Per tone echo cancellation for DMT-based systems[1]

Katleen Van Acker[2], Marc Moonen[2], Thierry Pollet[3]

October 2000
Submitted to IEEE Transactions on Communications

[1]This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/SISTA/vanacker/reports/00-119.ps.gz*

[2]K.U.Leuven, Dept. of Electrical Engineering (ESAT), Research group SISTA, Kardinaal Mercierlaan 94, 3001 Leuven, Belgium, Tel. 32/16/32 19 27, Fax 32/16/32 19 70, WWW: *http://www.esat.kuleuven.ac.be/sista*. E-mail: *katleen.vanacker@esat.kuleuven.ac.be*. 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 - IUAP P4-02 (1997-2001): Modeling, Identification, Simulation and Control of Complex Systems, the Concerted Research Action GOA-MEFISTO-666 (Mathematical Engineering for Information and Communication Systems Technology) of the Flemish Government and the IT-projects in the ITA-bis program of the Flemish Institute for Scientific and Technological Research in Industry (I.W.T.) ('IRMUT: Integration of Reconfigurable Multimedia terminals' (980271) and 'Advanced Internet Access'(980316)) and was partially funded by Alcatel. The scientific responsibility is assumed by its authors.

[3]Access to Networks, Corporate Research Centre, ALCATEL, 2018 Antwerpen, Belgium, E-mail: *Thierry.Pollet@alcatel.be*

Per tone echo cancellation for DMT-based systems

Katleen Van Acker[1], Marc Moonen[1], Thierry Pollet[2]

[1] Katholieke Universiteit Leuven

ESAT/SISTA

Kasteelpark Arenberg 10, 3001 Leuven - Belgium

tel 32/16/32 19 27, fax 32/16/32 19 70

katleen.vanacker@esat.kuleuven.ac.be

marc.moonen@esat.kuleuven.ac.be

[2] Access to Networks

Corporate Research Center

ALCATEL

2018 Antwerpen, Belgium

Thierry.Pollet@alcatel.be

Abstract

A new echo cancellation structure for discrete multitone (DMT) systems is presented where each used tone has its own per tone echo canceller in addition to a per tone equalizer [8] [9] which provides an alternative to current employed time domain and time/frequency domain approaches. The per tone approach enables us to optimize the Signal-to-Noise Ratio (SNR) for each tone separately by solving a Minimum Mean Square Error (MMSE) problem for each tone, with implicit so-called joint shortening. Complexity during data transmission is compared for time domain echo cancellation and per tone echo cancellation. Structures with reduced complexity are derived for an interpolated and a decimated rate set-up. Finally, simulation results for an ADSL setting demonstrate improved performance over time domain (or time/frequency domain) echo cancellation.

Keywords

Discrete Multitone, Echo Cancellation

I. Introduction

DMT-modulation [3] has become an important transmission method, for instance for asymmetric digital subscriber line (ADSL). DMT divides the available bandwidth into a large number of bands. The information sequence is split into parallel bitstreams that are used to QAM-modulate a carrier in each band. After modulation, which is implemented by means of an IFFT, a cyclic prefix is added to each symbol before transmission. If this cyclic prefix is longer than the channel impulse response, reception can be carried out by first removing the portion corresponding to the prefix, then demodulating by an FFT and finally equalizing the channel by a 1-taps frequency domain equalizer for each tone. A long prefix however results in a large overhead. Therefore, usually a T-taps *time domain equalizer (TEQ)* is added and applied to the received signal, such that the combined effect of channel and TEQ is sufficiently short [2] [6] [10].

In [8] [9] an alternative receiver structure, based on *'per tone equalization'* (PT-EQ), has been derived, where each tone then has its own (complex) T-taps equalizer. With a comparable complexity during data transmission, it becomes possible to optimize the SNR for each tone separately while at the same time the sensitivity to the decision delay is decreased compared to the TEQ-approach. In this paper, we will focus on the echo cancellation problem for full duplex or frequency division duplex systems, and develop a similar 'per tone' approach.

In [1] so-called 'joint shortening' of the far end channel impulse response and the echo impulse response by a TEQ was introduced. By first shortening the echo impulse response, one can use shorter echo cancellers. However the optimization criterion based on channel shortening has no direct relation to the resulting SNR of the system. An echo canceller which is implemented partly in the time domain and partly in the frequency domain was developed in [4] [5]. The advantages of this structure are an efficient updating scheme consisting of a 1-tap LMS algorithm for each tone separately and an efficient echo emulation.

In this paper, *'per tone echo cancellation'* (PT-EC) is introduced. The time domain echo canceller (TEC) is moved to the frequency domain to result in a (complex) echo canceller for each tone separately. By minimizing an MMSE criterion, the SNR is optimized in a per tone fashion, in fact with an implicit joint shortening. This approach is shown to give improved performance, compared to time domain (or time/frequency domain) EC in an ADSL setting.

The paper is organized as follows. Section II describes the data model. PT-EC is introduced in section III. Section IV explains joint initialization of the PT-EQ and the PT-EC by means of an MMSE cost function per tone. In sections V and VI efficient schemes are derived for PT-EC in an interpolated and a decimated rate set-up. In section VII complexity computations are given. Simulation results are presented in section VIII. Finally conclusions are drawn in section IX.

II. Data model

The following notation is adopted in the description of the DMT-system, analogous to [8] [9]. N is the symbol size of the far end signal expressed in samples, k is the time index of a symbol, $X_i^{(k)}$ is a complex subsymbol for tone i $(i = 1 \dots N)$ of the far end signal transmitted at symbol period k. Note that $X_i^{(k)} = X_{N-(i-2)}^{*(k)} \quad i = 2 \dots \frac{N}{2}$. Further, ν denotes the length of the cyclic prefix of the far end signal, $s = N + \nu$ the length of a far end symbol including prefix, $\bar{\mathbf{h}} = [h_L \dots h_0 \dots h_{-K}]$ the far end channel impulse response in reverse order, n_l additive channel noise and y_l the received signal with l the sample index.

To describe the data model, we consider three successive symbols[1] $X_{1:N}^{(c)}$ transmitted at time $c = k-1, k, k+1$ respectively. The kth symbol is the symbol of interest, the previous and the next symbol are used to include interferences with neighboring symbols in our model.

