ON TIME-DOMAIN AND FREQUENCY-DOMAIN MMSE-BASED TEQ DESIGNS FOR DMT TRANSMISSION

Koen Vanbleu, Geert Ysebaert, Gert Cuypers, Marc Moonen

K.U. Leuven, Dept. ESAT-SCD, Kasteelpark Arenberg 10, B-3001 Leuven, Belgium
{vanbleu,ysebaert,cuypers,moonen}@esat.kuleuven.ac.be

ABSTRACT

We reconsider the MMSE-based time-domain equalizer (TEQ), bitrate maximizing TEQ (BM-TEQ) and per-tone equalizer design for DMT transmission. The MMSE-TEQ criterion can be formulated as a least-squares (LS) criterion that minimizes a time-domain (TD) error energy. Based on this LS-based TD-MMSE-TEQ, we derive new LS-based frequency-domain (FD) MMSE-TEQ criteria that are intermediate in terms of computational complexity and performance between the TD-MMSE-TEQ and the BM-TEQ. In addition, we show that the BM-TEQ design itself is equivalent to a so-called iteratively-reweighted separable nonlinear LS-based FD-MMSE-TEQ design. As a side result, the considered LS-based equalizer designs, although at first sight very different in nature- appear closely related when turning them into generalized eigenvalue problems.

1. INTRODUCTION

In discrete multitone (DMT) based systems, such as asymmetric digital subscriber lines (ADSL), channel impulse responses can be very long, hence a long cyclic prefix (CP, length ν) would be required. A solution to avoid this overhead is to insert a (real) T-tap *time domain equalizer* $\mathbf{w}$ (TEQ) before demodulation, which then shortens the channel impulse response to $\nu+1$ samples. Among the numerous TEQ designs, we will focus in this paper on the so-called minimum mean-square error (MMSE)-based TEQ design [1] with a unit energy constraint (UEC) [2, 3] and the recently proposed bitrate maximizing TEQ design (BM-TEQ) [4]. In [5], the alternative *per-tone equalizer* (PTEQ) scheme is proposed that always performs at least as well as - and usually better than - a TEQ based receiver while keeping complexity during data transmission at the same level. The PTEQ is a complex MMSE equalizer designed for each tone separately.

The classical MMSE-TEQ criterion can be formulated as a constrained linear least-squares (CLLS) criterion that minimizes a time-domain (TD) error energy. Starting from this CLLS-based TD-MMSE-TEQ criterion, we derive new LS-based MMSE-TEQ criteria, that minimize a sum-square of frequency-domain (FD) error energies (i.e., after DFT demodulation), rather than a TD error energy; especially the so-called separable nonlinear LS (SNLLS)-based FD-MMSE-TEQ appears a reasonable intermediate in terms of complexity and performance between the TD-MMSE-TEQ and the BM-TEQ. Remarkably, the BM-TEQ criterion itself is found to be equivalent to a so-called iteratively-reweighted SNLLS-based FD-MMSE-TEQ criterion. As a side result, the LS-based formulations of the TD-MMSE-TEQ, FD-MMSE-TEQ, BM-TEQ and PTEQ design cost functions appear to be closely related, especially when turning each of them into a generalized eigenvalue (GEV) problem

$$\mathbf{B}\mathbf{w} = \lambda\mathbf{A}\mathbf{w} \tag{1}$$

where, loosely speaking, $\mathbf{A}$ is an autocorrelation metric of the received signal y_l and $\mathbf{B}$ depends on a crosscorrelation metric between transmitted (TX) and received (RX) signal x_l and y_l. For an extended version of this paper, we refer to [6].

Notation. The DMT symbol index is k. $\mathscr{S}_a$ is the set of N_a active tones; n is a tone index; N is the (I)DFT size; $\mathscr{F}_{\mathscr{S}_a}$ is an $N_a \times N$ submatrix of the full DFT matrix $\mathscr{F}_N$ with only the N_a rows of the active tones $\mathscr{S}_a$; the n-th DFT row is $\mathscr{F}_n$. Vectors are typeset in bold lowercase while matrices are in bold uppercase. A tilde over a variable distinguishes frequency-domain (FD) symbols from time-domain (TD) symbols, e.g. the $N_a \times 1$ TX symbol vector at time k, $\tilde{\mathbf{x}}_k$. FD vectors or matrices only account for the N_a active tones $\mathscr{S}_a$ unless a subscript N is added (e.g., the $N \times 1$ TX symbol vector, $\tilde{\mathbf{x}}_{k,N}$). The entry for tone n of a FD vector is denoted with a subscript, e.g., $\tilde{x}_{k,n}$. A subscript with the number of data points, e.g., L samples or K DMT symbols, is used to distinguish between a (deterministic) correlation *estimate*, e.g., $\mathbf{\Sigma}^2_{L,\mathbf{y}} = \frac{1}{L}\sum_{l=1}^{L}\mathbf{y}_l\mathbf{y}_l^T$ or $\boldsymbol{\sigma}_{K,n,\tilde{x}\tilde{\mathbf{y}}} = \frac{1}{K}\sum_{k=1}^{K}\tilde{x}^*_{k,n}\tilde{\mathbf{y}}_{k,n}$, and the *true* (stochastic) correlation, e.g., $\mathbf{\Sigma}^2_{\mathbf{y}} = \mathscr{E}\left\{\mathbf{y}_l^T\mathbf{y}_l\right\}$ or $\boldsymbol{\sigma}_{n,\tilde{x}\tilde{\mathbf{y}}} = \mathscr{E}\left\{\tilde{x}^*_{k,n}\tilde{\mathbf{y}}_{k,n}\right\}$. Throughout the text we only define the stochastic correlations.

