kept constant to 18, and the sample size is set equal to $K = L_e + 1 = 19$ OFDM symbols (corresponding to $K_c = K\,P = 1292$ sampling intervals). Let us first consider the average SSINR curves of the considered receivers, which are reported in Fig. 2. Remarkably, it can be observed that, except for very low values of the SNR, the exact version of the proposed blind TEQ essentially achieves the same performance of the ideal nonblind maximum-SINR channel shortener, by outperforming the "B-ZF (exact)" equalizer for all the considered SNR values. This performance loss of the "B-ZF (exact)" equalizer is due to the fact that its synthesis is based on a ZF criterion, which does not explicitly take into account the presence of noise. Furthermore, it is seen that, with respect to our channel shortener, the performance loss of the method proposed in [25] increases when the equalizers are estimated in batch mode from the received data, since in this case the performance of the "B-ZF (batch)" equalizer is additionally affected by channel-order estimation errors; in particular, in the high SNR region, the performance gap between the "Proposed (batch)" TEQ and the "B-ZF (batch)" one is about 2.5 dB. Regarding the average BER performances of the considered receivers, which are reported in Fig. 3, it is immediately apparent that, in the considered scenario, the detection process is completely unreliable without a TEQ. Overall, results show that the "Proposed (batch)" receiver largely outperforms the "B-ZF (batch)" one: indeed, the former exhibits only a slight performance degradation with respect to the exact receivers which, in their turns, achieve substantially the same performances for SNR > 14 dB; on the other hand, the latter exhibits a high BER floor of about $2 \cdot 10^{-2}$, as the SNR goes up. It should be observed that, compared with the proposed data-estimated channel shortener, the performance penalty paid by the "B-ZF (batch)" TEQ is more evident in terms of average BER than in terms of average SSINR. This result stems from the fact that there is a highly nonlinear relationship between the SSINR at the TEQ output and the BER at the FEQ output.

Example 2—Average SSINR and BER Versus K (NBI-Free Environment): In this example, we investigate the performances of the receivers under comparison in the absence of NBI, as a function of the sample size K, in terms of both average SSINR (see Fig. 4) and average BER (see Fig. 5). For the purpose of

[8]Herein, the SNR is defined as the ratio between the average energy per symbol $E[\|\boldsymbol{W}_{\text{IDFT}}\boldsymbol{s}(n)\|^2]/M = \sigma_s^2$ expended by the transmitter and the noise variance σ_v^2, and it should not be confused with the SNR at the TEQ input. According to (3), the SNR at the TEQ input is defined as $\text{SNR}_{\text{teq-in}} \triangleq E\left[\|\sum_{i=0}^{L_h} \boldsymbol{h}(i) u(k-i)\|^2\right] / E\left[\|\boldsymbol{w}(k)\|^2\right]$, which, for the considered simulation setting, is given by $\text{SNR}_{\text{teq-in}} = \text{SNR}[\sigma_h^2(L_h+1)/N] = 15 \cdot \text{SNR}$. Thus, it results that $(\text{SNR}_{\text{teq-in}})_{\text{dB}} = (\text{SNR})_{\text{dB}} + 11.76$.

[9]As already pointed out, we consider method A, whose implementation is detailed in [25, steps 1–7, pp. 3258–3259].

[10]After performing blind channel shortening, the shortened channel impulse response can be estimated at the TEQ output by using standard blind [16]–[18] and/or nonblind [41] techniques.

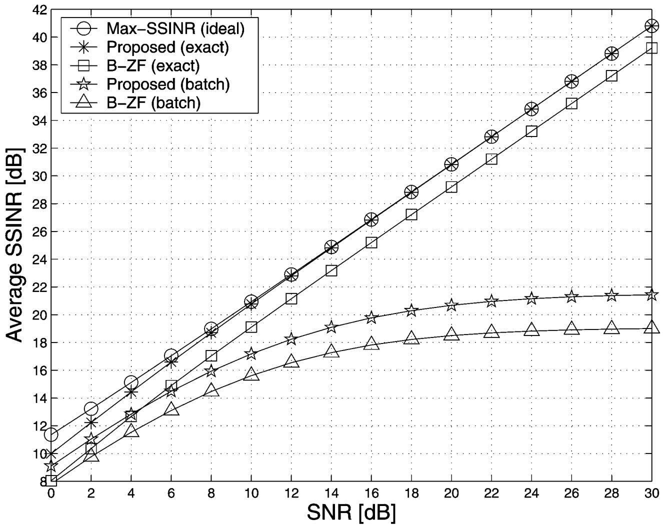


Fig. 2. Average SSINR versus SNR ($L_e = 18$, $K = 19$ OFDM symbols, NBI-free environment).

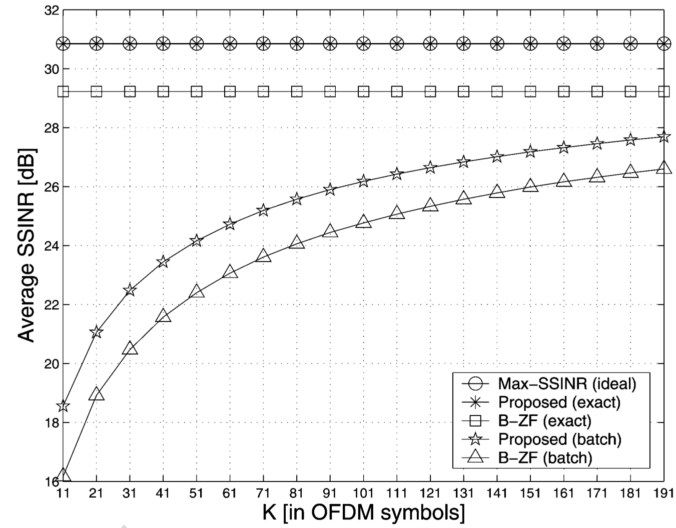


Fig. 4. Average SSINR versus K ($L_e = 18$, SNR $= 20$ dB, NBI-free environment).