Whenever echo is present, one can model the echo in a similar way as the far end signal. $U_i^{(k)}$ is a complex subsymbol for tone i $(i = 1 \dots N)$ of the echo signal transmitted at symbol period k. $\bar{\mathbf{h}}_E$ is the echo channel impulse response in reverse order. For the sake of a compact notation, we assume a symmetric rate set-up, i.e. the same symbol and prefix size for echo and far end signal. Extensions for asymmetric rate set-ups are straightforward. The received signal then becomes

[1] $X_{1:N}^{(c)}$ denotes vector $[X_1^{(c)} \dots X_N^{(c)}]^{\mathrm{T}}$.

$$\overbrace{\begin{bmatrix} y_{k\cdot s+\nu-T+2+\delta_1} \\ \vdots \\ y_{(k+1)\cdot s+\delta_1} \end{bmatrix}}^{\mathbf{y}} = \left[\begin{array}{c|ccc|c} \mathbf{O}_{(1)} & \boxed{\bar{\mathbf{h}}} & 0 & \cdots & \mathbf{O}_{(2)} \\ & \ddots & \ddots & & \\ & 0 & \cdots & \boxed{\bar{\mathbf{h}}} & \end{array}\right] \cdot \hat{\mathbf{P}} \cdot \hat{\mathcal{I}}_N \cdot \overbrace{\begin{bmatrix} X_{1:N}^{(k-1)} \\ X_{1:N}^{(k)} \\ X_{1:N}^{(k+1)} \end{bmatrix}}^{\hat{\mathbf{x}}} +$$

$$\left[\begin{array}{c|ccc|c} \mathbf{O}_{(3)} & \boxed{\bar{\mathbf{h}}_E} & 0 & \cdots & \mathbf{O}_{(4)} \\ & \ddots & \ddots & & \\ & 0 & \cdots & \boxed{\bar{\mathbf{h}}_E} & \end{array}\right] \cdot \hat{\mathbf{P}} \cdot \hat{\mathcal{I}}_N \cdot \overbrace{\begin{bmatrix} U_{1:N}^{(k-1)} \\ U_{1:N}^{(k)} \\ U_{1:N}^{(k+1)} \end{bmatrix}}^{\hat{\mathbf{u}}} + \overbrace{\begin{bmatrix} n_{k\cdot s+\nu-T+2+\delta_1} \\ \vdots \\ n_{(k+1)\cdot s+\delta_1} \end{bmatrix}}^{\mathbf{n}}$$

$$= \mathbf{H} \cdot \hat{\mathbf{x}} + \mathbf{H}_E \cdot \hat{\mathbf{u}} + \mathbf{n} \tag{1}$$

Here, T is the length of the equalization filters. $\mathbf{O}_{(1)}$ and $\mathbf{O}_{(2)}$ are zero matrices of size $(N{+}T{-}1)\times(N{+}\nu{-}T{+}1{-}L{+}\nu{+}\delta_1)$ and $(N{+}T{-}1)\times(N{+}\nu{-}K{-}\delta_1)$ respectively. Matrix $\hat{\mathbf{P}} = \text{diag}(\mathbf{P}, \mathbf{P}, \mathbf{P})$ is a block diagonal matrix where $\mathbf{P} = \left[\begin{array}{c|c} \mathbf{O} & \mathbf{I}_\nu \\ \hline \multicolumn{2}{c}{\mathbf{I}_N} \end{array}\right]$ adds in the cyclic prefix. Matrix $\hat{\mathcal{I}}_N = \text{diag}(\mathcal{I}_N, \mathcal{I}_N, \mathcal{I}_N)$ is a block diagonal matrix where $\mathcal{I}_N$ is an $N \times N$ IDFT matrix which modulates the input symbols. The zero reference delay of the far end signal corresponds to the head $[h_{-K} \dots h_{-1}]$ and the tail $[h_{\nu+1} \dots h_L]$ that maximize the energy in $[h_0 \dots h_\nu]$. Finally, δ_1 is the (relative) decision delay of the far end signal. Analogous definitions hold for the corresponding echo parameters $\mathbf{O}_{(3)}$, $\mathbf{O}_{(4)}$, $\bar{\mathbf{h}}_E$, L_E, K_E and δ_2.

In case of a 2-fold oversampled receiver, one has a vector $\mathbf{y}$ consisting of even and odd samples. This vector can be split into a vector of 'even numbered' samples $\mathbf{y}_e$ and a vector of 'odd numbered' samples $\mathbf{y}_o$, both of which may be specified by means of a formula of the form of (1). The corresponding impulse responses are $\mathbf{h}_e$, $\mathbf{h}_{E,e}$ and $\mathbf{h}_o$, $\mathbf{h}_{E,o}$ and the noise vectors are $\mathbf{n}_e$ and

$\mathbf{n}_o$.

The input to the echo canceller (echo reference signal) is modeled by

$$\overbrace{\begin{bmatrix} u_{k\cdot s+\nu-T_E+2+\delta_3} \\ \vdots \\ u_{(k+1)\cdot s+\delta_3} \end{bmatrix}}^{\mathbf{u}} = \left[\mathbf{O}_{(5)} \middle| \mathbf{I} \middle| \mathbf{O}_{(6)} \right] \cdot \hat{\mathbf{P}} \cdot \hat{\mathcal{I}}_N \cdot \hat{\mathbf{u}} \tag{2}$$
$$= \mathbf{H}_r \cdot \hat{\mathbf{u}}$$

with T_E the length of the echo cancellation filter, $\mathbf{O}_{(5)}$ and $\mathbf{O}_{(6)}$ zero matrices of size $(N+T_E-1) \times (N+\nu-T_E+1+\nu+\delta_3)$ and $(N+T_E-1) \times (N+\nu-\delta_3)$ respectively. There are three delays in the data model: δ_1 for the far end signal, δ_2 for the echo signal and δ_3 for the reference signal. The delays δ_2 and δ_3 are related to each other.

III. Combination of EQ and EC in a DMT receiver

In [8] [9] it was demonstrated how a TEQ can be moved to the frequency domain to result in a PT-EQ for each tone separately. The advantage of the latter structure is that optimizing the SNR for each tone separately results in an MMSE problem per tone. The same idea can also be applied to a TEC. By moving the TEC to the frequency domain, one obtains an EC for each tone separately which can be initialized by solving an MMSE problem.

The initial transceiver with TEQ and TEC is shown in **Fig. 1**. For each tone i, it is based on the following operation

$$Z_i^{(k)} = D_i \cdot \text{row}_i \overbrace{(\mathcal{F}_N) \cdot (\mathbf{Y} \cdot \mathbf{w} - \mathbf{U} \cdot \mathbf{w}_E)}^{1 \text{ FFT}} \tag{3}$$

where D_i is the (complex) 1-taps FEQ for tone i, $\mathcal{F}_N$ is an $N \times N$ DFT matrix, $\mathbf{w} = [w_0\, w_1 \ldots w_{T-1}]^{\mathrm{T}}$ is the (real) T-taps TEQ, and $\mathbf{Y}$ an $N \times T$ Toeplitz matrix of received samples. Similarly, $\mathbf{w}_E$ is the (real) T_E-taps TEC and $\mathbf{U}$ an $N \times T_E$ Toeplitz matrix of samples of the echo reference signal. Note that $\mathbf{Y}$ and $\mathbf{U}$ contain the same samples as vector $\mathbf{y}$ in formula (1) and $\mathbf{u}$ in formula (2) respectively, i.e.

$$\mathbf{Y} = \begin{bmatrix} y_{k\cdot s+\nu+1+\delta_1} & \cdots & y_{k\cdot s+\nu-T+2+\delta_1} \\ \vdots & \ddots & \vdots \\ y_{(k+1)\cdot s+\delta_1} & \cdots & y_{(k+1)\cdot s-T+1+\delta_1} \end{bmatrix} \text{ and } \mathbf{U} = \begin{bmatrix} u_{k\cdot s+\nu+1+\delta_3} & \cdots & u_{k\cdot s+\nu-T_E+2+\delta_3} \\ \vdots & \ddots & \vdots \\ u_{(k+1)\cdot s+\delta_3} & \cdots & u_{(k+1)\cdot s-T_E+1+\delta_3} \end{bmatrix} \quad (4)$$

