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 $L_{eff} \leq L_{cp} \leq L_{h}$.

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 $L_{eff} + 1$. 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
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, 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, 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 $L_{\text{eff}} + 1$. 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., $L_{\text{eff}} = 0$), 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 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.

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

2Two 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 by $\mathbb{C}$, $\mathbb{R}$, and $\mathbb{Z}$, respectively; matrices [vectors] are denoted with upper [lower] case boldface letters (e.g., $\mathbf{A}$ or $\mathbf{a}$); the field of $m \times n$ complex [real] matrices is denoted as $\mathbb{C}^{m \times n}$ [$\mathbb{R}^{m \times n}$], with $\mathbb{C}^{\mathbb{R}^{m \times n}}$ used as a shorthand for $\mathbb{R}^{m \times n}$; $\{\mathbf{A}\}_{i,j}$ or $\mathbf{A}_{i,j}$ indicates the $(i+1,j+1)$th element of matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$, with $i \in \{0, 1, \ldots, m-1\}$ and $j \in \{0, 1, \ldots, n-1\}$; the superscripts $*$, $T$, $I$, $H$, $\cdot$, 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; $\mathbf{0}_m \in \mathbb{R}^{m \times m}$, $\mathbf{I}_m \in \mathbb{R}^{m \times m}$ 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 $\|\mathbf{A}\|$, $\mathbb{R}(\mathbf{A})$, and $\mathbb{R}^\perp(\mathbf{A})$ denote the null space, the range (column space), and the orthogonal complement of the column space of $\mathbf{A} \in \mathbb{C}^{m \times n}$ in $\mathbb{C}^{\mathbb{R}^n}$; when applied to a vector $\mathbf{a} = \text{diag}(\mathbf{a})$ is the diagonal matrix with $\mathbf{A}_{i,i} = a_i$, whereas when applied to a matrix $\mathbf{a} = \text{diag}(\mathbf{A})$ is the vector with $a_i = A_{i,i}$; $\min[x,y]$ is the smallest of the two real numbers $x$ and $y$; the subscript $c$ stands for continuous-time signals, $\mathbb{E}[\cdot]$ denotes ensemble averaging and, finally, $[\cdot]$ and $\lfloor\cdot\rfloor$ denote the ceiling integer and the imaginary unit, respectively.
nth data block \( s(n) = [s_0(n), s_1(n), \ldots, s_{M-1}(n)]^T \in \mathbb{C}^M \) is first subject to the IDFT and, then, a CP of length \( L_{\text{CP}} \ll M \) is inserted at the beginning of the resulting vector, obtaining the time-domain block \( u(n) = [u_0(n), u_1(n), \ldots, u_{P-1}(n)]^T \in \mathbb{C}^P \), with \( P \triangleq M + L_{\text{CP}} \). This block can compactly be expressed as \( u(n) = T_p \mathbf{W}_{\text{IDFT}} s(n) \), where \( \mathbf{W}_{\text{IDFT}} \triangleq \left( 1/\sqrt{M} \right)^{\frac{1}{2}} \), with \( i_1, i_2 \in \{0, 1, \ldots, M-1\} \), represents the unitary symmetric IDFT matrix, \( T_p \triangleq \left[ T_{\text{CP}}, \mathbf{I}_M \right]^T \in \mathbb{R}^{P \times M} \) and \( \mathbf{I}_M \in \mathbb{R}^{M \times M} \) is obtained from \( \mathbf{I}_M \) by picking its last \( L_{\text{CP}} \) rows. Vector \( u(n) \) undergoes parallel-to-serial (P/S) conversion, and the resulting sequence \( u(nP + p) \equiv u_p(n) \), 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

\[
r_c(t) = \sum_{n=-\infty}^{+\infty} \sum_{p=0}^{P-1} u_p(n) h_c(t - nT_c - pT) + w_c(t)
\]

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 \( t \not\in [0, L_hT_c] \), whereas \( w_c(t) = w_c(nT_c + t) + w_c(\text{noise}(t)) \) 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

\[
r_q(k) \triangleq r_q(t_kq) = \sum_{i=0}^{L_h} h_c(i) u(k-i) + w_c(k)
\]

for \( q \in \{0, 1, \ldots, N-1\} \), with \( t_kq \triangleq kT_c + q(T_c/N) \), where \( h_c(k) \triangleq h_c(t_kq) \) and \( w_q(k) \triangleq w_c(t_kq) \) 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 \( \mathbf{r}(k) \triangleq [r_0(k), r_1(k), \ldots, r_{N-1}(k)]^T \in \mathbb{C}^N \), \( \mathbf{h}(k) \triangleq [h_0(k), h_1(k), \ldots, h_{N-1}(k)]^T \in \mathbb{C}^N \), and \( \mathbf{w}(k) \triangleq [w_0(k), w_1(k), \ldots, w_{N-1}(k)]^T \in \mathbb{C}^N \), one obtains the SIMO representation

\[
\mathbf{r}(k) = \sum_{i=0}^{L_h} \mathbf{h}(i) \mathbf{u}(k-i) + \mathbf{w}(k).
\] (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., \( L_{\text{CP}} < L_h \), a TEQ can be employed to suitably shorten the channel impulse response (see Fig. 1). Denoting with \( L_a \) the equalizer order (expressed in sampling intervals), the input-output relationship of a FIR TEQ can be expressed as \( \mathbf{y}(k) = \mathbf{f}^T \mathbf{z}(k) \), where \( \mathbf{f} \in \mathbb{C}^{N(L_a+1)} \) is the weight vector to be designed according to a specified optimization criterion and, accounting for (3), the input vector \( \mathbf{z}(k) \triangleq [\mathbf{r}(k), \mathbf{r}(k-1), \ldots, \mathbf{r}(k-L_e)]^T \in \mathbb{C}^{N(L_a+1)} \) assumes the form

\[
\mathbf{z}(k) = \mathbf{H} \tilde{\mathbf{u}}(k) + \mathbf{v}(k)
\] (4)

where, by setting \( L_g \triangleq L_e + L_h \)

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

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

\[
\tilde{\mathbf{u}}(k) \triangleq [u(k), u(k-1), \ldots, u(k-L_g)]^T \in \mathbb{C}^{L_g+1}
\] (6)
\( \mathbf{v}(k) \triangleq [\mathbf{w}^T(k), \mathbf{w}^T(k-1), \ldots, \mathbf{w}^T(k-L_{ef})]^T \in \mathbb{C}^{N(L_{ef}+1)}. \)  

(7)

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

\[ \mathbf{y}(k) = \mathbf{g}^H \tilde{\mathbf{u}}(k) + \bar{d}(k) \]  

(8)

where

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

(9)

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

\[ \mathbf{g}_{\text{target}} \triangleq [0, \ldots, 0, g_{\text{target}}(\Delta), g_{\text{target}}(\Delta+1), \ldots, g_{\text{target}}(\Delta + L_{ef}), 0, \ldots, 0]^T \]  

(10)

where \( 0 \leq \Delta \leq L_g - L_{ef} \) is a suitable shortening delay at the designer’s disposal, and \( L_{ef} \), which represents the effective order of the combined channel-TEQ impulse response, obeys \( L_{ef} \leq L_{cp} < L_g \) 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 \( \mathbf{y}(k) \), which is defined as

\[ \text{SSINR}(\mathbf{f}) \triangleq \frac{E\left[|g_{\text{win}}^H \tilde{\mathbf{u}}_{\text{win}}(k)|^2\right]}{E\left[|g_{\text{wall}}^H \tilde{\mathbf{u}}_{\text{wall}}(k) + \bar{d}(k)|^2\right]} \]  

(11)

where

\[ g_{\text{win}} \triangleq [g(\Delta), g(\Delta+1), \ldots, g(\Delta + L_{ef})]^T \in \mathbb{C}^{L_{ef}+1} \]  

(12)

\[ g_{\text{wall}} \triangleq [g(0), \ldots, g(\Delta-1), g(\Delta + L_{ef} + 1), \ldots, g(L_g)]^T \in \mathbb{C}^{L_g-L_{ef}} \]  

(13)

\[ \tilde{\mathbf{u}}_{\text{win}}(k) \triangleq [u(k-\Delta), u(k-\Delta-1), \ldots, u(k-L_{ef})]^T \in \mathbb{C}^{L_{ef}+1}, \]  

