Katholieke Universiteit Leuven

Departement Elektrotechniek

ESAT-SISTA/TR 2003-49

Linear and Decision-Feedback Per Tone Equalization for DMT-based Transmission over IIR Channels [1]

Koen Vanbleu, Marc Moonen [23]
and Geert Leus[4]

February 2005.
Accepted for publication in IEEE Transactions on Signal Processing.

[1]This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/sista/vanbleu/reports/03-49.pdf*

[2]K.U.Leuven, Dept. of Electrical Engineering (ESAT), Research group SCD, Kasteelpark Arenberg 10, 3001 Leuven-Heverlee, Belgium, Tel. 32/16/32 18 41, Fax 32/16/32 19 70, WWW: *http://www.esat.kuleuven.ac.be/sista*. E-mail: *koen.vanbleu@esat.kuleuven.ac.be*.

[3]This research work was carried out at the ESAT laboratory of the Katholieke Universiteit Leuven, in the frame of the Belgian State, Prime Minister's Office - Federal Office for Scientific, Technical and Cultural Affairs - Interuniversity Poles of Attraction Programme (2002-2007) - IUAP P5/22 ('Dynamical Systems and Control: Computation, Identification and Modelling') and P5/11 ('Mobile multimedia communication systems and networks'), the Concerted Research Action GOA-MEFISTO-666 (Mathematical Engineering for Information and Communication Systems Technology) of the Flemish Government, Research Project FWO nr.G.0196.02 ('Design of efficient communication techniques for wireless time-dispersive multi-user MIMO systems') and was partially sponsored by Alcatel-Bell. The scientific responsibility is assumed by its authors.

[4]Delft University of Technology, Faculty of Electrical Engineering, Mathematics and Computer Science, Mekelweg 4, 2628 CD Delft, The Netherlands, Tel. 31/15/278 43 27. E-mail: *leus@cas.et.tudelft.nl*.

Linear and Decision-Feedback Per Tone Equalization for DMT-based Transmission over IIR Channels

Koen Vanbleu[1]*, *IEEE Student Member,* Marc Moonen[1], *IEEE Member*, Geert Leus[2], *IEEE Member*

[1]*Address*:
Katholieke Universiteit Leuven, Dept. Elektrotechniek, ESAT-SCD, Kasteelpark Arenberg 10,
B-3001 Leuven-Heverlee, Belgium
Phone +32 16 32 18 41 - *Fax* +32 16 32 19 70
E-mail: {vanbleu,moonen}@esat.kuleuven.ac.be

[2]*Adress*:
Delft University of Technology, Faculty of Electrical Engineering, Mathematics and Computer Science,
Mekelweg 4, 2628 CD Delft, The Netherlands
Phone +31 15 278 43 27
E-mail: leus@cas.et.tudelft.nl

EDICS: 3-TDSL Telephone Networks and Digital Subscriber Loops, 3-COMM Signal Processing for Communications

This research work was carried out at the ESAT laboratory of the Katholieke Universiteit Leuven, in the frame of the Belgian State, Prime Minister's Office - Federal Office for Scientific, Technical and Cultural Affairs - Interuniversity Poles of Attraction Programme (2002-2007) - IUAP P5/22 ('Dynamical Systems and Control: Computation, Identification and Modelling') and P5/11 ('Mobile multimedia communication systems and networks'), the Concerted Research Action GOA-MEFISTO-666 (Mathematical Engineering for Information and Communication Systems Technology)of the Flemish Government and Research Project FWO nr.G.0196.02 ('Design of efficient communication techniques for wireless time-dispersive multi-user MIMO systems') and was partially sponsored by Alcatel-Bell. The scientific responsibility is assumed by its authors.

Abstract

The per-tone equalizer (PTEQ) has been presented as an attractive alternative for the classical time-domain equalizer (TEQ) in discrete multitone (DMT) based systems, such as ADSL systems. The PTEQ is based on a linear minimum mean-square-error (L-MMSE) equalizer design for each separate tone. In this paper, we reconsider DMT modulation and equalization in the ADSL context under the realistic assumption of an infinite impulse response (IIR) model for the wireline channel. First, optimum *linear zero-forcing (L-ZF) block equalizers* for arbitrary IIR model orders and cyclic prefix (CP) lengths are developed. It is shown that these L-ZF block equalizers can be decoupled per tone, hence they lead to an *L-ZF PTEQ*. Then, based on the L-ZF PTEQ, low-complexity *L-MMSE PTEQ extensions* are developed: the linear PTEQ extension exploits frequency-domain transmit redundancy from pilot and unused tones; alternatively, a closely related decision-feedback PTEQ extension can be applied. The PTEQ extensions then add flexibility to a DMT-based system design: the CP overhead can be reduced by exploiting frequency-domain transmit redundancy instead, so that a similar bitrate as with the original PTEQ is achieved at a lower memory and computational cost or, alternatively, a higher bitrate is achieved without a considerable cost increase. Both PTEQ extensions are also shown to improve the receiver's robustness to narrow-band interference.

Index Terms

Discrete Multitone, Digital Subscriber Lines, Per-Tone Equalization, Linear Equalization, Decision-Feedback Equalization, Narrow-Band Interference Suppression

I. INTRODUCTION

Discrete multitone (DMT) modulation and orthogonal frequency division multiplexing (OFDM) are all-digital multicarrier modulation schemes. DMT modulation is adopted as the transmission format for asymmetric digital subscriber lines (ADSL) and very high bit rate digital subscriber lines (VDSL); OFDM is adopted for wireless local area applications, e.g., IEEE 802.11/a and HiperLAN/2.

DMT schemes divide the available bandwidth into parallel subchannels or tones. The incoming bit-stream is split into parallel symbol streams that are used to QAM-modulate the different tones. An N-point inverse discrete Fourier transform (IDFT) is used for modulation. Before transmission of a DMT symbol, a cyclic prefix (CP) of ν samples is added. If the channel impulse response length is smaller than or equal to $\nu + 1$, intersymbol interference (ISI) between and intercarrier interference (ICI) within DMT symbols are avoided. Demodulation can then be done by means of a DFT, followed by a (complex-valued) 1-tap frequency domain equalizer (FEQ) per tone to compensate for channel amplitude and phase effects.

In this paper, we consider DMT modulation and equalization in the ADSL context. Practical ADSL channel impulse responses can be very long, hence a long CP would be required. However, a long CP introduces a large overhead, resulting in a reduced bitrate. An existing solution for this problem is to insert a (real-valued) T-tap *time-domain equalizer* (TEQ) before demodulation that shortens the channel

impulse response to $\nu + 1$ samples, where ν is only a fraction of the DFT-size N (e.g., $N = 512$ and $\nu = 32$ in ADSL). The TEQ design objective in ADSL is then to minimize ISI/ICI so that the aggregate number of bits transmitted over all tones, hence the bitrate, is maximized. In the past, many - in this respect suboptimum - TEQ design procedures have been developed (e.g., see [1], [2], [3], [4], [5]). Recently, a truly bitrate maximizing TEQ has been presented in [6], that closely approaches the performance of the so-called *per-tone equalizer* (PTEQ). The PTEQ has been presented in [7] as an attractive alternative equalizer scheme that always performs at least as well as - and usually better than - a TEQ-based receiver in terms of bitrate while keeping complexity during data transmission at the same level. A complex-valued linear minimum mean-square-error (L-MMSE) T-tap equalizer is then designed for each tone separately. In [8], [9], efficient, direct L-MMSE PTEQ design algorithms have been proposed, which are based on an adaptive RLS or a hybrid RLS/LMS algorithm; they owe their low computational and memory cost to the RLS processing of a set of common PTEQ inputs, the so-called difference terms, which is shared by all tones. In [10], it has been shown that a PTEQ-based DMT receiver with a sufficient number of taps has an increased robustness to narrow-band interference (NBI), when compared to a TEQ-based receiver, even if the latter includes a receiver window.

In this paper, we revisit and extend the results that we presented in [11]. As the ADSL transmission channel impulse response typically has a long tail, a parsimonious infinite impulse response (IIR) or pole-zero model has been previously adopted in, e.g., [12], [13]. We reconsider DMT equalization under the assumption of such an IIR channel model. Based on a corresponding DMT block data model, we first derive necessary and sufficient conditions for the existence of a *linear zero-forcing (L-ZF) block equalizer* which include conditions on the IIR channel order (i.e., the numerator and denominator order) and the required amount of transmit (TX) redundancy per DMT symbol block. Then, we derive *optimum L-ZF block equalizers* for arbitrary IIR channel order and CP length, which appear to allow for a computationally advantageous decoupling per tone, hence they lead to a so-called *L-ZF PTEQ*, i.e., with only one tone-dependent input. The development of these L-ZF block and per-tone equalizers forms the theoretical basis to propose *low-complexity extensions of the original L-MMSE PTEQ*, which accommodates an arbitrary IIR channel order and CP length. The so-called *linear PTEQ (L-PTEQ) extension* makes use of the frequency-domain (FD) TX redundancy that is provided in DMT transmission by unused tones and pilot tones: by adding a few DFT outputs of unused and pilot tones (and the corresponding pilot symbols), containing ISI/ICI from the active data-carrying tones, as common inputs to the PTEQ, the equalization of these active tones is enhanced. The so-called *decision-feedback PTEQ (DF-PTEQ) extension* is based on the observation that each decision on an FD data symbol can be treated in the same way as the a priori knowledge of a pilot symbol: the ISI/ICI in the DFT output of the considered active tone can again be exploited to enhance the equalization of the remaining active tones. The DF-PTEQ extension then feeds back a few symbol decisions within the DMT symbol and uses them, together with the corresponding DFT

outputs, as extra, common PTEQ inputs. The RLS-based and hybrid RLS/LMS-based design algorithms of [8], [9] are then straightforwardly extended with shared RLS processing for these extra, common inputs. The simulations show that DMT systems, employing an appropriate L-PTEQ or DF-PTEQ extension, have extra flexibility: the CP overhead can be reduced so that the same bitrate is achieved at a lower computational and memory cost or, alternatively, a higher bitrate is reached without considerable cost increase. Moreover, the L-PTEQ and DF-PTEQ extension have an increased robustness to NBI: as has been shown in [14], linearly combining the DFT outputs of unused tones that are affected by NBI allows to estimate and suppress spectral leakage of this NBI on neighbouring tones. Hence, in addition to the above mentioned NBI robustness, the L-PTEQ and DF-PTEQ extension enhance the NBI cancellation if the included unused and/or feedback tones are affected by NBI.

In the context of DMT/OFDM transmission, the introduction of TX redundancy has been previously studied and exploited under different forms (e.g., at bit level: through channel coding; in the time-domain (TD), i.e., after the symbol mapping: by means of a CP, zero padding or known-symbol padding; in the frequency-domain (FD): by means of pilot symbols) and for several purposes (e.g., to improve symbol detection (and bit error rate), to simplify equalization, to guarantee perfect ZF equalization, for blind channel estimation or for blind direct equalizer design). Specifically, we have noted in [11], and it was observed independently in [15], [16], [17], that FD TX redundancy from unused and pilot tones can enhance the equalization performance in the case of an insufficiently long or even absent CP. Apart from giving a more thorough and accurate description than in [11], this paper extends the results of [11] in two ways: in addition to an L-PTEQ, a low-complexity DF-PTEQ extension is developed and the NBI suppression capability of the L-PTEQ and DF-PTEQ extensions are motivated and investigated. Throughout the paper, we will indicate how our work relates to and is a generalization of the results obtained in [15], [16], [17]. In [18], [19], decision-feedback equalization (DFE) structures for OFDM transmission with an insufficiently long CP have already been presented. In [18], an OFDM system without cyclic prefix is considered; the ISI from the previous ODFM symbol is first removed in a DF fashion, followed by a linear equalization of the ISI-free OFDM symbol to remove ICI. In [19], a ZF-DFE and an MMSE-DFE are presented: in both DFEs, all decisions on the previous and current OFDM symbol are fed back; in addition, the MMSE-DFE uses three consecutive receive (RX) DMT symbols (the current, the previous and the next symbol) in the forward path. All tones are equalized in a joint, block-wise fashion, resulting in a computational and memory complexity of $\mathcal{O}(N^2)$, which is excessively high for ADSL. On the other hand, the complexity of the DF-PTEQ extension developed in this paper is $\mathcal{O}(N)$.

Section II develops a DMT block data model based on an IIR channel model and summarizes the original L-MMSE PTEQ design, including the adaptive RLS-based design algorithm. Based on the data model, two cases are considered. **Section III** deals with the case of a numerator order that is smaller

than or equal to the CP length; **Section IV** deals with the case of a numerator order that is larger than the CP length. For both cases, necessary and sufficient L-ZF conditions are derived, the optimum L-ZF block equalizer is developed and its reduction to an L-ZF PTEQ is discussed. Based on the L-ZF block equalizer and PTEQ, the low-complexity L-PTEQ and DF-PTEQ extensions are presented in **Section V**. The NBI suppression capability of the PTEQ extensions is also discussed. **Section VI** shows simulation results for different scenarios (several loops with and without NBI), different amounts of exploited FD TX redundancy, CP lengths and numbers of equalizer taps. **Section VII** concludes the paper.

Notation : A tilde is added over an FD symbol, to distinguish it from a TD symbol. Vectors are typeset in bold lowercase while matrices are in bold uppercase. $\mathcal{E}\{\cdot\}$ is the expectation operator. The transpose, Hermitian and complex conjugate operator are denoted by $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^*$, respectively. $\Re\{\cdot\}$ and $\Im\{\cdot\}$ are the real and imaginary operator. The m-th entry of a vector $\mathbf{a}$ is denoted as $\mathbf{a}[m]$, where the index m starts at zero. A diagonal matrix with $\mathbf{a}$ on the diagonal is denoted as $\mathrm{diag}(\mathbf{a})$. n is a tone index; N is the (I)DFT size; $\mathcal{F}_N$ is a unitary DFT matrix of size N; the n-th DFT row is $\mathcal{F}_n$, $n = 0, \ldots, N-1$; ν is the CP length; k is the DMT symbol time index. The $m \times m$ identity matrix is denoted as $\mathbf{I}_m$. The $l \times m$ all-zero matrix is denoted as $\mathbf{0}_{l\times m}$.

The complex-valued FD vector $\tilde{\mathbf{x}}_{k,N} = \begin{bmatrix} \tilde{x}_{k,0} & \cdots & \tilde{x}_{k,N-1} \end{bmatrix}^T$ is the k-th $N \times 1$ TX symbol vector that is fed to the modulating IDFT. The k-th TX symbol on tone n is $\tilde{x}_{k,n}$. As this paper deals with DMT-based systems, we assume baseband transmission, hence $\tilde{\mathbf{x}}_{k,N}$ has complex conjugate symmetry: $\tilde{x}_{k,0}$ and $\tilde{x}_{k,\frac{N}{2}}$ are real-valued and $\begin{bmatrix} \tilde{x}_{k,1} & \cdots & \tilde{x}_{k,\frac{N}{2}-1} \end{bmatrix} = \begin{bmatrix} \tilde{x}^*_{k,N-1} & \cdots & \tilde{x}^*_{k,\frac{N}{2}+1} \end{bmatrix}$. In the derivations, we will assume for simplicity that $\tilde{\mathbf{x}}_{k,N}$ can be partitioned as $\tilde{\mathbf{x}}_{k,N} = \begin{bmatrix} \tilde{\mathbf{x}}^H_{k,a} & \tilde{\mathbf{x}}^H_{k,p} & \tilde{\mathbf{x}}^H_{k,u} \end{bmatrix}^H$ where $\tilde{\mathbf{x}}_{k,a}$, $\tilde{\mathbf{x}}_{k,p}$ and $\tilde{\mathbf{x}}_{k,u} = \mathbf{0}$ are TX symbol vectors for the set of N_a active data-carrying tones $\mathcal{S}_a = \{0, 1, \ldots, N_a - 1\}$, the set of N_p pilot tones $\mathcal{S}_p = \{N_a, N_a + 1, \ldots, N_a + N_p - 1\}$ and the set of N_u unused tones $\mathcal{S}_u = \{N_a + N_p, \ldots, N_a + N_p + N_u - 1\}$, respectively (note that $N = N_a + N_p + N_u$). In practice, a different partitioning will be used, also because tones belonging to a certain set usually appear in complex conjugate pairs. E.g., the active tones 38 to 255 in ADSL downstream transmission give rise to a tone set, which includes the complex conjugate tones, i.e., $\mathcal{S}_a = \{38, \ldots, 255, 257, \ldots, 474\}$. Such alternative partitionings can be accommodated by redefining the DFT matrix $\mathcal{F}_N$ (and hence also the DFT operation) as a row-permuted version of the original DFT matrix. As the above notation suggests, the subscript N denotes FD vectors and matrices that take all N DFT bins into account, while the subscripts a, p and u are used in connection with the tone sets $\mathcal{S}_a$, $\mathcal{S}_p$ and $\mathcal{S}_u$. E.g., the $N_a \times N$ submatrix of $\mathcal{F}_N$ with the N_a rows that correspond to the active tone set $\mathcal{S}_a$ is denoted as $\mathcal{F}_a$; likewise, $\mathcal{F}_p$ is an $N_p \times N$ submatrix of $\mathcal{F}_N$ with rows corresponding to the pilot tone set $\mathcal{S}_p$. For the sake of conciseness, we will only consider active data-carrying tones $\mathcal{S}_a$ and pilot tones $\mathcal{S}_p$ in the derivations; the unused tones $\mathcal{S}_u$ can then be seen as a special case of pilot tones where the pilot symbols are zero.

