Katholieke Universiteit Leuven

Departement Elektrotechniek

ESAT-SISTA/TR 2004-10a

On the relation between time-domain equalizers and per-tone equalizers for DMT-based systems [1]

Koen Vanbleu, Geert Ysebaert, Gert Cuypers and Marc Moonen [2]

In the proceeding of the IEEE Benelux Signal Processing Symposium, Hilvarenbeek, the Netherlands, 16-17 April 2004.

[1] This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/sista/vanbleu/reports/04-10a.pdf*

[2] K.U.Leuven, Dept. of Electrical Engineering (ESAT), Research group SISTA, 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*. 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.

ON THE RELATION BETWEEN TIME-DOMAIN EQUALIZERS AND PER-TONE EQUALIZERS FOR DMT-BASED SYSTEMS

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 MMSE-based time-domain equalizer (TEQ), bitrate maximizing TEQ and per-tone equalizer design for DMT transmission. It is shown that they can all be formulated either as a constrained linear or a separable nonlinear least-squares problem. These least-squares formulations are then turned into generalized eigenvalue problems that allow to establish a remarkably close relationship between these -at first sight very different- equalizer designs.

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.

At first sight, the MMSE-TEQ, BM-TEQ and PTEQ design cost functions are very different in nature. The objective of this paper is to establish a remarkable relationship between these cost functions. Each of the considered cost functions will be restated either as a constrained linear least-squares (CLLS) based cost function or as a separable nonlinear least-squares (SNLLS) based cost function. The close relations between all LS cost functions become apparent 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. $\mathcal{S}_a$ is the set of N_a active tones; n is a tone index; N is the (I)DFT size; $\mathcal{F}_{\mathcal{S}_a}$ is an $N_a \times N$ submatrix of the full DFT matrix $\mathcal{F}_N$ with only the N_a rows of the active tones $\mathcal{S}_a$; the n-th DFT row is $\mathcal{F}_n$. 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 $\mathcal{S}_a$ 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 $\boldsymbol{\sigma}_{K,n,\tilde{\mathbf{x}}\tilde{\mathbf{y}}} = \frac{1}{K}\sum_{k=1}^{K}\tilde{x}^*_{k,n}\tilde{\mathbf{y}}_{k,n}$, and the *true* (stochastic) correlation, e.g., $\mathbf{\Sigma}^2_{\mathbf{y}} = \mathcal{E}\left\{\mathbf{y}_l^T\mathbf{y}_l\right\}$ or $\boldsymbol{\sigma}_{n,\tilde{x}\tilde{\mathbf{y}}} = \mathcal{E}\left\{\tilde{x}^*_{k,n}\tilde{\mathbf{y}}_{k,n}\right\}$. Throughout the text we only define the stochastic correlations.

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.

2. MMSE-TEQ, PTEQ AND BM-TEQ: 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 $\mathbf{y}_l = \begin{bmatrix} y_l & \cdots & y_{l-T+1} \end{bmatrix}^T$ 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 $\mathbf{x}_l = \begin{bmatrix} x_l & \cdots & x_{l-\nu} \end{bmatrix}^T$:

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

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

$$\min_{\mathbf{w}} \mathbf{w}^T\Big[\mathbf{\Sigma}^2_{L,\mathbf{y}} - \mathbf{\Sigma}^T_{L,\mathbf{xy}}\big(\mathbf{\Sigma}^2_{L,\mathbf{x}}\big)^{-1}\mathbf{\Sigma}_{L,\mathbf{xy}}\Big]\mathbf{w} \text{ s.t. } \mathbf{w}^T\mathbf{\Sigma}^2_{L,\mathbf{y}}\mathbf{w} = 1 \tag{5}$$

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

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

matrix pair

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

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*. Although these new designs are meant as an intermediate step to relate the TD-MMSE-TEQ with the PTEQ and the BM-TEQ, they also perform reasonably well (see Section 3).