2. MMSE-TEQ, BM-TEQ AND PTEQ: LS PROBLEMS

2.1 CLLS-based TD-MMSE-TEQ design

One of the earliest presented TEQ designs is the MMSE-based TEQ [1]: it minimizes the *time-domain* (TD) MSE between the output of the TEQ, $y_{l,\mathbf{w}} = \mathbf{y}_l^T\mathbf{w}$, with $\mathbf{w}$ the T-tap TEQ, l the sample index and $\mathbf{y}_l = [\begin{array}{ccc} y_l & \cdots & y_{l-T+1}\end{array}]^T$ a vector of RX samples[1], and the output $\mathbf{x}_l^T\mathbf{b}$ of a virtual FIR channel, the so-called target impulse response (TIR) $\mathbf{b}$ of length $\nu+1$ (with ν the CP length), which is fed with a vector of TX samples $\mathbf{x}_l = [\begin{array}{ccc} x_l & \cdots & x_{l-\nu}\end{array}]^T$:

$$\min_{\mathbf{w},\mathbf{b}}\mathscr{E}\left\{|e_l|^2\right\} = \min_{\mathbf{w},\mathbf{b}}\mathscr{E}\left\{\left|\mathbf{y}_l^T\mathbf{w} - \mathbf{x}_l^T\mathbf{b}\right|^2\right\} \tag{2}$$

To avoid the trivial solution $\mathbf{w} = \mathbf{0}$, $\mathbf{b} = \mathbf{0}$, a nontriviality constraint is added [2]. We focus on the particular choice of a so-called unit energy constraint (UEC) on $\mathbf{w}$ [3]:

$$\mathbf{w}^T\mathbf{\Sigma}^2_{\mathbf{y}}\mathbf{w} = 1 \tag{3}$$

with the autocorrelation matrix $\mathbf{\Sigma}^2_{\mathbf{y}} = \mathscr{E}\left\{\mathbf{y}_l^T\mathbf{y}_l\right\}$. This constrained TD-MMSE-TEQ criterion (2) forces the joint channel-TEQ impulse response to have a main energy window of $\nu+1$ samples. A deterministic **constrained linear least-squares (CLLS) based TD-MMSE-TEQ** criterion, equivalent to (2), is given by:

$$\boxed{\min_{\mathbf{w},\mathbf{b}}\frac{1}{L}\sum_{l=1}^{L}\left|\mathbf{y}_l^T\mathbf{w} - \mathbf{x}_l^T\mathbf{b}\right|^2 \text{ s.t. } \mathbf{w}^T\mathbf{\Sigma}^2_{L,\mathbf{y}}\mathbf{w} = 1} \tag{4}$$

with L the total number of available data samples and $\mathbf{\Sigma}^2_{L,\mathbf{y}}$ an estimate of $\mathbf{\Sigma}^2_{\mathbf{y}}$ as clarified earlier on this page in the paragraph on the adopted notation. Using the so-called orthogonality condition

This research work was carried out at the ESAT laboratory of the K.U. Leuven, in the frame of IUAP P5/22 and P5/11, GOA-MEFISTO-666, Research Project FWO nr.G.0196.02 and was partially sponsored by Alcatel-Bell. The scientific responsibility is assumed by its authors.

[1] The RX signal y_l and vector $\mathbf{y}_l$ depend on a synchronization delay Δ, which we do not mention explicitly here.

[1, 2] eliminating $\mathbf{b}$ and defining $\boldsymbol{\Sigma}_{\mathbf{x}}^2 = \mathscr{E}\left\{\mathbf{x}_l\mathbf{x}_l^T\right\}$ and $\boldsymbol{\Sigma}_{L,\mathbf{xy}} = \mathscr{E}\left\{\mathbf{x}_l\mathbf{y}_l^T\right\}$, (4) reduces to:

$$\min_{\mathbf{w}} \mathbf{w}^T\left[\boldsymbol{\Sigma}_{L,\mathbf{y}}^2 - \boldsymbol{\Sigma}_{L,\mathbf{xy}}^T\left(\boldsymbol{\Sigma}_{L,\mathbf{x}}^2\right)^{-1}\boldsymbol{\Sigma}_{L,\mathbf{xy}}\right]\mathbf{w} \text{ s.t. } \mathbf{w}^T\boldsymbol{\Sigma}_{L,\mathbf{y}}^2\mathbf{w} = 1 \quad (5)$$

The solution is seen to be the dominant GEV of (1) with the matrix pair

$$(\mathbf{B},\mathbf{A}) = \left(\boldsymbol{\Sigma}_{L,\mathbf{xy}}^T\left(\boldsymbol{\Sigma}_{L,\mathbf{x}}^2\right)^{-1}\boldsymbol{\Sigma}_{L,\mathbf{xy}}, \boldsymbol{\Sigma}_{L,\mathbf{y}}^2\right) \quad (6)$$

2.2 CLLS-based FD-MMSE-TEQ design

The TD-MMSE-TEQ (2) is *sample*-based and minimizes a *TD MSE*. In this and the next section, we develop new frequency domain (FD) MMSE-TEQ criteria that account for the DMT *block* transmission structure, including the CP, and minimize a sum of *FD MSEs*. Especially the FD-MMSE-TEQ criterion, developed in Section 2.3, appears a useful intermediate in terms of complexity and performance between the TD-MMSE-TEQ on one hand, and the PTEQ and the BM-TEQ on the other hand (see Section 3).