II. DMT DATA MODEL AND MOTIVATION

A. Data model

The transmission channel impulse response in a wired communication system, such as ADSL, typically has a long tail. An IIR model then offers a parsimonious representation of the channel, as observed in [12], [13]. In this section, we develop a DMT block data model based on an IIR channel model.

Assume that the IIR channel model $H(z)$ has a propagation delay δ_0, a numerator $B(z)$ of order L_b and a denominator $A(z)$ of order L_a:

$$H(z) = \frac{B(z)}{A(z)} = \frac{z^{-\delta_0} \sum_{m=0}^{L_b} b_m z^{-m}}{\sum_{m=0}^{L_a} a_m z^{-m}} \tag{1}$$

with $a_0 = 1$. Then, the channel model (1) leads to the following relation between TX samples x_l, RX samples y_l and additive noise samples n_l (l is a sample index):

$$\sum_{m=0}^{L_a} a_m (y_{\delta_0+l-m} - n_{\delta_0+l-m}) = \sum_{m=0}^{L_b} b_m x_{l-m} \tag{2}$$

Without loss of generality, we will assume from now on that the propagation delay δ_0 is equal to 0. The input-output relation (2) gives rise to the following block-based description of N equations:

$$\underbrace{\begin{bmatrix} [& \bar{\mathbf{a}}^T &] & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & & \ddots & \\ 0 & \cdots & 0 & [& \bar{\mathbf{a}}^T &] \end{bmatrix}}_{\mathbf{A}_{\mathrm{T}}} (\mathbf{y}_{k,-L_a:N-1} - \mathbf{n}_{k,-L_a:N-1}) = \underbrace{\begin{bmatrix} [& \bar{\mathbf{b}}^T &] & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & & \ddots & \\ 0 & \cdots & 0 & [& \bar{\mathbf{b}}^T &] \end{bmatrix}}_{\mathbf{B}_{\mathrm{T}}} \mathbf{x}_{k,-L_b:N-1} \tag{3}$$

where $\mathbf{A}_{\mathrm{T}}$ (of size $N \times (N + L_a)$) and $\mathbf{B}_{\mathrm{T}}$ (of size $N \times (N + L_b)$) are Toeplitz convolution matrices, which are built with the vectors of denominator and numerator coefficients $\mathbf{a} = \begin{bmatrix} a_0 & \cdots & a_{L_a} \end{bmatrix}^T$ and $\mathbf{b} = \begin{bmatrix} b_0 & \cdots & b_{L_b} \end{bmatrix}^T$, respectively, in reverse order (hence denoted with a bar), and where we adopt the following notation for the k-th sample vectors, $\mathbf{y}_{k,-L_a:N-1}$, $\mathbf{n}_{k,-L_a:N-1}$ and $\mathbf{x}_{k,-L_b:N-1}$:

$$\mathbf{v}_{k,i:j} = \begin{bmatrix} v_{k,i} & \cdots & v_{k,j} \end{bmatrix}^T \text{ with } v_{k,m} = v_{(k-1)(N+\nu)+\nu+m} \tag{4}$$

The choice of N equations in (3) will give rise to an elegant data model as N corresponds to the DFT size. It relates the k-th TX and RX sample vectors, $\mathbf{x}_{k,-L_b:N-1}$ and $\mathbf{y}_{k,-L_a:N-1}$, respectively, which both include all or part of the CP samples, depending on L_b and L_a.

In the following steps, we show how the linear convolutions with $\mathbf{a}$ and $\mathbf{b}$, in (3), can be turned into circular convolutions with appropriate correction terms at the interval edges. First of all, $\mathbf{A}_{\mathrm{T}}$, $\mathbf{B}_{\mathrm{T}}$, $\mathbf{y}_{k,-L_a:N-1}$ and $\mathbf{x}_{k,-L_b:N-1}$ in (3) are split as follows:

$$\underbrace{\left[\begin{array}{c|c} \mathbf{A}_\Delta & \mathbf{A}'_{\mathrm{T}} \end{array}\right]}_{\mathbf{A}_{\mathrm{T}}} \underbrace{\begin{bmatrix} \mathbf{y}_{k,-L_a:-1} \\ \mathbf{y}_{k,0:N-1} \end{bmatrix}}_{\mathbf{y}_{k,-L_a:N-1}} - \check{\mathbf{n}}_k = \underbrace{\left[\begin{array}{c|c} \mathbf{B}_\Delta & \mathbf{B}'_{\mathrm{T}} \end{array}\right]}_{\mathbf{B}_{\mathrm{T}}} \underbrace{\begin{bmatrix} \mathbf{x}_{k,-L_b:-1} \\ \mathbf{x}_{k,0:N-1} \end{bmatrix}}_{\mathbf{x}_{k,-L_b:N-1}} \tag{5}$$

The tall $N \times L_a$ Toeplitz matrix $\mathbf{A}_\Delta$ is given by:

$$\mathbf{A}_\Delta = \left[\begin{array}{ccc} a_{L_a} & \cdots & a_1 \\ \vdots & \ddots & \vdots \\ 0 & 0 & a_{L_a} \\ \hline \multicolumn{3}{c}{\mathbf{0}_{(N-L_a)\times L_a}} \end{array} \right] \tag{6}$$

The tall $N \times L_b$ Toeplitz matrix $\mathbf{B}_\Delta$ is similarly obtained. $\mathbf{A}_{\mathrm{T}}^{'}$ and $\mathbf{B}_{\mathrm{T}}^{'}$ are lower triangular Toeplitz matrices of size $N \times N$ with appropriately zero-padded vectors $\mathbf{a}$ and $\mathbf{b}$, respectively, as their first column vectors. The vector $\check{\mathbf{n}}_k = \mathbf{A}_{\mathrm{T}} \mathbf{n}_{k,-L_a:N-1}$ denotes an $N \times 1$ sample vector of channel noise n_l, coloured by the denominator $\mathbf{a}$. Equation (5) can be transformed into:

$$\mathbf{A}_{\mathrm{C}} \mathbf{y}_{k,0:N-1} + \mathbf{A}_\Delta \Delta \mathbf{y}_k = \mathbf{B}_{\mathrm{C}} \mathbf{x}_{k,0:N-1} + \mathbf{B}_\Delta \Delta \mathbf{x}_k + \check{\mathbf{n}}_k \tag{7}$$

where $\mathbf{A}_{\mathrm{C}}$ and $\mathbf{B}_{\mathrm{C}}$ are $N \times N$ circulant matrices with appropriately zero-padded vectors $\mathbf{a}$ and $\mathbf{b}$; they are given by $\mathbf{A}_{\mathrm{C}} = \mathbf{A}_{\mathrm{T}}^{'} + \mathbf{A}_\Delta \left[\begin{array}{c|c} \mathbf{0}_{L_a \times (N-L_a)} & \mathbf{I}_{L_a} \end{array} \right]$ and $\mathbf{B}_{\mathrm{C}} = \mathbf{B}_{\mathrm{T}}^{'} + \mathbf{B}_\Delta \left[\begin{array}{c|c} \mathbf{0}_{L_b \times (N-L_b)} & \mathbf{I}_{L_b} \end{array} \right]$. The TX and RX difference terms $\Delta \mathbf{x}_k$ and $\Delta \mathbf{y}_k$, which are real-valued in the case of baseband transmission, are defined as:

$$\Delta \mathbf{x}_k = \left[\begin{array}{ccc} \Delta x_{k,-L_b} & \cdots & \Delta x_{k,-1} \end{array} \right]^T = \left[\begin{array}{ccc} (x_{k,-L_b} - x_{k,N-L_b}) & \cdots & (x_{k,-1} - x_{k,N-1}) \end{array} \right]^T \tag{8}$$

$$\Delta \mathbf{y}_k = \left[\begin{array}{ccc} \Delta y_{k,-L_a} & \cdots & \Delta y_{k,-1} \end{array} \right]^T = \left[\begin{array}{ccc} (y_{k,-L_a} - y_{k,N-L_a}) & \cdots & (y_{k,-1} - y_{k,N-1}) \end{array} \right]^T \tag{9}$$

The sample vector $\mathbf{x}_{k,0:N-1}$ in (7) is the TX IDFT output, i.e., $\mathbf{x}_{k,0:N-1} = \mathcal{F}_N^H \tilde{\mathbf{x}}_{k,N}$ with $\mathcal{F}_N$ the $N \times N$ (unitary) DFT matrix and $\tilde{\mathbf{x}}_{k,N}$ the TX symbol vector. Similarly, $\tilde{\mathbf{y}}_{k,N} = \mathcal{F}_N \mathbf{y}_{k,0:N-1}$ is the DFT of the k-th unequalized RX sample vector.

When deriving equalizers based on (7) in the next sections, we ignore the fact that a receiver typically includes a decision delay δ, an equalizer design parameter that allows for a (slightly) acausal equalizer to optimize performance[1]. We will assume that $\delta = 0$ to keep the derivations tractable, but the extensions to $\delta > 0$ and $\delta < 0$ follow a similar reasoning, based on an appropriately modified data model (7).

The final data model and starting point for the equalizer derivations in the next sections is obtained by taking the N-point DFT of the set of equations (7) and exploiting a DFT-based decomposition of the circulant matrices $\mathbf{A}_{\mathrm{C}}$ and $\mathbf{B}_{\mathrm{C}}$, e.g., $\mathbf{B}_{\mathrm{C}} = \mathcal{F}_N^H \tilde{\mathbf{B}}_{N,\mathrm{D}} \mathcal{F}_N$, where $\tilde{\mathbf{B}}_{N,\mathrm{D}} = \mathrm{diag}(\tilde{\mathbf{b}}_N)$ with $\tilde{\mathbf{b}}_N = \left[\begin{array}{ccc} \tilde{b}_0 & \cdots & \tilde{b}_{N-1} \end{array} \right]^T = \mathcal{F}_N \left[\begin{array}{cc} \mathbf{b}^T & \mathbf{0}_{(N-L_b-1)\times 1}^T \end{array} \right]^T$, i.e., $\tilde{\mathbf{B}}_{N,\mathrm{D}}$ is a diagonal matrix with on the diagonal the DFT of the zero-padded vector $\mathbf{b}$ ($\mathbf{A}_{\mathrm{C}}$ can be decomposed in the same way):

$$\tilde{\mathbf{A}}_{N,\mathrm{D}} \tilde{\mathbf{y}}_{k,N} + \mathcal{F}_N \mathbf{A}_\Delta \Delta \mathbf{y}_k = \tilde{\mathbf{B}}_{N,\mathrm{D}} \tilde{\mathbf{x}}_{k,N} + \mathcal{F}_N \mathbf{B}_\Delta \Delta \mathbf{x}_k + \mathcal{F}_N \check{\mathbf{n}}_k \tag{10}$$

[1]Often, the aforementioned propagation delay δ_0 and the decision delay δ are combined into a single synchronization design parameter $\Delta = \delta_0 + \delta$.

The first right-hand side term of (10) can then be split into a contribution from the data symbols, $\tilde{\mathbf{x}}_{k,a}$, and the pilot symbols, $\tilde{\mathbf{x}}_{k,p}$:

$$\tilde{\mathbf{B}}_{N,\mathrm{D}}\tilde{\mathbf{x}}_{k,N} = \begin{bmatrix} \tilde{\mathbf{B}}_{a,\mathrm{D}} \\ \mathbf{0}_{N_p\times N_a} \end{bmatrix} \tilde{\mathbf{x}}_{k,a} + \begin{bmatrix} \mathbf{0}_{N_a\times N_p} \\ \tilde{\mathbf{B}}_{p,\mathrm{D}} \end{bmatrix} \tilde{\mathbf{x}}_{k,p} \tag{11}$$

where $\tilde{\mathbf{B}}_{a,\mathrm{D}} = \mathrm{diag}(\tilde{\mathbf{b}}_a)$ and $\tilde{\mathbf{B}}_{p,\mathrm{D}} = \mathrm{diag}(\tilde{\mathbf{b}}_p)$ and where $\tilde{\mathbf{b}}_a$ and $\tilde{\mathbf{b}}_p$ are obtained from $\tilde{\mathbf{b}}_N$ by selection of the entries corresponding to $\mathcal{S}_a$ and $\mathcal{S}_p$, respectively. Rearranging (10), so that all known data (i.e., the RX signal y_l and the pilot symbols $\tilde{\mathbf{x}}_{k,p}$) are grouped on the left-hand side, results in:

$$\boxed{\tilde{\mathbf{z}}_{k,N} = \begin{bmatrix} \tilde{\mathbf{B}}_{a,\mathrm{D}} \\ \mathbf{0} \end{bmatrix} \tilde{\mathbf{x}}_{k,a} + \mathcal{F}_N\mathbf{B}_\Delta\Delta\mathbf{x}_k + \mathcal{F}_N\check{\mathbf{n}}_k} \tag{12}$$

with

$$\boxed{\tilde{\mathbf{z}}_{k,N} = \begin{bmatrix} \tilde{\mathbf{z}}_{k,a} \\ \tilde{\mathbf{z}}_{k,p} \end{bmatrix} = \tilde{\mathbf{A}}_{N,\mathrm{D}}\tilde{\mathbf{y}}_{k,N} + \mathcal{F}_N\mathbf{A}_\Delta\Delta\mathbf{y}_k - \begin{bmatrix} \mathbf{0} \\ \tilde{\mathbf{B}}_{p,\mathrm{D}} \end{bmatrix} \tilde{\mathbf{x}}_{k,p}} \tag{13}$$

This **IIR-channel-based DMT block data model** (12-13) relates the DFT output vector $\tilde{\mathbf{y}}_{k,N}$ with the TX symbol vectors $\tilde{\mathbf{x}}_{k,a}$ and $\tilde{\mathbf{x}}_{k,p}$ using some (correcting) TX and RX difference terms (8-9). Figure 1