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

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

The Toeplitz matrix $\mathbf{Y}_k$ has size $N\times T$; its first column and row are given by $\begin{bmatrix} y_{k,0} & \cdots & y_{k,N-1}\end{bmatrix}^T$ and $\begin{bmatrix} y_{k,0} & \cdots & y_{k,-T+1}\end{bmatrix}$, 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 $\begin{bmatrix} x_{k,0} & \cdots & x_{k,N-1}\end{bmatrix}^T$ 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 $\mathbf{X}_{k,\mathrm{C}} = \mathcal{F}_N^H\tilde{\mathbf{X}}_{k,N,\mathrm{D}}\mathcal{F}_N$, 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:

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

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

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

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

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

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

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, $\tilde{\mathbf{Y}}_k=\mathcal{F}_{\mathcal{S}_a}\mathbf{Y}_k$ (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

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

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]:

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

where $\sigma^2_{n,\tilde{x}}=\mathcal{E}\left\{|\tilde{x}_{k,n}|^2\right\}$ is the variance of $\tilde{x}_{k,n}$ and where the denominator is the crosscorrelation $\mathcal{E}\left\{\tilde{x}^*_{k,n}\tilde{y}_{k,n,\mathbf{w}}\right\}$ 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 $\boldsymbol{\sigma}_{n,\tilde{x}\tilde{\mathbf{y}}}=\mathcal{E}\left\{\tilde{x}^*_{k,n}\tilde{\mathbf{y}}_{k,n}\right\}$ the $1\times T$ crosscorrelation vector of $\tilde{x}_{k,n}$ and the n-th sliding DFT output $\tilde{\mathbf{y}}_{k,n}=\mathcal{F}_n\mathbf{Y}_k$ (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:

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

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):

$$(\mathbf{B},\mathbf{A}) = \left(\Re\left\{\boldsymbol{\Sigma}^H_{K,\tilde{\mathbf{x}}\tilde{\mathbf{Y}}}\left(\boldsymbol{\Sigma}^2_{K,\tilde{\mathbf{x}}}\right)^{-1}\boldsymbol{\Sigma}_{K,\tilde{\mathbf{x}}\tilde{\mathbf{Y}}}\right\}, \Re\left\{\boldsymbol{\Sigma}^2_{K,\tilde{\mathbf{Y}}}\right\}\right) \quad (15)$$

$$= \left(\Re\left\{\sigma^{-2}_{K,n,\tilde{x}}\sum_{n\in\mathcal{S}_a}\boldsymbol{\sigma}^H_{K,n,\tilde{x}\tilde{\mathbf{y}}}\boldsymbol{\sigma}_{K,n,\tilde{x}\tilde{\mathbf{y}}}\right\}, \Re\left\{\sum_{n\in\mathcal{S}_a}\boldsymbol{\Sigma}^2_{K,n,\tilde{\mathbf{y}}}\right\}\right)$$

The N_a rows of $\boldsymbol{\Sigma}_{\tilde{\mathbf{x}}\tilde{\mathbf{Y}}}=\mathcal{E}\left\{\tilde{\mathbf{X}}^*_{k,\mathrm{D}}\tilde{\mathbf{Y}}_k\right\}$ are the above defined crosscorrelation vectors $\boldsymbol{\sigma}_{n,\tilde{x}\tilde{\mathbf{y}}}$; $\boldsymbol{\Sigma}^2_{\tilde{\mathbf{Y}}}=\sum_{n\in\mathcal{S}_a}\boldsymbol{\Sigma}^2_{n,\tilde{\mathbf{y}}}$ with $\boldsymbol{\Sigma}^2_{n,\tilde{\mathbf{y}}}=\mathcal{E}\left\{\tilde{\mathbf{y}}^H_{k,n}\tilde{\mathbf{y}}_{k,n}\right\}$ the autocorrelation matrix of the n-th sliding DFT output; $\boldsymbol{\Sigma}^2_{\tilde{\mathbf{x}}}=\mathcal{E}\left\{\tilde{\mathbf{x}}_k\tilde{\mathbf{x}}_k^H\right\}$ 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 $\sigma^2_{n,\tilde{x}}$; the $\Re$-operators ensure a real TEQ.