Formula (3) may be rewritten as follows

$$\begin{aligned} Z_i^{(k)} &= \mathrm{row}_i\,(\mathcal{F}_N) \cdot (\mathbf{Y} \cdot \mathbf{w} \cdot D_i - \mathbf{U} \cdot \mathbf{w}_E \cdot D_i) \\ &= \mathrm{row}_i \underbrace{(\mathcal{F}_N \cdot \mathbf{Y})}_{T \text{ FFT's}} \cdot \mathbf{w}_i - \mathrm{row}_i \underbrace{(\mathcal{F}_N \cdot \mathbf{U})}_{T_E \text{ FFT's}} \cdot \mathbf{w}_{E,i} \end{aligned} \quad (5)$$

By putting D_i to the right one obtains a complex T-taps PT-EQ $\mathbf{w}_i$ and a complex T_E-taps PT-EC $\mathbf{w}_{E,i}$ for each tone separately. In principle, we the aim at designing a $\mathbf{w}_i$ and a $\mathbf{w}_{E,i}$ for each tone separately, which will provide additional flexibility and improved performance compared to the time domain approach. The inputs of these filters are seen to come from T and T_E successive FFT's respectively. This however is found to result in a complex implementation and calls for complexity reduction. As demonstrated in [9], T respectively T_E successive FFT's can be calculated efficiently by means of a sliding FFT. Then only one 'full' FFT is actually computed and the other FFT outputs are derived as a complex linear combination of this one FFT and $T-1$ respectively T_E-1 (real) difference terms of the form $(y_{l+N} - y_l)$. The final modem set-up has one FFT operation for the received signal and one FFT operation for the reference echo signal together with equalization/echo filters directly acting upon difference terms. The FFT's have as their inputs the elements in the first column of $\mathbf{Y}$ and $\mathbf{U}$ respectively. For further details we refer to [9].

In **Fig. 2** a per tone filter (PTF) block is defined, including an N-point FFT operation and T-taps per tone filters $\mathbf{v}_i$ for all used tones. The notation $\mathbf{v}_i$ instead of $\mathbf{w}_i$ is adopted for the PTF's because we work here with the efficient implementation of these filters, i.e. their inputs are one FFT-output and $T-1$ difference terms. A complete DMT-receiver with PT-EQ and PT-EC is then shown in **Fig. 3**.

Remark that, if the modem works frame-synchronously, the only FFT needed in the PT-EC, reproduces the transmitted echo symbol $U_{1:N}^{(k)} = \mathcal{F}_N \cdot \mathbf{U}(:, \mathbf{1})$. Hence, in that case, the FFT in the echo branch becomes superfluous. Furthermore, in frame-synchronous mode, the first $\nu+1$ columns of $\mathbf{U}$ are equal up to a rotation, so the first ν difference terms are zero.

We will also study the performance of two-fold oversampled PT-EQ combined with PT-EC. If the received signal is sampled at twice the original sample rate, then the TEQ can be rewritten by means of its polyphase components which results in an even (e) and odd (o) TEQ at the sample rate. When transforming such a receiver structure to a per tone structure, one obtains the receiver shown in **Fig. 4**.

IV. Joint initialization of PT-EQ and PT-EC

In the sequel, we immediately treat the oversampled PT-EQ, which is more general. In case of sampling at the original sample rate, all 'odd' parts in the following formulas can be set to zero.

As in [8] [9], one finds the PTF's by solving an MMSE problem for each tone separately. The MMSE problem for joint initialization of the PT-EQ and the PT-EC, gives rise to an extended version of the MMSE problem for initialization of the PT-EQ only. For each tone i, one minimizes the following cost function:

$$J(\mathbf{v}_{i,e}, \mathbf{v}_{i,o}, \mathbf{v}_{E,i}) = \mathcal{E}\left\{ \left| \begin{bmatrix} \bar{\mathbf{v}}_{i,e}^{\mathrm{T}} & \bar{\mathbf{v}}_{i,o}^{\mathrm{T}} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{F}_{i,T} & \mathbf{O} \\ \mathbf{O} & \mathbf{F}_{i,T} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{y}_e \\ \mathbf{y}_o \end{bmatrix} - \underbrace{\bar{\mathbf{v}}_{E,i}^{\mathrm{T}} \cdot \mathbf{F}_{i,T_E} \cdot \mathbf{u}}_{\text{echo cancellation}} - X_i^{(k)} \right|^2 \right\} \tag{6}$$

with $\mathbf{F}_{i,T}$ equal to $\left[\begin{array}{c|c|c} \mathbf{I}_{T-1} & \mathbf{O} & -\mathbf{I}_{T-1} \\ \hline \mathbf{O} & \multicolumn{2}{c}{\mathcal{F}_N(i,:)} \end{array}\right]$, where the first block row extracts the difference terms and the bottom row corresponds to the FFT-output for tone i. Vectors $\mathbf{v}_{i,e}$ and $\mathbf{v}_{i,o}$ are respectively the even and the odd PT-EQ for tone i. $\mathbf{v}_{E,i}$ is the PT-EC for tone i. Formula (6) can be rewritten by including formulas (1) and (2), resulting in

$$\begin{aligned} & J(\mathbf{v}_{i,e}, \mathbf{v}_{i,o}, \mathbf{v}_{E,i}) \\ & = \mathcal{E}\left\{ \left| \begin{bmatrix} \bar{\mathbf{v}}_{i,e}^{\mathrm{T}} & \bar{\mathbf{v}}_{i,o}^{\mathrm{T}} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{F}_{i,T} & \mathbf{O} \\ \mathbf{O} & \mathbf{F}_{i,T} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{H}_e\hat{\mathbf{x}} + \mathbf{H}_{E,e}\hat{\mathbf{u}} + \mathbf{n}_e \\ \mathbf{H}_o\hat{\mathbf{x}} + \mathbf{H}_{E,o}\hat{\mathbf{u}} + \mathbf{n}_o \end{bmatrix} - \bar{\mathbf{v}}_{E,i}^{\mathrm{T}} \mathbf{F}_{i,T_E}(\mathbf{H}_r\hat{\mathbf{u}}) - \mathbf{e}_i^{(k)}\hat{\mathbf{x}} \right|^2 \right\} \\ & = \left\| \left[\begin{array}{cc|c} \mathbf{R}_{\hat{\mathbf{x}}}^{1/2}\mathbf{H}_e^{\mathrm{H}}\mathbf{F}_{i,T}^{\mathrm{H}} & \mathbf{R}_{\hat{\mathbf{x}}}^{1/2}\mathbf{H}_o^{\mathrm{H}}\mathbf{F}_{i,T}^{\mathrm{H}} & \mathbf{O} \\ \mathbf{R}_{\hat{\mathbf{u}}}^{1/2}\mathbf{H}_{E,e}^{\mathrm{H}}\mathbf{F}_{i,T}^{\mathrm{H}} & \mathbf{R}_{\hat{\mathbf{u}}}^{1/2}\mathbf{H}_{E,o}^{\mathrm{H}}\mathbf{F}_{i,T}^{\mathrm{H}} & -\mathbf{R}_{\hat{\mathbf{u}}}^{1/2}\mathbf{H}_r^{\mathrm{H}}\mathbf{F}_{i,T_E}^{\mathrm{H}} \\ \multicolumn{2}{c|}{\mathbf{R}_{\mathbf{n}}^{1/2}\begin{bmatrix} \mathbf{F}_{i,T}^{\mathrm{H}} & \mathbf{O} \\ \mathbf{O} & \mathbf{F}_{i,T}^{\mathrm{H}} \end{bmatrix}} & \mathbf{O} \end{array}\right] \cdot \begin{bmatrix} \bar{\mathbf{v}}_{i,e}^{*} \\ \bar{\mathbf{v}}_{i,o}^{*} \\ \bar{\mathbf{v}}_{E,i}^{*} \end{bmatrix} - \begin{bmatrix} \mathbf{R}_{\hat{\mathbf{x}}}^{1/2}\mathbf{e}_i^{(k)\mathrm{H}} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \right\|_2^2 \end{aligned} \tag{7}$$