(14)

\[ \tilde{\mathbf{u}}_{\text{wall}}(k) \triangleq [u(k), u(k-1), \ldots, u(k-\Delta+1), u(k-L_{ef}+1), \ldots, u(k-L_g)]^T \in \mathbb{C}^{L_g-L_{ef}}. \]  

(15)

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

\[ \mathbf{y}(n) \triangleq \mathbf{R}_{cp} \mathbf{y}(n) = \mathbf{G} \mathbf{T}_{cp} \mathbf{W}_{\text{IDFT}} \mathbf{s}(n) + \bar{\mathbf{d}}(n) \]  

(16)

where the matrix \( \mathbf{R}_{cp} \triangleq \begin{bmatrix} \mathbf{I}_{M \times L_{cp}}, \mathbf{1}_M \end{bmatrix} \in \mathbb{R}^{M \times P} \) discards the first \( L_{cp} \) entries of \( \mathbf{y}(n) \), \( \mathbf{G} \in \mathbb{C}^{M \times P} \) is the Toeplitz channel matrix, whose first column and row are given by \([\mathbf{g}_{\text{target}}(\Delta + L_{ef}), 0, \ldots, 0]^T \) and \([\mathbf{g}_{\text{target}}(\Delta), 0, \ldots, 0]^T \), respectively, and \( \bar{\mathbf{d}}(n) \triangleq \mathbf{R}_{cp} \bar{\mathbf{d}}(n) \), with the disturbance vector \( \bar{\mathbf{d}}(n) \triangleq [\bar{d}_0(n), \bar{d}_1(n), \ldots, \bar{d}_{P-1}(n)]^T \in \mathbb{C}^P \) collecting \( \bar{d}_p(n) \triangleq \bar{d}(nP + p + \Delta) \). Moreover, the last equality in (16) comes from the fact that, since \( \mathbf{G} \mathbf{T}_{cp} \) is a circulant matrix [34], by using standard eigenstructure concepts, it can equivalently be expressed [2] as \( \mathbf{G} \mathbf{T}_{cp} = \mathbf{W}_{\text{IDFT}} \mathbf{G} \mathbf{W}_{\text{IDFT}} \), where the diagonal entries of \( \mathbf{G} \triangleq \mathrm{diag}(G(z_0), G(z_1), \ldots, G(z_{M-1})) \in \mathbb{C}^{M \times M} \) are the values of the (conjugate) combined channel-TEQ transfer function \( G(z) \triangleq \sum_{l=0}^{L_{ef}} g_{\text{target}}(\Delta + l)^{-1} z^{-l} \) evaluated at the subcarriers \( z_m \triangleq e^{(2\pi/\alpha)\imath m}. \) If \( G(z) \) has no zero located at \( \{z_m, m = 1, \ldots, M\} \), the matrix \( \mathbf{G} \) 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 \( \mathbf{x}(n) = \mathbf{G}^{-1} \mathbf{W}_{\text{IDFT}} \mathbf{G} \mathbf{s}(n) \). After FEQ, the th\( \text{th} \) entry of \( \mathbf{x}(n) \) is quantized to the nearest (in Euclidean distance) information symbol to form the estimate of the symbol \( \mathbf{s}_m(n) \) belonging to the th\( \text{th} \) data substream. In the sequel, the following customary assumptions are considered: (a1) the information symbols \( \mathbf{s}(k) \) are modeled as a sequence of zero-mean independent and identically distributed (i.i.d.) circular random variables, with variance \( \sigma^2_0 \triangleq E[|\mathbf{s}(k)|^2] \); (a2) the disturbance vector \( \mathbf{v}(k) \) in (4) is modeled as a zero-mean complex circular wide-sense stationary (WSS) random vector, statistically independent of \( \mathbf{s}(k) \), with autocorrelation matrix \( \mathbf{R}_{nn} \triangleq \mathbf{E}[\mathbf{v}(k)\mathbf{v}^H(k)] \in \mathbb{C}^{(N(L_{ef}+1)) \times (N(L_{ef}+1))} \).

### 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., \( \mathbf{v}(k) = \mathbf{0}_{N(L_{ef}+1)} \), the optimum channel-shortening approach consists of solving the system of linear equations \( \mathbf{H}^H \mathbf{f} = \mathbf{g}_{\text{target}} \), which allows one to perform perfect or ZF channel shortening [22]. Mathematically, this system is consistent [32] if and only if (iff) \( \mathbf{H}^H \mathbf{H} = \mathbf{g}_{\text{target}} \), which leads to perfect channel shortening in the absence of disturbance if the channel matrix (5) is full-column rank, i.e., \( \text{rank}(\mathbf{H}) = L_g + 1 \), in which case, one obtains \( \mathbf{f}_{\text{ZF}} = \mathbf{H}^H(\mathbf{H}^H \mathbf{H})^{-1} \mathbf{g}_{\text{target}} \). It is interesting to observe that, if \( \mathbf{H} \) is full-column rank, one has \( \mathbf{H}^H(\mathbf{H}^H)^{-1} = \mathbf{I}_{L_g+1} \) and, then, the system \( \mathbf{H}^H \mathbf{f} = \mathbf{g}_{\text{target}} \) turns out to be consistent independently of \( \mathbf{g}_{\text{target}} \). Apparently, the synthesis of \( \mathbf{f}_{\text{ZF}} \) requires a priori knowledge or estimation of the channel matrix \( \mathbf{H} \).
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 \( f \) in order to maximize the output SSINR, which accounting for (11) and assumptions (a1) and (a2), can be written as

\[
\text{SSINR}(f) = \frac{f^H (H_{\text{win}} \tilde{R}_{\text{win}} H_{\text{win}}^H) f}{f^H (H_{\text{wall}} \tilde{R}_{\text{wall}} H_{\text{wall}}^H + R_{\text{ BW}}) f}
\]  

(17)

where

\[
H_{\text{win}} \triangleq [h_{\Delta}, \ldots, h_{\Delta+L_{\text{eff}}} ] \in \mathbb{C}^{N(L_e+1) \times (L_{\text{eff}}+1)}
\]  

(18)

\[
\tilde{R}_{\text{win}} \triangleq E[\tilde{u}_{\text{win}} \tilde{u}_{\text{win}}^H] \in \mathbb{C}^{(L_{\text{eff}}+1) \times (L_{\text{eff}}+1)}
\]  

(19)

\[
H_{\text{wall}} \triangleq [h_{0}, \ldots, h_{\Delta-1}; h_{\Delta+L_{\text{eff}}} \ldots, h_{L_g}] \in \mathbb{C}^{N(L_e+1) \times (L_{\text{wall}}-L_{\text{eff}})}
\]  

(20)

\[
\tilde{R}_{\text{wall}} \triangleq E[\tilde{u}_{\text{wall}} \tilde{u}_{\text{wall}}^H] \in \mathbb{C}^{(L_{\text{wall}}-L_{\text{eff}}) \times (L_{\text{wall}}-L_{\text{eff}})}
\]  

(21)

It should be observed that (17) is different from the shortening SNR (SSNR) considered in [9]. Specifically, the SSNR given by (17) boils down to the SSNR defined in [9] if: 1) there is no disturbance, i.e., \( \nu(k) = 0 \); 2) the vector \( \tilde{u}(k) \) is white, which implies \( \tilde{R}_{\text{win}} = \sigma^2 \tilde{I}_{L_{\text{eff}}+1} \) and \( \tilde{R}_{\text{wall}} = \sigma^2 \tilde{I}_{L_{\text{wall}}-L_{\text{eff}}} \). The vector \( f_{\text{max}} \) maximizing (17) is given [35] by

\[
f_{\text{max}} = (g_{\text{max}} \in \mathbb{C}, \alpha_{\text{max}} \in \mathbb{C}^*)^T
\]  

(22)

Let [see (10)]

\[
g_{\text{target}, \text{win}} \triangleq \begin{bmatrix} g_{\text{target}}(\Delta), g_{\text{target}}(\Delta+1), \ldots, g_{\text{target}}(\Delta+L_{\text{eff}}) \end{bmatrix}^T \in \mathbb{C}^{L_{\text{eff}}+1}
\]  

(23)

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

\[
\text{MSE}(f; g_{\text{target}, \text{win}}) \triangleq E \left[ \| g(k) - g_{\text{target}, \text{win}} \tilde{u}(k) \|^2 \right]
\]  

(24)

with respect to both \( f \) and \( g_{\text{target}, \text{win}} \), subject to \( g_{\text{target}, \text{win}} \) having \( g_{\text{target}, \text{win}} \) rows and \( L_{\text{eff}} \) columns. It is worth observing that \( f_{\text{max}} \) ends up as the ZF solution \( f_{\text{ZF}} \), as the mean power \( E[\| g(k) \|^2] \) of the disturbance approaches zero.