First, we rewrite (2) on a *per-DMT-symbol* basis:

$$\min_{\mathbf{w},\mathbf{b}} \underbrace{\mathscr{E}\left\{\|\mathbf{Y}_k\mathbf{w} - \mathbf{X}_k\mathbf{b}\|^2\right\}}_{\mathscr{E}\{\|\mathbf{e}_k\|^2\}} \text{ s.t. } \mathbf{w}^T\boldsymbol{\Sigma}_{\mathbf{Y}}^2\mathbf{w} = 1 \quad (7)$$

The Toeplitz matrix $\mathbf{Y}_k$ has size $N \times T$; its first column and row are given by $[y_{k,0} \ \cdots \ y_{k,N-1}]^T$ and $[y_{k,0} \ \cdots \ y_{k,-T+1}]$, respectively, with $y_{k,i} = y_{k(N+\nu)+i}$. The matrix $\mathbf{X}_k$, which incorporates the CP, has size $N \times (\nu+1)$ and is columnwise circulant with first column $[x_{k,0} \ \cdots \ x_{k,N-1}]^T$ and $x_{k,i} = x_{k(N+\nu)+i}$. The first term $\mathbf{Y}_k\mathbf{w}$ in (7) convolves the k-th DMT RX symbol with the TEQ and is the $N \times 1$ TEQ output vector that is fed to the RX DFT. The second term $\mathbf{X}_k\mathbf{b}$ is the convolution of the k-th DMT TX symbol and the TIR $\mathbf{b}$.

In a second step, $\mathbf{X}_k$ is extended with $N-\nu-1$ columns to an $N \times N$ circulant matrix $\mathbf{X}_{k,\mathrm{C}}$, $\mathbf{b}$ is zero-padded accordingly and the DFT-based decomposition of the circulant matrix $\mathbf{X}_{k,\mathrm{C}} = \mathscr{F}_N^H\tilde{\mathbf{X}}_{k,N,\mathrm{D}}\mathscr{F}_N$, with $\tilde{\mathbf{X}}_{k,N,\mathrm{D}} = \mathrm{diag}(\tilde{\mathbf{x}}_{k,N})$ and $\tilde{\mathbf{x}}_{k,N}$ the $N \times 1$ DMT TX symbol vector, is plugged in:

$$\mathbf{X}_k\mathbf{b} = \mathbf{X}_{k,\mathrm{C}}\begin{bmatrix}\mathbf{b}\\ \mathbf{0}\end{bmatrix} = \mathscr{F}_N^H\tilde{\mathbf{X}}_{k,N,\mathrm{D}}\underbrace{\mathscr{F}_N\begin{bmatrix}\mathbf{b}\\ \mathbf{0}\end{bmatrix}}_{\tilde{\mathbf{b}}_N} \quad (8)$$

Thirdly, the cost function and constraint (7) are transformed to the FD and only the active tones $\mathscr{S}_a$ are considered:

$$\min_{\mathbf{w},\mathbf{b}} \underbrace{\mathscr{E}\left\{\left\|\mathscr{F}_{\mathscr{S}_a}\mathbf{e}_k\right\|^2\right\}}_{\mathscr{E}\{\|\tilde{\mathbf{e}}_k\|^2\}} = \min_{\mathbf{w},\mathbf{b}} \mathscr{E}\left\{\left\|\tilde{\mathbf{Y}}_k\mathbf{w} - \tilde{\mathbf{X}}_{k,\mathrm{D}}\tilde{\mathbf{b}}\right\|^2\right\} \quad (9)$$

$$\text{s.t. } \mathbf{w}^T\boldsymbol{\Sigma}_{\tilde{\mathbf{Y}}}^2\mathbf{w} = 1,\ \tilde{\mathbf{b}} = \mathscr{F}_{\mathscr{S}_a}\begin{bmatrix}\mathbf{b}\\ \mathbf{0}\end{bmatrix} \text{ and real } \mathbf{w} \text{ and } \mathbf{b} \quad (10)$$

where $\tilde{\mathbf{Y}}_k = \mathscr{F}_{\mathscr{S}_a}\mathbf{Y}_k$, $\tilde{\mathbf{X}}_{k,\mathrm{D}} = \mathrm{diag}(\tilde{\mathbf{x}}_k)$ (with $\tilde{\mathbf{x}}_k$ the $N_a \times 1$ DMT TX symbol vector) and $\boldsymbol{\Sigma}_{\tilde{\mathbf{Y}}}^2 = \mathscr{E}\left\{\tilde{\mathbf{Y}}_k^H\tilde{\mathbf{Y}}_k\right\}$. The first term of the error vector $\tilde{\mathbf{e}}_k$ in (9) corresponds to the RX DFT output at the tones $\mathscr{S}_a$:

$$\underbrace{\mathscr{F}_{\mathscr{S}_a}(\mathbf{Y}_k\mathbf{w})}_{\{1\}} = \underbrace{(\mathscr{F}_{\mathscr{S}_a}\mathbf{Y}_k)\mathbf{w}}_{\{2\}} = \tilde{\mathbf{Y}}_k\mathbf{w} = \tilde{\mathbf{y}}_{k,\mathbf{w}} \quad (11)$$

which can either be computed as $\{1\}$ the DFT of the TEQ output $\mathbf{Y}_k\mathbf{w}$ or $\{2\}$ as a linear combination $\mathbf{w}$ of the sliding DFT of the k-th DMT RX symbol, $\tilde{\mathbf{Y}}_k = \mathscr{F}_{\mathscr{S}_a}\mathbf{Y}_k$ (see [4] for details). The second constraint in (10) comes from the original TD-MMSE-TEQ design that imposes channel shortening by means of a TIR $\mathbf{b}$ of length $\nu+1$. If we drop the constraints on $\tilde{\mathbf{b}}$ in (10) and instead optimize

$$\min_{\mathbf{w},\tilde{\mathbf{b}}} \mathscr{E}\left\{\left\|\tilde{\mathbf{Y}}_k\mathbf{w} - \tilde{\mathbf{X}}_{k,\mathrm{D}}\tilde{\mathbf{b}}\right\|^2\right\} \text{s.t.} \mathbf{w}^T\boldsymbol{\Sigma}_{\tilde{\mathbf{Y}}}^2\mathbf{w} = 1 \text{ and real } \mathbf{w} \quad (12)$$