with $\mathbf{e}_i^{(k)} = [0 \ldots 0\,1\,0 \ldots 0]$ a $1 \times 3N$ vector with the non-zero entry in the $(N+i)$th column. Matrices $\mathbf{R}_{\hat{\mathbf{x}}} = \mathcal{E}\{\hat{\mathbf{x}}\hat{\mathbf{x}}^{\mathrm{H}}\}$ and $\mathbf{R}_{\hat{\mathbf{u}}} = \mathcal{E}\{\hat{\mathbf{u}}\hat{\mathbf{u}}^{\mathrm{H}}\}$ are the autocorrelation matrices of vectors $\hat{\mathbf{x}}$ and $\hat{\mathbf{u}}$ respectively. Matrix $\mathbf{R}_{\mathbf{n}} = \mathcal{E}\{\mathbf{n}\mathbf{n}^{\mathrm{H}}\}$ is the autocorrelation of vector $\mathbf{n} = \begin{bmatrix} \mathbf{n}_e \\ \mathbf{n}_o \end{bmatrix}$.

This set of equations has a particular structure. The first block row is constructed with the far end impulse responses and the second block row with the echo impulse responses. The third block row consists of noise data. The first and second block column corresponds to the even and the odd PT-EQ respectively. The third block column corresponds to the PT-EC, so it consists of the echo reference signal.

Cost function (7) corresponds to MMSE-optimal equalizer and echo filter design, and hence incorporates optimal *per tone joint shortening* for the general case of a T-taps equalizer and a T_E-taps echo canceller. The per tone joint shortening improves upon the original time domain joint shortening [1], [6], [11] because it maximizes the SNR for each tone separately.

Cost function (6) forms the basis for practical recursive/adaptive training data based parameter estimation procedures, such as RLS, which enable the filters to converge to their optimal settings. In contrast with formula (7), such recursive/adaptive procedures do not need prior knowledge of the signal statistics and the channel impulse responses, see also [7].

V. Interpolated echo cancellation at RT

In the previous sections, it is assumed that the transmit IFFT and receive FFT both have size N. However, when the transmit IFFT is γ times smaller (as is typically the case in an ADSL remote terminal (RT)), one can show that there exist a more efficient implementation.

Assume an RT with $\frac{N}{\gamma}$-point IFFT transmission and N-point FFT reception. The echo cancellation on tone i at time k, $C_i^{(k)}$, corresponds to the second term in formula (5)

$$C_i^{(k)} = \text{row}_i \underbrace{(\mathcal{F}_N \cdot \mathbf{U})}_{T_E \text{ FFT's}} \cdot \mathbf{w}_{E,i} \quad \text{for} \quad i = 1 \ldots \frac{N}{2} + 1 \tag{8}$$

The $\frac{N}{\gamma}$-point IFFT is followed by a prefix insertion, a γ-fold upsampling and interpolation filtering. The filters are part of the echo path, hence the echo reference signal is the upsampled

IFFT output signal (with prefix) and so $\mathbf{U}$ is an $N \times T_E$ Toeplitz matrix with in each column $\gamma - 1$ zeros in between every 2 non-zero entries. This also means that γ successive FFT's, i.e. γ FFT's of successive columns of $\mathbf{U}$, are equal up to a phase shift.

Let us first assume an alignment of the echo symbols such that the FFT's of the first γ columns in matrix $\mathbf{U}$ are equal up to a phase shift (the same property holds for the following groups of γ columns). Hence, the FFT of all the columns of $\mathbf{U}$ can be obtained by computing the FFT of a selection of columns, namely $\mathbf{U}(:, 1 : \gamma : \text{end})$ and then multiplying these FFT's with the appropriate phase shifts, namely $\boldsymbol{\alpha} = \begin{bmatrix} 1 & \alpha^{i-1} & \dots & \alpha^{(i-1)(\gamma-1)} \end{bmatrix}$ with $\alpha = e^{-j2\pi\frac{1}{N}}$. If T_E is not a multiple of γ, then the last columns of $\mathbf{U}(:, 1 \! : \! \gamma \! : \! \text{end})$ will only be multiplied with a subvector of $\boldsymbol{\alpha}$. A T_E-taps EC can then be represented as shrunk into an equivalent $\lceil \frac{T_E}{\gamma} \rceil$-taps EC

$$C_i^{(k)} = \text{row}_i \underbrace{\left(\mathcal{F}_N \cdot \mathbf{U}(:, 1 \! : \! \gamma \! : \! \text{end}) \right)}_{\lceil \frac{T_E}{\gamma} \rceil \text{ FFT's}} \cdot \underbrace{\begin{bmatrix} \boldsymbol{\alpha} & \mathbf{0} & \dots \\ & \ddots & \\ \dots & \mathbf{0} & \boldsymbol{\alpha} \end{bmatrix} \cdot \mathbf{w}_{E,i}}_{\tilde{\mathbf{w}}_{E,i}} \tag{9}$$

$$\boldsymbol{\alpha} = \begin{bmatrix} 1 & \alpha^{i-1} & \dots & \alpha^{(i-1)(\gamma-1)} \end{bmatrix} \text{ with } \alpha = e^{-j2\pi\frac{1}{N}}$$

where $\tilde{\mathbf{w}}_{E,i}$ is a $\lceil \frac{T_E}{\gamma} \rceil$-taps PT-EC for tone i. Remark that, when T_E is not divisible by γ, the phase matrix has an additional diagonal element in the lower right corner, which is a subvector of $\boldsymbol{\alpha}$.

Second, if the alignment of the echo symbols is such that only the first $\acute{\gamma}$ columns of $\mathbf{U}$ have their FFT's equal up to a phase shift, with $\acute{\gamma} < \gamma$, a formula similar to (9) can be obtained. Depending on the alignment of the echo symbols, the echo filter $\tilde{\mathbf{w}}_{E,i}$ may have $\lceil \frac{T_E}{\gamma} \rceil + 1$ taps and the phase matrix may have additional diagonal elements in left upper and the right lower corner.