Some comments are now in order about the assumption \( \text{rank}(H) = L_g + 1 \). The matrix \( H \) 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 \( Z \)-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].

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 \( E[\hat{z}(k) \hat{z}^H(k+i)] \) of the received vector \( \hat{z}(k) \), with \( i \in \{0, \Delta, \ldots, L_{\text{eff}}+1\} \). 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 \( N(L_e-L_g) \), 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 MOE(f) defined as

\[
f_{\text{max}} \triangleq E[\| g(k) \|^2] = f^H R_{\text{ZZ}} f
\]

(25)

where \( R_{\text{ZZ}} \triangleq E[\hat{z}(k) \hat{z}^H(k)] \in \mathbb{C}^{N(L_e+1) \times N(L_e+1)} \) denotes the autocorrelation matrix of \( \hat{z}(k) \), 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

\[
\hat{y}(k) = \sum_{d=0}^{L_g} g^H(d) u(k-d) + \hat{d}(k)
\]

(26)

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

\[
h_d = \Theta_d \xi_d,
\]

(27)

with \( \Theta_d \in \mathbb{R}^{N(L_e+1) \times N(L_e+1)} \) and \( \xi_d \in \mathbb{C}^{N(L_e+1)} \), where, depending on the relative values of \( d, L_e \) and \( L_h \), the following cases hold.

Case 1: \( 0 \leq d \leq \min \{L_e, L_h\} \). In this case

\[
\Theta_d = \left[ I_{N(d+1) \times N(d+1)} \right]_d^T
\]

(28)

and

\[
\xi_d = \left[ h^T(d), h^T(d-1), \ldots, h^T(0) \right]^T
\]

(29)

with \( L_d = d \).

Case 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

\[
\Theta_d = \left[ O_{N(L_d-1) \times N(L_h+1)}, I_{N(L_h+1)} \times I_{N(L_h+1)} \right]^T
\]

and

\[
\xi_d = \left[ h^T(L_h), h^T(L_h - 1), \ldots, h^T(0) \right]^T
\]

with \( L_d = L_h \).

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

\[
\Theta_d = \left[ O_{N(L_d-1) \times N(L_g-d+1)}, I_{N(L_g-d+1)} \times N(L_g-d+1) \right]^T
\]

and

\[
\xi_d = \left[ h^T(L_h), h^T(L_h - 1), \ldots, h^T(d - I_e) \right]^T
\]

with \( L_d = L_g - d \).

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

\[
\overline{y}(k) = f^H \Theta_6 \xi_6 u(k - \delta) + \sum_{d=0}^{L_g} \sum_{d'=0}^{d} f^H \Theta_{d'} \xi_{d'} u(k - d) + \hat{\alpha}(k)
\]

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 \( \Delta \) strictly depends on the choice of \( \delta \). Equation (33) shows that, to blindly preserve the signal contribution \( u(k - \delta) \) while minimizing the object function \( \text{MOE}^2(f) \), one can impose the set of \( N(L_d+1) \) linear constraints \( \Theta_d^T f = \gamma_d \), where \( \gamma_d \in \mathbb{C}^{N(L_d+1)} \) 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 \( \gamma_d^T \xi_6 \). As a first remark, it is worth noting that unit-norm or unit-tap constraints on \( f \), 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 \( \Theta_d^T f = \gamma_d \) is designed to preserve the signal contribution associated with the \( (\delta + 1) \)th column of \( H \), more generally, it has the effect of preserving the signal contributions associated with all those columns of \( H \) belonging to the column space of \( \Theta_6 \). Indeed, if \( h_d \in \mathcal{R}(\Theta_6) \), for some \( d \in \{0, \ldots, L_g\} \setminus \{\delta\} \), there exists a nonzero vector \( \alpha_d \in \mathbb{C}^{N(L_d+1)} \) such that \( h_d = \Theta_6 \alpha_d \) and, consequently, the signal contribution \( u(k - d) \) is passed to the TEQ output with gain \( f^H h_d = f^H \Theta_6 \alpha_d = \gamma_d^H \alpha_d \). On the other hand, if none of the columns \( h_d \) of \( H \) belongs to \( \mathcal{R}(\Theta_6) \), for \( d \in \{0, 1, \ldots, L_g\} \setminus \{\delta\} \), 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 \( \delta \), besides the vector \( h_d \), some other columns of \( H \) also belong to \( \mathcal{R}(\Theta_6) \), and, thus, the constraint \( \Theta_d^T f = \gamma_d \) preserves more than one signal contribution: in this case, the TEQ forces the overall system impulse response to have a nonzero order \( L_{\text{eff}} \), where \( L_{\text{eff}} \) 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

\[
\mathbf{f}_{\text{AMMO}} = \arg \min_{f \in \mathbb{C}^{N(L_e+1)}} f^H R_{zz} f
\]

subject to \( \Theta_6^T f = \gamma_6 \)

\[
f_{\text{AMMO}} = F_{\text{AMMO}} \gamma_6
\]

with

\[
F_{\text{AMMO}} = R_{zz}^{-1} \Theta_6 \left( \Theta_6^T R_{zz}^{-1} \Theta_6 \right)^{-1} \in \mathbb{C}^{N(L_e+1) \times N(L_e+1)}
\]

It is interesting to point out that solution \( f_{\text{AMMO}} \) 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 \( f_{\text{AMMO}} \), according to the generalized sidelobe canceler (GSC) scheme [26], into two orthogonal components as

\[
f_{\text{AMMO}} = f^{(0)}_{\text{SYND}} - \Pi_6 f^{(a)}_{\text{AMMO}}
\]

where the fixed quiescent vector \( f^{(0)}_{\text{SYND}} = \Theta_6 \left( \Theta_6^T \Theta_6 \right)^{-1} \gamma_6 = \Theta_6 \gamma_6 \) belongs to the constraint subspace \( \mathcal{R}(\Theta_6) \), the signal blocking matrix \( \Pi_6 \in \mathbb{R}^{N(L_e+1) \times N(L_e+L_6)} \) is chosen so that its columns form a basis for the null space of \( \Theta_6^T \), i.e., \( \mathcal{R}(\Pi_6) \equiv \mathcal{R}(\Theta_6) \). and the optimal solution of the data-dependent component \( f^{(a)}_{\text{AMMO}} \) is given by

\[
f^{(a)}_{\text{AMMO}} = \left( \Pi_6^T R_{zz} \Pi_6 \right)^{-1} \Pi_6^T R_{zz} f^{(0)}_{\text{SYND}}
\]

As a direct consequence of the GSC decomposition given by (37), it can easily be verified by direct inspection that \( F_{\text{AMMO}} \) defined in (35) can be equivalently decomposed as

\[
F_{\text{AMMO}} = \Theta_6 - \Pi_6 \left( \Pi_6^T R_{zz} \Pi_6 \right)^{-1} \Pi_6^T R_{zz} \Theta_6
\]

where

\[
F_{\text{AMMO}} = \Theta_6 - \Pi_6 f^{(a)}_{\text{AMMO}}
\]

It is worth noting that, for a given vector \( \gamma_6 \), synthesis of \( F_{\text{AMMO}} \) requires only the computation in real-time of \( f^{(a)}_{\text{AMMO}} \), since both the matrices \( \Theta_6 \) and \( \Pi_6 \) can be precomputed off-line.

\[\text{(39)}\]
In its turn, the synthesis of the matrix $F_{\text{AMMEO}}^{(d)}$ does not involve EVD/SDV or Cholesky decompositions and only requires the inversion of $\Pi_6^T R_{zz} \Pi_6 \in \mathbb{C}^{N(L_e-L_d) \times N(L_e-L_d)}$, 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 $\mathcal{O}[N^3(L_e-L_d)^2]$ floating point operations (flips) [35]. However, the matrix $F_{\text{AMMEO}}^{(d)}$ 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 $\mathcal{O}[N^2(L_e-L_d)^2]$ flips per iteration.