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:

$$\min_{\mathbf{w},\tilde{\mathbf{d}}} \underbrace{\mathcal{E}\left\{\left\|\mathrm{diag}(\tilde{\mathbf{d}})\tilde{\mathbf{Y}}_k\mathbf{w}-\tilde{\mathbf{x}}_k\right\|^2\right\}}_{\mathcal{E}\left\{\|\tilde{\mathbf{e}}_{k,\boldsymbol{\theta}}\|^2\right\}} \text{ with real } \mathbf{w} \quad (16)$$

where the FEQ output error vector $\tilde{\mathbf{e}}_{k,\boldsymbol{\theta}}$ depends on both the real TEQ and complex FEQ parameters, $\boldsymbol{\theta}=\begin{bmatrix}\mathbf{w}^H & \tilde{\mathbf{d}}^H\end{bmatrix}^H$. The UEC constraint of (12) has been dropped as the criterion (16) has no trivial solution anymore. This criterion corresponds to an SNLLS criterion [8, 9]:

$$\boxed{\min_{\mathbf{w},\tilde{\mathbf{d}}} \frac{1}{K}\sum_{k=1}^{K}\left\|\mathrm{diag}(\tilde{\mathbf{d}})\tilde{\mathbf{Y}}_k\mathbf{w}-\tilde{\mathbf{x}}_k\right\|^2 \text{ with real } \mathbf{w}} \quad (17)$$

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 (16) 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]:

$$\tilde{d}_n^{\text{MMSE}} = \frac{\mathbf{w}^T \boldsymbol{\sigma}_{n,\tilde{x}\tilde{\mathbf{y}}}^H}{\mathbf{w}^T \boldsymbol{\Sigma}_{n,\tilde{\mathbf{y}}}^2 \mathbf{w}} \tag{18}$$

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., $\mathcal{E}\left\{|\tilde{y}_{k,n,\mathbf{w}}|^2\right\} = \mathbf{w}^T \boldsymbol{\Sigma}_{n,\tilde{\mathbf{y}}}^2 \mathbf{w}$. It will be shown in Section 2.5 that the SNLLS problem (17) can be solved iteratively with a sequence of GEV problems (1).

2.4. LS-based MMSE PTEQ design

We recall for convenience that the **LS-based MMSE-PTEQ** $\mathbf{w}_n$ for tone n is the solution to the following linear LS problem [5]:

$$\boxed{\min_{\mathbf{w}_n} \frac{1}{K} \sum_{k=1}^{K} |\tilde{\mathbf{y}}_{k,n} \mathbf{w}_n - \tilde{x}_{k,n}|^2} \tag{19}$$

where the *complex* $\mathbf{w}_n$ combines the task of the FEQ and the TEQ. The solution to (19) is given by

$$\mathbf{w}_n = \left(\boldsymbol{\Sigma}_{K,n,\tilde{\mathbf{y}}}^2\right)^{-1} \boldsymbol{\sigma}_{K,n,\tilde{x}\tilde{\mathbf{y}}}^H \tag{20}$$

The MMSE-PTEQ is closely related to the FD-MMSE-TEQ criteria of Section 2.2 and 2.3 when only 1 tone is considered. The SNLLS-based FD-MMSE-TEQ criterion (16) for a single tone n can be shown to reduce to the following nonlinear problem in $\mathbf{w}_n$ by eliminating the linearly appearing biased MMSE-based FEQ $\tilde{d}_n^{\text{MMSE}}$ (18):

$$\max_{\mathbf{w}_n} \frac{\sigma_{n,\tilde{\mathbf{x}}}^{-2} \mathbf{w}_n^T \Re\left\{\boldsymbol{\sigma}_{n,\tilde{x}\tilde{\mathbf{y}}}^H \boldsymbol{\sigma}_{n,\tilde{x}\tilde{\mathbf{y}}}\right\} \mathbf{w}_n}{\mathbf{w}_n^T \Re\left\{\boldsymbol{\Sigma}_{K,n,\tilde{\mathbf{y}}}^2\right\} \mathbf{w}_n} \tag{21}$$

The solution to the deterministic version of (21) is the dominant GEV of (1) with

$$(\mathbf{B}, \mathbf{A}) = \left(\sigma_{K,n,\tilde{\mathbf{x}}}^{-2} \Re\left\{\boldsymbol{\sigma}_{K,n,\tilde{x}\tilde{\mathbf{y}}}^H \boldsymbol{\sigma}_{K,n,\tilde{x}\tilde{\mathbf{y}}}\right\}, \Re\left\{\boldsymbol{\Sigma}_{K,n,\tilde{\mathbf{y}}}^2\right\}\right) \tag{22}$$

where the $\Re$-operators give rise to a **real PTEQ** $\mathbf{w}_n$. For a **complex PTEQ** $\mathbf{w}_n$, the $\Re$-operators should be dropped and the matrix $\mathbf{B}$ is rank 1 (for a real TEQ, its rank is 2). Then, there is only one strictly positive GEV (the other GEVs are zero), and the dominant generalized eigenvector $\mathbf{w}_n$ can be shown to indeed have the closed form PTEQ solution (20) (up to a scaling), meaning that the (complex valued) solution for (17) for a single tone is the PTEQ (20) (up to a scaling). It also follows straightforwardly from (15) that criterion (12) for a single tone is equivalent to the same GEV problem (22). The *real or complex* PTEQ must be multiplied with an LS-based uMMSE or MMSE FEQ, (13) or (18). It follows from (18) that the biased MMSE FEQ for a *complex* PTEQ is equal to one, meaning that a (complex) PTEQ effectively incorporates a biased MMSE FEQ.