The N-point FFT's in formula (9), can be written as γ times repeated $\frac{N}{\gamma}$-point FFT's because of the zero-structure in each column of $\mathbf{U}$. For every tone i, the echo to be cancelled is then

$$\begin{aligned} C_i^{(k)} &= \mathrm{row}_i \left(\mathbf{R} \cdot \mathcal{F}_{N/\gamma} \cdot \overbrace{\mathbf{U}(1:\gamma:N, 1:\gamma:\mathrm{end})}^{\tilde{\mathbf{U}}} \right) \cdot \tilde{\mathbf{w}}_{E,i} \\ &= \mathrm{row}_{\mathrm{mod}(\frac{i-1}{N/\gamma})+1} \underbrace{\left(\mathcal{F}_{N/\gamma} \cdot \tilde{\mathbf{U}} \right)}_{\lceil \frac{T_E}{\gamma} \rceil \text{ FFT's}} \cdot \tilde{\mathbf{w}}_{E,i} \end{aligned} \tag{10}$$

with $\mathbf{R} = [\mathbf{I}_{N/\gamma} \ldots \mathbf{I}_{N/\gamma}]^{\mathrm{T}}$, $\tilde{\mathbf{U}}$ is an $\frac{N}{\gamma} \times \lceil \frac{T_E}{\gamma} \rceil$ Toeplitz matrix with the $\frac{N}{\gamma}$-point IFFT output samples (with prefix) and 'mod' the remainder after division. Note that, if $\acute{\gamma} < \gamma$, a similar formula can be obtained. A SFG of the interpolated (N, T, γ)-PTF is shown in **Fig. 5**, where the notation $\mathbf{v}_i$ instead of $\mathbf{w}_i$ is adopted because we work with the efficient implementation of these filters. Remark that a derivation of the interpolated PT-EC can also be given as a function of the filter $\mathbf{v}_i$.

Overall, it is seen that roughly a γ-fold complexity reduction is obtained in the EC part.

VI. Decimated echo cancellation at CO

In this secion we consider a transceiver with N-point IFFT and $\frac{N}{\gamma}$-point FFT as is typically the case in an ADSL central office and again derive an efficient implementation. Similar to equation (8), on may write the echo to be cancelled on tone i as

$$C_i^{(k)} = \mathrm{row}_i \underbrace{\left(\mathcal{F}_{N/\gamma} \cdot \mathbf{U}(1:\gamma:N,:) \right)}_{T_E \text{ FFT's}} \cdot \mathbf{w}_{E,i} \quad \text{for} \quad i = 1 \ldots \frac{N}{2\gamma} + 1 \tag{11}$$

with $\mathbf{U}$ an $N \times T_E$ Toeplitz matrix. The echo canceller has T_E taps at high rate but there are less tones now, which will result in a similar γ-fold complexity reduction as in the interpolated case.

To derive the cost effective PT implementation, we consider $\mathbf{U}(1\!:\!\gamma\!:\!N,:)$

$$\mathbf{U}(1\!:\!\gamma\!:\!N,:) = \begin{bmatrix} u_{\nu+1} & u_{\nu} & u_{\nu-1} & \cdots & u_{\nu-\gamma+1} & \cdots & u_{\nu-T_E+2} \\ u_{\nu+\gamma+1} & u_{\nu+\gamma} & u_{\nu+\gamma-1} & \cdots & u_{\nu+1} & \cdots & u_{\nu+\gamma-T_E+2} \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ u_{\nu+N-\gamma+1} & u_{\nu+N-\gamma} & u_{\nu+N-\gamma-1} & \cdots & u_{\nu+N-2\gamma+1} & \cdots & \end{bmatrix} \quad (12)$$

where u_l is the echo reference signal with l the sample index. The echo reference data model is given in formula (2). The terms $k\cdot s$ and δ_3 are omitted in the sample index for the sake of a compact notation. The first γ columns of $\mathbf{U}(1\!:\!\gamma\!:\!N,:)$ do not have any samples in common. The $(\gamma+1)$st column is obtained by downshifting the first column, adding a sample at the beginning and deleting one at the end. Or stated in other words, columns $1, \gamma+1, 2\gamma+1, \ldots$ form a Toeplitz matrix, i.e. $\mathbf{U}(1\!:\!\gamma\!:\!N, 1\!:\!\gamma\!:\!\text{end})$, the same holds for matrices $\mathbf{U}(1\!:\!\gamma\!:\!N, 2:\gamma:\text{end}), \ldots, \mathbf{U}(1:\gamma\!:\!N, \gamma:\gamma:\text{end})$. If T_E is a multiple of γ, these matrices have $\frac{T_E}{\gamma}$ columns. Otherwise the last $\text{mod}(\frac{T_E}{\gamma})$ columns of $\mathbf{U}(1\!:\!\gamma\!:\!N,:)$ are added to the appropriate Toeplitz matrices.

As a result, one may mathematically split the T_E-taps PTF in γ $\frac{T_E}{\gamma}$-taps PTF's: $\mathbf{w}_{E,i}(1:\gamma:\text{end}), \ldots, \mathbf{w}_{E,i}(\gamma:\gamma:\text{end})$ and add their outputs for each tone i. The latter filters each individually can be implemented cost effectively by means of difference terms. A SFG of the decimated $(N, T, \frac{1}{\gamma})$-PTF is shown in **Fig. 6**, where we adopt again the notation $\mathbf{v}_i$ instead of $\mathbf{w}_i$ because we work here with the efficient implemenation of these filters. The SFG consists of γ parallel $(\frac{N}{\gamma}, \frac{T}{\gamma}, 1)$-PTF's which correspond to the so-called polyphase components of $\mathbf{w}_{E,i}$. Roughly this again corresponds to a γ-fold complexity reduction.

VII. Complexity during data transmission

We compute the complexity during data transmission for the TEC and PT-EC by counting the number of real multiplications and additions. By 'complexity during data transmission', we mean the filtering operation itself. Hence, adopted to echo cancellation, 'complexity during data transmission' is similar to 'echo emulation'. We do not consider the filter implementation and adaptation here.

The complexity for the two-fold oversampled TEQ and PT-EQ can be derived from the complexity figures given in [9].

A. Symmetric rate set-up

It is shown in [9] that the complexity during data transmission for a T-taps TEQ is roughly the same as for $\frac{N}{2}$ T-taps PT-EQ's. This is due to the fact that PTEQ's work at an N-fold downsampled rate with respect to the TEQ. When not all the tones are used or when a smaller equalizer size is sufficient for some tones, complexity of the PT method can be reduced even further.

The complexity of the TEC consists only of the T_E taps in the time domain, as presented in **Table I**. F_s is the sample rate. There is neither an FFT nor 1-taps PT-EQ's involved in this EC.

The PT-EC has the same complexity as the PT-EQ but now with the number of taps equal to T_E. This complexity is shown in **Table II** where it is assumed that the number of echo cancelled tones is $\frac{N}{2}$ and the symbol duration $\frac{N+\nu}{F_s} \approx \frac{N}{F_s}$. It is seen that the number of multiplications and additions for PT-EC remains more or less the same as for TEC when all tones are echo cancelled. However in reality not all tones are echo cancelled and for these tones the number of taps can be set to zero which corresponds to an extra complexity saving. Additional complexity however is introduced by the extra FFT in the PT-EC-scheme. This complexity can be decreased if the modem has an asymmetric rate set-up. In the symmetric rate synchronous case, the FFT overhead is non-existent.