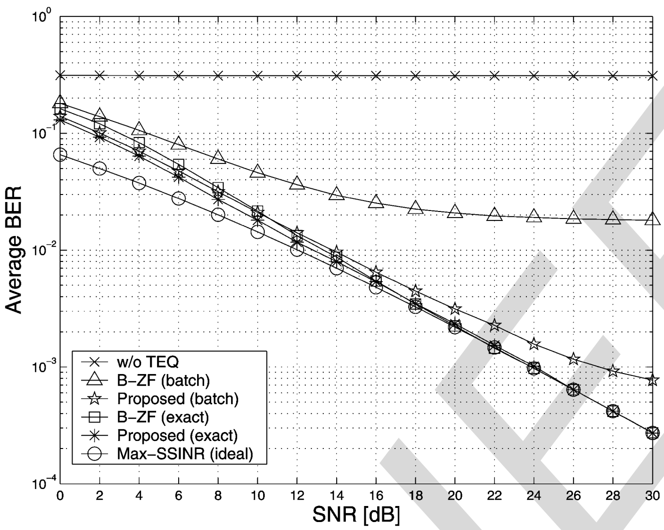


Fig. 3. Average BER versus SNR ($L_e = 18$, $K = 19$ OFDM symbols, NBI-free environment).

comparison, we also report the performances of the exact and ideal receivers, which do not depend on K. The order of all the considered equalizer is $L_e = 18$ and the SNR is 20 dB. It can be seen from Fig. 4 that the average SSINR performances of both the "Proposed (batch)" and "B-ZF (batch)" equalizers improve as the sample size K increases, by tending to the curves of their corresponding exact counterparts; in particular, at 24 dB average SSINR, the "B-ZF (batch)" TEQ requires a sample size of about 30 symbols larger than that required by the proposed channel shortener. Additionally, results of Fig. 5 evidence that, in terms of average BER, the "Proposed (batch)" receiver pays only a slight performance penalty with respect to both its exact counterpart and the ideal nonblind maximum-SSINR equalizer, by achieving almost the same performance for $K > 51$ OFDM symbols. In contrast, as K increases, the performance of the "B-ZF (batch)" equalizer improves very slowly, without significantly approaching the curve of the "B-ZF (exact)" receiver.

Example 3—Average SSINR and BER Versus SIR (NBI-Contaminated Environment): At this point, we evaluate the performances of all the considered receivers in the presence of NBI, as a function of the SIR. The SNR is 20 dB, the equalizers' order is $L_e = 24$, and the sample size is set equal to $K = L_e + 1 = 25$ OFDM symbols (corresponding to $K_c = K\,P = 1700$ sampling intervals). For a fair comparison, we simulate a *modified* version of the "B-ZF (exact)" equalizer, for which the dimension $L_g + 1$ of the signal subspace (denoted with N' in [25]) is augmented by the rank $R_{\rm nbi}$ of the NBI (i.e., the equalizer is synthesized by setting $N' = L_g + R_{\rm nbi} + 1$).[11] with $R_{\rm nbi} = 6$. With reference to the average SSINR, it can be noted from Fig. 6 that the exact version of the proposed equalizer pays an acceptable performance penalty with respect to the ideal nonblind maximum-SINR receiver, while outperforming the "B-ZF (exact)" TEQ, for all the considered SIR values. When the channel shorteners are estimated from the received data, in comparison with our receiver, the performance loss of the equalizer [25] becomes more significant for low-to-moderate values of the SIR; specifically, at 10 dB SIR, the performance gap between the "Proposed (batch)" TEQ and the "B-ZF (batch)" one is about 5 dB. The robustness of the proposed TEQ against NBI is further corroborated in terms of average BER by results of Fig. 7, which show that, even in the presence of a strong NBI signal, the "Proposed (batch)" channel shortener performs close to its exact counterpart.

Example 4—Average SSINR Versus Number of Iterations k (NBI-Free Environment): To assess the learning capabilities of the adaptive version of the proposed method in the absence of NBI, we report in Fig. 8 its average SSINR performance as a function of the number of iterations k (expressed in sampling intervals), for different values of the SNR, with L_e set to 18. It should be stressed that, since the method of [25] does not straightforwardly lend itself to an adaptive implementation, it

[11]Since the "B-ZF (batch)" equalizer estimates N' directly from the received data, it automatically uses an estimate of the dimension of the signal-plus-interference subspace, and, thus, its synthesis does not require any modification.

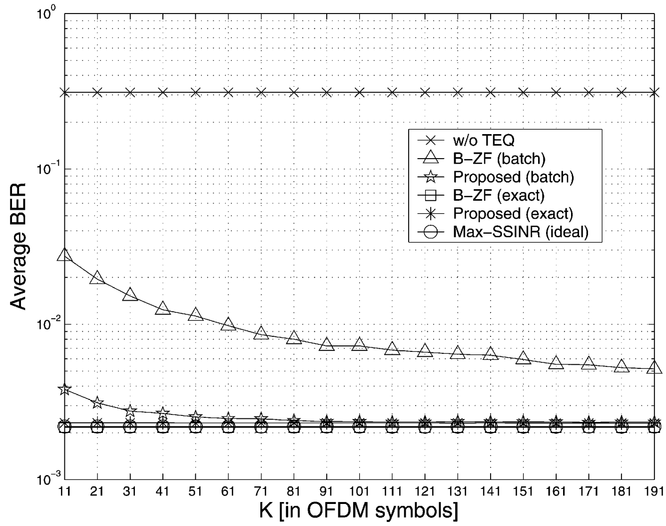


Fig. 5. Average BER versus K ($L_e = 18$, SNR $= 20$ dB, NBI-free environment).