In Section IV-A, we analyze the expression of the combined channel-TEQ response vector $g_{\text{AMMEO}}^T = H^H f_{\text{AMMEO}}^T \in \mathbb{C}^{L_o+1}$ and evaluate the mean power $f_{\text{AMMEO}}^T R_{uu} f_{\text{AMMEO}}^H$, 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 a combined channel-TEQ response vector $g_{\text{AMMEO}}$ 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 (40) is composed of two terms $v(k) = v_{\text{IM}}(k) + v_{\text{Noise}}(k)$, where $v_{\text{IM}}(k)$ and $v_{\text{Noise}}(k)$ account for NBI and noise, respectively. In addition to assumption (a3), we assume that: (a3) the first $R_{uu} \ll N(L_e+1)$ eigenvalues of the NBI autocorrelation matrix $R_{uu} \triangleq E[v_{\text{IM}}(k) v_{\text{IM}}^H(k)]$ are significantly different from zero, whereas the remaining ones are vanishingly small; (a4) the vector $v_{\text{IM}}(k)$ is a white random process, statistically independent of $v_{\text{IM}}(k)$, with autocorrelation matrix $R_{\text{noise}} \triangleq E[v_{\text{noise}}(k) v_{\text{noise}}^H(k)] = \sigma^2_{\text{v}} I_{N(L_e+1)}$. 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 $R_{\text{uu}} = [N(L_e + 1) I_{L_o} W_{\text{dBH}}] + 1$, where $W_{\text{dBH}}$ 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 $W_{\text{dBH}}$ is significantly smaller, i.e., $T_{\text{c}} W_{\text{dBH}} \ll 1$, and, thus, it turns out that $R_{\text{uu}} \ll N(L_e + 1)$. Under assumption (a3), the NBI autocorrelation matrix can be well modeled by the following full-rank factorization (see [32])

$$R_{\text{uu}} = J^H J$$

where the matrix $J \in \mathbb{C}^{N(L_e+1) \times R_{\text{uu}}}$ is full-column rank, i.e., rank$(J) = R_{\text{uu}}$. By virtue of this fact, the NBI suppression capability of the MMOE channel shortener can be quantified by evaluating the mean power $\mathcal{E}_{\text{AMMEO}} \triangleq E[f_{\text{AMMEO}}^H v_{\text{IM}}(k)[k]^2] = f_{\text{AMMEO}}^H J^H J f_{\text{AMMEO}}^H = \| J^H f_{\text{AMMEO}}^H \|^2$ of the NBI at the TEQ output, which depends on the expression of the vector $\ell_{\text{AMMEO}} \triangleq J H f_{\text{AMMEO}} \in \mathbb{C}^{R_{\text{uu}}}$.

Therefore, the channel shortening and the NBI suppression capabilities of the proposed TEQ can jointly be characterized by the vector $b_{\text{AMMEO}} \triangleq H^H f_{\text{AMMEO}}$, with

$$H \triangleq [H^T J^T] \in \mathbb{C}^{N(L_o+1) \times (L_o + R_{\text{uu}})}$$. On the basis of (37) and (38), the vector $b_{\text{AMMEO}}$ can be explicitly written as follows:

$$b_{\text{AMMEO}}^T = \begin{bmatrix} g_{\text{AMMEO}}^T \\ \ell_{\text{AMMEO}}^T \end{bmatrix}$$

$$= H^H \left[ I_{N(L_o+1)} - \Pi_6 \left( \Pi_6^T R_{zz} \Pi_6 \right)^{-1} \Pi_6^T R_{zz} \right] \times \Theta_6 \gamma_6$$

where, accounting for (4) and invoking assumptions (a1)–(a4),

$$R_{zz} = H \tilde{R}_{uu} H^H + J J^H + \sigma^2_{\text{v}} I_{N(L_o+1)}$$

$$= \mathcal{H} \begin{bmatrix} \tilde{R}_{uu} & O_{L_o+1 \times R_{\text{uu}}} \\ O_{R_{\text{uu}} \times (L_o+1)} & I_{R_{\text{uu}}} \end{bmatrix} H^H$$

$$+ \sigma^2_{\text{v}} I_{N(L_o+1)}$$

$$= \mathcal{H} \Delta \mathcal{H}^H + \sigma^2_{\text{v}} I_{N(L_o+1)}$$

$$= \mathcal{H} \Delta \Delta^H + \sigma^2_{\text{v}} I_{N(L_o+1)}$$

with $\tilde{R}_{uu} \triangleq [\tilde{V}(k) \tilde{V}^H(k)] \in \mathbb{C}^{(L_o+1) \times (L_o+1)}$ representing the nonsingular autocorrelation matrix of the vector $\tilde{v}(k)$ defined in (4). As a consequence of the nonsingularity of $\tilde{R}_{uu}$, the block diagonal matrix $\Delta$ is nonsingular, too. Substituting (41) in (40) and recalling that $\Pi_6^T \Pi_6 = I_{N(L_o-L_d)}$ and $\Pi_6^T \Theta_6 = O_{N(L_o-L_d) \times N(L_o+1)}$, one obtains

$$b_{\text{AMMEO}}^T = H^H \left[ I_{N(L_o+1)} - \Pi_6 \right]$$

$$\cdot \left[ \left( \Pi_6^T \mathcal{H} \Delta^{1/2} \right) \cdot \left( \Pi_6^T \mathcal{H} \Delta^{1/2} \right)^H + \sigma^2_{\text{v}} I_{N(L_o+1)} \right]^{-1} \cdot \left( \Pi_6^T \mathcal{H} \Delta^{1/2} \right)^H \Theta_6 \gamma_6$$

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

$$b_{\text{AMMEO}} \triangleq \lim_{\sigma^2_{\text{v}} \to 0} b_{\text{AMMEO}} = \begin{bmatrix} \tilde{g}_{\text{AMMEO}}^T \\ \ell_{\text{AMMEO}}^T \end{bmatrix}$$

$$= H^H \left[ I_{N(L_o+1)} - \Pi_6 \left( \Delta^{1/2} \mathcal{H} \Pi_6 \right)^H \Delta^{1/2} \mathcal{H}^H \right] \times \Theta_6 \gamma_6$$

$$= \left[ I_{L_o+R_{\text{uu}}+1} - \mathcal{H}^H \Pi_6 \left( \mathcal{H}^H \Pi_6 \right)^H \right] \mathcal{H}^H \Theta_6 \gamma_6$$

$$= P_{\delta} \mathcal{H}^H \Theta_6 \gamma_6$$

where $P_{\delta} \in \mathbb{C}^{(L_o+R_{\text{uu}}+1) \times (L_o+R_{\text{uu}}+1)}$.
where, because of the nonsingularity of $\Delta$, we have used in the second equality the fact that
\[(\Delta^{1/2}H^H\Pi_{\delta})^\dagger = (H^H\Pi_{\delta})^\dagger \Delta^{-1/2}.\]
As a first remark, it is important to note that the asymptotic expression of $b_{\text{symm},\delta}$ does not depend on $\Delta$ and, hence, on the structure of $R_{\text{UBL}}$. Furthermore, since $P_{\delta}$ represents the orthogonal projector [32] onto the null space of $\Pi_{\delta}^H H$, it is apparent from (43) that $b_{\text{symm},\delta}$ belongs to the subspace $\mathcal{N}(\Pi_{\delta}^H H) \equiv \mathbb{R}^\perp (H^H\Pi_{\delta})$. 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 $H_{\delta} \in \mathbb{C}^{N(L_e+1) \times R_{\delta}}$ 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 $\mathcal{R}(\Theta_{\delta})$ and $\mathcal{R}(H_{\delta})$ are nonoverlapping or disjoint, that is
\[\mathcal{R}(\Theta_{\delta}) \cap \mathcal{R}(H_{\delta}) = \{0_{N(L_e+1)}\},\] (45) then, for $\sigma_n^2 \to 0$, the combined channel-TEQ vector $g_{\text{symm},\delta}$ assumes the target form (10) where: for $0 \leq \delta \leq \min\{L_e, L_{\delta}\}$ (case P1), $\Delta = 0$, $L_{\text{eff}} = \delta$ and $g_{\text{target}}(e) = \tilde{h}_e^H \Theta_{\delta} \gamma_{\delta}$, with $e \in \{0, 1, \ldots, \delta\}$; for $L_{\delta} < \delta < L_e$ (case P2), $\Delta = \delta$, $L_{\text{eff}} = 0$ and $g_{\text{target}}(e) = \tilde{h}_e^H \Theta_{\delta} \gamma_{\delta}$, with $e \in \{0, 1, \ldots, \delta\}$; for $L_e < \delta < L_{\delta}$ (case P3), $\Delta = \delta - 1$, $L_{\text{eff}} = L_{\delta} - \delta$ and $g_{\text{target}}(e) = \tilde{h}_e^H \Theta_{\delta} \gamma_{\delta}$, with $e \in \{0, 1, \ldots, \delta\}$. Furthermore, if condition (45) is fulfilled, perfect NBI suppression is achieved in the limiting case of vanishingly small noise, i.e., $\hat{e}_{\text{symm},\delta} \rightarrow 0_{R_{\text{UBL}}}$ (and, thus, $E_{\text{NBI}} \rightarrow 0$) as $\sigma_n^2 \to 0$.