B. Interpolated echo cancellation at RT

The complexity of TEC versus PT-EC is shown in **Table III**. Note that F_s is the high rate sample frequency, i.e. according to N. In comparison with the set-up with an IFFT of size N, the number of multiplications and additions for the TEC and PT-EC is roughly divided by γ, and only one $\frac{N}{\gamma}$-point FFT is needed in the PT implementation.

C. Decimated echo cancellation at CO

The complexity during data transmission can be found by multiplying by γ the complexity of an $(\frac{N}{\gamma}, \frac{T}{\gamma}, 1)$-PTF. In the assumption that $\frac{N}{2\gamma}$ tones are used and that $\mathrm{mod}(\frac{T_E}{\gamma}) = 0$, the complexity is given in **Table IV**, with F_s the high rate sample frequency. In comparison with the interpolated EC, one needs γ $\frac{N}{\gamma}$-point FFT's instead of 1 $\frac{N}{\gamma}$-point FFT and γ times more difference terms, but in this case the number of tones that have to be demodulated is significantly smaller.

VIII. Simulation results

Simulation results are presented for a 4km 26 AWG downstream and upstream channel with white noise of -140dBm/Hz. For downstream, the IFFT and FFT have size $N = 512$. For upstream, the IFFT and FFT have size $N = 128$. Hence, the interpolation factor (in RT) and the decimation factor (in CO) are equal to $\gamma = 4$. The corresponding cyclic prefix lengths are $\nu = 32$ and $\nu/\gamma = 8$. An FDM set-up for the used tones is adopted. The upstream tones are tones 8 till 30 and the downstream tones are 39 till 256. Due to relaxed transmit and receive filter order specifications, downstream symbols experience echo from upstream symbols and vice versa. Downstream and upstream PSD on the used tones are -40dBm/Hz and -38dBm/Hz respectively.

The bit rate is calculated by:

$$rate = (\sum_{i=used\ tone} b_i) \cdot \frac{F_s}{N+\nu} \tag{13}$$

with $F_s = 2.208$MHz the (high) sample rate (corresponding to symbol size N). The zero reference delay of the far end and the echo impulse response is determined by the interval (with length equal to the prefix length plus one) with highest energy in the far end and the echo impulse response respectively. Bit rates are calculated for relative delays $\delta_1 = 0$ and $\delta_2 = \delta_3 = 20$ at both RT and CO.

Fig. 7 shows the downstream bit rate as a function of the number of taps T_E of the PT-EC. We compare 32-taps PT-EQ (at original sample rate) with 2 times 16-taps PT-EQ (hence, with 2-fold oversampling). The upper bounds are the bit rates for both set-ups when no echo is present. They are 2.89 Mbits/s and 3.23 Mbits/s respectively. For T_E equal to 150, the upper bounds are reached. Remark that, as far as complexity is concerned, the number of EC taps can effectively be divided by $\gamma = 4$ (see section V). To compare these bit rates with bit rates obtained by T_E-taps TEC, we have simulated 'ideal' TEC performance by zeroing T_E successive elements with highest energy in the echo impulse response. A 400-taps TEC is then needed to obtain the upper bound bit rates.

Fig. 8 shows analogous simulation results for upstream. The upper bounds for the bit rates are 1.12 Mbits/s for 32-taps PT-EQ and 1.11 Mbits/s for 2 times 16-taps PT-EQ. In this case the the sample rate set-up has slightly better performance compared to the oversampled set-up. As far as complexity is concerned, each T_E taps PT-EC may effectively be replaced by $\gamma = 4$

$\frac{T_E}{\gamma}$-taps PT-EC's (see section VI). TEC performance has also been simulated by zeroing T_E successive elements in the echo impulse response. A 400-taps TEC is then needed to obtain 1.10 Mbits/s with 32-taps PT-EQ and a 500-taps TEC is needed to obtain 1.10 Mbits/s with 2 times 16-taps PT-EQ.

IX. Conclusions

Per tone echo cancellation (PT-EC) is proposed as an alternative to time domain (or time/frequency domain) echo cancellation (TEC). The resulting receiver structure enables us to optimize the SNR for each tone separately by solving an MMSE problem. This MMSE problem is the basis for adaptive algorithms which truly optimize capacity, unlike the TEC-based approaches. PT-EC is shown to have a comparable complexity during data transmission as TEC. Extra complexity however is introduced by the FFT in the PT-EC. Reduced complexity structures are derived for an interpolated and a decimated rate set-up.

References

[1] N. Al-Dhahir. Joint Channel and Echo Impulse Response Shortening on Digital Subscriber Lines. *IEEE Signal Processing Letters*, 3(10):280–282, Oct. 1996.

[2] N. Al-Dhahir and J. M. Cioffi. Optimum Finite-Length Equalization for Multicarrier Transceivers. *IEEE Trans. on Commun.*, 44(1):56–64, Jan. 1996.

[3] J. A. C. Bingham. Multicarrier Modulation for Data Transmission: An Idea Whose Time Has Come. *IEEE Commun. Mag.*, 28(5):5–14, May 1990.

[4] M. Ho. *Multicarrier Echo Cancellation and Multichannel Equalization*. PhD thesis, Stanford University, 1995.

[5] M. Ho, J. M. Cioffi, and J. A. C. Bingham. Discrete Multitone Echo Cancellation. *IEEE Trans. on Commun.*, 44(7):817–825, July 1996.

[6] P. J. W. Melsa, R. C. Younce, and C. E. Rohrs. Impulse Response Shortening for Discrete Multitone Transceivers. *IEEE Trans. on Commun.*, 44(12):1662–1672, Dec. 1996.

[7] K. Van Acker, G. Leus, M. Moonen, and T. Pollet. RLS-based Initialization for Per Tone Equalizers in DMT-Receivers. In *Proc. European Signal Processing Conf. (Eusipco)*, Tampere, Finland, Sep. 2000.

[8] K. Van Acker, G. Leus, M. Moonen, O. van de Wiel, and T. Pollet. Per Tone Equalization for DMT Receivers. In *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, pages 2311–2315, Rio de Janeiro, Brazil, Dec. 1999.

[9] K. Van Acker, G. Leus, M. Moonen, O. van de Wiel, and T. Pollet. Per Tone Equalization for DMT-based Systems. Accepted for publication in IEEE Trans. on Commun., 2000.

[10] M. Van Bladel and M. Moeneclaey. Time-Domain Equalization for Multicarrier Communication. In *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, pages 167–171, Singapore, Nov. 1995.

[11] B. Wang and T. Adali. Joint Impulse Response Shortening for Discrete Multitone Systems. In *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, pages 2508–2512, Rio de Janeiro, Brazil, Dec. 1999.

Acknowledgments

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 - IUAP P4-02 (1997-2001): Modeling, Identification, Simulation and Control of Complex Systems, the Concerted Research Action GOA-MEFISTO-666 (Mathematical Engineering for Information and Communication Systems Technology) of the Flemish Government and the IT-projects in the ITA-bis program of the Flemish Institute for Scientific and Technological Research in Industry (I.W.T.) ('IRMUT: Integration of Reconfigurable Multimedia terminals' (980271) and 'Advanced Internet Access'(980316)) and was partially funded by Alcatel. The scientific responsibility is assumed by its authors.