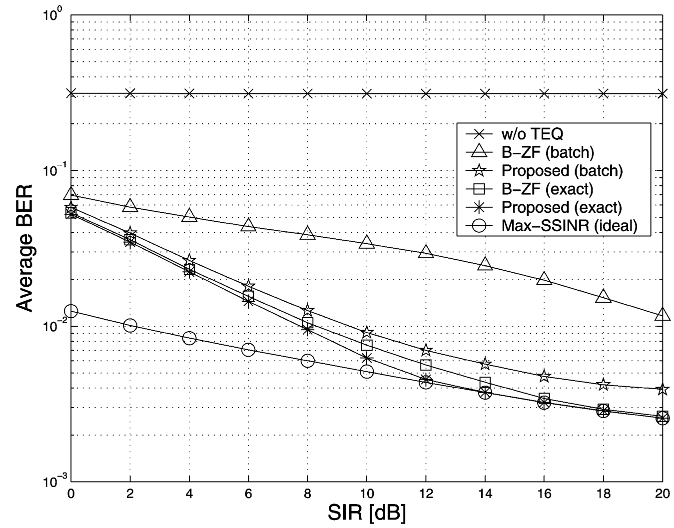


Fig. 7. Average BER versus SIR (SNR $= 20$, $L_e = 24$, $K = 25$ OFDM symbols, NBI-contaminated environment).

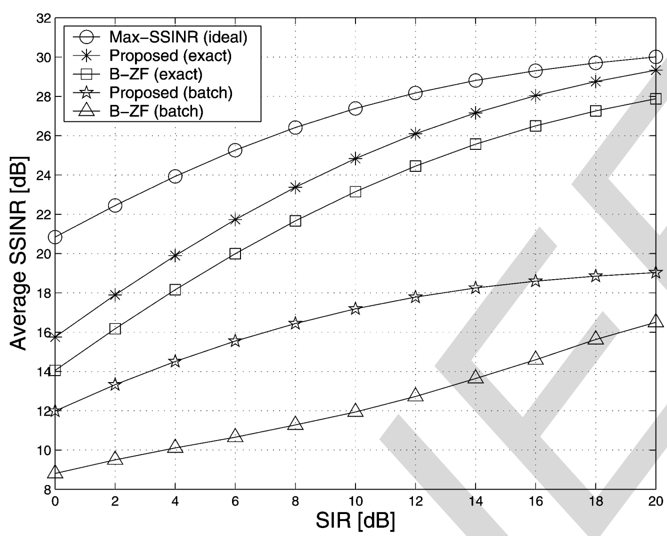


Fig. 6. Average SSINR versus SIR (SNR $= 20$, $L_e = 24$, $K = 25$ OFDM symbols, NBI-contaminated environment).

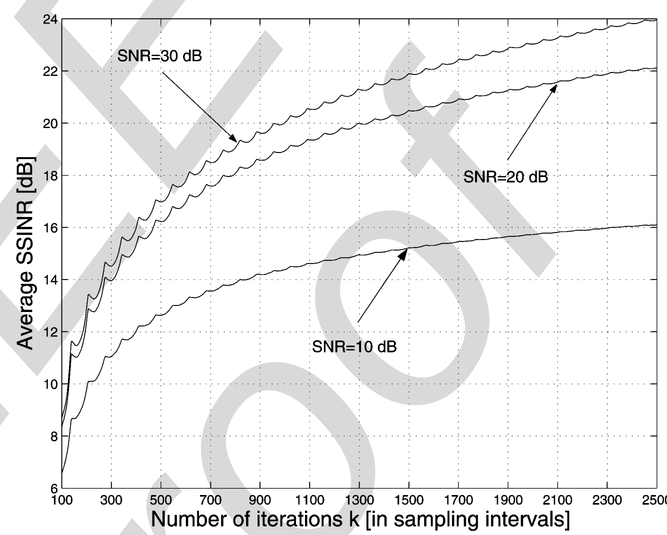


Fig. 8. Average SSINR of the proposed adaptive TEQ versus number of iterations k ($L_e = 18$, NBI-free environment).

will not be considered anymore. Regarding the RLS implementation of our channel shortener, we choose $\lambda_1 = \lambda_2 = 1$ and $\delta = 1$. It is apparent that the proposed adaptive channel-shortening algorithm exhibits a satisfactory convergence speed for all the considered SNR values, by achieving very good steady-state performance. In particular, at 20 dB SNR, about 1300 iterations (corresponding to about 19 OFDM symbols) are required in order to achieve an average SSINR of 20 dB.

Example 5—Average SSINR Versus Number of Iterations k (NBI-Contaminated Environment): In this last example, we study the learning capabilities of proposed adaptive method in the presence of NBI, as a function of the number of iterations k (expressed in sampling intervals), for different values of the SIR, with SNR kept constant to 20 dB and $L_e = 24$. As in the previous example, we choose $\lambda_1 = \lambda_2 = 1$ and $\delta = 1$. Results of Fig. 9 show that the presence of a strong NBI signal does not significantly affect the convergence speed of the proposed adaptive algorithm, which rapidly achieves satisfactory steady-state performance in terms of average SSINR.

VI. Conclusion

The problem of channel shortening for multicarrier systems has been considered in this paper. Specifically, a blind channel-shortening design has been introduced based on the MMOE criterion. The blindness of the proposed TEQ relies on the fact that oversampling of the received signal confers to the channel matrix a quite rich structure. Such a structure can be exploited to preserve the desired signal contribution, without requiring neither the *a priori* knowledge of the channel impulse response nor the estimation of its order. A theoretical analysis has also been presented in order to enlighten both

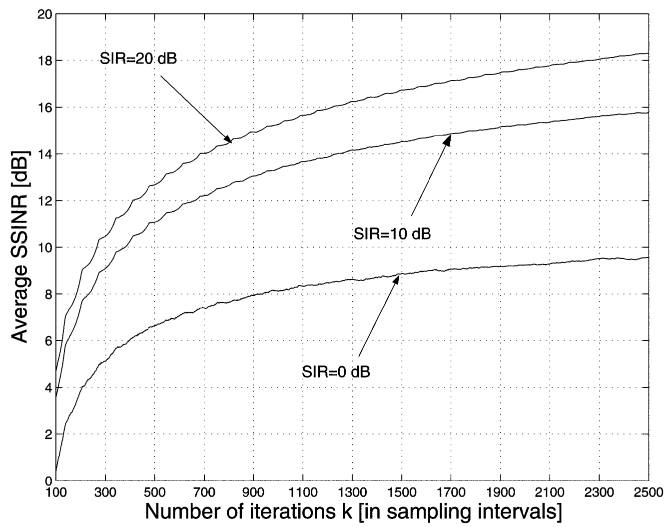


Fig. 9. Average SSINR of the proposed adaptive TEQ versus number of iterations k ($\text{SNR} = 20$, $L_e = 24$, NBI-contaminated environment).