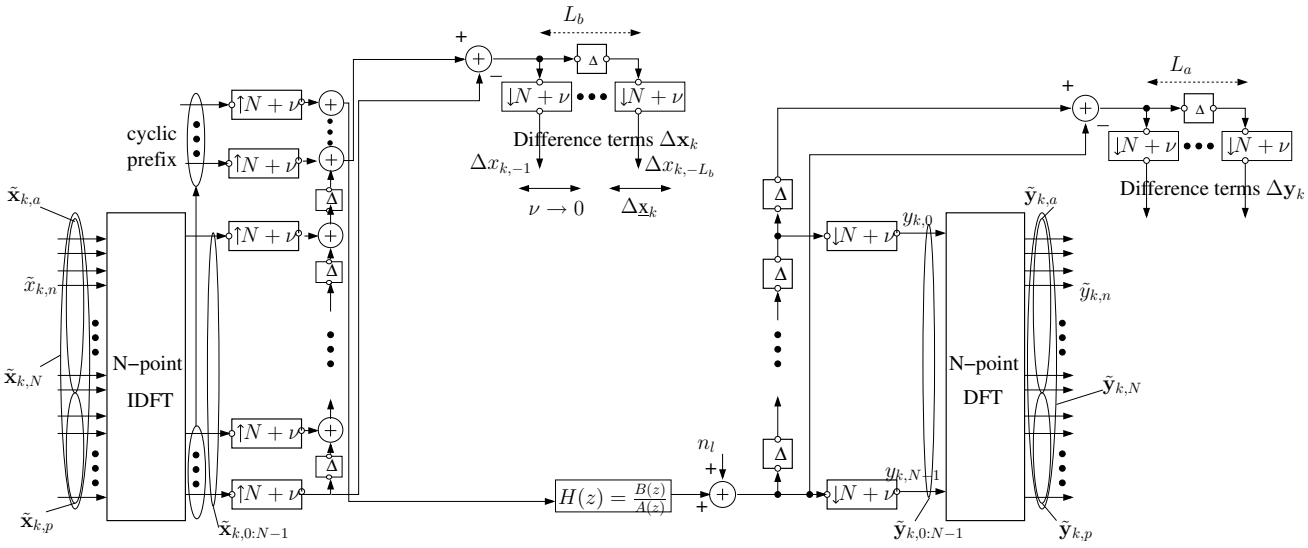


Fig. 1. Block diagram with the key signal samples and vectors in (12)-(13)

gives a block diagram that allows to interprete the different signal vectors in (12)-(13). The data model allows to isolate terms that cause ISI/ICI in each DFT output. After left multiplication of (12)-(13) with $\tilde{\mathbf{A}}_{N,\mathrm{D}}^{-1}$ and some rearranging, the DFT output $\tilde{\mathbf{y}}_{k,N}$ can be expressed as:

$$\tilde{\mathbf{y}}_{k,N} = \tilde{\mathbf{A}}_{N,\mathrm{D}}^{-1}\left(\begin{bmatrix} \tilde{\mathbf{B}}_{a,\mathrm{D}} \\ \mathbf{0} \end{bmatrix} \tilde{\mathbf{x}}_{k,a} + \begin{bmatrix} \mathbf{0} \\ \tilde{\mathbf{B}}_{p,\mathrm{D}} \end{bmatrix} \tilde{\mathbf{x}}_{k,p} + \mathcal{F}_N\mathbf{B}_\Delta\Delta\mathbf{x}_k - \mathcal{F}_N\mathbf{A}_\Delta\Delta\mathbf{y}_k + \mathcal{F}_N\check{\mathbf{n}}_k\right) \tag{14}$$

When considering the n-th DFT output, $\tilde{y}_{k,n} = \tilde{\mathbf{y}}_{k,N}[n]$, the diagonal nature of $\tilde{\mathbf{B}}_{a,\mathrm{D}}$, $\tilde{\mathbf{B}}_{p,\mathrm{D}}$ and $\tilde{\mathbf{A}}_{N,\mathrm{D}}$ makes that the first two right-hand side terms do not cause ISI/ICI: they cause a contribution from a TX symbol $\tilde{x}_{k,n}$, scaled with $\frac{\tilde{b}_n}{\tilde{a}_n}$, which either corresponds to a data symbol ($n \in \mathcal{S}_a$) or a pilot symbol ($n \in \mathcal{S}_p$). The third and fourth term do cause ISI/ICI, as both $\Delta\mathbf{x}_k$ and $\Delta\mathbf{y}_k$ are built with TD samples, hence they are a superposition of a desired signal contribution from $\tilde{x}_{k,n}$ and ISI/ICI, caused by the channel $H(z)$, from all other TX symbols $\tilde{x}_{l,m}$ with $l \neq k$ or $m \neq n$. The fifth term is the additive noise.

B. The per-tone equalizer (PTEQ)

The data model (12-13) is of special interest for DMT equalization as it suggests a relation with the PTEQ. The PTEQ has been presented in [7] as an attractive alternative for the TEQ to equalize a long FIR channel: the PTEQ follows from the observation that the DFT demodulation and the TEQ can be swapped, i.e., the equalizer can be moved behind the DFT and combined with the FEQ. Whereas a T-tap TEQ equalizes all tones with a single filter in a joint fashion, the T-tap PTEQ minimizes the mean-square-error (MSE) for each tone separately, hence the PTEQ optimizes the signal-to-noise ratio (SNR) and thus the bitrate for each tone. As a consequence, a PTEQ-based receiver always performs at least as well as - and usually much better than - a TEQ-based receiver in terms of bitrate, while keeping the data transmission computational cost at the same level. The T-tap PTEQ $\mathbf{v}_n^{\mathrm{PTEQ}}$ for tone n is the solution of the following L-MMSE design criterion [7]:

$$\boxed{\mathbf{v}_n^{\mathrm{PTEQ}} = \arg\min_{\mathbf{v}_n} \mathcal{E}\left\{ \left| \begin{bmatrix} \Delta\mathbf{y}_k^T & \tilde{y}_{k,n} \end{bmatrix} \mathbf{v}_n - \tilde{x}_{k,n} \right|^2 \right\}} \tag{15}$$

hence the **L-MMSE PTEQ for tone** n linearly combines one single tone-dependent DFT output, $\tilde{y}_{k,n}$, with $T-1$ RX difference terms $\Delta\mathbf{y}_k$. If $L_a = T-1$, these difference terms are the same as those defined in (9). These difference terms are thus common PTEQ inputs for all tones. The fact that there is only one tone-dependent PTEQ input decouples the equalizer design per tone and, at the same time, has a beneficial impact on the computational and memory cost. An efficient direct PTEQ design algorithm has been presented in [8], based on an adaptive square-root recursive-least-squares (RLS) algorithm. Using the so-called ADSL medley signal, i.e., a stream of DMT training symbols that is transmitted during connection set-up, an RLS-update of the PTEQ for tone n is given by[2]:

$$\mathbf{v}_n \leftarrow \mathbf{v}_n + (1-\lambda)\mathbf{R}_n^{-1} \begin{bmatrix} \Delta\mathbf{y}_k \\ \tilde{y}_{k,n}^* \end{bmatrix} \left(\tilde{x}_{k,n} - \begin{bmatrix} \Delta\mathbf{y}_k^T & \tilde{y}_{k,n} \end{bmatrix} \mathbf{v}_n \right) \tag{16}$$

with

$$\mathbf{R}_n \leftarrow \lambda\mathbf{R}_n + (1-\lambda) \begin{bmatrix} \Delta\mathbf{y}_k \\ \tilde{y}_{k,n}^* \end{bmatrix} \begin{bmatrix} \Delta\mathbf{y}_k^T & \tilde{y}_{k,n} \end{bmatrix} \tag{17}$$

[2]By including an exponential weighting λ, one allows for tracking of a changing environment.

As suggested in [8], it is preferable to store and update the inverse transpose of the Cholesky factor (square-root) of $\mathbf{R}_n$, i.e., the lower triangular matrix $\mathbf{S}_n$ with $\mathbf{R}_n^{-1} = \mathbf{S}_n^H \mathbf{S}_n$, rather than to store and update $\mathbf{R}_n$ itself for all active tones $\mathcal{S}_a$. By construction, the $T-1$ first rows of $\mathbf{S}_n$ constitute a triangular matrix, denoted by $\mathbf{S}_\Delta$ (such that $\left(\mathcal{E}\left\{\Delta\mathbf{y}_k \Delta\mathbf{y}_k^T\right\}\right)^{-1} = \mathbf{S}_\Delta^T \mathbf{S}_\Delta$), which is real-valued and common for all tones, and hence tone-independent; one update requires $\mathcal{O}(T^2)$ computations and the memory cost is also $\mathcal{O}(T^2)$. Only the last row of $\mathbf{S}_n$, denoted by $\mathbf{s}_n$, is complex-valued and tone-dependent and its update requires $\mathcal{O}(T)$ computations and coefficients per active tone. The partitioning $\mathbf{S}_n = \begin{bmatrix} \mathbf{S}_\Delta^T & \mathbf{s}_n^H \end{bmatrix}^H$ then leads to an efficient computation of the update (16), based on the Kalman gain vector:

$$\mathbf{R}_n^{-1} \begin{bmatrix} \Delta\mathbf{y}_k \\ \tilde{y}_{k,n}^* \end{bmatrix} = \mathbf{S}_\Delta^T \mathbf{S}_\Delta \Delta\mathbf{y}_k + \mathbf{s}_n^H \left(\mathbf{s}_n \begin{bmatrix} \Delta\mathbf{y}_k \\ \tilde{y}_{k,n}^* \end{bmatrix} \right) \tag{18}$$

The first term is tone-independent and requires $\mathcal{O}(T^2)$ computations, while the second, tone-dependent term requires $\mathcal{O}(T)$ computations per active tone. In ADSL downstream where $N_a \gg T$, the overall memory and computational cost is then dominated by a cost term $\mathcal{O}(N_a T)$, which depends linearly on N_a and T. For details, we refer to [8]. This cost can be further reduced with the so-called hybrid RLS/LMS-based PTEQ design algorithm, presented in [9].

C. Motivation

In Section III and IV, necessary and sufficient L-ZF conditions are derived and optimum L-ZF block equalizers are developed, based on the IIR-channel-based DMT block data model (12-13). In Section III, we consider the case of a numerator order L_b that is smaller than or equal to the CP length ($L_b \leq \nu$); Section IV deals with the case of a numerator order L_b that is larger than the CP length ($L_b > \nu$). In both cases, the L-ZF block equalizer turns out to reduce to an L-ZF PTEQ, i.e., with only one tone-dependent input. Based on the L-ZF PTEQ, we are then able to present low-complexity linear and decision-feedback extensions of the original L-MMSE PTEQ design (15) in Section V.

III. DMT EQUALIZATION OF AN IIR CHANNEL WITH $L_b \leq \nu$

In general, if all tones are active and data-carrying, i.e. $\mathcal{S}_a = \{0, \ \ldots, N-1\}$, $N = N_a$ and $\mathcal{S}_p = \emptyset$, the data model (12-13) does not have an L-ZF block equalizer. In the noiseless case, (12-13) then corresponds to an underdetermined set of N (real) equations in $N + L_b$ (real) unknown variables $\begin{bmatrix} \tilde{\mathbf{x}}_{k,a}^H & \Delta\mathbf{x}_k^T \end{bmatrix}^H$: the tones 0 and $\frac{N}{2}$ each result in one real equation and one real-valued variable $\tilde{x}_{k,n}$; the $\frac{N}{2}-1$ (complex-conjugate) tones each add two complex conjugate equations and two complex conjugate variables $\tilde{x}_{k,n}$ and $\tilde{x}_{k,n}^*$, which correspond to two real equations and two real-valued variables $\Re\{\tilde{x}_{k,n}\}$ and $\Im\{\tilde{x}_{k,n}\}$; finally, there are L_b unknown TX difference terms $\Delta\mathbf{x}_k$. However, the cyclic prefix renders ν difference terms $\Delta x_{k,l}$ equal to zero and there are oftentimes unused tones and/or pilot tones, so that the actual number of unknowns is typically smaller than $N + L_b$.

If $L_b \leq \nu$, all L_b TX difference terms $\Delta\mathbf{x}_k$ vanish, due to the cyclic prefix ($x_{k,N-l} = x_{k,-l}$ for $l = 1, \ldots, \nu$, see (8)), and so the data model (12-13) becomes:

$$\boxed{\tilde{\mathbf{z}}_{k,N} = \tilde{\mathbf{A}}_{N,\mathrm{D}}\tilde{\mathbf{y}}_{k,N} + \mathcal{F}_N\mathbf{A}_\Delta\Delta\mathbf{y}_k - \begin{bmatrix} \mathbf{0} \\ \tilde{\mathbf{B}}_{p,\mathrm{D}} \end{bmatrix}\tilde{\mathbf{x}}_{k,p} = \begin{bmatrix} \tilde{\mathbf{B}}_{a,\mathrm{D}} \\ \mathbf{0} \end{bmatrix}\tilde{\mathbf{x}}_{k,a} + \mathcal{F}_N\check{\mathbf{n}}_k} \tag{19}$$

Now, only the term with RX difference terms $\Delta\mathbf{y}_k$ causes ISI/ICI.

A. Optimum L-ZF block equalization

If $L_b \leq \nu$, the number of equations N in (19) is always larger than or equal to the number of unknowns N_a. Provided that a necessary and sufficient condition (see further) is met, there exist one (if $N = N_a$) or multiple (if $N > N_a$) L-ZF block equalizers. The MMSE L-ZF block equalizer, i.e., with the lowest MSE among all L-ZF block equalizers, leads to the following L-ZF block estimate [20], [21], [22]:

$$\hat{\tilde{\mathbf{x}}}_{k,a}^{\mathrm{L-ZF}} = \left(\underbrace{\begin{bmatrix} \tilde{\mathbf{B}}_{a,\mathrm{D}}^H & \mathbf{0} \end{bmatrix}\left(\mathcal{F}_N\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2\mathcal{F}_N^H\right)^{-1}\begin{bmatrix} \tilde{\mathbf{B}}_{a,\mathrm{D}} \\ \mathbf{0} \end{bmatrix}}_{\{1\}}\right)^{-1} \underbrace{\begin{bmatrix} \tilde{\mathbf{B}}_{a,\mathrm{D}}^H & \mathbf{0} \end{bmatrix}\left(\mathcal{F}_N\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2\mathcal{F}_N^H\right)^{-1}}_{\{2\}}\tilde{\mathbf{z}}_{k,N} \tag{20}$$

where $\mathcal{F}_N\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2\mathcal{F}_N^H$ is the autocorrelation matrix of the FD noise vector $\mathcal{F}_N\check{\mathbf{n}}_k$ with $\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2 = \mathcal{E}\left\{\check{\mathbf{n}}_k\check{\mathbf{n}}_k^T\right\} = \mathbf{A}_{\mathrm{T}}\mathcal{E}\left\{\mathbf{n}_{k,-L_a:N-1}\mathbf{n}_{k,-L_a:N-1}^T\right\}\mathbf{A}_{\mathrm{T}}^T$. In the derivations below, it is assumed that $\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2$ is non-singular. The factors $\{1\}$ and $\{2\}$ in (20) can be simplified to:

$$\begin{aligned} \{1\} \quad \begin{bmatrix} \tilde{\mathbf{B}}_{a,\mathrm{D}}^H & \mathbf{0} \end{bmatrix}\left(\mathcal{F}_N\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2\mathcal{F}_N^H\right)^{-1}\begin{bmatrix} \tilde{\mathbf{B}}_{a,\mathrm{D}} \\ \mathbf{0} \end{bmatrix} &= \tilde{\mathbf{B}}_{a,\mathrm{D}}^H\mathcal{F}_a\left(\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2\right)^{-1}\mathcal{F}_a^H\tilde{\mathbf{B}}_{a,\mathrm{D}} \\ \{2\} \quad \begin{bmatrix} \tilde{\mathbf{B}}_{a,\mathrm{D}}^H & \mathbf{0} \end{bmatrix}\left(\mathcal{F}_N\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2\mathcal{F}_N^H\right)^{-1} &= \tilde{\mathbf{B}}_{a,\mathrm{D}}^H\mathcal{F}_a\left(\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2\right)^{-1}\left[\begin{array}{c|c} \mathcal{F}_a^H & \mathcal{F}_p^H \end{array}\right] \end{aligned} \tag{21}$$

where the last equality makes use of the fact that $\mathcal{F}_N = \left[\begin{array}{c|c} \mathcal{F}_a^H & \mathcal{F}_p^H \end{array}\right]^H$. Combining (21) and (20) leads to the **optimum L-ZF block estimate**:

$$\boxed{\hat{\tilde{\mathbf{x}}}_{k,a}^{\mathrm{L-ZF}} = \left[\begin{array}{c|c} \tilde{\mathbf{B}}_{a,\mathrm{D}}^{-1} & \underbrace{\tilde{\mathbf{B}}_{a,\mathrm{D}}^{-1}\left(\mathcal{F}_a\left(\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2\right)^{-1}\mathcal{F}_a^H\right)^{-1}\overbrace{\mathcal{F}_a\left(\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2\right)^{-1}\mathcal{F}_p^H}^{\{1\}}}_{\mathbf{M}_1} \end{array}\right]\tilde{\mathbf{z}}_{k,N} = \tilde{\mathbf{B}}_{a,\mathrm{D}}^{-1}\tilde{\mathbf{z}}_{k,a} + \mathbf{M}_1\tilde{\mathbf{z}}_{k,p}} \tag{22}$$

with $\tilde{\mathbf{B}}_{a,\mathrm{D}}^{-1}$ a diagonal matrix and with $\mathbf{M}_1$ a non-sparse $N_a \times N_p$ matrix. It follows from (22) that the only condition for the existence of the L-ZF block equalizer if $L_b \leq \nu$ requires that $\tilde{\mathbf{b}}_a$ does not have zero entries, i.e., the channel $H(z)$, and hence the numerator $B(z)$, does not have zeros on the N-point DFT grid at frequencies that correspond to the tones $\mathcal{S}_a$. This condition is always fulfilled this condition as it is not possible to transmit data on the corresponding tones.