we obtain an *FD-MMSE-TEQ criterion* in the (typically) real TEQ $\mathbf{w}$ and the complex vector $\tilde{\mathbf{b}}$ instead of $\mathbf{b}$. The optimum solution for the unconstrained $\tilde{\mathbf{b}}$ follows from the so-called orthogonality condition and is a vector with as entries b_n the inverses of the unbiased MMSE-based (uMMSE) FEQs $\tilde{d}_n^{\mathrm{uMMSE}}$, which are in fact the optimal choice of FEQs for a given $\mathbf{w}$ [4, 7]:

$$\tilde{d}_n^{\mathrm{uMMSE}} = \frac{\sigma_{n,\tilde{x}}^2}{\boldsymbol{\sigma}_{n,\tilde{x}\tilde{\mathbf{y}}}\mathbf{w}} = \frac{1}{\tilde{b}_n} \quad (13)$$

where $\sigma_{n,\tilde{x}}^2 = \mathscr{E}\left\{\left|\tilde{x}_{k,n}\right|^2\right\}$ is the variance of $\tilde{x}_{k,n}$ and where the denominator is the crosscorrelation $\mathscr{E}\left\{\tilde{x}_{k,n}^*\tilde{y}_{k,n,\mathbf{w}}\right\}$ between the RX DFT output and the TX symbol on tone n. It follows from (11) that this crosscorrelation is equal to $\boldsymbol{\sigma}_{n,\tilde{x}\tilde{\mathbf{y}}}\mathbf{w}$, with $\boldsymbol{\sigma}_{n,\tilde{x}\tilde{\mathbf{y}}} = \mathscr{E}\left\{\tilde{x}_{k,n}^*\tilde{\mathbf{y}}_{k,n}\right\}$ the $1 \times T$ crosscorrelation vector of $\tilde{x}_{k,n}$ and the n-th sliding DFT output $\tilde{\mathbf{y}}_{k,n} = \mathscr{F}_n\mathbf{Y}_k$ (see [4] for details). Solving (12) then optimizes the sum-square energy between the DFT outputs $\tilde{y}_{k,n,\mathbf{w}}$ and the scaled desired symbols $\frac{\tilde{x}_{k,n}}{\tilde{d}_n^{\mathrm{uMMSE}}}$. A deterministic **CLLS-based FD-MMSE-TEQ** criterion, equivalent with (12) is given by:

$$\boxed{\min_{\mathbf{w},\tilde{\mathbf{b}}} \frac{1}{K}\sum_{k=1}^{K}\left\|\tilde{\mathbf{Y}}_k\mathbf{w} - \tilde{\mathbf{X}}_{k,\mathrm{D}}\tilde{\mathbf{b}}\right\|^2 \text{s.t.} \mathbf{w}^T\boldsymbol{\Sigma}_{K,\tilde{\mathbf{Y}}}^2\mathbf{w} = 1 \text{ and real } \mathbf{w}} \quad (14)$$

where K is the number of available DMT symbols. Due to the similarity between the CLLS-based FD-MMSE-TEQ criterion (12) and the CLLS-based TD-MMSE-TEQ (4), it comes as no surprise that (12) reduces to a GEV problem (1) that is closely related to (6):

$$\begin{aligned}(\mathbf{B},\mathbf{A}) &= \left(\Re\left\{\boldsymbol{\Sigma}_{K,\tilde{\mathbf{x}}\tilde{\mathbf{Y}}}^H\left(\boldsymbol{\Sigma}_{K,\tilde{\mathbf{x}}}^2\right)^{-1}\boldsymbol{\Sigma}_{K,\tilde{\mathbf{x}}\tilde{\mathbf{Y}}}\right\}, \Re\left\{\boldsymbol{\Sigma}_{K,\tilde{\mathbf{Y}}}^2\right\}\right) \quad (15)\\ &= \left(\Re\left\{\sigma_{K,n,\tilde{x}}^{-2}\sum_{n\in\mathscr{S}_a}\boldsymbol{\sigma}_{K,n,\tilde{x}\tilde{\mathbf{y}}}^H\boldsymbol{\sigma}_{K,n,\tilde{x}\tilde{\mathbf{y}}}\right\}, \Re\left\{\sum_{n\in\mathscr{S}_a}\boldsymbol{\Sigma}_{K,n,\tilde{\mathbf{y}}}^2\right\}\right)\end{aligned}$$

The N_a rows of $\boldsymbol{\Sigma}_{\tilde{\mathbf{x}}\tilde{\mathbf{Y}}} = \mathscr{E}\left\{\tilde{\mathbf{X}}_{k,\mathrm{D}}^*\tilde{\mathbf{Y}}_k\right\}$ are the above defined crosscorrelation vectors $\boldsymbol{\sigma}_{n,\tilde{x}\tilde{\mathbf{y}}}$; $\boldsymbol{\Sigma}_{\tilde{\mathbf{Y}}}^2 = \sum_{n\in\mathscr{S}_a}\boldsymbol{\Sigma}_{n,\tilde{\mathbf{y}}}^2$ with $\boldsymbol{\Sigma}_{n,\tilde{\mathbf{y}}}^2 = \mathscr{E}\left\{\tilde{\mathbf{y}}_{k,n}^H\tilde{\mathbf{y}}_{k,n}\right\}$ the autocorrelation matrix of the n-th sliding DFT output; $\boldsymbol{\Sigma}_{\tilde{\mathbf{x}}}^2 = \mathscr{E}\left\{\tilde{\mathbf{x}}_k\tilde{\mathbf{x}}_k^H\right\}$ is the autocorrelation matrix of the DMT TX symbol vector; the second equality assumes independent symbols $\tilde{x}_{k,n}$ such that $\boldsymbol{\Sigma}_{\tilde{\mathbf{x}}}$ is diagonal with diagonal elements $\sigma_{n,\tilde{x}}^2$; the $\Re$-operators ensure a real TEQ.