2.5. Bitrate maximizing FD-MMSE-TEQ design

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

$$\max_{\boldsymbol{\theta}} \sum_{n \in \mathcal{S}_a} \log_2\left(1 + \frac{\text{SNR}_{n,\boldsymbol{\theta}_n}}{\Gamma_n}\right) \tag{23}$$

$$\text{with } \text{SNR}_{n,\boldsymbol{\theta}_n} = \frac{\sigma_{n,\tilde{x}}^2}{\mathcal{E}\left\{|\check{e}_{k,n,\boldsymbol{\theta}_n}|^2\right\}} = \frac{\sigma_{n,\tilde{x}}^2}{\mathcal{E}\left\{|\tilde{d}_n \tilde{\mathbf{y}}_{k,n} \mathbf{w} - \tilde{x}_{k,n}|^2\right\}} \tag{24}$$

$$\text{subject to } \tilde{d}_n = \frac{\sigma_{n,\tilde{x}}^2}{\boldsymbol{\Sigma}_{n,\tilde{x}\tilde{\mathbf{y}}} \mathbf{w}}, \ \forall n \in \mathcal{S}_a \tag{25}$$

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 (25) (see also (13)), which render the subchannel SNR model in (24) exact [4]. It has been shown in [6], based on (23-25), 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]:

$$\boxed{\min_{\boldsymbol{\theta}} \frac{1}{K} \sum_{k=1}^{K} \left\| \text{diag}\left(\sqrt{\check{\boldsymbol{\gamma}}_{K,\boldsymbol{\theta}_{\text{prev}}}}\right) \mathbf{e}_{k,\boldsymbol{\theta}} \right\|^2} \tag{26}$$

with

$$\check{e}_{k,n,\boldsymbol{\theta}_n} = \tilde{d}_n \tilde{\mathbf{y}}_{k,n} \mathbf{w} - \tilde{x}_{k,n} \tag{27}$$

$$\check{\gamma}_{K,n,\boldsymbol{\theta}_n} = \frac{\left(\text{SNR}_{K,n,\boldsymbol{\theta}_n} + 1\right)^2}{\sigma_{K,n,\tilde{x}}^2 \left(\text{SNR}_{K,n,\boldsymbol{\theta}_n} + \Gamma_n\right)} \tag{28}$$

$$\text{SNR}_{K,n,\boldsymbol{\theta}_n} = \frac{\sigma_{K,n,\tilde{x}}^2}{\frac{1}{K}\sum_{k=1}^{K} \|\check{e}_{k,n,\boldsymbol{\theta}_n}\|^2} = \frac{1}{\rho_{K,n,\boldsymbol{\theta}_n}^{-2} - 1} \tag{29}$$

$$\rho_{K,n,\boldsymbol{\theta}_n}^2 = \frac{|\boldsymbol{\sigma}_{K,n,\tilde{x}\tilde{\mathbf{y}}} \mathbf{w}|^2}{\sigma_{K,n,\tilde{x}}^2 \left(\mathbf{w}^T \boldsymbol{\Sigma}_{K,n,\tilde{\mathbf{y}}}^2 \mathbf{w}\right)} \tag{30}$$

The SNLLS-based FD-MMSE-TEQ (16) is indeed an unweighted version of, hence closely related to the IR-SNLLS-based BM-FD-MMSE-TEQ (26).