the channel-shortening process and the NBI suppression capability of the MMOE-based TEQ. A direct consequence of this analysis is the synthesis of a blind procedure for optimally preserving the desired signal contribution. Simulation results show that, when the SOS of received data are exactly known, the proposed blind TEQ performs very close to the ideal nonblind maximum-SSINR equalizer. On the other hand, when SOS are estimated in batch mode from a finite number of samples of the received signal, the MMOE-based channel shortener outperforms the competing alternative [25], especially in terms of average BER at the FEQ output, by requiring a lower computational complexity. Finally, an effective adaptive implementation of the proposed TEQ has been also developed to allow it to be employed in real-time applications.

APPENDIX
PROOF OF THEOREM 1

First, a characterization of the structure of $\mathbf{\Pi}_\delta^T \boldsymbol{\mathcal{H}} = \left[\mathbf{\Pi}_\delta^T \boldsymbol{H}, \mathbf{\Pi}_\delta^T \boldsymbol{J}\right] \in \mathbb{C}^{N(L_e-L_\delta)\times(L_g+R_{\rm nbi}+1)}$ is needed. To this end, we observe that, accounting for (26), the $(d+1)$th column of $\mathbf{\Pi}_\delta^T \boldsymbol{H}$ is given by $\mathbf{\Pi}_\delta^T \boldsymbol{h}_d = \mathbf{\Pi}_\delta^T \mathbf{\Theta}_d \boldsymbol{\xi}_d$, for $d \in \{0, 1, \ldots, L_g\}$, and, hence, by remembering that $\mathcal{R}(\mathbf{\Pi}_\delta) \equiv \mathcal{R}^\perp(\mathbf{\Theta}_\delta)$, we can argue that the null space of $\mathbf{\Pi}_\delta^T \boldsymbol{H}$ strongly depends on the particular parameterization used, i.e., on the values of the delay δ. For this reason, we distinguish the following three different cases. See (64)–(66), shown at the bottom of the page. In the first case, when the parameterization P1 is used, that is, $\delta \in \{0, 1, \ldots, \min[L_e, L_h]\}$, it is readily verified that $\mathcal{R}^\perp(\mathbf{\Theta}_\delta) \supseteq \mathcal{R}^\perp(\mathbf{\Theta}_{\delta-i})$, for any $i \in \{1, 2, \ldots, \delta\}$. Let the channel matrix be partitioned as $\boldsymbol{H} = [\boldsymbol{H}_{\rm left}, \boldsymbol{H}_{\rm right}]$, with $\boldsymbol{H}_{\rm left} \in \mathbb{C}^{N(L_e+1)\times(\delta+1)}$ and $\boldsymbol{H}_{\rm right} \in \mathbb{C}^{N(L_e+1)\times(L_g-\delta)}$, the previous subspace inclusion implies that all the entries of the first $\delta+1$ columns of $\mathbf{\Pi}_\delta^T \boldsymbol{H}$ are zero, thus yielding $\mathbf{\Pi}_\delta^T \boldsymbol{\mathcal{H}} = \left[\mathbf{O}_{N(L_e-L_\delta)\times(\delta+1)}, \mathbf{\Pi}_\delta^T \boldsymbol{\mathcal{H}}_{\rm right}\right]$, with $\boldsymbol{\mathcal{H}}_{\rm right} \triangleq [\boldsymbol{H}_{\rm right}, \boldsymbol{J}] \in \mathbb{C}^{N(L_e+1)\times(L_g+R_{\rm nbi}-\delta)}$. Henceforth, using the fact [32] that $\left[\mathbf{O}_{N(L_e-L_\delta)\times(\delta+1)}, \mathbf{\Pi}_\delta^T \boldsymbol{\mathcal{H}}_{\rm right}\right]^\dagger = \left[\mathbf{O}_{N(L_e-L_\delta)\times(\delta+1)}, \{\left(\mathbf{\Pi}_\delta^T \boldsymbol{\mathcal{H}}_{\rm right}\right)^\dagger\}^T\right]^T$, after some straightforward algebraic manipulations, the asymptotic expression (43) of $\boldsymbol{b}_{\delta,\rm mmoe}$ is given by (64), where the second equality comes from the fact that, if the matrix $\mathbf{\Pi}_\delta^T \boldsymbol{\mathcal{H}}_{\rm right}$ is full-column rank (in this case, it turns out that $\boldsymbol{\mathcal{H}}_\delta = \boldsymbol{\mathcal{H}}_{\rm right}$, with $R_\delta = L_g + R_{\rm nbi} - \delta$), then $\boldsymbol{\mathcal{H}}_{\rm right}^H \mathbf{\Pi}_\delta (\boldsymbol{\mathcal{H}}_{\rm right}^H \mathbf{\Pi}_\delta)^\dagger = \mathbf{I}_{L_g+R_{\rm nbi}-\delta}$. From (64), it turns out that $\overline{\boldsymbol{g}}_{\delta,\rm mmoe} = \left[(\boldsymbol{H}_{\rm left}^H \mathbf{\Theta}_\delta \boldsymbol{\gamma}_\delta)^T, \mathbf{0}_{L_g-\delta}^T\right]^T$ and $\overline{\boldsymbol{\ell}}_{\delta,\rm mmoe} = \mathbf{0}_{R_{\rm nbi}}$.