Using the definition of $\tilde{\mathbf{z}}_{k,N}$ (13), the **optimum L-ZF symbol estimate for tone** n, $\hat{\tilde{x}}_{k,n}^{\mathrm{L-ZF}}$, is given by:

$$\hat{\tilde{x}}_{k,n}^{\mathrm{L-ZF}} = \underbrace{\left[\begin{array}{c|c|c|c} \tilde{b}_n^{-1}\mathcal{F}_n\mathbf{A}_\Delta + \mathbf{m}_{n,1}\mathcal{F}_p\mathbf{A}_\Delta & \mathbf{m}_{n,1}\tilde{\mathbf{A}}_{p,\mathrm{D}} & -\mathbf{m}_{n,1}\tilde{\mathbf{B}}_{p,\mathrm{D}} & \tilde{b}_n^{-1}\tilde{a}_n \end{array}\right]}_{\left(\mathbf{v}_n^{\mathrm{L-ZF}}\right)^T} \left[\begin{array}{c} \Delta\mathbf{y}_k \\ \tilde{\mathbf{y}}_{k,p} \\ \tilde{\mathbf{x}}_{k,p} \\ \tilde{y}_{k,n} \end{array}\right], \; n \in \mathcal{S} \qquad (23)$$

where $\tilde{\mathbf{A}}_{p,\mathrm{D}} = \mathrm{diag}(\tilde{\mathbf{a}}_p)$ and $\tilde{\mathbf{a}}_p$ is obtained from $\tilde{\mathbf{a}}_N$ by selecting the entries that correspond to the pilot tone set $\mathcal{S}_p$, where $\mathbf{m}_{n,1}$ is the row of $\mathbf{M}_1$ that corresponds to tone $n \in \mathcal{S}_a$ and where $\mathbf{v}_n^{\mathrm{L-ZF}}$ is the optimum L-ZF equalizer for tone n.

B. Discussion and interpretation

One can effectively say that *the optimum L-ZF equalizer for tone n, $\mathbf{v}_n^{\mathrm{L-ZF}}$ in (23), is a PTEQ*, as it decouples the equalization per active tone n: the $(L_a + 1 + 2N_p)$-tap PTEQ $\mathbf{v}_n^{\mathrm{L-ZF}}$ linearly combines one single tone-dependent DFT output $\tilde{y}_{k,n}$ with tone-independent PTEQ inputs that are common for all tones, namely L_a real-valued RX difference terms $\Delta\mathbf{y}_k$, the N_p DFT outputs of the pilot tones $\tilde{\mathbf{y}}_{k,p}$ and the N_p pilot symbols $\tilde{\mathbf{x}}_{k,p}$. The overall computational and memory cost is then dominated by a term that depends linearly on N_a, L_a and N_p (see below). The concept of block equalization and per-tone equalization are depicted in Figure 2.

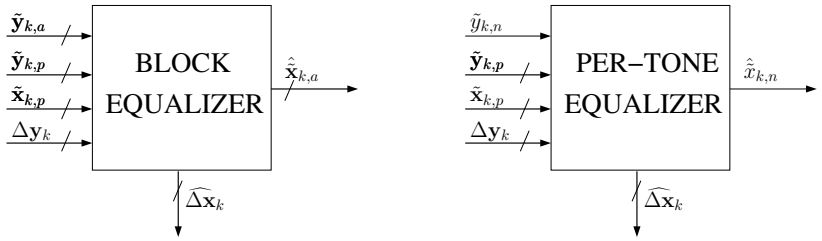


Fig. 2. Concept of a block equalizer and a per-tone equalizer.

If $N = N_a$ and $N_p = 0$, i.e., the number of equations and unknowns in (19) is equal, hence there are no pilot tones, $\mathbf{v}_n^{\mathrm{L-ZF}}$ in (23) reduces to an $(L_a + 1)$-tap L-ZF PTEQ, which linearly combines the n-th DFT output with L_a RX difference terms $\Delta\mathbf{y}_k$:

$$\boxed{\hat{\tilde{x}}_{k,n}^{\mathrm{L-ZF}} = \tilde{b}_n^{-1}\tilde{z}_{k,n} = \underbrace{\left[\begin{array}{c|c} \tilde{b}_n^{-1}\mathcal{F}_n\mathbf{A}_\Delta & \tilde{b}_n^{-1}\tilde{a}_n \end{array}\right]}_{\left(\mathbf{v}_n^{\mathrm{L-ZF}}\right)^T} \left[\begin{array}{c} \Delta\mathbf{y}_k \\ \tilde{y}_{k,n} \end{array}\right]} \qquad (24)$$

The first L_a coefficients of the PTEQ $\mathbf{v}_n^{\mathrm{L-ZF}}$ in (24) compensate for the ISI/ICI, caused by the denominator $A(z)$ of the IIR channel, while the last coefficient compensates for the amplitude and phase distortion $\frac{\tilde{b}_n}{\tilde{a}_n}$; the CP takes care of the time dispersion by the numerator $B(z)$, so that $B(z)$ does not cause ISI/ICI.

In fact, (24) is nothing but an alternative implementation of the early TEQ design of [12], which is based on an IIR channel identification of $H(z)$ in (1): the TEQ is then an FIR filter with transfer function $A(z)$ and the FEQs compensate for the discrete frequency response $\tilde{b}_n$ of $B(z)$, provided that $L_b \leq \nu$. The importance of the alternative formulation (24) lies in its close relation with the original PTEQ (15): *in the noiseless case, the L-ZF PTEQ* $\mathbf{v}_n^{\mathrm{L-ZF}}$ *(24) is equivalent to the* T*-tap PTEQ* $\mathbf{v}_n^{\mathrm{PTEQ}}$ *in (15), provided that* $L_b \leq \nu$ *and* $L_a + 1 \leq T$.

If $N > N_a$, (24) is a valid L-ZF PTEQ, be it that it does not provide the *MMSE* L-ZF symbol estimate: (23) suggests that the MMSE L-ZF PTEQ should have the DFT outputs of the pilot tones and the pilot symbols as additional, common inputs. However, it can be shown that *the MMSE L-ZF PTEQ (23) only loosely depends on the pilot tones and reduces for* $N \to \infty$ *to the simplified L-ZF PTEQ (24)*. The reasoning, which is developed below, is based on the asymptotic equivalence of Toeplitz and circulant matrices, as described in [23]. The $N \times N$ noise autocorrelation matrix $\mathbf{\Sigma}^2_{\check{\mathbf{n}}}$, which shows up in (22), is Toeplitz. If only a band of d diagonals around the main diagonal of $\mathbf{\Sigma}^2_{\check{\mathbf{n}}}$ is significantly different from zero, with $d \ll N$, $\mathbf{\Sigma}^2_{\check{\mathbf{n}}}$ can be well approximated by a circulant matrix $\mathbf{\Sigma}^2_{\check{\mathbf{n}},\mathrm{C}}$. It is said that $\mathbf{\Sigma}^2_{\check{\mathbf{n}}}$ and $\mathbf{\Sigma}^2_{\check{\mathbf{n}},\mathrm{C}}$ are asymptotically equivalent, i.e., for $N \to \infty$, $\mathbf{\Sigma}^2_{\check{\mathbf{n}}} = \mathbf{\Sigma}^2_{\check{\mathbf{n}},\mathrm{C}}$, which will be denoted as $\mathbf{\Sigma}^2_{\check{\mathbf{n}}} \sim \mathbf{\Sigma}^2_{\check{\mathbf{n}},\mathrm{C}}$ [23]. For larger N, smaller denominator order L_a and less coloured noise (i.e., with a shorter support of the noise autocorrelation function around lag zero), d/N is smaller and the approximation is better. Invoking a DFT-based decomposition, $\mathbf{\Sigma}^2_{\check{\mathbf{n}},\mathrm{C}} = \mathcal{F}_N^H \mathbf{\Sigma}^2_{N,\tilde{\check{\mathbf{n}}},\mathrm{D}} \mathcal{F}_N$, where $\mathbf{\Sigma}^2_{N,\tilde{\check{\mathbf{n}}},\mathrm{D}}$ is diagonal with the DFT of the first column of $\mathbf{\Sigma}^2_{\check{\mathbf{n}},\mathrm{C}}$ on the diagonal, $\mathbf{M}_1$ in (22) tends to zero for $N \to \infty$ because the following asymptotical equivalence holds true for the factor $\{1\}$ in (22):

$$\begin{aligned} \mathcal{F}_a \left(\mathbf{\Sigma}^2_{\check{\mathbf{n}}}\right)^{-1} \mathcal{F}_p^H \sim \mathcal{F}_a \left(\mathbf{\Sigma}^2_{\check{\mathbf{n}},\mathrm{C}}\right)^{-1} \mathcal{F}_p^H &= \mathcal{F}_a \mathcal{F}_N^H \left(\mathbf{\Sigma}^2_{N,\tilde{\check{\mathbf{n}}},\mathrm{D}}\right)^{-1} \mathcal{F}_N \mathcal{F}_p^H \\ &= \begin{bmatrix} \mathbf{I}_{N_a} & \mathbf{0}_{N_a \times N_p} \end{bmatrix} \left(\mathbf{\Sigma}^2_{N,\tilde{\check{\mathbf{n}}},\mathrm{D}}\right)^{-1} \begin{bmatrix} \mathbf{0}_{N_a \times N_p} \\ \mathbf{I}_{N_p} \end{bmatrix} = \mathbf{0}_{N_a \times N_p} \end{aligned} \tag{25}$$

Hence, for $N \to \infty$, $\mathbf{M}_1$ becomes an all-zero matrix and the MMSE L-ZF PTEQ in (23) reduces to the simpler L-ZF PTEQ of (24).

IV. DMT EQUALIZATION OF AN IIR CHANNEL WITH $L_b > \nu$

If $L_b > \nu$, not all L_b TX difference terms in (12) vanish. There are ν difference terms $\Delta\mathbf{x}_k$ equal to zero because of the cyclic prefix: $\Delta x_{k,-l} = x_{k,-l} - x_{k,N-l} = 0$ with $l = 1, \ldots, \nu$. The remaining $L_b - \nu$ nonzero TX difference terms cause ISI/ICI. We denote the vector of nonzero TX difference terms as $\Delta\underline{\mathbf{x}}_k = \begin{bmatrix} \Delta x_{k,-L_b} & \cdots & \Delta x_{k,-\nu-1} \end{bmatrix}^T$. $\underline{\mathbf{B}}_\Delta$ denotes the corresponding $N \times (L_b - \nu)$ submatrix of $\mathbf{B}_\Delta$:

$$\underline{\mathbf{B}}_{\Delta} = \left[\begin{array}{ccc} b_{L_b} & \cdots & b_{\nu+1} \\ \vdots & \ddots & \vdots \\ 0 & 0 & b_{L_b} \\ \hline \multicolumn{3}{c}{\mathbf{0}_{(N-L_b+\nu)\times(L_b-\nu)}} \end{array}\right] \tag{26}$$

The data model (12) then reduces to:

$$\boxed{\tilde{\mathbf{z}}_{k,N} = \left[\begin{array}{c|c} \tilde{\mathbf{B}}_{a,\mathrm{D}} & \mathcal{F}_N\underline{\mathbf{B}}_{\Delta} \\ \hline \mathbf{0}_{N_p\times N_a} & \end{array}\right] \left[\begin{array}{c} \tilde{\mathbf{x}}_{k,a} \\ \Delta\underline{\mathbf{x}}_k \end{array}\right] + \mathcal{F}_N\check{\mathbf{n}}_k} \tag{27}$$

A. Optimum L-ZF block equalization

The data model (27) forms a set of N equations in $N_a + (L_b - \nu)$ unknowns $\left[\begin{array}{cc} \tilde{\mathbf{x}}_{k,a}^H & \Delta\underline{\mathbf{x}}_k^T \end{array}\right]^H$. A necessary condition for the existence of an L-ZF block equalizer for (27) is then given by $N \geq N_a + (L_b - \nu)$ or $N - N_a = \boxed{N_p \geq L_b - \nu}$. Hence, the insufficient *time-domain* TX redundancy in the form of a CP can be compensated for by *frequency-domain* TX redundancy in the form of pilot (and unused) tones. At this point, we should remind the reader that we assume baseband transmission, so each (pilot) tone goes together with a complex conjugate (pilot) tone, hence $\frac{N_p}{2}$ pairs of (complex conjugate) pilot tones suffice for L-ZF equalization. In [15], [16], [17], a similar L-ZF condition has been derived: for an *FIR channel* of order L, the authors conclude that $L - \nu$ unused tones are required, without considering the particular case of baseband transmission; pilot tones are claimed to result in a slightly more stringent condition (namely $L - \nu + 1$ pilot tones are required), while we have shown here that pilot tones and unused tones can be treated in the same way and give rise to the same L-ZF condition.

If $N_p \geq L_b - \nu$, there exist again one (if $N_p = L_b - \nu$) or multiple (if $N_p > L_b - \nu$) L-ZF block equalizers. The MMSE L-ZF block equalizer leads to the following block estimate [20], [21], [22]:

$$\left[\begin{array}{c} \hat{\tilde{\mathbf{x}}}_{k,a}^{\mathrm{L-ZF}} \\ \widehat{\Delta\underline{\mathbf{x}}}_k^{\mathrm{L-ZF}} \end{array}\right] = \underbrace{\left(\left[\begin{array}{c|c} \tilde{\mathbf{B}}_{a,\mathrm{D}}^H & \mathbf{0} \\ \hline \underline{\mathbf{B}}_{\Delta}^T\mathcal{F}_N^H & \end{array}\right] \left(\mathcal{F}_N\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2\mathcal{F}_N^H\right)^{-1} \left[\begin{array}{c|c} \tilde{\mathbf{B}}_{a,\mathrm{D}} & \mathcal{F}_N\underline{\mathbf{B}}_{\Delta} \\ \hline \mathbf{0} & \end{array}\right]\right)}_{\{1\}}{}^{-1} \times \underbrace{\left[\begin{array}{c|c} \tilde{\mathbf{B}}_{a,\mathrm{D}}^H & \mathbf{0} \\ \hline \underline{\mathbf{B}}_{\Delta}^T\mathcal{F}_N^H & \end{array}\right] \left(\mathcal{F}_N\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2\mathcal{F}_N^H\right)^{-1}\tilde{\mathbf{z}}_{k,N}}_{\{2\}} \tag{28}$$

The derivation in the appendix shows that (28) leads naturally to a two-step procedure, where $\Delta\underline{\mathbf{x}}_k$ is estimated in a first step and used in a second step to estimate $\tilde{\mathbf{x}}_{k,a}$.

1) The TX difference terms $\Delta\underline{\mathbf{x}}_k$ can be estimated first from:

$$\underline{\mathbf{B}}_{\Delta}\widehat{\Delta\underline{\mathbf{x}}}_k^{\mathrm{L-ZF}} = \left[\begin{array}{c} \mathbf{M}_4 \\ \hline \mathbf{0}_{(N-L_b+\nu)\times N_p} \end{array}\right] \tilde{\mathbf{z}}_{k,p} \tag{29}$$

where $\mathbf{M}_4$ is the following non-sparse $(L_b-\nu)\times N_p$ matrix:

$$\mathbf{M}_4 = \left[\begin{array}{c|c} \mathbf{I}_{L_b-\nu} & \mathbf{0} \end{array}\right] \mathbf{M}_3^{-1} \left[\begin{array}{c} \mathbf{I}_{L_b-\nu} \\ \hline \mathbf{0}^T \end{array}\right] \left[\begin{array}{c|c} \mathbf{I}_{L_b-\nu} & \mathbf{0} \end{array}\right] \mathbf{M}_3 \mathcal{F}_p^H \tag{30}$$

with $\mathbf{M}_3 = \left(\mathbf{\Sigma}_{\check{\mathbf{n}}}^2\right)^{-1} - \left(\mathbf{\Sigma}_{\check{\mathbf{n}}}^2\right)^{-1} \mathcal{F}_a^H \left(\mathcal{F}_a \left(\mathbf{\Sigma}_{\check{\mathbf{n}}}^2\right)^{-1} \mathcal{F}_a^H\right)^{-1} \mathcal{F}_a \left(\mathbf{\Sigma}_{\check{\mathbf{n}}}^2\right)^{-1}$ an $N\times N$ matrix (which can be decomposed into a concatenation of a noise whitening with $\left(\mathbf{\Sigma}_{\check{\mathbf{n}}}^2\right)^{-1/2}$, a projection on the orthogonal complement of $\left(\mathbf{\Sigma}_{\check{\mathbf{n}}}^2\right)^{-1/2} \mathcal{F}_a^H$ and another noise whitening) and with $\mathbf{0} = \mathbf{0}_{(L_b-\nu)\times(N-L_b+\nu)}$. The L-ZF estimate $\widehat{\Delta\underline{\mathbf{x}}_k}^{\mathrm{L-ZF}}$ is thus a linear combination of $\tilde{\mathbf{z}}_{k,p}$, which is in itself a linear combination of $\tilde{\mathbf{y}}_{k,p}$, $\tilde{\mathbf{x}}_{k,p}$ and $\Delta\mathbf{y}_k$ (see (13)). Information about the ISI/ICI, caused by the nonzero TX difference terms $\Delta\underline{\mathbf{x}}_k$, is extracted from the FD TX redundancy (pilot tones) and the RX difference terms.