**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 $\delta$ strongly influences both the actual values of the shortening delay $\Delta$ and the effective order $L_{\text{eff}}$ of $g_{\text{target}}$, 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 $\{h_{\ell}(k)\}_{k=0}^{N-1}$ exhibit common zeros in their $\mathbb{Z}$-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 $\Pi_{\delta}^H H_{\delta} \in \mathbb{C}^{N(L_e-L_{\delta}) \times R_{\delta}}$ be full-column rank, i.e., $\text{rank} (\Pi_{\delta}^H H_{\delta}) = R_{\delta}$. A necessary condition for satisfying this rank requirement is that $N(L_e - L_{\delta}) \geq R_{\delta}$, 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 $L_{\delta}$ does not exceed the maximum value $L_{\delta,\text{max}}(\delta) = (N-1)(L_e-\delta) - R_{\text{UBL}}$, for $0 \leq \delta \leq \min\{L_e, L_{\delta}\}$, 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 $L_{\text{cp}} \geq \delta$. 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 $R_{\text{UBL}}$, the upper bound $L_{\delta,\text{max}}(\delta)$ achieves its maximum value $(N-1) L_e - R_{\text{UBL}}$, 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 $\sigma_n^2 \to 0$, but it is not always the best choice in noisy environments: indeed, using too small values of $\delta$ does not allow one to exploit the entire channel energy, since the vectors $\tilde{h}_0, \tilde{h}_1, \ldots, \tilde{h}_\delta$, which determine the nonzero entries of $g_{\text{symm},\delta}$, 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 $\delta = L_{\text{eff}} = L_{\text{cp}}$ and, thus, leaving to CP removal the duty of suppressing the residual IBI. In this case,
the channel order must obey \( I_h \leq (N-1)(L_e - L_{cp}) - R_{nbi} \), which can be equivalently expressed as
\[
I_e \geq L_{cp} + \left\lceil \frac{L_h + R_{nbi}}{N - 1} \right\rceil.
\] (47)

In the absence of NBI (i.e., \( R_{nbi} = 0 \)), when the received signal is sampled at rate \( 2/T_c \), inequality (47) becomes \( I_e \geq L_{cp} + L_h \).

This lower bound on \( I_e \) is more restrictive than condition (c2), in the sense that, for a given channel order, a larger TEQ order \( L_s \) 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
\[
L_h \geq \left\lceil \frac{N - 1}{N - 2}(L_{cp} + 1) \right\rceil.
\] (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 \( \Theta_6 \) 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 \( \Theta_6 \) 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., \( L_{eff} = 0 \), provided that the channel order \( L_h \) does not exceed the maximum value
\[
L_{2,max} = \left\lceil \frac{(N-1)L_e - R_{nbi}}{N + 1} \right\rceil.
\] (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 \( \xi_6 \), which determines the nonzero entry \( g(\delta) \) of \( g_{\text{Ammpoe}}^\delta \), 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 \( \Delta + L_{eff} > L_p \), the \( L_{eff} + 1 \) nonzero entries of \( g_{\text{Ammpoe}}^\delta \) are not consecutive. Moreover, similarly to the P2 case, the structure of \( \Theta_6 \) 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 \( \gamma_6 \) 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 \( \gamma_6 \) 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 SSINR(\( f_{\text{Ammpoe}}^\delta \)) \( \to +\infty \) as \( \sigma_p^2 \to 0 \), regardless of the choice of \( \gamma_6 \). 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., \( \| g_{\text{wall}} \|^2 \approx 0 \), and almost perfect NBI suppression, i.e., \( \epsilon_{nbi} \approx 0 \). Consequently, accounting for (35) and assumptions (a3)–(a4), the SSINR at the output of the MIMO-based TEQ can approximately be written as

\[
\text{SSINR}(f_{\text{Ammpoe}}^\delta) = \frac{\gamma_6^H (F_{\text{Ammpoe}}^\delta H_{\text{win}} \hat{R}_{\text{win}} H_{\text{win}}^H f_{\text{Ammpoe}}^\delta) \gamma_6}{\sigma_w^2 \gamma_6^H F_{\text{Ammpoe}}^\delta H_{\text{win}} \hat{R}_{\text{win}} H_{\text{win}}^H F_{\text{Ammpoe}}^\delta \gamma_6}.
\] (50)

Therefore, on the basis of (50), maximizing SSINR(\( f_{\text{Ammpoe}}^\delta \)) amounts to choose \( \gamma_6 \) as the solution \( \gamma_{\text{opt}} \) of the following maximization problem:

\[
\arg \max_{\gamma_6 \in \mathbb{C}^{L_e+1}} \frac{\gamma_6^H (F_{\text{Ammpoe}}^\delta H_{\text{win}} \hat{R}_{\text{win}} H_{\text{win}}^H f_{\text{Ammpoe}}^\delta) \gamma_6}{\gamma_6^H F_{\text{Ammpoe}}^\delta H_{\text{win}} \hat{R}_{\text{win}} H_{\text{win}}^H F_{\text{Ammpoe}}^\delta \gamma_6}.
\] (51)

The solution of (51) is given [35] by \( \gamma_{\text{opt}} = (\mathbf{Q}_{\text{Ammpoe}}^\delta)^{-1} \), where \( \varrho \in \mathbb{C} \) and \( \mathbf{Q}_{\text{Ammpoe}}^\delta \) is the generalized eigenvector associated with the largest generalized eigenvalue of the two matrices \( F_{\text{Ammpoe}}^\delta H_{\text{win}} \hat{R}_{\text{win}} H_{\text{win}}^H \). Since \( F_{\text{Ammpoe}}^\delta H_{\text{win}} \hat{R}_{\text{win}} H_{\text{win}}^H \) is full-column rank for all the SNR values of practical interest, i.e., \( \text{rank}(F_{\text{Ammpoe}}^\delta H_{\text{win}} \hat{R}_{\text{win}} H_{\text{win}}^H) = N(L_\delta + 1) \), 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 \( F_{\text{Ammpoe}}^\delta \) can be factorized as follows:

\[
F_{\text{Ammpoe}}^\delta = Q_{\text{Ammpoe}}^\delta R_{\text{Ammpoe}}^\delta
\] (52)

where \( R_{\text{Ammpoe}}^\delta \in \mathbb{C}^{N(L_e+1)\times N(L_e+1)} \) is a nonsingular upper triangular matrix, whereas the columns of the matrix \( Q_{\text{Ammpoe}}^\delta \in \mathbb{C}^{N(L_e+1)\times N(L_e+1)} \) are orthonormal, i.e., \( Q_{\text{Ammpoe}}^\delta H_{\text{win}} \hat{R}_{\text{win}} H_{\text{win}}^H Q_{\text{Ammpoe}}^\delta = I_{N(L_e+1)} \). Relying on this decomposition, it is readily seen that the solution of (51) can equivalently be expressed as

\[
\gamma_{\text{opt}} = R_{\text{Ammpoe}}^{-1} \xi_{\text{Ammpoe}}^\delta \gamma_{\text{opt}}
\] (53)

where, in its turn, \( \xi_{\text{Ammpoe}}^\delta \in \mathbb{C}^{N(L_e+1)} \) is determined by the maximization problem

\[
\xi_{\text{Ammpoe}}^\delta = \arg \max_{\xi \in \mathbb{C}^{N(L_e+1)}} \frac{\xi^H M_{\text{Ammpoe}}^\delta \xi}{\xi^H \xi}
\] (54)

with \( M_{\text{Ammpoe}}^\delta \triangleq Q_{\text{Ammpoe}}^\delta H_{\text{win}} \hat{R}_{\text{win}} H_{\text{win}}^H Q_{\text{Ammpoe}}^\delta \), whose solution [34] is the eigenvector associated with the largest eigenvalue of \( M_{\text{Ammpoe}}^\delta \). Accounting for the QRD of \( F_{\text{Ammpoe}}^\delta \) and (53), the weight vector (35) reduces to
\[ f_{\text{OPT}} = F_{\text{AMMOE}} \gamma_{\text{OPT}} = Q_{\text{AMMOE}} \zeta_{\text{OPT}} \]  