The *complex* LS-based MMSE-PTEQ [5] is closely related to the CLLS-based FD-MMSE-TEQ (14) when only 1 tone n is considered. It follows from (15) that the *real* PTEQ for tone n, $\mathbf{w}_n$, is the dominant eigenvector of

$$(\mathbf{B},\mathbf{A}) = \left(\sigma_{K,n,\tilde{\mathbf{x}}}^{-2}\Re\left\{\boldsymbol{\sigma}_{K,n,\tilde{x}\tilde{\mathbf{y}}}^H\boldsymbol{\sigma}_{K,n,\tilde{x}\tilde{\mathbf{y}}}\right\}, \Re\left\{\boldsymbol{\Sigma}_{K,n,\tilde{\mathbf{y}}}^2\right\}\right) \quad (16)$$

In case of a *complex* $\mathbf{w}_n$, the $\Re$-operators should be dropped. In this case the matrix $\mathbf{B}$ becomes rank-one and the dominant eigenvector of (16) (up to a scaling) is seen to be given by [6]

$$\mathbf{w}_n = \left(\boldsymbol{\Sigma}_{K,n,\tilde{\mathbf{y}}}^2\right)^{-1}\boldsymbol{\sigma}_{K,n,\tilde{x}\tilde{\mathbf{y}}}^H \quad (17)$$

This is exactly the solution of the **LS-based MMSE-PTEQ** criterion of [5]:

$$\boxed{\min_{\mathbf{w}_n} \frac{1}{K}\sum_{k=1}^{K}\left|\tilde{\mathbf{y}}_{k,n}\mathbf{w}_n - \tilde{x}_{k,n}\right|^2} \quad (18)$$

2.3 SNLLS-based FD-MMSE-TEQ design

An alternative (suboptimal) FD criterion is obtained by minimizing the sum-square energies at the *FEQ output* instead of the DFT output:

$$\min_{\mathbf{w},\tilde{\mathbf{d}}} \underbrace{\mathscr{E}\left\{\left\|\mathrm{diag}(\tilde{\mathbf{d}})\tilde{\mathbf{Y}}_k\mathbf{w}-\tilde{\mathbf{x}}_k\right\|^2\right\}}_{\mathscr{E}\left\{\left\|\tilde{\mathbf{e}}_{k,\boldsymbol{\theta}}\right\|^2\right\}} \text{ with real } \mathbf{w} \tag{19}$$

where the FEQ output error vector $\tilde{\mathbf{e}}_{k,\boldsymbol{\theta}}$ depends on both the real TEQ and complex FEQ parameters, $\boldsymbol{\theta} = \left[\mathbf{w}^H \;\; \tilde{\mathbf{d}}^H\right]^H$. The UEC constraint of (12) has been dropped as the criterion (19) has no trivial solution anymore. This criterion corresponds to an SNLLS criterion [8, 9]:

$$\boxed{\min_{\mathbf{w},\tilde{\mathbf{d}}} \frac{1}{K}\sum_{k=1}^{K}\left\|\mathrm{diag}(\tilde{\mathbf{d}})\tilde{\mathbf{Y}}_k\mathbf{w}-\tilde{\mathbf{x}}_k\right\|^2 \text{ with real } \mathbf{w}} \tag{20}$$

which we call an **SNLLS-based FD-MMSE-TEQ** criterion. The separability property follows from the fact that the error $\tilde{\mathbf{e}}_{k,\boldsymbol{\theta}}$ is nonlinear in $\boldsymbol{\theta}$, whereas the TEQ $\mathbf{w}$ and FEQs $\tilde{\mathbf{d}}$ appear linearly. Solving (19) as a linear problem in $\tilde{\mathbf{d}}$, while keeping $\mathbf{w}$ fixed, results in the (biased) MMSE FEQs for the given $\mathbf{w}$ [4, 7]:

$$\tilde{d}_n^{\mathrm{MMSE}} = \frac{\mathbf{w}^T\boldsymbol{\sigma}^H_{n,\tilde{x}\tilde{\mathbf{y}}}}{\mathbf{w}^T\boldsymbol{\Sigma}^2_{n,\tilde{\mathbf{y}}}\mathbf{w}} \tag{21}$$

where the numerator is equal to the complex conjugate of the denominator of the uMMSE FEQ (13) and where the denominator is the autocorrelation of the DFT output, i.e., $\mathscr{E}\left\{\left|\tilde{y}_{k,n,\mathbf{w}}\right|^2\right\} = \mathbf{w}^T\boldsymbol{\Sigma}^2_{n,\tilde{\mathbf{y}}}\mathbf{w}$. It will be shown in Section 2.4 that the SNLLS problem (20) can be solved iteratively with a sequence of GEV problems (1). As will be shown in the simulations of Section 3, this SNLLS-based FD-MMSE-TEQ design consistently outperforms the CLLS-based TD-MMSE-TEQ design and closely approaches the BM-TEQ performance.

2.4 Bitrate maximizing FD-MMSE-TEQ design

The bitrate maximizing TEQ (BM-TEQ), originally presented in [4], is the solution to the following constrained nonlinear optimization problem in $\boldsymbol{\theta} = \left[\mathbf{w}^H \;\; \tilde{\mathbf{d}}^H\right]^H$:

$$\max_{\boldsymbol{\theta}} \textstyle\sum_{n\in\mathscr{S}_a}\log_2\left(1+\frac{\mathrm{SNR}_{n,\boldsymbol{\theta}_n}}{\Gamma_n}\right) \tag{22}$$

$$\text{with } \mathrm{SNR}_{n,\boldsymbol{\theta}_n} = \frac{\sigma^2_{n,\tilde{x}}}{\mathscr{E}\left\{\left|\tilde{e}_{k,n,\boldsymbol{\theta}_n}\right|^2\right\}} = \frac{\sigma^2_{n,\tilde{x}}}{\mathscr{E}\left\{\left|\tilde{d}_n\tilde{\mathbf{y}}_{k,n}\mathbf{w}-\tilde{x}_{k,n}\right|^2\right\}} \tag{23}$$