2) The estimation of $\hat{\tilde{\mathbf{x}}}_{k,a}$ can now make use of $\widehat{\Delta\underline{\mathbf{x}}_k}^{\mathrm{L-ZF}}$ and is shown in the appendix to lead to the following **optimum L-ZF block estimate**:

$$\boxed{\hat{\tilde{\mathbf{x}}}_{k,a}^{\mathrm{L-ZF}} = \left[\begin{array}{c|c} \tilde{\mathbf{B}}_{a,\mathrm{D}}^{-1} & \underbrace{\tilde{\mathbf{B}}_{a,\mathrm{D}}^{-1} \left(\mathcal{F}_a \left(\mathbf{\Sigma}_{\check{\mathbf{n}}}^2\right)^{-1} \mathcal{F}_a^H\right)^{-1} \mathcal{F}_a \left(\mathbf{\Sigma}_{\check{\mathbf{n}}}^2\right)^{-1} \mathcal{F}_p^H}_{\mathbf{M}_1} \end{array}\right] \left(\tilde{\mathbf{z}}_{k,N} - \mathcal{F}_N \underline{\mathbf{B}}_\Delta \widehat{\Delta\underline{\mathbf{x}}_k}^{\mathrm{L-ZF}}\right)} \tag{31}$$

or

$$\boxed{\hat{\tilde{\mathbf{x}}}_{k,a}^{\mathrm{L-ZF}} = \tilde{\mathbf{B}}_{a,\mathrm{D}}^{-1} \tilde{\mathbf{z}}_{k,a} + \mathbf{M}_5 \tilde{\mathbf{z}}_{k,p}} \tag{32}$$

where $\mathbf{M}_5 = \mathbf{M}_1 + \left(\tilde{\mathbf{B}}_{a,\mathrm{D}}^{-1}\mathcal{F}_a + \mathbf{M}_1\mathcal{F}_p\right) \left[\begin{array}{c} \mathbf{M}_4 \\ \hline \mathbf{0} \end{array}\right]$ is a non-sparse $N_a\times N_p$ matrix. The **optimum L-ZF symbol estimate for tone** n, $\hat{\tilde{x}}_{k,n}^{\mathrm{L-ZF}}$, is then a linear combination of $\tilde{y}_{k,n}$, with the common inputs $\tilde{\mathbf{y}}_{k,p}$, $\tilde{\mathbf{x}}_{k,p}$ and $\Delta\mathbf{y}_k$:

$$\hat{\tilde{x}}_{k,n}^{\mathrm{L-ZF}} = \underbrace{\left[\begin{array}{c|c|c|c} \tilde{b}_n^{-1}\mathcal{F}_n\mathbf{A}_\Delta + \mathbf{m}_{n,5}\mathcal{F}_p\mathbf{A}_\Delta & \mathbf{m}_{n,5}\tilde{\mathbf{A}}_{p,\mathrm{D}} & -\mathbf{m}_{n,5}\tilde{\mathbf{B}}_{p,\mathrm{D}} & \tilde{b}_n^{-1}\tilde{a}_n \end{array}\right]}_{\left(\mathbf{v}_n^{\mathrm{L-ZF}}\right)^T} \left[\begin{array}{c} \Delta\mathbf{y}_k \\ \tilde{\mathbf{y}}_{k,p} \\ \tilde{\mathbf{x}}_{k,p} \\ \tilde{y}_{k,n} \end{array}\right] \tag{33}$$

where $\mathbf{m}_{n,5}$ is the row of $\mathbf{M}_5$ that corresponds to tone $n\in\mathcal{S}_a$.

It follows from the derivation above that the only **necessary and sufficient condition** for the existence of an L-ZF equalizer if $L_b>\nu$ is given by $N_p \geq L_b-\nu$; note that we assume again that $\tilde{b}_n \neq 0$ as no data can be transferred on tones that coincide with a channel zero.

B. Discussion and interpretation

The L-ZF block estimates $\hat{\tilde{\mathbf{x}}}_{k,a}^{\mathrm{L-ZF}}$ in (31-32) for $L_b>\nu$ and in (22) for $L_b\leq\nu$ look very similar, as they both reveal a linear combination of $\tilde{y}_{k,n}$, with the common inputs $\tilde{\mathbf{y}}_{k,p}$, $\tilde{\mathbf{x}}_{k,p}$ and $\Delta\mathbf{y}_k$. The

difference is in the compensation for the ISI/ICI from $\mathcal{F}_N \underline{\mathbf{B}}_\Delta \Delta\underline{\mathbf{x}}_k$ (see (31)) using the L-ZF estimate $\widehat{\Delta\underline{\mathbf{x}}_k}^{\mathrm{L-ZF}}$ in (29). As a consequence, the matrices $\mathbf{M}_1$ in (22) and $\mathbf{M}_5$ in (32) are considerably different. It has been shown in Section III-B that $\mathbf{M}_1$ tends to an all-zero matrix for $N \to \infty$, hence, for the case $L_b \leq \nu$, $\hat{\tilde{\mathbf{x}}}_{k,a}^{\mathrm{L-ZF}}$ only loosely depends on $\tilde{\mathbf{y}}_p$ and $\tilde{\mathbf{x}}_p$. For $N \to \infty$, $\mathbf{M}_5$ tends to $\tilde{\mathbf{B}}_{a,\mathrm{D}}^{-1} \mathcal{F}_a \left[\begin{array}{c} \mathbf{M}_4 \\ \hline \mathbf{0} \end{array}\right]$, which is non-zero as $\mathbf{M}_4$ does not tend to zero for $N \to \infty$. Still, *the optimum L-ZF equalizer for tone* n, $\mathbf{v}_n^{\mathrm{L-ZF}}$ *in (33), is a PTEQ*, as it decouples the equalization per active tone n: the complex-valued $(L_a + 1 + 2N_p)$-tap PTEQ $\mathbf{v}_n^{\mathrm{L-ZF}}$ linearly combines one tone-dependent input $\tilde{y}_{k,n}$ and the L_a common, real-valued inputs $\Delta\mathbf{y}_k$ with additional, common inputs $\tilde{\mathbf{y}}_{k,p}$ and $\tilde{\mathbf{x}}_{k,p}$, exploiting the FD TX redundancy (pilot tones).

It is important to note that these additional inputs come in *pairs*: each DFT output of a pilot tone, $\tilde{y}_{k,m}$ with $m \in \mathcal{S}_p$, appears together with the corresponding pilot symbol, $\tilde{x}_{k,m}$. This allows the PTEQ to (implicitly) remove the contribution of $\tilde{x}_{k,m}$ from $\tilde{y}_{k,m}$, so that only the ISI/ICI, introduced by the $L_b - \nu$ nonzero TX difference terms $\Delta\mathbf{x}_k$, and noise are left. The ISI/ICI information then allows to enhance the equalization of the active tones $\mathcal{S}_a$. Of course, the DFT output of an *unused* tone only has ISI/ICI and noise, as then $\tilde{x}_{k,m} = 0$. This result is a generalization of the results obtained in [15], [16], [17]: there, it has been shown that an L-ZF equalizer for tone n in case of an FIR channel of order L is obtained by linearly combining the n-th DFT output with $L - \nu$ DFT outputs of unused tones, and is, as such, an alternative for a TEQ-based receiver. Here, we show that both pilot and unused tones can be incorporated, as unused tones are a special case of pilot tones; moreover, we present the exploitation of FD TX redundancy as an extension of the capability of the L-ZF PTEQ (24) to eliminate ISI/ICI, rather than an equalization solution on its own.

V. Low-complexity extensions of the L-MMSE PTEQ

Based on the optimum L-ZF block equalizers derived in Section III and IV, we present in this section low-complexity extensions of the original L-MMSE PTEQ (15). In Section V-A, a linear PTEQ (L-PTEQ) extension is presented that exploits the FD TX redundancy as offered by pilot (and unused) tones. In Section V-B, a decision-feedback PTEQ (DF-PTEQ) extension is presented, which is based on the observation that the decision on a FD data symbol can be treated in the same way as the a priori knowledge of a pilot symbol. *A particular MMSE PTEQ design, based on a CP length* ν, $T - 1$ *RX difference terms* $\Delta\mathbf{y}_k$, N_p *pilot tones and* N_f *feedback tones then implicitly assumes an IIR channel model with* $L_b = \nu + N_p + N_f$ *and* $L_a = T - 1$. *Moreover, both the L-PTEQ and DF-PTEQ extension can be designed using a modified version of the RLS-based adaptive algorithm, presented in Section II-B.* Section V-C explains why the L-PTEQ and DF-PTEQ extensions increase the robustness of a DMT receiver to NBI.

A. The L-PTEQ extension

Based on the L-ZF PTEQs (23) and (33), we propose the following **L-PTEQ extension** for (15), which takes the DFT outputs on the pilot tones $\tilde{\mathbf{y}}_{k,p}$, as well as the pilot symbols $\tilde{\mathbf{x}}_{k,p}$, as additional inputs:

$$\boxed{\mathbf{v}_n^{\mathrm{L-PTEQ}} = \arg\min_{\mathbf{v}_n} \mathcal{E}\left\{\left|\begin{bmatrix} \Delta\mathbf{y}_k^T & \tilde{\mathbf{y}}_{k,p}^T & \tilde{\mathbf{x}}_{k,p}^T & \tilde{y}_{k,n} \end{bmatrix}\mathbf{v}_n - \tilde{x}_{k,n}\right|^2\right\}} \tag{34}$$

The RLS-based algorithm (16-17), described in Section II-B, can be extended in a straightforward way to solve (34) with the additional, common inputs $\tilde{\mathbf{y}}_{k,p}$ and $\tilde{\mathbf{x}}_{k,p}$ in the same way as the RX difference terms $\Delta\mathbf{y}_k$: the common tone-independent part $\mathbf{S}_\Delta$ in the inverse transpose of the Cholesky factor of $\mathbf{R}_n$ in (17) should be replaced with a tone-independent matrix $\mathbf{S}_{\Delta,p}$ such that:

$$\left(\mathcal{E}\left\{\begin{bmatrix} \Delta\mathbf{y}_k \\ \tilde{\mathbf{y}}_{k,p}^* \\ \tilde{\mathbf{x}}_{k,p}^* \end{bmatrix}\begin{bmatrix} \Delta\mathbf{y}_k^T & \tilde{\mathbf{y}}_{k,p}^T & \tilde{\mathbf{x}}_{k,p}^T \end{bmatrix}\right\}\right)^{-1} = \mathbf{S}_{\Delta,p}^H\mathbf{S}_{\Delta,p} \tag{35}$$

Whereas we have assumed until now that all available FD TX redundancy are exploited, it is in fact only useful to exploit pilot and unused tones that have a sufficiently high ISI/ICI-to-noise ratio (INR). As will be argued below, it then typically holds in an ADSL scenario that the actual number of exploited pilot (and unused) tones obeys $N_p < N - N_a$. To determine the optimal choice of pilot and unused tones is far from trivial. However, we give below a few rules of thumb regarding which TD and FD TX redundancy to exploit.

- In ADSL, the CP length is standardized to be 32 samples. However, the number of nonzero TX difference terms $\Delta\underline{\mathbf{x}}_k$, $L_b - \nu$, hence the CP length, affects the ISI/ICI level and the INR: a shorter CP will result in a higher ISI/ICI level, hence a higher INR. It turns out (see Section VI) that it can be beneficial to increase the ISI/ICI level by reducing the CP length, e.g., to 16 samples, as this renders the exploitation of FD TX redundancy more effective. In fact, by reducing the CP length and using pilot and unused tones instead, *TD TX redundancy (CP overhead) is exchanged for FD TX redundancy (pilot and unused tones)*.
- One could introduce extra FD TX redundancy by turning off one or more equally spreaded tones to enhance the ISI/ICI estimation and suppression, be it that this results in capacity loss that is not necessarily recovered by the enhanced equalization. It is therefore often preferred to make use of already available pilot and unused tones. E.g., in frequency-division duplexing (FDD) ADSL transmission, the available tones are assigned either to upstream or downstream with some unused guard tones (e.g., $n = 33, \ldots, 37$) in between. Hence, in downstream, the lower tones, up to tone 37, are unused. However, most of these tones are contaminated, e.g., by the echo of the upstream signal, and only a few of them are useful. Other candidates are the tones that are unused because of an insufficient SNR caused by a channel zero or the upper downstream tones in case of a long channel, as long as the INR on these tones is sufficiently high.

- Through extensive simulations, we have observed (see Section VI) that there is no need to incorporate all available pilot and unused tones: only a few, well-chosen pilot and unused tones suffice to attain good performance at a low complexity. Also, performance improves with increasing interspacing between the pilot/unused tones (ideally, the pilot/unused tones should be equidistant): the ISI/ICI on tones that lie close together is correlated, hence it is better to choose available tones that are farther apart and less coherent wherever possible. This is confirmed by the simulations where the introduction of two tones that are far apart (tones 37 and 64) result in the largest performance improvement.
- It has been observed in [24] that, in an ADSL context, the front-end filtering (e.g., to separate upstream and downstream transmission), rather than the transmission channel itself, is oftentimes the main cause of ISI/ICI; ISI/ICI then primarily shows up near the transition band of the front-end filters. The front-end filters add extra degrees to the numerator and the denominator of the overall IIR channel representation, and therefore make equalization generally more difficult. In the simulations of Section VI, we will take this into account when choosing $\mathcal{S}_p$ and $\mathcal{S}_u$.

As we deal with baseband transmission, both $\tilde{\mathbf{y}}_{k,p}$ and $\tilde{\mathbf{x}}_{k,p}$ have complex conjugate entries (e.g., $\tilde{y}_{k,m}$ and $\tilde{y}^*_{k,m}$ with $m \in \mathcal{S}_p$), which are linearly combined using complex-valued coefficients (e.g., $\mathbf{v}_n[i]\tilde{y}_{k,m} + \mathbf{v}_n[j]\tilde{y}^*_{k,m}$). One can save on computations (two multiplications per complex conjugate pair of inputs) by linearly combining the real and imaginary part instead (i.e., $(\mathbf{v}_n[i] + \mathbf{v}_n[j])\,\Re\{\tilde{y}_{k,m}\} + j\,(\mathbf{v}_n[i] - \mathbf{v}_n[j])\,\Im\{\tilde{y}_{k,m}\}$). In the sequel, we keep the notation based on the complex-valued variables $\tilde{\mathbf{y}}_{k,p}$ and $\tilde{\mathbf{x}}_{k,p}$ for the sake of conciseness.

In baseband transmission, only half of the N_a active tones need to be demodulated (the other tones are complex conjugate). The overall memory cost then amounts to $\frac{N_a}{2}(L_a + 1 + 2N_p)$ complex-valued equalizer coefficients $\mathbf{v}_n$, while the computational cost amounts to $N_a(L_a + 2 + 2N_p)$ real multiplications per DMT symbol. If N_u unused tones are incorporated, the overall memory cost amounts to $\frac{N_a}{2}(L_a + 1 + 2N_p + N_u)$ complex-valued coefficients and the computational cost amounts to $N_a(L_a + 2 + 2N_p + N_u)$ real multiplications per DMT symbol.