(55)

which shows that the synthesis of \( f_{\text{OPT}} \) does not need neither evaluation nor inversion of \( R_{\text{AMMOE}} \), but only computation of \( Q_{\text{AMMOE}} \). On the other hand, it is apparent that evaluation of \( \zeta_{\text{OPT}} \) requires a priori knowledge of the channel-dependent matrix \( H_{\text{WIN}} \), which is unknown at the receiver. Interestingly, the computation of \( \zeta_{\text{OPT}} \) can blindly be carried out, without requiring knowledge or estimation of \( H_{\text{WIN}} \). To this end, let us consider the linear transformation \( \mathbf{j}(k) = Q_{\text{AMMOE}}^H \mathbf{z}(k) \in \mathbb{C}^{N(L_e+1)} \) of the received vector \( \mathbf{z}(k) \), which, accounting for (4) and invoking Theorem 1, for moderate-to-high values of the SNR, can be approximatively decomposed as

\[ \mathbf{j}(k) = Q_{\text{AMMOE}}^H \mathbf{H}_{\text{WIN}} \mathbf{u}_{\text{WIN}}(k) + Q_{\text{AMMOE}}^H \mathbf{v}_{\text{NOISE}}(k) \]  

(56)

whose autocorrelation matrix \( R_{\mathbf{j}} = E[\mathbf{j}(k)\mathbf{j}^H(k)] \in \mathbb{C}^{N(L_e+1) \times N(L_e+1)} \) is given by

\[ R_{\mathbf{j}} = M_{\text{AMMOE}} + \sigma^2_{\text{AMMOE}} \mathbf{I}_{N(L_e+1)} \]  

(57)

which shows that \( \zeta_{\text{OPT}} \) corresponding to the largest eigenvalue of \( M_{\text{AMMOE}} \), can blindly be evaluated as the eigenvector corresponding to the largest eigenvalue of \( R_{\mathbf{j}} \), which can consistently be estimated from the received data. From a computational point of view, it turns out that the evaluation of \( \zeta_{\text{OPT}} \) is dominated by the QRD of \( F_{\text{AMMOE}} \), which involves \( O\left[N^3(L_e+1)^2(L_e+1)\right] \) flops if computed from scratch, and, additionally, by the calculus of the dominant eigenvector of \( R_{\mathbf{j}} \), which requires \( O\left[N^5(L_e+1)^3\right] \) flops if one resorts to batch algorithms.

In Section IV-C, we will show that \( \zeta_{\text{OPT}} \) can adaptively be estimated by means of a fast recursive rule, which only takes \( O\left(N^2(L_e-L_\delta)\min\{L_e-L_\delta, (L_\delta+1)^2\}\right) \) 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 \( K_e \) consecutive sample blocks of the received data vectors \( \{z(k)\}_{k=0}^{K_e-1} \), produce a reliable estimate \( \hat{F}_{\text{AMMOE}}(\zeta) \) of the matrix \( F_{\text{AMMOE}}(\zeta) = \left( \Pi_\delta^H R_{ZZ} \Pi_\delta \right)^{-1} \Pi_\delta^H R_{ZZ} \Theta_\delta \), and build the matrix \( \hat{F}_{\text{AMMOE}}(\zeta) = \Theta_\delta - \Pi_\delta \hat{F}_{\text{AMMOE}}(\zeta) \) [see (39)].

Step 2) Calculate the QRD of \( \hat{F}_{\text{AMMOE}}(\zeta) \), thus obtaining an estimate \( \hat{Q}_{\text{AMMOE}}(\zeta) \) of \( Q_{\text{AMMOE}}(\zeta) \) [see (52)].

Step 3) Form an estimate \( \hat{\zeta}_{\text{OPT}} \) of \( \zeta_{\text{OPT}} \) by evaluating the dominant eigenvector of the autocorrelation matrix of vector \( \hat{Q}_{\text{AMMOE}}^H \mathbf{z}(k) \) [see (56) and (57)].

Step 4) Build the weight vector of the proposed TEQ as \( \hat{f}_{\text{OPT}} = \hat{Q}_{\text{AMMOE}}(\zeta) \hat{\zeta}_{\text{OPT}} \) [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 \( \hat{Q}_{\text{AMMOE}}(\zeta) \) in step 2) is obtained by computing the QRD of \( \hat{F}_{\text{AMMOE}}(\zeta) \) from scratch, the synthesis of the weight vector \( \hat{f}_{\text{OPT}} \) requires \( O\left[N^3 \left\{ (L_e-L_{CP})^3 + (L_{CP}+1)^3 + (L_{CP}+1)^2 \right\} \right] \) 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 \( O(54572) \) flops, whereas the proposed technique requires only \( O(26752) \) 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 \( \hat{f}_{\text{OPT}} \) 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 \( F_{\text{AMMOE}}(\zeta) \) can be estimated from the incoming data through sample-by-sample updating. Specifically, after some lengthy calculations, it can be shown [31] that, for \( k \in \{0, 1, \ldots, K_e-1\} \), the recursion for updating \( \hat{F}_{\text{AMMOE}}(\zeta) \) is

\[ \hat{F}_{\text{AMMOE}}(\zeta)(k-1) = \hat{F}_{\text{AMMOE}}(\zeta)(k) \]

\[ + \left[ p(k) \right] \left[ \hat{F}_{\text{AMMOE}}(\zeta)(k-1) - \hat{z}(k) \right]^H \hat{q}(k) \in \mathbb{C}^{N(L_e+1)} \]

\[ = \hat{F}_{\text{AMMOE}}(\zeta)(k-1) + p(k) \hat{q}(k)^H \]  

(58)

where \( \hat{F}_{\text{AMMOE}}(\zeta)(k-1) = \Theta_\delta - \Pi_\delta \hat{F}_{\text{AMMOE}}(\zeta)(k-1) \) and

\[ \hat{q}(k) \triangleq \frac{\Phi_{ZZ}(k-1) \Pi_\delta^H \hat{z}(k)}{\lambda_1 + \hat{z}(k)^H \Pi_\delta \hat{F}_{\text{AMMOE}}(\zeta)(k-1) \Pi_\delta^H \hat{z}(k)} \in \mathbb{C}^{N(L_e-L_\delta)} \]  

(59)

is the overall gain vector, with \( \Phi_{ZZ}(k-1) \) and \( \lambda_1 \in (0,1] \) denoting the estimate of \( \left[ \Pi_\delta^H \Pi_\delta \right]^{-1} \) 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 \( \Phi_{ZZ}(-1) = \delta^{-1} I_{N(L_e-L_\delta)} \) and \( \hat{F}_{\text{AMMOE}}(-1) = \left[ I_{N(L_e+1)} \otimes (N(L_e+1) \times N(L_e-2L_\delta+1))^{-1} \right] \), 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 \( \hat{F}_{\text{AMMOE}}(\zeta)(k) \) converges in mean square to the optimal matrix \( F_{\text{AMMOE}}(\zeta) \), regardless of the eigenstructure of \( R_{ZZ} \). It is worth noting that the recursive (58) requires \( O\left[N^2(L_e-L_\delta)^2\right] \) flops per iteration.