$$\text{subject to } \tilde{d}_n = \frac{\sigma^2_{n,\tilde{x}}}{\boldsymbol{\Sigma}_{n,\tilde{x}\tilde{\mathbf{y}}}\mathbf{w}}, \forall n \in \mathscr{S}_a \tag{24}$$

with $\boldsymbol{\theta}_n = \left[\mathbf{w}^H \;\; \tilde{d}_n^*\right]^H$, i.e., maximizing the number of bits per DMT symbol (given a certain SNR gap Γ_n between SNR_n and the SNR required to achieve Shannon capacity, typically assumed to be independent of the equalizer [4]), over the joint TEQ-FEQ parameters $\boldsymbol{\theta}$, subject to the use of uMMSE FEQs (24) (see also (13)), which render the subchannel SNR model in (23) exact [4]. It has been shown in [6], based on (22-24), that this optimization criterion is equivalent to the following ***iteratively reweighted*** **SNLLS-based bitrate maximizing FD-MMSE-TEQ (IR-SNLLS-based BM-FD-MMSE-TEQ)** criterion (explained below) [10]:

$$\boxed{\min_{\boldsymbol{\theta}} \frac{1}{K}\sum_{k=1}^{K}\left\|\mathrm{diag}\left(\sqrt{\check{\boldsymbol{\gamma}}_{K,\boldsymbol{\theta}_{\mathrm{prev}}}}\right)\mathbf{e}_{k,\boldsymbol{\theta}}\right\|^2} \tag{25}$$

with

$$\check{e}_{k,n,\boldsymbol{\theta}_n} = \tilde{d}_n\tilde{\mathbf{y}}_{k,n}\mathbf{w}-\tilde{x}_{k,n} \tag{26}$$

$$\check{\gamma}_{K,n,\boldsymbol{\theta}_n} = \frac{\left(\mathrm{SNR}_{K,n,\boldsymbol{\theta}_n}+1\right)^2}{\sigma^2_{K,n,\tilde{x}}\left(\mathrm{SNR}_{K,n,\boldsymbol{\theta}_n}+\Gamma_n\right)} \tag{27}$$

$$\mathrm{SNR}_{K,n,\boldsymbol{\theta}_n} = \frac{\sigma^2_{K,n,\tilde{x}}}{\frac{1}{K}\sum_{k=1}^{K}\left\|\check{e}_{k,n,\boldsymbol{\theta}_n}\right\|^2} = \frac{1}{\rho^{-2}_{K,n,\boldsymbol{\theta}_n}-1} \tag{28}$$

$$\rho^2_{K,n,\boldsymbol{\theta}_n} = \frac{\left|\boldsymbol{\sigma}_{K,n,\tilde{x}\tilde{\mathbf{y}}}\mathbf{w}\right|^2}{\sigma^2_{K,n,\tilde{x}}\left(\mathbf{w}^T\boldsymbol{\Sigma}^2_{K,n,\tilde{\mathbf{y}}}\mathbf{w}\right)} \tag{29}$$

The SNLLS-based FD-MMSE-TEQ (19) is indeed an unweighted version of, hence closely related to the IR-SNLLS-based BM-FD-MMSE-TEQ (25).

IR-LS problems such as (25) are weighted LS problems where the weights $\check{\boldsymbol{\gamma}}_{K,\boldsymbol{\theta}_{\mathrm{prev}}}$ depend on the LS errors $\mathbf{e}_{k,\boldsymbol{\theta}}$ (here: via the subchannel SNRs (28)), hence on the optimization parameters $\boldsymbol{\theta}$. They are typically solved as a sequence of weighted LS problems (here: a SNLLS problem) where the weights in each iteration are computed with the parameter estimates from the previous iteration, $\boldsymbol{\theta}_{\mathrm{prev}}$. According to [10], convergence occurs provided that the weights are bounded and non-increasing in the (absolute value) of the LS errors. For a non-convex cost function, the IR-LS algorithm leads to a local optimum.

According to [8, 9], an SNLLS problem, such as the FD-MMSE-TEQ criteria (19) and (25), are -as the IR-LS problem-also solved iteratively by alternately updating the parameters $\mathbf{w}$ and $\tilde{\mathbf{d}}$. An iteration step for the IR-SNLLS-based BM-FD-MMSE-TEQ criterion then consists of the computation of (1) the weights, $\check{\boldsymbol{\gamma}}_{K,\boldsymbol{\theta}_{\mathrm{prev}}}$, (2) estimates of the biased MMSE FEQs (21), $\tilde{\mathbf{d}}_K$, which are the solutions of (25) for a fixed $\mathbf{w}_{\mathrm{prev}}$ and (3) a new BM-TEQ estimate $\mathbf{w}$:

$$\mathbf{w} = \underbrace{\Re\left\{\sum_{n\in\mathscr{S}_a}\check{\gamma}_n\left|\tilde{d}_n\right|^2\boldsymbol{\Sigma}^2_{K,n,\tilde{\mathbf{y}}}\right\}^{-1}}_{\left(\boldsymbol{\Sigma}^2_{K,\tilde{\mathbf{y}},\gamma}\right)^{-1}}\underbrace{\Re\left\{\sum_{n\in\mathscr{S}_a}\check{\gamma}_n\tilde{d}_n\boldsymbol{\sigma}^H_{K,n,\tilde{x}\tilde{\mathbf{y}}}\right\}}_{\boldsymbol{\sigma}_{K,\tilde{x}\tilde{\mathbf{y}},\gamma}} \tag{30}$$

with $\check{\gamma}_n = \check{\gamma}_{K,n,\boldsymbol{\theta}_{n,\mathrm{prev}}}$ and $\tilde{d}_n = \tilde{d}_{K,n}$, which is (similar to the PTEQ $\mathbf{w}_n$ (17)) the solution of a GEV problem with rank-one matrix $\mathbf{B}$:

$$(\mathbf{B},\mathbf{A}) \;=\; \left(\boldsymbol{\sigma}_{K,\tilde{x}\tilde{\mathbf{y}},\gamma}\boldsymbol{\sigma}^H_{K,\tilde{x}\tilde{\mathbf{y}},\gamma},\boldsymbol{\Sigma}^2_{K,\tilde{\mathbf{y}},\gamma}\right) \tag{31}$$