In the second case, when the parameterization P2 is used, that is, $\delta \in \{L_h+1, L_h+2, \ldots, L_e-1\}$, contrary to the previous case, there are no relationships between the orthogonal complements of the column spaces of the parameterization matrices. Therefore, only the entries of the $(\delta+1)$th column of $\mathbf{\Pi}_\delta^T \boldsymbol{H} = \mathbf{\Pi}_\delta^T [\boldsymbol{H}_{\rm left}, \mathbf{0}_{N(L_e+1)}, \boldsymbol{H}_{\rm right}]$ are zero, with $\boldsymbol{H}_{\rm left} \in \mathbb{C}^{N(L_e+1)\times\delta}$ and $\boldsymbol{H}_{\rm right} \in \mathbb{C}^{N(L_e+1)\times(L_g-\delta)}$, thus leading to the following partition $\mathbf{\Pi}_\delta^T \boldsymbol{\mathcal{H}} = \left[\mathbf{\Pi}_\delta^T \boldsymbol{H}_{\rm left}, \mathbf{0}_{N(L_e-L_\delta)}, \mathbf{\Pi}_\delta^T \boldsymbol{\mathcal{H}}_{\rm right}\right]$, where $\boldsymbol{\mathcal{H}}_{\rm right} \triangleq [\boldsymbol{H}_{\rm right}, \boldsymbol{J}] \in \mathbb{C}^{N(L_e+1)\times(L_g+R_{\rm nbi}-\delta)}$. By using some results on the Moore–Penrose inverse of partitioned matrices [32], it can be proven that $\left[\mathbf{\Pi}_\delta^T \boldsymbol{H}_{\rm left}, \mathbf{0}_{N(L_e-L_\delta)}, \mathbf{\Pi}_\delta^T \boldsymbol{\mathcal{H}}_{\rm right}\right]^\dagger =$

$$\overline{\boldsymbol{b}}_{\delta,\rm mmoe} = \begin{bmatrix} \boldsymbol{H}_{\rm left}^H \mathbf{\Theta}_\delta \boldsymbol{\gamma}_\delta \\ \left\{\mathbf{I}_{L_g+R_{\rm nbi}-\delta} - \boldsymbol{\mathcal{H}}_{\rm right}^H \mathbf{\Pi}_\delta (\boldsymbol{\mathcal{H}}_{\rm right}^H \mathbf{\Pi}_\delta)^\dagger\right\} \boldsymbol{\mathcal{H}}_{\rm right}^H \mathbf{\Theta}_\delta \boldsymbol{\gamma}_\delta \end{bmatrix} = \begin{bmatrix} \boldsymbol{H}_{\rm left}^H \mathbf{\Theta}_\delta \boldsymbol{\gamma}_\delta \\ \mathbf{0}_{L_g+R_{\rm nbi}-\delta} \end{bmatrix} \tag{64}$$

$$\overline{\boldsymbol{b}}_{\delta,\rm mmoe} = \begin{bmatrix} \left\{\mathbf{I}_\delta - \boldsymbol{H}_{\rm left}^H \mathbf{\Pi}_\delta (\boldsymbol{H}_{\rm left}^H \mathbf{\Pi}_\delta)^\dagger\right\} \boldsymbol{H}_{\rm left}^H \mathbf{\Theta}_\delta \boldsymbol{\gamma}_\delta - \boldsymbol{H}_{\rm left}^H \mathbf{\Pi}_\delta (\boldsymbol{\mathcal{H}}_{\rm right}^H \mathbf{\Pi}_\delta)^\dagger \boldsymbol{\mathcal{H}}_{\rm right}^H \mathbf{\Theta}_\delta \boldsymbol{\gamma}_\delta \\ \boldsymbol{\xi}_\delta^H \boldsymbol{\gamma}_\delta \\ \left\{\mathbf{I}_{L_g+R_{\rm nbi}-\delta} - \boldsymbol{\mathcal{H}}_{\rm right}^H \mathbf{\Pi}_\delta (\boldsymbol{\mathcal{H}}_{\rm right}^H \mathbf{\Pi}_\delta)^\dagger\right\} \boldsymbol{\mathcal{H}}_{\rm right}^H \mathbf{\Theta}_\delta \boldsymbol{\gamma}_\delta - \boldsymbol{\mathcal{H}}_{\rm right}^H \mathbf{\Pi}_\delta (\boldsymbol{H}_{\rm left}^H \mathbf{\Pi}_\delta)^\dagger \boldsymbol{H}_{\rm left}^H \mathbf{\Theta}_\delta \boldsymbol{\gamma}_\delta \end{bmatrix} \tag{65}$$

$$\overline{\boldsymbol{b}}_{\delta,\rm mmoe} = \begin{bmatrix} \left\{\mathbf{I}_\delta - \boldsymbol{H}_{\rm left}^H \mathbf{\Pi}_\delta (\boldsymbol{H}_{\rm left}^H \mathbf{\Pi}_\delta)^\dagger\right\} \boldsymbol{H}_{\rm left}^H \mathbf{\Theta}_\delta \boldsymbol{\gamma}_\delta - \boldsymbol{H}_{\rm left}^H \mathbf{\Pi}_\delta (\boldsymbol{J}^H \mathbf{\Pi}_\delta)^\dagger \boldsymbol{J}^H \mathbf{\Theta}_\delta \boldsymbol{\gamma}_\delta \\ \boldsymbol{H}_{\rm right}^H \mathbf{\Theta}_\delta \boldsymbol{\gamma}_\delta \\ \left\{\mathbf{I}_{R_{\rm nbi}} - \boldsymbol{J}^H \mathbf{\Pi}_\delta (\boldsymbol{J}^H \mathbf{\Pi}_\delta)^\dagger\right\} \boldsymbol{J}^H \mathbf{\Theta}_\delta \boldsymbol{\gamma}_\delta - \boldsymbol{J}^H \mathbf{\Pi}_\delta (\boldsymbol{H}_{\rm left}^H \mathbf{\Pi}_\delta)^\dagger \boldsymbol{H}_{\rm left}^H \mathbf{\Theta}_\delta \boldsymbol{\gamma}_\delta \end{bmatrix} \tag{66}$$