The second step involves the QRD of \( \hat{F}_{\text{AMMOE}}(\zeta)(k) \) in order to obtain the following factorization:

\[ \hat{F}_{\text{AMMOE}}(\zeta)(k) = \hat{Q}_{\text{AMMOE}}(\zeta)(k) \hat{R}_{\text{AMMOE}}(\zeta)(k) \]  

(60)

where \( \hat{R}_{\text{AMMOE}}(\zeta)(k) \in \mathbb{C}^{N(L_e+1) \times N(L_e+1)} \) is a nonsingular upper triangular matrix, whereas the columns of
the matrix \( \hat{Q}_{\text{QMPD}}(k) \in \mathbb{C}^{N(L_\delta+1) \times N(L_\delta+1)} \) are orthonormal, i.e., \( \hat{Q}_{\text{QMPD}}(k) \hat{Q}_{\text{QMPD}}(k) = \mathbf{I}_{N(L_\delta+1)} \). By observing that

\[
\hat{P}_{\text{QMPD}}(k) = \Theta_\delta - \Pi_{\hat{A}(k)}^{(a)}
\]

the matrix \( \hat{Q}_{\text{QMPD}}(k) \) can be conveniently obtained from the QRD of \( \hat{P}_{\text{QMPD}}(k) \) rather than from the QRD of \( \hat{F}_{\text{QMPD}}(k) \). Indeed, the matrix \( \hat{P}_{\text{QMPD}}(k) \) can be rewritten as shown by (61), shown at the bottom of the page, where

\[
\hat{Q}_{\text{QMPD}}(k) \in \mathbb{C}^{N(L_\delta+1) \times N(L_\delta+1)}
\]

be a nonsingular upper triangular matrix and \( \hat{Q}_{\text{QMPD}}(k) \in \mathbb{C}^{N(L_\delta+1) \times N(L_\delta+1)} \) satisfying

\[
\begin{bmatrix}
\hat{Q}_{\text{QMPD}}(k)
\end{bmatrix}^H \hat{Q}_{\text{QMPD}}(k) = \mathbf{I}_{N(L_\delta+1)}.
\]

It is important to observe that, by construction, the columns of \( \hat{Q}_{\text{QMPD}}(k) \) are orthonormal, i.e.,

\[
\begin{bmatrix}
\hat{Q}_{\text{QMPD}}(k)
\end{bmatrix}^H \hat{Q}_{\text{QMPD}}(k) = \mathbf{I}_{N(L_\delta+1)}.
\]

Capitalizing on (61), the QRD of \( \hat{P}_{\text{QMPD}}(k) \) is obtained by performing \( N(L_\delta+1) \) successive Givens or Householder rotations [35] \( \mathbf{Q}_\ell(k) \in \mathbb{C}^{2(N(L_\delta+1)) \times 2(N(L_\delta+1)), \ell \in \{1, 2, \ldots, N(L_\delta+1)\} \), thus getting

\[
\begin{bmatrix}
\hat{Q}_{\text{QMPD}}(k) \\
\mathbf{Q}_1(k)^H
\end{bmatrix}^H \begin{bmatrix}
\mathbf{Q}_{N(L_\delta+1)}(k) \\
\mathbf{Q}_1(k)^H
\end{bmatrix} \begin{bmatrix}
\hat{R}_{\text{QMPD}}(k) \\
\mathbf{Q}_1(k)
\end{bmatrix} = \mathbf{Q}_N(L_\delta+1) \times N(L_\delta+1) \mathbf{Q}_1(k)^H.
\]

Therefore, given the QRD of \( \hat{F}_{\text{QMPD}}(k) \), the calculation of \( \hat{Q}_{\text{QMPD}}(k) \) requires \( O[N^2(L_\delta+1)] \) flops if standard Givens/Householder transformations are used [35]. Let us now focus attention on the evaluation of the QRD of \( \hat{F}_{\text{QMPD}}(k) \), whose computation from scratch takes \( O[N^2(L_\delta+1)] \) flops. To develop an alternative procedure for computing the QRD of \( \hat{F}_{\text{QMPD}}(k) \), we observe from (58) that \( \hat{F}_{\text{QMPD}}(k) \) comes from a rank-one change of \( \hat{F}_{\text{QMPD}}(k-1) \). If the complete QRD of \( \hat{F}_{\text{QMPD}}(k-1) \) is known and

\[
\text{rank} \left[ \hat{F}_{\text{QMPD}}(k-1) \right] = \text{rank} \left[ \hat{F}_{\text{QMPD}}(k) \right] = N(L_\delta+1),
\]

there exists [35] a simple procedure to compute the QRD of the rank-one update \( \hat{F}_{\text{QMPD}}(k) \) of \( \hat{F}_{\text{QMPD}}(k-1) \), which involves a computational burden of \( O[N^2(L_\delta+1)] \) flops. For the sake of conciseness, we defer directly to [35] for details regarding the structure of the rotations matrices \( \{Q_r(k)\}_{k=1}^{N(L_\delta+1)} \) and the rank-updating of the QRD. However, it is worth noting that, if \( L_\delta+1 \) is larger than \( (L_\delta+1)^2 \), updating the QRD of \( \hat{F}_{\text{QMPD}}(k) \) from \( \hat{F}_{\text{QMPD}}(k-1) \) turns out to be more expensive than computing the QRD of \( \hat{F}_{\text{QMPD}}(k) \) from scratch. Hence, we can conclude that computation of \( \hat{Q}_{\text{QMPD}}(k) \) requires an overall complexity of \( O \left[ N^2(L_\delta+1) \right] \) flops.

In the third step, starting from the samples of \( \mathbf{J}(k) \), we obtain \( \hat{Q}_N(L_\delta+1) \times N(L_\delta+1) \mathbf{J}(k) \), for \( k \in \{0, 1, \ldots, K_c-1\} \), tracking of the most dominant eigenvector of the autocorrelation matrix of \( \mathbf{J}(k) \) 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 \( \hat{\mathbf{g}}_{\text{opt}} \) of \( \mathbf{R}_\mathbf{D} \). The PAST algorithm for updating \( \hat{\mathbf{g}}_{\text{opt}} \) is given by

\[
\hat{\mathbf{g}}_{\text{opt}}(k) = \hat{\mathbf{g}}_{\text{opt}}(k-1) + \mu(k) \mathbf{J}(k) - \hat{\mathbf{g}}_{\text{opt}}(k-1) \psi(k) \psi^H(k),
\]

where \( \psi(k) = \hat{\mathbf{g}}_{\text{opt}}(k-1) - \hat{\mathbf{g}}_{\text{opt}}(k-1) \mathbf{J}(k-1) \mathbf{J}(k-1)^H \), with \( \psi(k) = (0, 1, 0, \ldots, 0)^T \), and \( \hat{\mathbf{g}}_{\text{opt}}(k) \) denotes the forgetting factor of the recursive algorithm. As suggested in [40], the simplest way of choosing the initial values \( \mu(-1) \) and \( \hat{\mathbf{g}}_{\text{opt}}(-1) \) is to set \( \mu(-1) = 1 \) and \( \hat{\mathbf{g}}_{\text{opt}}(-1) = [1, 0, \ldots, 0]^T \). Remarkably, the recursive rule (63) requires only \( O[N(L_\delta+1)] \) flops per iteration.

In conclusion, we can state that, in practical settings wherein \( \delta = L_{\text{cp}} \) (which implies that \( L_\delta = L_{\text{cp}} \) and \( L_\delta \) is significantly larger than \( L_{\text{cp}} \), the overall adaptive implementation of the proposed method involves a computational burden of only \( O \left[ N^2(L_\delta+1) \right] \) flops per iteration. For instance, for the representative parameters used in Section V, wherein \( N = 2, L_{\text{cp}} = 4 \) and \( L_\delta = 18 \), the proposed adaptive technique requires \( O(784) \) 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 \( L_{\text{cp}} = 4 \),

\[
\hat{P}_{\text{QMPD}}(k) = \begin{bmatrix} \Theta_\delta & -\Pi_{\hat{A}(k)}^{(a)} \end{bmatrix} \begin{bmatrix} \mathbf{O}(N(L_\delta+1) \times N(L_\delta+1)) & \mathbf{I}_{N(L_\delta+1)} \\ \hat{Q}_{\text{QMPD}}(k)^H & \hat{Q}_{\text{QMPD}}(k) \end{bmatrix} \begin{bmatrix} \mathbf{O}(N(L_\delta+1) \times N(L_\delta+1)) & \mathbf{I}_{N(L_\delta+1)} \end{bmatrix} \hat{F}_{\text{QMPD}}(k-1) \end{bmatrix} \end{bmatrix}
\]
which works at half-sampling spacing $T_c/2$ (i.e., $N = 2$). The system operates over a random FIR channel of order $L_d = 14$, whose taps $h_0(0), h_1(0), h_2(0), \ldots, h_{14}(0), h_{15}(14)$ are modeled as i.i.d. complex proper zero-mean Gaussian random variables, with variance $\sigma^2 = 2$. We consider either an NBI-free scenario, wherein $n(k) = \eta_{\text{noise}}(k)$, or an NBI-contaminated environment, wherein $n(k) = \eta_{\text{NBI}}(k) + \eta_{\text{noise}}(k)$. In the latter case, the continuous-time NBI signal is modeled as $w_{\text{NBI}}(t) = i(t)e^{-j2\pi f_b t}$, where $i(t)$ is a zero-mean WSS circular Gaussian process, with autocorrelation function $r_{ii}(\tau) \triangleq E\{i(t)\bar{i}(t-\tau)\} = \sigma^2 e^{-|\tau|}$, where $\sigma^2$ is the NBI power, $a > 0$ can be related to the 3-dB NBI one-sided bandwidth $W_{\text{NBI}}$ by the relation $W_{\text{NBI}} \approx 0.1a$, and $f_b$ is the NBI carrier frequency-offset. According to assumption (a4), the additive noise $\eta_{\text{noise}}(k)$ is modeled as a zero-mean WSS circular white Gaussian random vector, with autocorrelation matrix $\mathbf{R}_{\text{noise}} = \sigma^2_a \mathbf{I}_N$. The SNR is defined as $\frac{\sigma^2}{\sigma^2_a}$, whereas, for the NBI-contaminated scenario, the SIR is defined as $\frac{\sigma^2_a}{\sigma^2}$, and the parameters $a$ and $f_b$ are set equal to 0.05($N/T_c$) (corresponding to $W_{\text{NBI}} \approx 0.01/T_c$) and 31.5 ($N/T_d$), respectively.