For a complex TEQ, the $\Re$-operators must be omitted. The iterations for solving the SNLLS-based FD-MMSE-TEQ (19) do not include the first step, i.e., the weights $\check{\gamma}_{K,n,\boldsymbol{\theta}_{n,\mathrm{prev}}}$ always equal 1. Note that other solution strategies for SNLLS problems exist: in [8, 9], it is argued that step (3), which solves for $\mathbf{w}$ keeping $\tilde{\mathbf{d}}$ fixed can be better replaced by, e.g., a much faster converging Gauss-Newton updating step of the joint parameter vector $\boldsymbol{\theta}$.

2.5 Relation between the LS cost functions

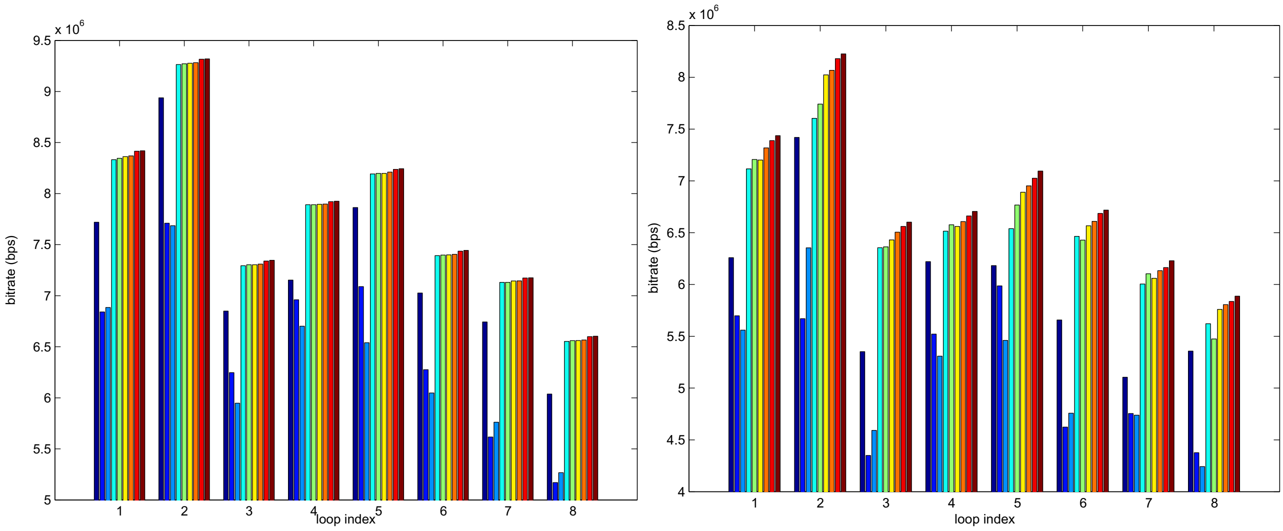
Throughout the text, each LS problem has been shown to be equivalent to a GEV problem (1), with the SNLLS-based criteria giving rise to an iterative sequence of GEV problems. **Table 1** summarizes the encountered matrix pairs $(\mathbf{B},\mathbf{A})$ (for real-valued TEQ and PTEQ designs) and shows that the $\mathbf{A}$ matrices are closely related autocorrelation matrices of the RX signal y_l, while the $\mathbf{B}$ matrices are closely related, often low-rank, matrices determined by a crosscorrelation metric between the RX and TX signal y_l and x_l, respectively. Complex TEQs or PTEQs are obtained by omitting the $\Re$-operators in Table 1.

3. SIMULATIONS

Figure 1 shows bitrate performance plots for the considered equalizer designs with 32 taps (both real and complex TEQs and PTEQs are considered). The FD-SNLLS-based TEQ and IR-SNLLS-based BM-TEQ have been computed using the iterative Gauss-Newton algorithm suggested in Section 2.4. The bitrate is depicted for 8 downstream CSA loops with strong front-end filtering to separate up- and downstream transmission (see [4] for details). All simulations use the same synchronization delay Δ, which is determined by the first sample index of the channel impulse response window of $\nu+1$ samples with maximum energy. The noise in Figure 1a is a superposition of AWG noise at -140dBm/Hz, residual echo and near-end crosstalk from 24 ADSL disturbers. In Figure 1b, severe RFI (7 RFIs with carrier frequencies 540, 650, 680, 760, 790, 840 and 1080kHz; the first two RFIs have a power of -30dBm, the remaining five have a power of -50dBm) is added. RFI, especially ingress from AM radio stations, can be an important interferer in ADSL. It is