B. The DF-PTEQ extension

As argued in the previous section, the number of useful pilot and unused tones is often limited, e.g., because of a too low ISI/ICI level. On the other hand, there could be active data-carrying tones that have a high ISI/ICI level. A symbol decision on such an active tone can then be treated in the same way as the a priori knowledge of a pilot symbol: the ISI/ICI in the DFT output of the considered active tone can again be exploited to enhance the equalization of the remaining active tones. The idea is then to feed back the symbol decisions of a few well-chosen tones within the DMT symbol and include the corresponding DFT outputs of these tones to use them in the very same way as pilot tones and symbols.

Similar rules of thumb as already given for pilot and unused tones apply for the choice of feedback tones. Denoting the DFT outputs and symbol decisions on the set of N_f (complex conjugate) feedback tones $\mathcal{S}_f$ by $\tilde{\mathbf{y}}_{k,f}$ and $\hat{\tilde{\mathbf{x}}}_{k,f}$, respectively, a **decision-feedback (DF) PTEQ extension**[3] is then based on the following MMSE design criterion:

$$\boxed{\mathbf{v}_n^{\mathrm{DF-PTEQ}} = \arg\min_{\mathbf{v}_n} \mathcal{E}\left\{\left|\left[\begin{array}{cccc} \Delta\mathbf{y}_k^T & \tilde{\mathbf{y}}_{k,f}^T & \hat{\tilde{\mathbf{x}}}_{k,f}^T & \tilde{y}_{k,n} \end{array}\right]\mathbf{v}_n - \tilde{x}_{k,n}\right|^2\right\}} \tag{36}$$

In order to avoid the need for an iterative solution method, (36) is only applied to the tones $n \in \mathcal{S}_a \setminus \mathcal{S}_f$. The equalization of the tones n in $\mathcal{S}_f$ is done serially, based on:

$$\min_{\mathbf{v}_n} \mathcal{E}\left\{\left|\left[\begin{array}{cccc} \Delta\mathbf{y}_k^T & \tilde{\mathbf{y}}_{k,0:n_f-1,f}^T & \hat{\tilde{\mathbf{x}}}_{k,0:n_f-1,f}^T & \tilde{y}_{k,n} \end{array}\right]\mathbf{v}_n - \tilde{x}_{k,n}\right|^2\right\} \text{ with } 0 \leq n_f \leq \frac{N_f}{2} - 1 \tag{37}$$

where $\tilde{\mathbf{y}}_{k,0:n_f-1,f} = \left[\begin{array}{ccc} \tilde{\mathbf{y}}_{k,f}[0] & \cdots & \tilde{\mathbf{y}}_{k,f}[n_f-1] \end{array}\right]^T$ and $\hat{\tilde{\mathbf{x}}}_{k,0:n_f-1,f} = \left[\begin{array}{ccc} \hat{\tilde{\mathbf{x}}}_{k,f}[0] & \cdots & \hat{\tilde{\mathbf{x}}}_{k,f}[n_f-1] \end{array}\right]^T$, hence the PTEQ of the n_f-th feedback tone in $\mathcal{S}_f$ only exploits the ISI/ICI on the feedback tones $0, \ldots, n_f - 1$ in $\mathcal{S}_f$[4]. This means indeed that the equalization of the N_f feedback tones $\mathcal{S}_f$ does not fully benefit from the exploited ISI/ICI information. However, it is found in Section VI that this has no detrimental effect on the performance, as only a small number of feedback tones, N_f, is required. Ignoring the fact that the equalizers for the feedback tones do not use all fed-back symbol decisions, the overall memory cost is equal to $\frac{N_a}{2}(L_a + 1 + 2N_f)$ complex-valued equalizer coefficients, while the computational cost amounts to $N_a(L_a + 2 + 2N_f)$ real multiplications per DMT symbol.

In [18], [19], DFE structures for OFDM transmission with an insufficiently long CP have already been presented. In [18], an OFDM system without cyclic prefix is considered; the ISI from the previous ODFM symbol is first removed in a DF fashion, followed by a linear equalization of the ISI-free OFDM symbol to remove ICI. In [19], a ZF-DFE and an MMSE-DFE are presented: in both DFEs, all decisions on the previous and current OFDM symbol are fed back; in addition, the MMSE-DFE uses three consecutive RX DMT symbols (the current, the previous and the next symbol) in the forward path. All tones are equalized in a joint, block-wise fashion. While acceptable for OFDM systems, such as HIPERLAN/2, which employs a DFT/IDFT size of $N = 64$ and a CP length of $\nu = 16$, these DFE structures are too complex for implementation in DMT-based systems, such as ADSL, which employs a DFT/IDFT size of $N = 512$ and a CP length of $\nu = 32$. The low-complexity DF-PTEQ extension (36) developed here only feeds back a few ($N_f \ll N$) symbol decisions *within* the current DMT symbol and uses the original PTEQ in its forward path. Its complexity depends linearly, rather than quadratically on N. Hence, the DF-PTEQ extension is a drastically simplified and truncated version of the MMSE-DFE of [19].

[3]Of course, the L-PTEQ and DF-PTEQ extensions can be combined.

[4]For simplicity, we ignore the complex conjugate tones in the reasoning and the notation $\tilde{\mathbf{y}}_{k,0:n_f-1,f}$ and $\hat{\tilde{\mathbf{x}}}_{k,0:n_f-1,f}$.

Again, (36-37) can be solved adaptively with an extended version of the RLS-based algorithm (16-17). By an appropriate choice of the ordering of the DF-PTEQ inputs, e.g.,

$$\left[\begin{array}{c|c|c} \Delta\mathbf{y}_k^T & \left(\left[\begin{array}{cc} 1 & 0 \end{array} \right] \otimes \tilde{\mathbf{y}}_{k,f}^T + \left[\begin{array}{cc} 0 & 1 \end{array} \right] \otimes \hat{\tilde{\mathbf{x}}}_{k,f}^T \right) & \tilde{y}_{k,n} \end{array} \right] \tag{38}$$

where $\otimes$ denotes a Kronecker product, which ensures that the DFT outputs $\tilde{\mathbf{y}}_{k,f}^T$ and the symbol decisions $\hat{\tilde{\mathbf{x}}}_{k,f}^T$ are interlaced, the inverse transpose $\mathbf{S}_{\Delta,f}$ of the Cholesky factor of the $(L_a + 2N_f) \times (L_a+2N_f)$ autocorrelation matrix of $\left[\begin{array}{c|c|c} \Delta\mathbf{y}_k^T & \left(\left[\begin{array}{cc} 1 & 0 \end{array} \right] \otimes \tilde{\mathbf{y}}_{k,f}^T + \left[\begin{array}{cc} 0 & 1 \end{array} \right] \otimes \hat{\tilde{\mathbf{x}}}_{k,f}^T \right) & \end{array} \right]$ is again common for all tones. For the feedback tone n_f, only the $(L_a + 2n_f)$ first rows of $\mathbf{S}_{\Delta,f}$, corresponding to $\left[\begin{array}{c|c|c} \Delta\mathbf{y}_k^T & \left(\left[\begin{array}{cc} 1 & 0 \end{array} \right] \otimes \tilde{\mathbf{y}}_{k,0:n_f-1,f}^T + \left[\begin{array}{cc} 0 & 1 \end{array} \right] \otimes \hat{\tilde{\mathbf{x}}}_{k,0:n_f-1,f}^T \right) & \end{array} \right]$, are needed. Each tone has its own tone-dependent last row of $\mathbf{S}_n$ with length $L_a + 2n_f$ for the feedback tone n_f and length $L_a + 2N_f$ for the other tones.

It is clear from the above discussion that the L-PTEQ and DF-PTEQ extensions add flexibility when designing DMT-based systems. CP overhead and FD TX redundancy can be traded off against each other. *A particular PTEQ design, based on a CP length ν, $T-1$ RX difference terms $\Delta\mathbf{y}_k$ and N_p+N_f pilot/feedback tones then implicitly assumes an IIR channel model with $L_b = \nu+N_p+N_f$ and $L_a = T-1$.* It will be shown in Section VI that a good choice of these design parameters allows to achieve the same performance as with the original PTEQ at a lower computational and memory cost or, alternatively, a better performance is achieved without considerable cost increase.

Note that the DF-PTEQ does not suffer from error propagation, as the feedback symbols are only used within a single DMT symbol. The impact of decision errors can be further reduced (without significant impact on performance) by increasing the margin on the feedback tones. Because of the small number of involved feedback tones, the impact of the ordering of the feedback tones on the performance is negligible.

C. NBI suppression capability

In [10], it has been shown that a PTEQ-based DMT receiver with a sufficient number of taps, i.e., incorporating a sufficient number of RX difference terms, has an increased robustness to NBI, such as radio-frequency interference (RFI), when compared to a TEQ-based receiver, even if the latter includes a receiver window [25]. The PTEQ implicitly includes a receiver window that is optimized per-tone, so that, for a tone that is affected by ISI/ICI as well as NBI, equalization and NBI suppression are traded off against each other.

In [26], an L-MMSE based NBI canceller has been presented that estimates and eliminates the spectral leakage of an NBI on individual tones by measuring the NBI on a few unused (or used) measurement tones (and their complex conjugates) around the NBI center frequency. It turns out that a well performing L-MMSE NBI canceller is obtained, based on a rough a priori knowledge of the power-spectral density

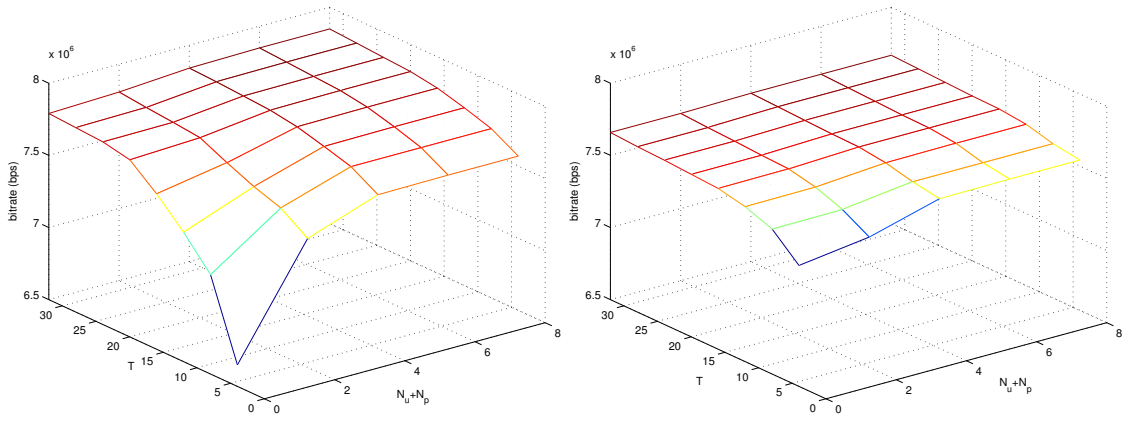
(PSD) of the NBI (i.e., an estimate of center frequency and bandwidth, and assuming a flat spectrum). Based on the time-bandwidth product of the NBI signal over one DMT symbol, a rule of thumb for the required number of measurement tones is equal to the NBI bandwidth (expressed as a multiple of the DFT tone spacing) plus one to three. The center frequency and bandwidth of the NBI can be estimated using the squared magnitude of the DFT outputs, as in a periodogram, searching for the tones with the strongest interference [26].

Here, we suggest to combine this NBI canceller with the PTEQ and design the equalizer and the NBI canceller in a joint and direct fashion. In fact, if the L-PTEQ extension (34) makes use of a sufficient number of DFT outputs of unused tones, $\tilde{\mathbf{y}}_{k,u}$, around the NBI center frequency, it already incorporates the proposed NBI canceller. Alternatively, one can make use of the DF-PTEQ extension (36) with feedback measurement tones that are centered around the NBI center frequency. If the NBI is present during connection set-up, the equalizer and NBI canceller can be designed jointly by adaptively solving (34) or (36) using the ADSL medley training signal. It will be shown in Section VI that both kinds of NBI suppresssion, namely based on adding RX difference terms [10] and exploiting unused tones around the NBI center frequency [26] can again be traded off against each other to optimize computational and memory cost versus performance.

It should be noted that the joint PTEQ/NBI canceller coefficients no longer have the frequency invariance property as reported in [26]. A change of the center frequency of the NBI then requires a re-training of the receiver coefficients. Hence, for fast changing HAM disturbers, an exclusive NBI canceller is preferred.

VI. Simulations

In this section, we provide extensive ADSL simulation results for the L-PTEQ and DF-PTEQ extension in different scenarios (with and without NBI) and with varying amounts of exploited FD TX redundancy (pilot tones $\mathcal{S}_p$, unused tones $\mathcal{S}_u$ and feedback tones $\mathcal{S}_f$), CP lengths ν and numbers of RX difference terms $T-1$. All L-PTEQ and DF-PTEQ designs have been obtained using the described *RLS-based adaptive algorithm*. We include plots for the FDD ADSL downstream CSA #1-8 loops [1], [2] with active tones 38 to 255 (and their complex conjugates). The loops are described by a channel impulse response of length N where $N = 512$ is the DFT size. The synchronization delay Δ is determined by the first sample index of the channel impulse response window of $\nu+1$ samples with maximum energy. The noise is a superposition of AWG noise at -140dBm/Hz, residual echo of the upstream signal and near-end crosstalk from 24 ADSL disturbers. Strong front-end filters to separate up- and downstream transmission are included in the channel; as already pointed out, they are known to be the major source of ISI/ICI [24]. Therefore, the unused tones, pilot tones and feedback tones are chosen within a tone range around the transition band of the front-end filters, i.e., around tone 38. Tone 33 to 37 are used as guard tones



(a) $\nu = 16$ (b) $\nu = 32$

Fig. 3. Bitrate performance of the L-PTEQ extension for the CSA #4 loop as a function of T and $N_u + N_p$ for $\nu = 16$ and standardized $\nu = 32$.

between upstream and downstream. In ADSL, tone 64 is used as a pilot tone for synchronization and can also be exploited for equalization.

Figure 3 shows the bitrate performance of the L-PTEQ extension for the CSA #4 loop as a function of T and $N_u + N_p$ for $\nu = 16$ (Figure 3a) and standardized $\nu = 32$ (Figure 3b). The values $T = \{4, 8, 12, 16, 20, 24, 28, 32\}$ are considered. FD TX redundancy is added in complex conjugate pairs of tones, hence N_u and N_p are even. For conciseness, we exclude the complex conjugate tones in the sequel, when referring to the used FD TX redundancy. We order the FD TX redundancy as $\{64, 37, 36, 35\}$, where tone 64 is a pilot tone and the other tones are unused tones. The effect of using FD TX redundancy is larger for a shorter CP length ν: the ISI/ICI level is higher when $L_b - \nu$ is larger, hence the exploitation of ISI/ICI information on pilot and unused tones is more effective. The bitrate saturates when tones 64 and 37 ($N_u + N_p = 4$) are included; extra unused tones 36 and 35 do not result in further performance improvement, presumably because they are next to tone 37, with which they have strongly correlated ISI/ICI; the residual echo signal also contaminates the guard tones 33 to 37. Note that the equalizer scheme, proposed in [17], would correspond to a curve for $T = 0$ and varying N_u, which is not depicted here as it gives considerably worse performance than with $T \geq 4$: this is probably due to the more difficult simulation scenarios adopted here, which include non-white crosstalker noise, residual echo, front-end filters, etc.

Based on Figure 3, we choose the following four L-PTEQ configurations to compare the performance on the CSA #1-8 loops in **Figure 4**: (1) $\nu = 32$, $T = 32$ and $N_u = N_p = 0$, (2) $\nu = 16$, $T = 8$, $N_p = 2$ and $N_u = 2$, (3) $\nu = 16$, $T = 16$, $N_p = 2$ and $N_u = 0$, (4) $\nu = 16$, $T = 32$, $N_p = 2$ and

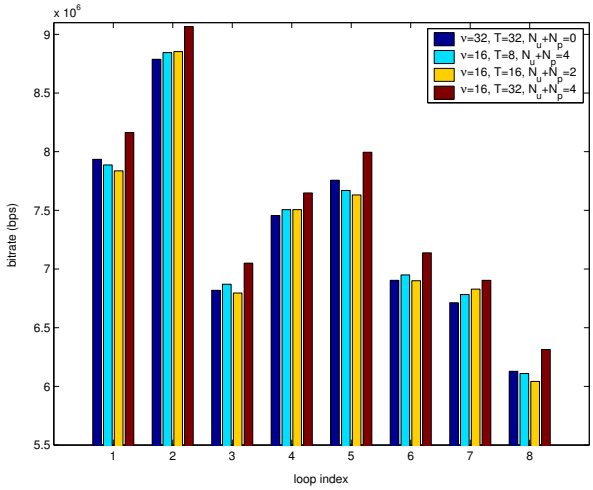


Fig. 4. Bitrate performance of the L-PTEQ extension on the CSA #1-8 loops for 4 scenarios.