Since the approach of [25] is implemented by 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 $\mathbf{H}$ and the disturbance autocorrelation matrix $\mathbf{R}_{\text{NBI}}$. 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 $f = [1, 0, \ldots, 0]^T$. For both the maximum-SSINR equalizer and the method of [25], we set $\Delta = 0$ and $I_{\text{eff}} = I_{\text{CR}}$. Regarding the proposed method, we consider parameterization P1 and we set $\delta = I_{\text{CR}}$, which, according to Theorem 1, leads to $\Delta = 0$ and $I_{\text{eff}} = I_{\text{CR}}$.

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 $\text{BER}_m$ is the output BER at the $m$th 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 $\mathbf{G}$ in (16) is assumed. For each Monte Carlo trial, after estimating the TEQ weights on the basis of the given data record length $K$ (expressed in OFDM symbols), an independent record of $K_{\text{alser}} = 200$ OFDM symbols is considered to evaluate the BER at the TEQ 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 $K_{\text{Cr}} = K_P = 1292$ 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-SSINR 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 TEQ 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

---

8Herein, the SNR is defined as the ratio between the average energy per symbol $\mathbf{E}\{||\mathbf{W}_{\text{TEQ}}(m, n)||^2\}/M = \sigma^2$ expended by the transmitter and the noise variance $\sigma^2_n$, 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 $\text{SNR}_{\text{TEQ-in}} = \mathbf{E}\{\sum_{k=0}^{L_d-1} h(k, n) w(k, n)\}/\mathbf{E}\{||w(k, n)||^2\}$, which, for the considered simulation setting, is given by $\text{SNR}_{\text{TEQ-in}} = \text{SNR}_{\text{TEQ}} \frac{L_d + 1}{N}$. Since $\text{SNR}_{\text{TEQ}} = 10 \log_{10} \text{SNR}_{\text{TEQ-in}}$, this results in $\text{SNR}_{\text{TEQ-in}} = 11.76$.

9As already pointed out, we consider method A, whose implementation is detailed in [25, steps 1–7, pp. 3258–3259].
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 $L_e + 1$ of the signal subspace (denoted with $N'$ in [25]) is augmented by the rank $R_{d3}$ of the NBI (i.e., the equalizer is synthesized by setting $N' = L_e + R_{d3} + 1$). 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-SSINR 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.
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 $k$ (SNR = 20, $L_e = 24$, NBI-contaminated environment).

In the first case, when the parameterization P1 is used, that is, $\delta \in \{0, 1, \ldots, \min(L_e, L_d)\}$, it is readily verified that $\mathcal{R}^\perp(\Theta_d) \supseteq \mathcal{R}^\perp(\Theta_{\delta-1})$, for any $\delta \in \{1, 2, \ldots, \delta\}$. Let the channel matrix be partitioned as $H = [H_{\text{left}}, H_{\text{right}}]$, with $H_{\text{left}} \in \mathbb{C}^{(L_e+1) \times (\delta+1)}$ and $H_{\text{right}} \in \mathbb{C}^{(L_e+1) \times (L_d-\delta)}$, the previous subspace inclusion implies that all the entries of the first $\delta + 1$ columns of $\Pi_i^T H$ are zero, thus yielding $\Pi_i^T H = \left[ O_{N(L_e-L_d) \times (\delta+1)}, \Pi_i^T H_{\text{right}} \right]$, with $H_{\text{right}} \triangleq [H_{\text{right}}, J] \in \mathbb{C}^{(L_e+1) \times (L_d-\delta)}$. Henceforth, using the fact [32] that $\left[ O_{N(L_e-L_d) \times (\delta+1)}, \Pi_i^T H_{\text{right}} \right] = \left[ O_{N(L_e-L_d) \times (\delta+1)} \right]$ $\Pi_i^T H_{\text{right}}$, after some straightforward algebraic manipulations, the asymptotic expression (43) of $b_{\text{thmoe}}$ is given by (64), where the second equality comes from the fact that, if the matrix $\Pi_i^T H_{\text{right}}$ is full-column rank (in this case, it turns out that $H_{\text{right}} = \Pi_i^T H_{\text{right}}$, with $\tilde{R}_0 = L_d + L_{\text{data}} - \delta$), then $H_{\text{right}} = \Pi_i^T H_{\text{right}}$, $\Pi_i^T H_{\text{right}}$, where $H_{\text{right}} = [H_{\text{right}}, J] \in \mathbb{C}^{(L_e+1) \times (L_d-\delta)}$. By using some results on the Moore–Penrose inverse of partitioned matrices [32], it can be proven that $\left[ O_{N(L_e-L_d) \times (\delta+1)}, \Pi_i^T H_{\text{right}} \right] = \left[ O_{N(L_e-L_d) \times (\delta+1)}, \Pi_i^T H_{\text{right}} \right] = $
\[
\left( \begin{bmatrix} \Pi_6^T \Pi_6 \end{bmatrix} \right)^T, \quad \left( \begin{bmatrix} \Pi_6^T \Pi_6 \end{bmatrix} \right) \right)^T \right] \quad \text{which, after some algebra, allows one to write the asymptotic expression (43) of } \beta_{\text{AMMOC}} \text{ as shown in (65). If the matrix }
\begin{bmatrix} \Pi_6^T \Pi_6 \end{bmatrix} \text{ is full-column rank (in this case, it turns out that } \Pi_6^T \Pi_6 = \Pi_6^T \Pi_6 \text{), with } R_6 = L_g + R_{\text{bla}}, \text{ it follows that the full-column rank matrices, which in their turns imply that }
\Pi_6^T \Pi_6 = \Pi_6^T \Pi_6 \text{ and } \Pi_6^T \Pi_6 \text{ are necessarily full-column rank matrices, which in their turns imply that }
\Pi_6^T \Pi_6 = \Pi_6^T \Pi_6 \text{ and } \Pi_6^T \Pi_6 \text{ is given by (66). Reasoning as previously done, it can be shown that, if the matrix } \Pi_6^T \Pi_6 \text{ is full-column rank (in this case, it turns out that } \Pi_6^T \Pi_6 = \Pi_6^T \Pi_6 \text{), with } R_6 = R_{\text{bla}} + \delta + 1, \text{ then }
H_6^T \Pi_6 \Pi_6 (H_6^T \Pi_6 \Pi_6)^T = I_{L_g},
\text{ and }
J_6^T \Pi_6 (J_6^T \Pi_6)^T = I_{R_{\text{bla}}} \quad \text{and } J_6^T \Pi_6 (J_6^T \Pi_6)^T = O_{R_{\text{bla}} \times R_{\text{bla}}},
\text{ and }
\Pi_6^T \Pi_6 = \Pi_6^T \Pi_6 \text{ is equivalent to impose that the null space of } \Pi_6^T \Pi_6 \text{ and the column space of } \Pi_6^T \Pi_6 \text{ intersect only trivially, i.e. } \mathcal{N}(\Pi_6^T \Pi_6) \cap \mathcal{R}(\Pi_6^T \Pi_6) = \{0_{N(L_g+1)}\}.
\text{At this point, observing that } \mathcal{N}(\Pi_6^T \Pi_6) \cap \mathcal{R}(\Pi_6^T \Pi_6) = \{0_{N(L_g+1)}\} \text{, it follows that the condition } \mathcal{R}(\Pi_6^T \Pi_6) \cap \mathcal{R}(\Pi_6^T \Pi_6) = \{0_{N(L_g+1)}\} \text{, 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