$\left[\{(\boldsymbol{\Pi}_\delta^T \boldsymbol{H}_{\mathrm{left}})^\dagger\}^T, \mathbf{0}_{N(L_e-L_\delta)}, \{(\boldsymbol{\Pi}_\delta^T \boldsymbol{\mathcal{H}}_{\mathrm{right}})^\dagger\}^T\right]^T$ which, after some algebra, allows one to written the asymptotic expression (43) of $\boldsymbol{b}_{\delta,\mathrm{mmoe}}$ as shown in (65). If the matrix $\boldsymbol{\Pi}_\delta^T[\boldsymbol{H}_{\mathrm{left}}, \boldsymbol{\mathcal{H}}_{\mathrm{right}}]$ is full-column rank (in this case, it turns out that $\boldsymbol{\mathcal{H}}_\delta = [\boldsymbol{H}_{\mathrm{left}}, \boldsymbol{\mathcal{H}}_{\mathrm{right}}]$, with $R_\delta = L_g + R_{\mathrm{nbi}}$), it follows that $\boldsymbol{\Pi}_\delta^T \boldsymbol{H}_{\mathrm{left}}$ and $\boldsymbol{\Pi}_\delta^T \boldsymbol{\mathcal{H}}_{\mathrm{right}}$ are necessarily full-column rank matrices, which in their turns imply that $\boldsymbol{H}_{\mathrm{left}}^H \boldsymbol{\Pi}_\delta \left(\boldsymbol{H}_{\mathrm{left}}^H \boldsymbol{\Pi}_\delta\right)^\dagger = \mathbf{I}_\delta$ and $\boldsymbol{\mathcal{H}}_{\mathrm{right}}^H \boldsymbol{\Pi}_\delta \left(\boldsymbol{\mathcal{H}}_{\mathrm{right}}^H \boldsymbol{\Pi}_\delta\right)^\dagger = \mathbf{I}_{L_g+R_{\mathrm{nbi}}-\delta}$. In addition, the full-column rank property of the partitioned matrix $[\boldsymbol{\Pi}_\delta^T \boldsymbol{H}_{\mathrm{left}}, \boldsymbol{\Pi}_\delta^T \boldsymbol{\mathcal{H}}_{\mathrm{right}}]$ also implies that $\mathcal{R}\left(\boldsymbol{\Pi}_\delta^T \boldsymbol{H}_{\mathrm{left}}\right) \cap \mathcal{R}\left(\boldsymbol{\Pi}_\delta^T \boldsymbol{\mathcal{H}}_{\mathrm{right}}\right) = \{\mathbf{0}_{N(L_e-L_\delta)}\} \Longleftrightarrow \left(\boldsymbol{\Pi}_\delta^T \boldsymbol{\mathcal{H}}_{\mathrm{right}}\right)^H \boldsymbol{\Pi}_\delta^T \boldsymbol{H}_{\mathrm{left}} = \mathbf{O}_{(L_g+R_{\mathrm{nbi}}-\delta)\times\delta}$, from which it follows that $\boldsymbol{H}_{\mathrm{left}}^H \boldsymbol{\Pi}_\delta \left(\boldsymbol{\mathcal{H}}_{\mathrm{right}}^H \boldsymbol{\Pi}_\delta\right)^\dagger = \mathbf{O}_{\delta\times(L_g+R_{\mathrm{nbi}}-\delta)}$ and $\boldsymbol{\mathcal{H}}_{\mathrm{right}}^H \boldsymbol{\Pi}_\delta \left(\boldsymbol{H}_{\mathrm{left}}^H \boldsymbol{\Pi}_\delta\right)^\dagger = \mathbf{O}_{(L_g+R_{\mathrm{nbi}}-\delta)\times\delta}$. On the basis of these results, (65) implies that $\bar{\boldsymbol{g}}_{\delta,\mathrm{mmoe}} = \left[\mathbf{0}_\delta^T, \boldsymbol{\xi}_\delta^H \boldsymbol{\gamma}_\delta, \mathbf{0}_{L_g-\delta}^T\right]^T$ and $\bar{\boldsymbol{\ell}}_{\delta,\mathrm{mmoe}} = \mathbf{0}_{R_{\mathrm{nbi}}}$.