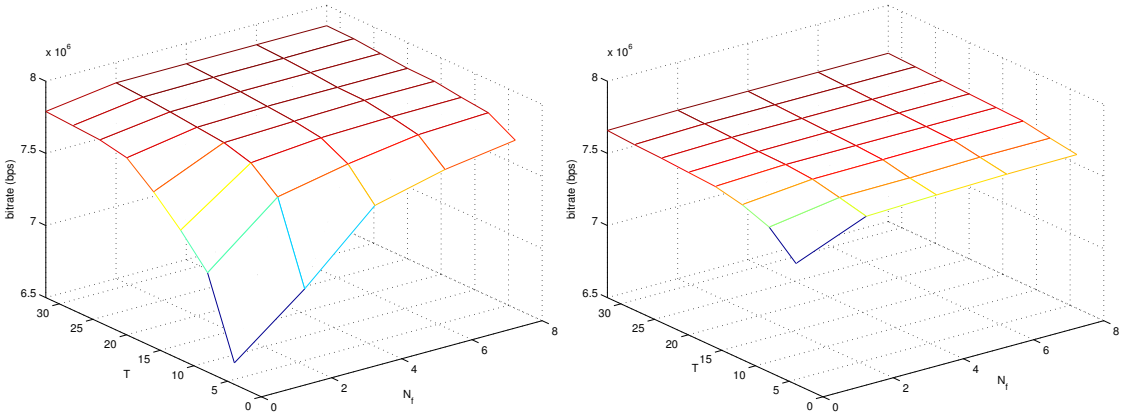
	L-PTEQ extension		DF-PTEQ extension	
	memory cost	computational cost	memory cost	computational cost
$\nu = 32,\ T = 32,\ N_u + N_p = N_f = 0$	6976	14388	6976	14388
$\nu = 16,\ T = 8,\ N_u + N_p = N_f = 4$	3052	6540	3476	7388
$\nu = 16,\ T = 16,\ N_u + N_p = N_f = 2$	4360	9156	4356	9148
$\nu = 16,\ T = 32,\ N_u + N_p = N_f = 4$	8284	17004	8708	17852

TABLE I

COMPUTATIONAL COST (IN # REAL MULTIPLICATIONS PER DMT SYMBOL) AND MEMORY COST (IN # COMPLEX COEFFICIENTS) OF THE SCENARIOS CONSIDERED IN FIGURE 4 (L-PTEQ EXTENSION) AND FIGURE 6 (DF-PTEQ EXTENSION).

$N_u = 2$. The computational and memory cost of the 4 configurations (with $N_a = 218$ active tones) is summarized in **Table I**. For all loops, the first three configurations give rise to similar performance, while configurations 2 and 3 (with $\nu = 16$) are 45 to 55 % less complex than configuration 1 (with $\nu = 32$): hence, by choosing an appropriate L-PTEQ extension, equalization complexity can be reduced, even if there is more ISI/ICI because of a shorter CP. The 4th configuration ($\nu = 16$) is slightly more complex (less than 20 %) than the 1st configuration, but reaches a bitrate that is consistently around 200 to 300 kbps higher, which is due to a better bandwidth efficiency because of the reduced CP.

Figure 5 shows the bitrate performance of the DF-PTEQ extension for the CSA #4 loop as a function of T and N_f ($N_u + N_p = 0$) for $\nu = 16$ (Figure 5a) and standardized $\nu = 32$ (Figure 5b). The values $T = \{4, 8, 12, 16, 20, 24, 28, 32\}$ are considered. The N_f feedback tones are taken from $\{38, 44, 50, 56\}$,



(a) $\nu = 16$ (b) $\nu = 32$

Fig. 5. Bitrate performance of the DF-PTEQ extension for the CSA #4 loop as a function of T and N_f for $\nu = 16$ and standardized $\nu = 32$.

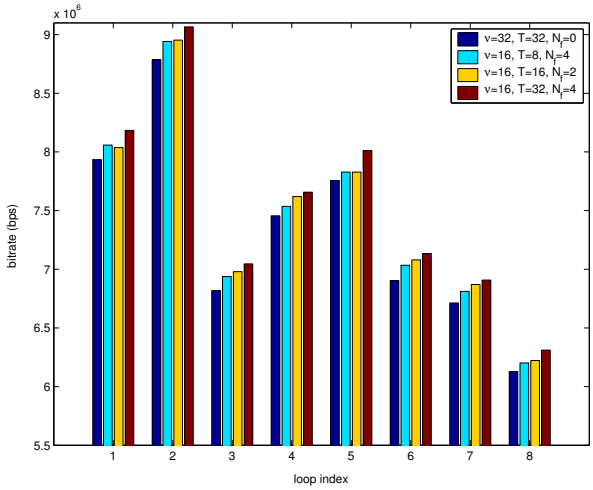


Fig. 6. Bitrate performance of the DF-PTEQ extension for the CSA #1-8 loops for 4 scenarios.

i.e., near the front-end filter transition band and sufficiently interspaced. As with pilot and unused tones, the impact of using feedback tones is larger for shorter CP length. For small T, including up to 4 feedback tones is beneficial, while for $T \geq 12$, 2 feedback tones suffice. Similar to Figure 4, **Figure 6** compares the performance on the CSA #1-8 loops for 4 DF-PTEQ configurations: (1) $\nu = 32$, $T = 32$ and $N_f = 0$, (2) $\nu = 16$, $T = 8$ and $N_f = 4$, (3) $\nu = 16$, $T = 16$ and $N_f = 2$, (4) $\nu = 16$, $T = 32$ and $N_f = 4$. The computational and memory cost is also given in **Table I**. The less complex 2nd and 3rd configuration ($\nu = 16$) consistently outperform the 1st configuration ($\nu = 32$). The 4th configuration ($\nu = 32$) reaches again a bitrate that is up to 300 kbps higher than with the 1st configuration. **Figure 7**

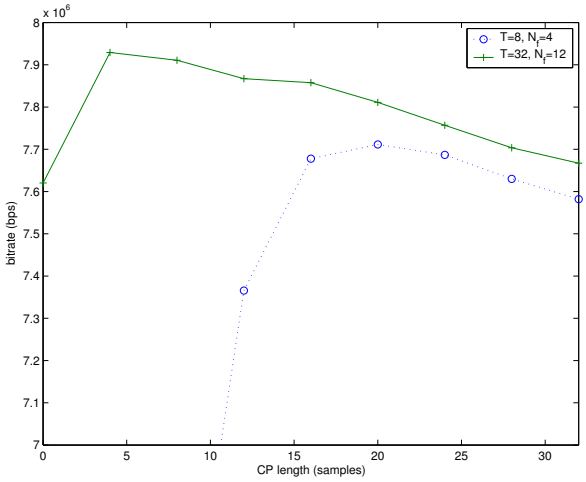
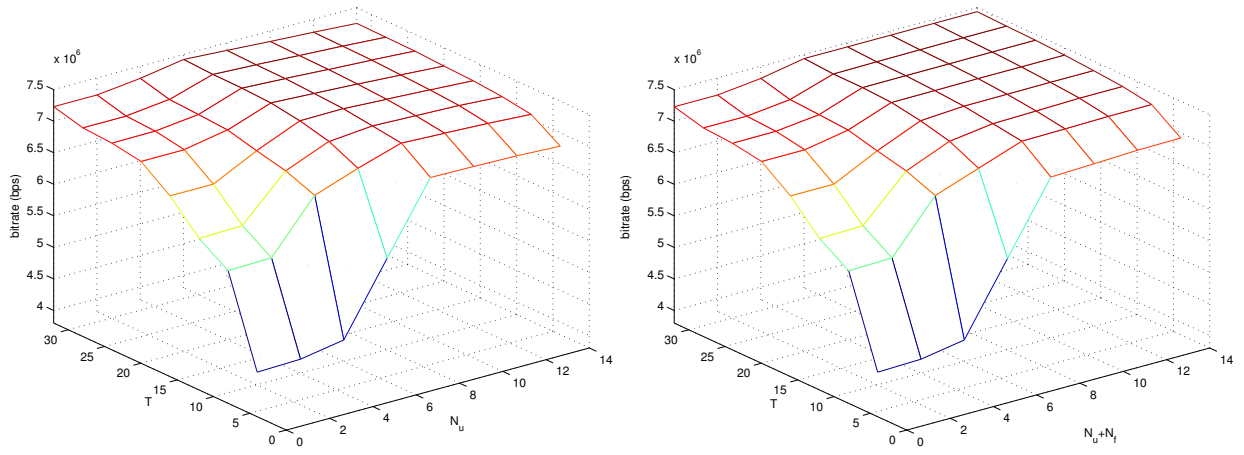


Fig. 7. Bitrate as a function of the CP length ν for two DF-PTEQ configurations with the CSA #4 loop.

shows the achieved bitrate as a function of the CP length ν for two DF-PTEQ configurations: (1) a low complexity configuration with $T = 8$ and $N_f = 4$ and (2) a high-complexity configuration with $T = 32$ and $N_f = 12$, which is expected to perform close to the upper bound that is achievable with the DF-PTEQ extension. The curves suggest that a good performance versus complexity trade-off is obtained with a value for ν between 16 and 20.

To illustrate the NBI suppression capability of the L-PTEQ and the DF-PTEQ extension, we include **Figure 8**, which shows the bitrate as a function of T and N_u or N_f on the CSA #4 loop ($\nu = 32$) with an NBI with a flat PSD of -40dBm/Hz, a center frequency at 430 kHz (i.e., between tone 100 and 101) and a bandwidth of 4.3 kHz (i.e., equal to the DFT tone spacing). The considered measurement tones are ordered as follows: $\{101, 99, 103, 97, 105, 95, 107\}$. The curve in **Figure 8a** for $N_u = 0$ illustrates that the robustness of the original PTEQ increases with T, as observed in [10]. In accordance with the proposed rule of thumb, 3 to 4 measurement tones (and their complex conjugates) result in a further performance improvement. Adding 6 measurement tones increases the performance for $T = 32$ with 200 kbps; the configuration with $T = 12$ and $N_u = 8$ measurement tones gives about the same performance as the configuration with $T = 32$ without measurement tones, be it at a 40% lower complexity. Adding 8 or more instead of 6 measurement tones reduces the performance as tones with a sufficient SNR to carry data are turned off. **Figure 8b** shows the performance when feedback tones are used as measurement tones. The curves in Figure 8a and Figure 8b coincide up to 6 measurement tones: the tones $\{101, 99, 103\}$ have a too low SNR after NBI suppression, so that they can not carry data, hence there is no need to use decision-feedback. For 8 or more measurement tones and $T = 32$, the bitrate does not decrease, as the extra measurement tones do carry data now. However, using 8 or more measurement tones appears not to lead to better performance. **Figure 9** shows the impact of different L-PTEQ extensions on the NBI



(a) L-PTEQ extension (b) DF-PTEQ extension

Fig. 8. Bitrate performance of the L-PTEQ and DF-PTEQ extension on the CSA #4 loop ($\nu = 32$) with one NBI as a function of T and N_u or N_f.

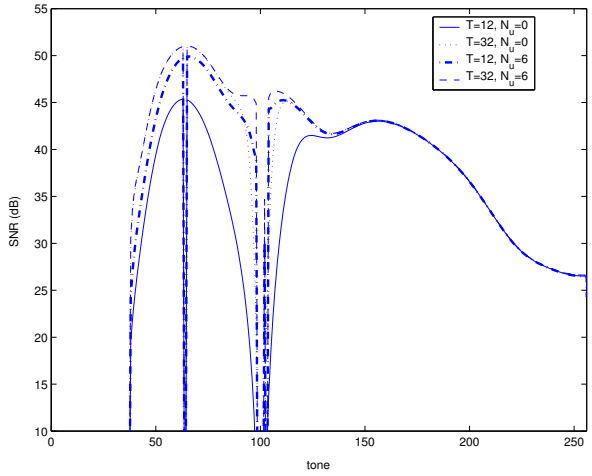


Fig. 9. SNR per tone on the CSA #4 loop with one NBI for several L-PTEQ configurations.

suppression capability ($N_p = 0$): the effect of a larger number of RX difference terms and the impact of including measurement tones are depicted. Both NBI suppression techniques ((1) increasing T from 12 to 32 with $N_u = 0$ and (2) including $N_u = 6$ unused measurement tones with $T = 12$) have a beneficial impact on all affected tones. In addition, the measurement tones appear to locally increase robustness in the neighbourhood of the NBI (compare $T = 32,\ N_u = 0$ with $T = 32,\ N_u = 6$), while a larger T results in a better ISI/ICI suppression at the lower tones (compare $T = 12,\ N_u = 6$ with $T = 32,\ N_u = 6$).

VII. Conclusions

Until now, most DMT receivers assumed only implicitly an infinite impulse response (IIR) model for the ADSL channel: the PTEQ-based receiver was originally presented as an alternative for TEQ-based receivers over FIR channels. In this paper, we have reconsidered DMT equalization under the explicit assumption that the wireline channel impulse response is well approximated by an IIR model. Based on the corresponding block data model and the optimum L-ZF block equalizers, we have shown that a PTEQ-based receiver is also suited for the equalization of an IIR channel. A low-complexity L-PTEQ extension has been developed to accommodate arbitrary IIR model orders and CP lengths: frequency-domain transmit redundancy as introduced by pilot and unused tones can be used to enhance the original PTEQ; alternatively, by treating symbol decisions in the same way as pilot symbols, a DF-PTEQ extension is obtained. These PTEQ extensions also improve the NBI suppression capability of the PTEQ. In the simulations, it has been shown that the PTEQ extensions add flexibility when designing DMT-based systems: they allow to trade off CP overhead and frequency-domain transmit redundancy against each other so that a similar bitrate as with the original PTEQ can be achieved at a lower memory and computational cost or, alternatively, a higher bitrate is achieved without a considerable cost increase.

References

[1] N. Al-Dhahir and J. M. Cioffi, "Efficiently computed reduced-parameter input-aided MMSE equalizers for ML detection: A unified approach," *IEEE Trans. Information Theory*, vol. 42, no. 3, pp. 903–915, May 1996.

[2] G. Arslan, B. L. Evans, and S. Kiaei, "Equalization for discrete multitone transceivers to maximize bit rate," *IEEE Trans. Signal Processing*, vol. 49, no. 12, pp. 3123–3135, Dec. 2001.

[3] P. J. W. Melsa, R. C. Younce, and C. E. Rohrs, "Impulse response shortening for discrete multitone transceivers," *IEEE Trans. Communications*, vol. 44, no. 12, pp. 1662–1672, Dec. 1996.

[4] M. Milosevic, L. F. C. Pessoa, B. L. Evans, and R. Baldick, "Optimal time domain equalization design for maximizing data rate of discrete multi-tone systems," *Submitted to IEEE Trans. Signal Processing*, 2002.

[5] R. K. Martin, J. Balakrishnan, W. A. Sethares, and C. R. Johnson, Jr., "A blind, adaptive TEQ for multicarrier systems," *IEEE Signal Processing Letters*, vol. 9, no. 11, pp. 341–343, Nov. 2002.

[6] K. Vanbleu, G. Ysebaert, G. Cuypers, M. Moonen, and K. Van Acker, "Bitrate maximizing time-domain equalizer design for DMT-based systems," *IEEE Trans. Communications*, vol. 52, no. 6, pp. 871–876, June 2004.

[7] K. Van Acker, G. Leus, M. Moonen, O. van de Wiel, and T. Pollet, "Per-tone equalization for DMT-based systems," *IEEE Trans. Communications*, vol. 49, no. 1, pp. 109–119, Jan. 2001.

[8] K. Van Acker, G. Leus, M. Moonen, and T. Pollet, "RLS-based initialization for per-tone equalizers in DMT receivers," *IEEE Trans. Communications*, vol. 51, no. 6, pp. 885–889, June 2003.