	# multiplications	# additions
TEC	$F_s T_E$	$F_s(T_E - 1)$

TABLE I
Complexity of TEC (in operations/s)

	# multiplications	# additions
FFT	$\frac{F_s}{N+\nu}\mathcal{O}(N \cdot \log(N))$	$\frac{F_s}{N+\nu}\mathcal{O}(N \cdot \log(N))$
PT-EC	$F_s(T_E+1)$	$F_sT_E + \frac{F_s}{N}(T_E-1)$

TABLE II
APPROXIMATED COMPLEXITY OF PT-EC (IN OPERATIONS/S)

	# multiplications	# additions
TEC	$F_s\lceil\frac{T_E}{\gamma}\rceil$	$F_s(\lceil\frac{T_E}{\gamma}\rceil - 1)$
FFT	$\frac{F_s}{N+\nu}\mathcal{O}(\frac{N}{\gamma} \cdot \log(\frac{N}{\gamma}))$	$\frac{F_s}{N+\nu}\mathcal{O}(\frac{N}{\gamma} \cdot \log(\frac{N}{\gamma}))$
PT-EC	$F_s(\lceil\frac{T_E}{\gamma}\rceil + 1)$	$F_s\lceil\frac{T_E}{\gamma}\rceil + \frac{F_s}{N}(\lceil\frac{T_E}{\gamma}\rceil - 1)$

TABLE III
COMPLEXITY COMPARISON OF TEC WITH PT-EC (IN OPERATIONS/S) FOR INTERPOLATED ECHO CANCELLATION (IN RT WITH $\frac{N}{\gamma}$-POINT IFFT AND N-POINT FFT)

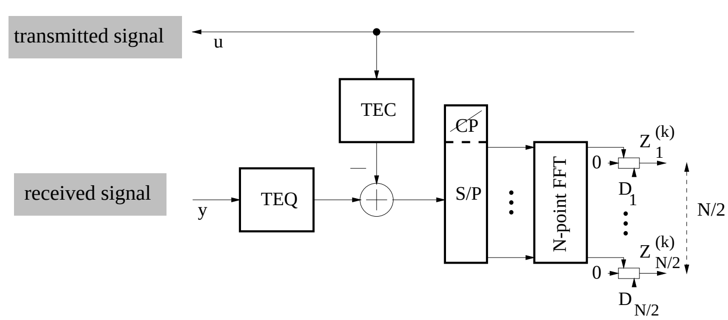

Fig. 1. Receiver with T-taps TEQ and T_E-taps TEC.

	# multiplications	# additions
TEC	$\frac{F_s}{\gamma} T_E$	$\frac{F_s}{\gamma}(T_E - 1)$
FFT	$\gamma \frac{F_s}{N+\nu} \mathcal{O}(\frac{N}{\gamma} \cdot \log(\frac{N}{\gamma}))$	$\gamma \frac{F_s}{N+\nu} \mathcal{O}(\frac{N}{\gamma} \cdot \log(\frac{N}{\gamma}))$
PT-EC	$\gamma \frac{F_s}{\gamma}(\frac{T_E}{\gamma} + 1)$	$\gamma \frac{F_s}{\gamma} \frac{T_E}{\gamma} + \gamma \frac{F_s}{N}(\frac{T_E}{\gamma} - 1)$

TABLE IV

COMPLEXITY COMPARISON OF TEC WITH PT-EC (IN OPERATIONS/S) FOR DECIMATED ECHO CANCELLATION (IN CO WITH N-POINT IFFT AND $\frac{N}{\gamma}$-POINT FFT)

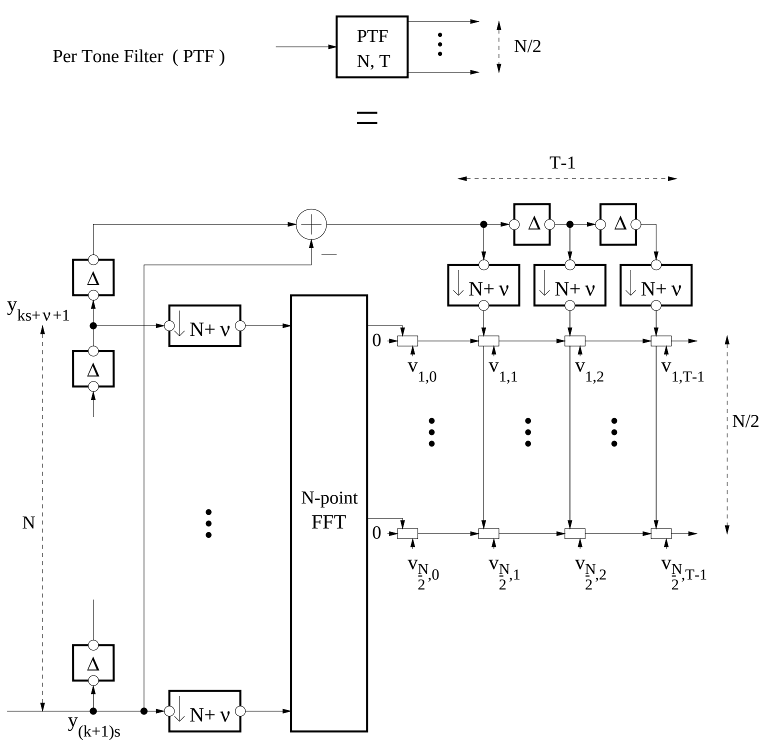

Fig. 2. An (N,T)-PTF block includes one N-point FFT operation and T-taps per tone filters $\mathbf{v}_i$ for all used tones.

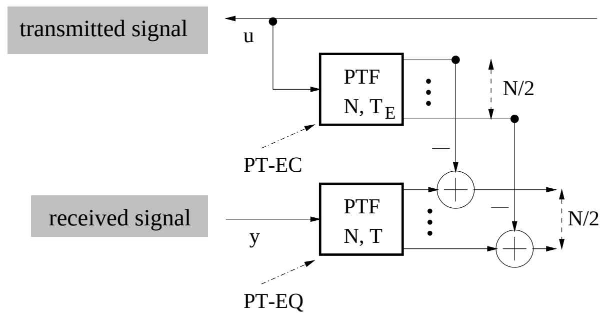

Fig. 3. Receiver with T-taps PT-EQ and T_E-taps PT-EC.

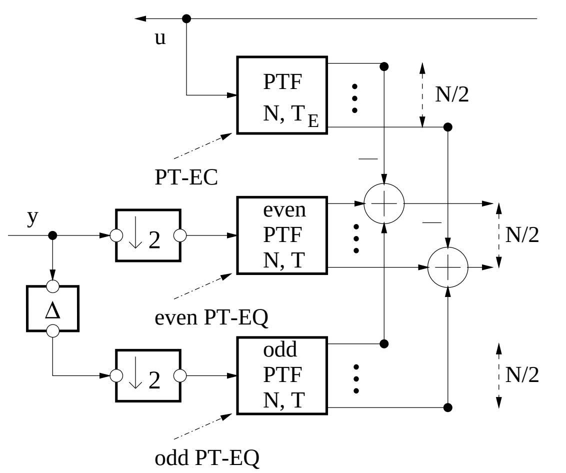

Fig. 4. Receiver with 2 times oversampled $2T$-taps PT-EQ and T_E-taps PT-EC.