IR-LS problems such as (26) are weighted LS problems where the weights $\check{\boldsymbol{\gamma}}_{K,\boldsymbol{\theta}_{\text{prev}}}$ depend on the LS errors $\mathbf{e}_{k,\boldsymbol{\theta}}$ (here: via the subchannel SNRs (29)), 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, $\boldsymbol{\theta}_{\text{prev}}$. 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 (16) and (26), 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, $\check{\boldsymbol{\gamma}}_{K,\boldsymbol{\theta}_{\text{prev}}}$, (2) estimates of the biased MMSE FEQs (18), $\tilde{\mathbf{d}}_K$, which are the solutions of (26) for a fixed $\mathbf{w}_{\text{prev}}$ and (3) a new BM-TEQ estimate $\mathbf{w}$:

$$\mathbf{w} = \underbrace{\Re\left\{\sum_{n \in \mathcal{S}_a} \check{\gamma}_n \left|\tilde{d}_n\right|^2 \boldsymbol{\Sigma}_{K,n,\tilde{\mathbf{y}}}^2\right\}^{-1}}_{\left(\boldsymbol{\Sigma}_{K,\tilde{\mathbf{y}},\gamma}^2\right)^{-1}} \underbrace{\Re\left\{\sum_{n \in \mathcal{S}_a} \check{\gamma}_n \tilde{d}_n \boldsymbol{\sigma}_{K,n,\tilde{x}\tilde{\mathbf{y}}}^H\right\}}_{\boldsymbol{\sigma}_{K,\tilde{x}\tilde{\mathbf{y}},\gamma}} \tag{31}$$

with $\check{\gamma}_n = \check{\gamma}_{K,n,\boldsymbol{\theta}_{n,\text{prev}}}$ and $\tilde{d}_n = \tilde{d}_{K,n}$, which is (similar to the PTEQ solution (20)) the solution of a GEV problem with rank-one matrix $\mathbf{B}$:

$$(\mathbf{B}, \mathbf{A}) = \left(\boldsymbol{\sigma}_{K,\tilde{x}\tilde{\mathbf{y}},\gamma} \boldsymbol{\sigma}_{K,\tilde{x}\tilde{\mathbf{y}},\gamma}^H, \boldsymbol{\Sigma}_{K,\tilde{\mathbf{y}},\gamma}^2\right) \tag{32}$$

For a complex TEQ, the $\Re$-operators must be omitted. The iterations for solving the SNLLS-based FD-MMSE-TEQ (16) do not include the first step, i.e., the weights $\check{\gamma}_{K,n,\boldsymbol{\theta}_{n,\text{prev}}}$ 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 $\boldsymbol{\theta}$.

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

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

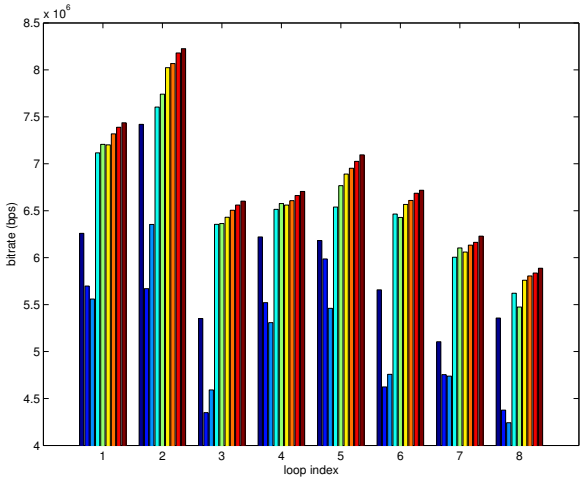


Figure 1: Bitrate performance for 8 CSA loops with RFI. From left to right: TD-MMSE-TEQ, real/complex CLLS-based FD-MMSE-TEQ, real/complex SNLLS-based FD-MMSE-TEQ, real/complex SNLLS-based BM-FD-MMSE-TEQ, real/complex MMSE-PTEQ.

2.6. 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). The bitrate is depicted for 8 downstream CSA loops with strong front-end filtering to separate up- and downstream transmission. The same synchronization delay Δ is used per loop and optimized to capture the main energy window of $\nu+1$ samples of the channel impulse response. The noise is a superposition of AWG noise at -140dBm/Hz, residual echo and near-end crosstalk from 24 ADSL disturbers and 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). See [4] for further details. The BM-FD-MMSE-TEQ and MMSE-PTEQ can effectively mitigate RFI and outperform the suboptimal MMSE-TEQ designs. The SNLLS-based FD-MMSE-TEQ consistently outperforms the CLLS-based TD-MMSE-TEQ and FD-MMSE-TEQ and closely approaches the BM-FD-MMSE-TEQ performance.

4. 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 design 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.

[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.