Let us consider at this point the last case, when the parameterization P3 is used, that is, $\delta \in \{\max[L_e, L_h], \max[L_e, L_h]+1, \ldots, L_g\}$. In this case, it can readily be demonstrated that $\mathcal{R}^\perp(\boldsymbol{\Theta}_\delta) \supseteq \mathcal{R}^\perp(\boldsymbol{\Theta}_{\delta+i})$, for any $i \in \{0, 1, \ldots, L_g - \delta\}$. Let the channel matrix be partitioned as $\boldsymbol{H} = [\boldsymbol{H}_{\mathrm{left}}, \boldsymbol{H}_{\mathrm{right}}]$, with $\boldsymbol{H}_{\mathrm{left}} \in \mathbb{C}^{N(L_e+1)\times\delta}$ and $\boldsymbol{H}_{\mathrm{right}} \in \mathbb{C}^{N(L_e+1)\times(L_g-\delta+1)}$, it turns out that all the entries of the last $L_g - \delta + 1$ columns of the matrix $\boldsymbol{\Pi}_\delta^T \boldsymbol{H}$ are zero, thus yielding $\boldsymbol{\Pi}_\delta^T \boldsymbol{\mathcal{H}} = \left[\boldsymbol{\Pi}_\delta^T \boldsymbol{H}_{\mathrm{left}}, \mathbf{O}_{N(L_e-L_\delta)\times(L_g-\delta+1)}, \boldsymbol{\Pi}_\delta^T \boldsymbol{J}\right]$. Henceforth, using the fact [32] that $\left[\boldsymbol{\Pi}_\delta^T \boldsymbol{H}_{\mathrm{left}}, \mathbf{O}_{N(L_e-L_\delta)\times(L_g-\delta+1)}, \boldsymbol{\Pi}_\delta^T \boldsymbol{J}\right]^\dagger = \left[\left\{\left(\boldsymbol{\Pi}_\delta^T \boldsymbol{H}_{\mathrm{left}}\right)^\dagger\right\}^T, \mathbf{O}_{N(L_e-L_\delta)\times(L_g-\delta+1)}, \left\{(\boldsymbol{\Pi}_\delta^T \boldsymbol{J})^\dagger\right\}^T\right]^T$, after some lengthy algebraic manipulations, the asymptotic expression (43) of $\boldsymbol{b}_{\delta,\mathrm{mmoe}}$ is given by (66). Reasoning as previously done, it can be shown that, if the matrix $\boldsymbol{\Pi}_\delta^T[\boldsymbol{H}_{\mathrm{left}}, \boldsymbol{J}]$ is full-column rank (in this case, it turns out that $\boldsymbol{\mathcal{H}}_\delta = [\boldsymbol{H}_{\mathrm{left}}, \boldsymbol{J}]$, with $R_\delta = R_{\mathrm{nbi}} + \delta + 1$), then $\boldsymbol{H}_{\mathrm{left}}^H \boldsymbol{\Pi}_\delta \left(\boldsymbol{H}_{\mathrm{left}}^H \boldsymbol{\Pi}_\delta\right)^\dagger = \mathbf{I}_\delta$, $\boldsymbol{J}^H \boldsymbol{\Pi}_\delta \left(\boldsymbol{J}^H \boldsymbol{\Pi}_\delta\right)^\dagger = \mathbf{I}_{R_{\mathrm{nbi}}}$, $\boldsymbol{H}_{\mathrm{left}}^H \boldsymbol{\Pi}_\delta \left(\boldsymbol{J}^H \boldsymbol{\Pi}_\delta\right)^\dagger = \mathbf{O}_{\delta\times R_{\mathrm{nbi}}}$ and $\boldsymbol{J}^H \boldsymbol{\Pi}_\delta (\boldsymbol{H}_{\mathrm{left}}^H \boldsymbol{\Pi}_\delta)^\dagger = \mathbf{O}_{R_{\mathrm{nbi}}\times\delta}$. Hence, (66) implies that $\bar{\mathbf{g}}_{\delta,\mathrm{mmoe}} = \left[\mathbf{0}_\delta^T, \left(\boldsymbol{H}_{\mathrm{right}}^H \boldsymbol{\Theta}_\delta \boldsymbol{\gamma}_\delta\right)^T\right]^T$ and $\bar{\boldsymbol{\ell}}_{\delta,\mathrm{mmoe}} = \mathbf{0}_{R_{\mathrm{nbi}}}$.

In order to conclude the proof, observe that, from a unified perspective, the assumption that $\boldsymbol{\Pi}_\delta^T \boldsymbol{\mathcal{H}}_\delta \in \mathbb{C}^{N(L_e-L_\delta)\times R_\delta}$ be full-column rank necessarily requires that $\boldsymbol{\mathcal{H}}_\delta$ be full-column rank, too. Thus, the condition $\mathrm{rank}(\boldsymbol{\Pi}_\delta^T \boldsymbol{\mathcal{H}}_\delta) = R_\delta$ is equivalent to impose that the null space of $\boldsymbol{\Pi}_\delta^T$ and the column space of $\boldsymbol{\mathcal{H}}_\delta$ intersect only trivially, i.e., $\mathcal{N}(\boldsymbol{\Pi}_\delta^T) \cap \mathcal{R}(\boldsymbol{\mathcal{H}}_\delta) = \{\mathbf{0}_{N(L_e+1)}\}$. At this point, observing that $\mathcal{N}(\boldsymbol{\Pi}_\delta^T) = \mathcal{R}^\perp(\boldsymbol{\Pi}_\delta)$ and recalling that $\mathcal{R}^\perp(\boldsymbol{\Pi}_\delta) = \mathcal{R}(\boldsymbol{\Theta}_\delta)$, it follows that the condition $\mathrm{rank}(\boldsymbol{\Pi}_\delta^T \boldsymbol{\mathcal{H}}_\delta) = R_\delta$ can be equivalently expressed as $\mathcal{R}(\boldsymbol{\Theta}_\delta) \cap \mathcal{R}(\boldsymbol{\mathcal{H}}_\delta) = \{\mathbf{0}_{N(L_e+1)}\}$, which is the condition advocated in (45).

ACKNOWLEDGMENT

The authors would like to thank Prof. G. Gelli for his careful reading of the manuscript.

REFERENCES

[1] J. A. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, vol. 28, no. 5, pp. 5–14, May 1990.

[2] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications—Where Fourier meets Shannon," *IEEE Signal Process. Mag.*, vol. 17, no. 5, pp. 29–48, May 2000.

[3] T. Starr, J. M. Cioffi, and P. J. Silvermann, *Understanding Digital Subscriber Line Technology*. Englewood Cliffs, NJ: Prentice-Hall, 1999.

[4] HomePlug 1.0 Specification [Online]. Available: http://www.homeplug.org

[5] IEEE 802.11 Working Group for Wireless Local Area Networks [Online]. Available: http://grouper.ieee.org/groups/802/11

[6] High Performance Radio Local Area Networks (HIPERLAN) Type 2 ETSI Normalization Committee [Online]. Available: http://www.etsi.org

[7] Radio Broadcasting Systems, Digital Audio Broadcasting (DAB) to Mobile, Portible, and Fixed Receivers ETSI Normalization Committee [Online]. Available: http://www.etsi.org

[8] Digital Video Broadcasting (DVB): Framing Structure, Channel Coding and Modulation for Digital Terrestrial Television ETSI Normalization Committee [Online]. Available: http://www.etsi.org

[9] P. J. W. Melsa, R. C. Younce, and C. E. Rohrs, "Impulse response shortening for discrete multitone transceivers," *IEEE Trans. Commun.*, vol. 44, no. 12, 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, no. 7, pp. 1069–1072, Jul. 2003.