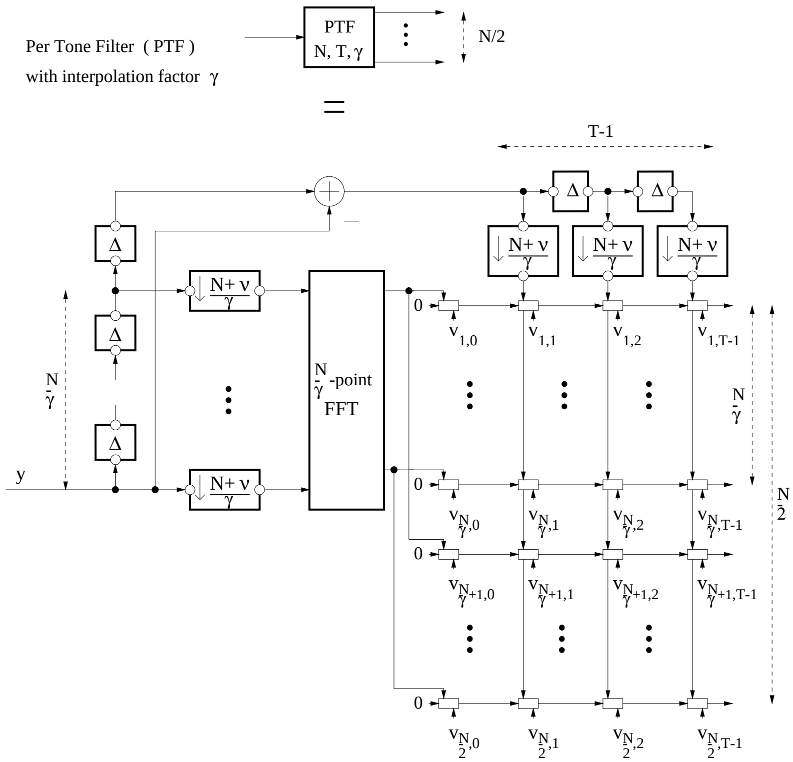

Fig. 5. An interpolated (N,T,γ)-PTF block includes one $\frac{N}{\gamma}$-point FFT operation and T-taps per tone filters.

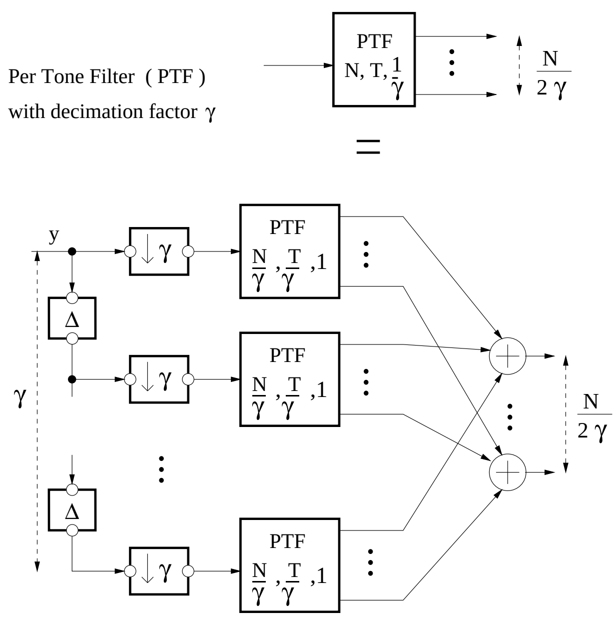

Fig. 6. A decimated $(N,T,\frac{1}{\gamma})$-PTF block includes γ $(\frac{N}{\gamma},\frac{T}{\gamma},1)$-PTF blocks each having one $\frac{N}{\gamma}$-point FFT operation and $\frac{T}{\gamma}$-taps per tone filters.

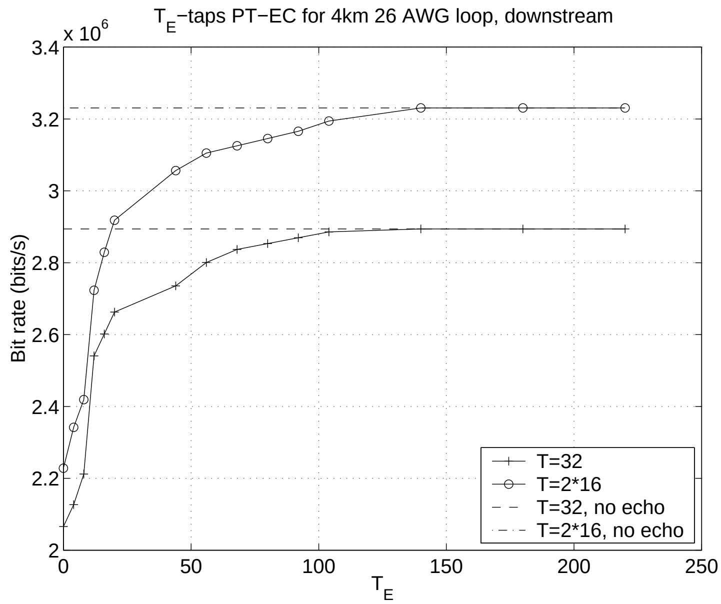

Fig. 7. Bit rate as a function of the number of taps T_E of the PT-EC for a 4km 26 AWG downstream loop for 32-taps PT-EQ (at original sample rate) and 2 times 16-taps PT-EQ (with 2-fold oversampling).

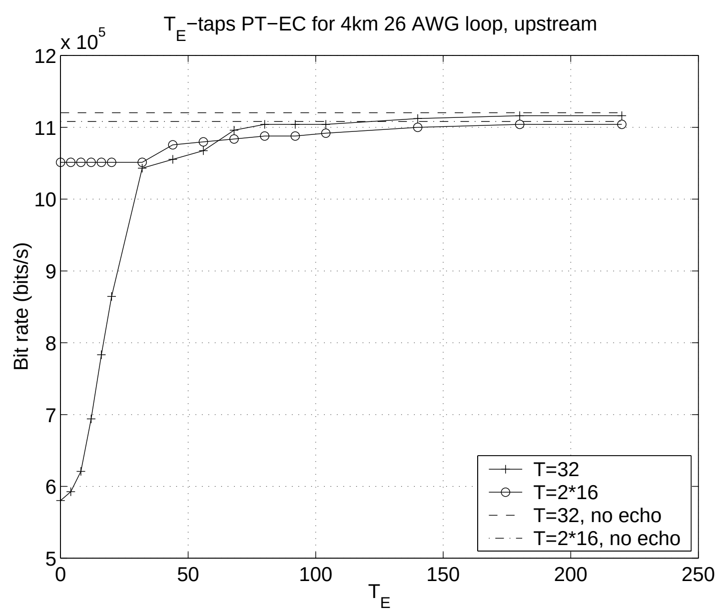

Fig. 8. Bit rate as a function of the number of taps T_E of the PT-EC for a 4km 26 AWG upstream loop.