[9] G. Ysebaert, K. Vanbleu, G. Cuypers, M. Moonen, and T. Pollet, "Combined RLS-LMS initialization for per tone equalizers in DMT-receivers," *IEEE Trans. Signal Processing*, vol. 51, no. 7, pp. 1916–1927, July 2003.

[10] K. Van Acker, T. Pollet, G. Leus, and M. Moonen, "Combination of per tone equalization and windowing in DMT-receivers," *Signal Processing*, vol. 81, no. 8, pp. 1571–1579, Aug. 2001.

[11] K. Vanbleu, G. Leus, and M. Moonen, "Per-tone equalization for DMT-based transmission over IIR channels," Nov. 2001, pp. 405–409.

[12] J. S. Chow, J. C. Tu, and J. M. Cioffi, "A discrete multitone transceiver system for HDSL applications," vol. 9, no. 6, pp. 895–907, Aug. 1991.

[13] A. Scaglione, S. Barbarossa, and G. B. Giannakis, "Filterbank transceivers optimizing information rate in block transmissions over dispersive channels," *IEEE Trans. Information Theory*, vol. 45, no. 3, pp. 1019–1032, Apr. 1999.

[14] R. Nilsson, F. Sjöberg, and J. P. LeBlanc, "A narrow-band interference canceller for OFDM-based systems," in *Proc. of 4th European Personal Mobile Communications Conference (EPMCC)*, Feb. 2001.

[15] S. Trautmann and N. J. Fliege, "A new equalizer for multitone sysems without guard time," vol. 6, no. 1, pp. 34–36, Jan. 2002.

[16] ——, "Perfect equalization for DMT systems without guard interval," vol. 20, no. 5, pp. 987–996, June 2002.

[17] S. Trautmann, T. Karp, and N. J. Fliege, "Frequency domain equalization of DMT/OFDM systems with insufficient guard interval," in *Proc. IEEE Int. Conf. Communications (ICC)*, Apr. 2002, pp. 1646–1650.

[18] Y. Sun, "Bandwidth-efficient wireless OFDM," vol. 19, no. 11, pp. 2267–2278, Nov. 2001.

[19] A. Stamoulis, G. B. Giannakis, and A. Scaglione, "Block FIR decision-feedback equalizers for filterbank precoded transmissions with blind channel estimation capabilities," *IEEE Trans. Communications*, vol. 49, no. 1, pp. 69–83, Jan. 2001.

[20] A. D. Whalen, *Detection of signals in noise*. New York: Academic, 1971.

[21] A. Klein, G. K. Kaleh, and B. W. Baier, "Zero forcing and minimum mean-square-error equalization for multiuser detection in code-division multiple-access channels," *IEEE Trans. Vehicular Technology*, vol. 45, no. 2, pp. 276–287, May 1996.

[22] E. de Carvalho and D. T. M. Slock, "Burst mode equalization: Optimal approach and suboptimal continuous-processing approximation," *Signal Processing*, vol. 80, no. 10, pp. 1999–2015, Oct. 2000.

[23] R. M. Gray, "Toeplitz and circulant matrices: A review," Stanford University, Tech. Rep. Available via http://www-ee.stanford.edu/g̃ray/toeplitz.pdf, Aug. 2002.

[24] T. Pollet, H. Steendam, and M. Moeneclaey, "Performance degradation of multi-carrier systems caused by an insufficient guard interval duration," in *Proc. Intern. Workshop on Copper Wire Access Systems "Bridging the Last Copper Drop" (CWAS)*, Budapest, Hungary, Oct. 1997, pp. 265–270.

[25] P. Spruyt, P. Reusens, and S. Braet, "Performance of improved DMT transceiver for VDSL," ANSI T1E1.4 Committee contribution 96-104, Colorado Springs, Apr. 1996.

[26] R. Nilsson, F. Sjöberg, and J. P. LeBlanc, "A rank-reduced LMMSE canceller for narrowband interference suppression in OFDM-based systems," *IEEE Trans. Communications*, vol. 51, no. 12, pp. 2126–2140, Dec. 2003.

APPENDIX

In this appendix, we show that the optimum L-ZF block equalizer for the case where $N_p \geq L_b - \nu$ gives rise to a two-step procedure: in a first step, an estimate of the $L_b - \nu$ TX difference terms are obtained; these are used in a second step to provide the L-ZF estimate of the TX symbol vector $\tilde{\mathbf{x}}_{k,a}$.

The MMSE L-ZF block equalizer leads to the following block estimate (see also (28)):

$$\begin{bmatrix} \hat{\tilde{\mathbf{x}}}_{k,a}^{\mathrm{L-ZF}} \\ \widehat{\Delta \underline{\mathbf{x}}}_k^{\mathrm{L-ZF}} \end{bmatrix} = \underbrace{\left(\left[\begin{array}{c|c} \tilde{\mathbf{B}}_{a,\mathrm{D}}^H & \mathbf{0} \\ \hline \underline{\mathbf{B}}_{\Delta}^T \mathcal{F}_N^H & \end{array} \right] \left(\mathcal{F}_N \boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2 \mathcal{F}_N^H \right)^{-1} \left[\begin{array}{c|c} \tilde{\mathbf{B}}_{a,\mathrm{D}} & \\ \hline \mathbf{0} & \mathcal{F}_N \underline{\mathbf{B}}_{\Delta} \end{array} \right] \right)}_{\{1\}}{}^{-1} \times \underbrace{\left[\begin{array}{c|c} \tilde{\mathbf{B}}_{a,\mathrm{D}}^H & \mathbf{0} \\ \hline \underline{\mathbf{B}}_{\Delta}^T \mathcal{F}_N^H & \end{array} \right] \left(\mathcal{F}_N \boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2 \mathcal{F}_N^H \right)^{-1}}_{\{2\}} \tilde{\mathbf{z}}_{k,N} \tag{39}$$

The derivation below leads naturally to a two-step procedure, where $\Delta \underline{\mathbf{x}}_k$ is estimated in a first step and used in a second step to estimate $\tilde{\mathbf{x}}_{k,a}$. We introduce the following block-based description for the factors $\{1\}$ and $\{2\}$ in (39):

$$\{1\} = \begin{bmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{F}^H & \mathbf{G} \end{bmatrix} = \left[\begin{array}{c|c} \tilde{\mathbf{B}}_{a,\mathrm{D}}^H \mathcal{F}_a \left(\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2 \right)^{-1} \mathcal{F}_a^H \tilde{\mathbf{B}}_{a,\mathrm{D}} & \tilde{\mathbf{B}}_{a,\mathrm{D}}^H \mathcal{F}_a \left(\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2 \right)^{-1} \underline{\mathbf{B}}_{\Delta} \\ \hline \underline{\mathbf{B}}_{\Delta}^T \left(\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2 \right)^{-1} \mathcal{F}_a^H \tilde{\mathbf{B}}_{a,\mathrm{D}} & \underline{\mathbf{B}}_{\Delta}^T \left(\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2 \right)^{-1} \underline{\mathbf{B}}_{\Delta} \end{array} \right] \tag{40}$$

$$\{2\} = \begin{bmatrix} \mathbf{K} \\ \mathbf{L} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{B}}_{a,\mathrm{D}}^H \mathcal{F}_a \left(\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2 \right)^{-1} \mathcal{F}_N^H \\ \underline{\mathbf{B}}_{\Delta}^T \left(\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2 \right)^{-1} \mathcal{F}_N^H \end{bmatrix} \tag{41}$$

The inverse of the factor $\{1\}$ in (39) and (40) is given by:

$$\begin{bmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{F}^H & \mathbf{G} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{E}^{-1} + \mathbf{E}^{-1}\mathbf{F}\mathbf{Q}^{-1}\mathbf{F}^H\mathbf{E}^{-1} & -\mathbf{E}^{-1}\mathbf{F}\mathbf{Q}^{-1} \\ -\mathbf{Q}^{-1}\mathbf{F}^H\mathbf{E}^{-1} & \mathbf{Q}^{-1} \end{bmatrix} \tag{42}$$

with

$$\mathbf{Q} = \mathbf{G} - \mathbf{F}^H \mathbf{E}^{-1} \mathbf{F} \tag{43}$$

$$= \underline{\mathbf{B}}_{\Delta}^T \left(\underbrace{\left(\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2 \right)^{-1} - \left(\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2 \right)^{-1} \mathcal{F}_a^H \left(\mathcal{F}_a \left(\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2 \right)^{-1} \mathcal{F}_a^H \right)^{-1} \mathcal{F}_a \left(\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2 \right)^{-1}}_{\mathbf{M}_3} \right) \underline{\mathbf{B}}_{\Delta} \tag{44}$$

Using (39-44) and (13), $\Delta \underline{\mathbf{x}}_k$ can be estimated first. It follows from (39-42) that:

$$\widehat{\Delta \underline{\mathbf{x}}}_k^{\mathrm{L-ZF}} = -\mathbf{Q}^{-1} \left(\mathbf{F}^H \mathbf{E}^{-1} \mathbf{K} - \mathbf{L} \right) \tilde{\mathbf{z}}_{k,N} \tag{45}$$

Using the definitions of the matrices $\mathbf{E}$, $\mathbf{F}$, $\mathbf{K}$ and $\mathbf{L}$ in (40-41) results in:

$$\widehat{\Delta \underline{\mathbf{x}}}_k^{\mathrm{L-ZF}} = -\mathbf{Q}^{-1} \underline{\mathbf{B}}_{\Delta}^T \left(\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2 \right)^{-1} \left(\underbrace{\mathcal{F}_a^H \left(\mathcal{F}_a \left(\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2 \right)^{-1} \mathcal{F}_a^H \right)^{-1} \mathcal{F}_a \left(\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2 \right)^{-1} \mathcal{F}_N^H - \mathcal{F}_N^H}_{\{1\}} \right) \tilde{\mathbf{z}}_{k,N} \tag{46}$$

As $\mathcal{F}_N = \begin{bmatrix} \mathcal{F}_a^H & \mathcal{F}_p^H \end{bmatrix}^H$, the factor $\{1\}$ in (46) can be split into two parts, where the first part appears to be an all-zero matrix:

$$\widehat{\Delta \underline{\mathbf{x}}}_k^{\mathrm{L-ZF}} = -\mathbf{Q}^{-1} \underline{\mathbf{B}}_{\Delta}^T \left(\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2 \right)^{-1} \left[\begin{array}{c|c} \mathbf{0}_{N \times N_a} & \mathcal{F}_a^H \left(\mathcal{F}_a \left(\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2 \right)^{-1} \mathcal{F}_a^H \right)^{-1} \mathcal{F}_a \left(\boldsymbol{\Sigma}_{\check{\mathbf{n}}}^2 \right)^{-1} \mathcal{F}_p^H - \mathcal{F}_p^H \end{array} \right] \tilde{\mathbf{z}}_{k,N} \tag{47}$$

From the definition of $\tilde{\mathbf{z}}_{k,N}$ (13), it follows that the L-ZF estimate $\widehat{\underline{\Delta\mathbf{x}}_k}^{\mathrm{L-ZF}}$ is a linear combination of $\tilde{\mathbf{z}}_{k,p}$:

$$\widehat{\underline{\Delta\mathbf{x}}_k}^{\mathrm{L-ZF}} = -\mathbf{Q}^{-1}\underline{\mathbf{B}}_{\Delta}^{T}\left(\mathbf{\Sigma}_{\check{\mathbf{n}}}^{2}\right)^{-1}\left(\mathcal{F}_a^H\left(\mathcal{F}_a\left(\mathbf{\Sigma}_{\check{\mathbf{n}}}^{2}\right)^{-1}\mathcal{F}_a^H\right)^{-1}\mathcal{F}_a\left(\mathbf{\Sigma}_{\check{\mathbf{n}}}^{2}\right)^{-1}\mathcal{F}_p^H - \mathcal{F}_p^H\right)\tilde{\mathbf{z}}_{k,p} \tag{48}$$

The definitions of $\mathbf{Q}$ and $\mathbf{M}_3$ in (43) then lead to:

$$\widehat{\underline{\Delta\mathbf{x}}_k}^{\mathrm{L-ZF}} = \left(\underline{\mathbf{B}}_{\Delta}^{T}\mathbf{M}_3\underline{\mathbf{B}}_{\Delta}\right)^{-1}\underline{\mathbf{B}}_{\Delta}^{T}\mathbf{M}_3\mathcal{F}_p^H\tilde{\mathbf{z}}_{k,p} \tag{49}$$

or

$$\underline{\mathbf{B}}_{\Delta}\widehat{\underline{\Delta\mathbf{x}}_k}^{\mathrm{L-ZF}} = \left[\begin{array}{c}\mathbf{M}_4\\ \hline \mathbf{0}\end{array}\right]\tilde{\mathbf{z}}_{k,p} \tag{50}$$

with $\mathbf{0} = \mathbf{0}_{(N-L_b+\nu)\times(L_b-\nu)}$ and

$$\mathbf{M}_4 = \left[\begin{array}{c|c} \mathbf{I}_{L_b-\nu} & \mathbf{0}^T\end{array}\right]\mathbf{M}_3^{-1}\left[\begin{array}{c}\mathbf{I}_{L_b-\nu}\\ \hline \mathbf{0}\end{array}\right]\left[\begin{array}{c|c} \mathbf{I}_{L_b-\nu} & \mathbf{0}^T\end{array}\right]\mathbf{M}_3\mathcal{F}_p^H \tag{51}$$

The estimation of $\hat{\tilde{\mathbf{x}}}_{k,a}$ can now make use of $\widehat{\underline{\Delta\mathbf{x}}_k}^{\mathrm{L-ZF}}$, as it follows from (39-42) that:

$$\begin{aligned}\hat{\tilde{\mathbf{x}}}_{k,a}^{\mathrm{L-ZF}} &= \mathbf{E}^{-1}\left(\mathbf{K}\tilde{\mathbf{z}}_{k,N} - \mathbf{F}\widehat{\underline{\Delta\mathbf{x}}_k}^{\mathrm{L-ZF}}\right) &(52)\\ &= \mathbf{E}^{-1}\tilde{\mathbf{B}}_{a,\mathrm{D}}^{H}\mathcal{F}_a\left(\mathbf{\Sigma}_{\check{\mathbf{n}}}^{2}\right)^{-1}\mathcal{F}_N^H\left(\tilde{\mathbf{z}}_{k,N} - \mathcal{F}_N\underline{\mathbf{B}}_{\Delta}\widehat{\underline{\Delta\mathbf{x}}_k}^{\mathrm{L-ZF}}\right) &(53)\end{aligned}$$

Using (50), the definition of $\mathbf{E}$ in (40) and $\mathcal{F}_N = \left[\begin{array}{c|c}\mathcal{F}_a^H & \mathcal{F}_p^H\end{array}\right]^H$, (53) reduces to the following **optimum L-ZF block estimate**:

$$\begin{aligned}\hat{\tilde{\mathbf{x}}}_{k,a}^{\mathrm{L-ZF}} &= \left[\begin{array}{c|c}\tilde{\mathbf{B}}_{a,\mathrm{D}}^{-1} & \underbrace{\tilde{\mathbf{B}}_{a,\mathrm{D}}^{-1}\left(\mathcal{F}_a\left(\mathbf{\Sigma}_{\check{\mathbf{n}}}^{2}\right)^{-1}\mathcal{F}_a^H\right)^{-1}\mathcal{F}_a\left(\mathbf{\Sigma}_{\check{\mathbf{n}}}^{2}\right)^{-1}\mathcal{F}_p^H}_{\mathbf{M}_1}\end{array}\right]\left(\tilde{\mathbf{z}}_{k,N} - \mathcal{F}_N\underline{\mathbf{B}}_{\Delta}\widehat{\underline{\Delta\mathbf{x}}_k}^{\mathrm{L-ZF}}\right) &(54)\\ &= \tilde{\mathbf{B}}_{a,\mathrm{D}}^{-1}\tilde{\mathbf{z}}_{k,a} + \mathbf{M}_5\tilde{\mathbf{z}}_{k,p} &(55)\end{aligned}$$

where $\mathbf{M}_5 = \mathbf{M}_1 + \left(\tilde{\mathbf{B}}_{a,\mathrm{D}}^{-1}\mathcal{F}_a + \mathbf{M}_1\mathcal{F}_p\right)\left[\begin{array}{c}\mathbf{M}_4\\ \hline \mathbf{0}\end{array}\right]$.