	$\mathbf{B}$	$\mathbf{A}$
CLLS-based TD-MMSE-TEQ	$\boldsymbol{\Sigma}^T_{L,\mathbf{xy}}\left(\boldsymbol{\Sigma}^2_{L,\mathbf{x}}\right)^{-1}\boldsymbol{\Sigma}_{L,\mathbf{xy}}$	$\boldsymbol{\Sigma}^2_{L,\mathbf{y}}$
CLLS-based FD-MMSE-TEQ	$\Re\left\{\boldsymbol{\Sigma}^H_{K,\tilde{\mathbf{x}}\tilde{\mathbf{Y}}}\left(\boldsymbol{\Sigma}^2_{\tilde{\mathbf{x}}}\right)^{-1}\boldsymbol{\Sigma}_{K,\tilde{\mathbf{x}}\tilde{\mathbf{Y}}}\right\} = \sum_{n\in\mathscr{S}_a}\sigma^{-2}_{K,n,\tilde{\mathbf{x}}}\Re\left\{\boldsymbol{\sigma}^H_{K,n,\tilde{x}\tilde{\mathbf{y}}}\boldsymbol{\sigma}_{K,n,\tilde{x}\tilde{\mathbf{y}}}\right\}$	$\Re\left\{\boldsymbol{\Sigma}^2_{K,\tilde{\mathbf{Y}}}\right\} = \sum_{n\in\mathscr{S}_a}\Re\left\{\boldsymbol{\Sigma}^2_{K,n,\tilde{\mathbf{y}}}\right\}$
SNLLS-based FD-MMSE-TEQ	$\Re\left\{\sum_{n\in\mathscr{S}_a}\tilde{d}^*_{K,n}\boldsymbol{\sigma}^H_{K,n,\tilde{x}\tilde{\mathbf{y}}}\right\}\times \Re\left\{\sum_{n\in\mathscr{S}_a}\tilde{d}_{K,n}\boldsymbol{\sigma}_{K,n,\tilde{x}\tilde{\mathbf{y}}}\right\}$	$\Re\left\{\sum_{n\in\mathscr{S}_a}\left\|\tilde{d}_{K,n}\right\|^2\boldsymbol{\Sigma}^2_{K,n,\tilde{\mathbf{y}}}\right\}$
IR-SNLLS-based BM-FD-MMSE-TEQ	$\Re\left\{\sum_{n\in\mathscr{S}_a}\check{\gamma}_{K,n,\boldsymbol{\theta}_{n,\text{prev}}}\tilde{d}^*_{K,n}\boldsymbol{\sigma}^H_{K,n,\tilde{x}\tilde{\mathbf{y}}}\right\}\times \Re\left\{\sum_{n\in\mathscr{S}_a}\check{\gamma}_{K,n,\boldsymbol{\theta}_{n,\text{prev}}}\tilde{d}_{K,n}\boldsymbol{\sigma}_{K,n,\tilde{x}\tilde{\mathbf{y}}}\right\}$	$\Re\left\{\sum_{n\in\mathscr{S}_a}\check{\gamma}_{K,n,\boldsymbol{\theta}_{n,\text{prev}}}\left\|\tilde{d}_{K,n}\right\|^2\boldsymbol{\Sigma}^2_{K,n,\tilde{\mathbf{y}}}\right\}$
LS-based MMSE-PTEQ	$\sigma^{-2}_{K,n,\tilde{\mathbf{x}}}\Re\left\{\boldsymbol{\sigma}^H_{K,n,\tilde{x}\tilde{\mathbf{y}}}\boldsymbol{\sigma}_{K,n,\tilde{x}\tilde{\mathbf{y}}}\right\}$	$\Re\left\{\boldsymbol{\Sigma}^2_{K,n,\tilde{\mathbf{y}}}\right\}$

Table 1: Real-valued TEQ/PTEQ designs as a GEV problem $\mathbf{Bw} = \lambda\mathbf{Aw}$. Complex equalizers are obtained by omitting $\Re$-operators.



(a) (b)

Figure 1: Bitrate performance of the considered TEQ and PTEQ designs for 8 CSA loops. From left to right: TD-MMSE-TEQ, real and complex CLLS-based FD-MMSE-TEQ, real and complex SNLLS-based FD-MMSE-TEQ, real and complex BM-TEQ, real and complex PTEQ. (a) Without RFI. (b) With RFI.

clear from Figure 1b that in this RFI case, the BM-TEQ and PTEQ can effectively mitigate RFI and outperform the suboptimal TEQ designs. The SNLLS-based FD-MMSE-TEQ consistently outperforms the CLLS-based FD-MMSE-TEQ and TD-MMSE-TEQ and closely approaches the BM-TEQ performance. The CLLS-based FD-MMSE-TEQ performs worse than the TD-MMSE-TEQ; apparently, it makes more sense to minimize the sum-square FEQ output energies than the sum-square FFT output energies.

REFERENCES

[1] J. S. Chow and J. M. Cioffi, "A cost-effective maximum likelihood receiver for multicarrier systems," *Proc. ICC*, 1992, vol. 2, pp. 948–952.

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

[3] G. Ysebaert, K. Van Acker, M. Moonen, and B. De Moor, "Constraints in channel shortening equalizer design for DMT-based systems,"*Signal Processing*, vol. 83, no. 3, pp. 641–648, 2003.

[4] K. Vanbleu, G. Ysebaert, G. Cuypers, M. Moonen, and K. Van Acker, "Bitrate maximizing time-domain equalizer design for DMT-based systems," Accepted for publication in *IEEE Trans. Comm.* Available at ftp://ftp.esat.kuleuven. ac.be/ pub/sista/vanbleu/reports/02-118.pdf.

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

[6] K. Vanbleu, G. Ysebaert, G. Cuypers, and M. Moonen, "On time-domain and frequency-domain MMSE-based TEQ designs for DMT transmission," Tech. Rep. ESAT-SISTA/TR 2004-02. Submitted, Jan. 2004. Available at ftp://ftp.esat.kuleuven. ac.be/pub/sista/vanbleu/reports/04-02.pdf.

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

[8] A. Bjorck, *Least-Squares Methods*, Elsevier, 1987.

[9] G. Golub and V. Pereyra, "Separable nonlinear least squares: the variable projection method and its applications," *Institute of Physics, Inverse Problems*, vol. 19, no. 2, pp. R1–R26, Apr. 2003, available at http://www.iop.org/ EJ/ article/ 0266-5611/19/2/201/ip32r1.pdf.

[10] R.H. Byrd and D.A. Payne, "Convergence of the iteratively reweighted least squares algorithm for robust regression," Tech. Rep. 313, The Johns Hopkins University, Baltimore, MD, June 1979.