[12] N. Al-Dhahir and J. M. Cioffi, "Optimum finite-length equalization for multicarrier transceivers," *IEEE Trans. Commun.*, vol. 44, no. 1, 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, no. 1, 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, no. 12, pp. 3123–3135, Dec. 2001.

[15] S. Celebi, "Interblock interference (IBI) minimizing time-domain equalizer (TEQ) for OFDM," *IEEE Signal Process. Lett.*, vol. 10, no. 8, pp. 232–234, Aug. 2003.

[16] B. Muquet, M. de Courville, and P. Duhamel, "Subspace-based blind and semi-blind channel estimation for OFDM systems," *IEEE Trans. Signal Process.*, vol. 50, no. 7, pp. 1699–1712, Jul. 2002.

[17] S. Roy and C. Li, "A subspace blind channel estimation method for OFDM systems without cyclic prefix," *IEEE Trans. Wireless Commun.*, vol. 1, no. 10, pp. 572–579, Oct. 2002.

[18] C. Li and S. Roy, "Subspace-based blind channel estimation method for OFDM by exploiting virtual carriers," *IEEE Trans. Wireless Commun.*, vol. 2, no. 1, pp. 141–150, Jan. 2003.

[19] 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, no. 11, pp. 341–343, Nov. 2002.

[20] 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, no. 12, pp. 3086–3093, Dec. 2003.

[21] R. K. Martin, J. M. Walsh, and C. R. Johnson, "Low-complexity MIMO blind adaptive channel shortening," *IEEE Trans. Signal Process.*, vol. 53, no. 4, pp. 1324–1334, Apr. 2005.

[22] D. Pal, G. N. Iyengar, and J. M. Cioffi, "A new method of channel shortening with applications to discrete multitone (DMT) systems," in *Proc. IEEE Int. Conf. Communication*, Atlanta, GA, Jun. 1998, pp. 763–768.

[23] M. de Courville, P. Duhamel, P. Madec, and J. Palicot, "Blind equalization of OFDM systems based on the minimization of a quadratic criterion," in *Proc. IEEE Int. Conf. Communication*, Dallas, TX, Jun. 1996, pp. 1318–1321.

[24] F. Romano and S. Barbarossa, "Non-data aided adaptive channel shortening for efficient multicarrier systems," in *Proc. IEEE Int. Conf. Acoustic, Speech, and Signal Processing*, Hong Kong, China, Apr. 2003, pp. 233–236.

[25] T. Miyajima and Z. Ding, "Second-order statistical approaches to channel shortening in multicarrier systems," *IEEE Trans. Signal Process.*, vol. 52, no. 11, pp. 3253–3264, Nov. 2004.

[26] H. L. Van Trees, *Optmimum Array Processing*. New York: Wiley, 2002.

[27] M. K. Tsatsanis and Z. D. Xu, "Constrained optimization methods for direct blind equalization," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 424–433, Mar. 1999.

[28] G. Gelli and F. Verde, "Two-stage interference-resistant adaptive periodically time-varying CMA blind equalization," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 662–672, Mar. 2002.

[29] M. Honig, U. Madhow, and S. Verdù, "Blind adaptive multiuser detection," *IEEE Trans. Inf. Theory*, vol. 41, no. 7, pp. 944–960, Jul. 1995.

[30] M. K. Tsatsanis and Z. D. Xu, "Performance analysis of minimum variance cdma receivers," *IEEE Trans. Signal Process.*, vol. 46, no. 11, pp. 3014–3022, Nov. 1998.

[31] G. Gelli, L. Paura, and F. Verde, "A two-stage cma-based receivers for blind joint equalization and multiuser detection in high data-rate DS-CDMA systems," *IEEE Trans. Wireless Commun.*, vol. 3, no. 7, pp. 1209–1223, Jul. 2004.

[32] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses*. New York: Springer-Verlag, 2002.

[33] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, New Jersey: Prentice-Hall, 1993.

[34] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press.

[35] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD: John Hopkins Univ. Press, 1996.

[36] E. Serpedin and G. B. Giannakis, "A simple proof of a known blind channel identifiability result," *IEEE Trans. Signal Process.*, vol. 47, no. 2, pp. 591–593, Feb. 1999.

[37] Z. Ding, "Characterictics of band-limited channels unidentifiable from second-order cyclostationary statistics," *IEEE Signal Process. Lett.*, vol. 3, no. 5, pp. 150–152, May 1996.

[38] S. Haykin, *Adaptive Filter Theory*. New York: Prentice-Hall, 1996.

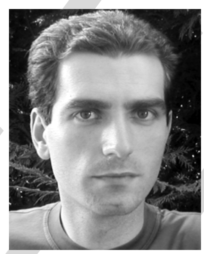
[39] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis, and uncertaintly—V: The discrete case," *Bell Syst. Tech. J.*, vol. 47, pp. 1371–1430, May 1978.

[40] B. Yang, "Projection approximation subspace tracking," *IEEE Trans. Signal Process.*, vol. 43, no. 1, pp. 95–107, Jan. 1995.

[41] M. Morelli and U. Mengali, "A comparison of pilot-aided channel estimation methods for OFDM systems," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3065–3073, Dec. 2001.

Donatella Darsena (M'06) was born in Napoli, Italy, on December 11, 1975. She received the Dr. Eng. degree *summa cum laude* in telecommunications engineering and the Ph.D. degree in electronic and telecommunications engineering from the University of Napoli Federico II, Napoli, Italy, in 2001 and 2005, respectively.

Since 2005, she has been an Assistant Professor with the Department of Technology, 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, and narrowband-interference suppression for multicarrier systems.



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

Since 2002, he has been an Assistant Professor of signal theory with the Department of Electronic and Telecommunication Engineering, University of Napoli Federico II. He also held teaching positions at the Second University of Napoli. His research activities lie in the area of statistical signal processing, digital communications, and communication systems. In particular, his current interests are focused on cyclostationarity-based techniques for blind identification, equalization and interference suppression for narrowband modulation systems, code-division multiple-access systems, and multicarrier modulation systems.