

Research Article

Bit Rate Maximising Per-Tone Equalisation with Adaptive Implementation for DMT-Based Systems

Suchada Sitjongsataporn and Peerapol Yuvapoositanon

Centre of Electronic Systems Design and Signal Processing (CESdSP), Department of Electronic Engineering, Mahanakorn University of Technology, 140 Cheumsamphan Road, Nong-chok, Bangkok 10530, Thailand

Correspondence should be addressed to Suchada Sitjongsataporn, ssuchada@mut.ac.th

Received 3 December 2008; Revised 9 July 2009; Accepted 19 September 2009

Recommended by Azzedine Zerguine

We present a bit rate maximising per-tone equalisation (BM-PTEQ) cost function that is based on an exact subchannel SNR as a function of per-tone equaliser in discrete multitone (DMT) systems. We then introduce the proposed BM-PTEQ criterion whose derivation for solution is shown to inherit from the methodology of the existing bit rate maximising time-domain equalisation (BM-TEQ). By solving a nonlinear BM-PTEQ cost function, an adaptive BM-PTEQ approach based on a recursive Levenberg-Marquardt (RLM) algorithm is presented with the adaptive inverse square-root (iQR) algorithm for DMT-based systems. Simulation results confirm that the performance of the proposed adaptive iQR RLM-based BM-PTEQ converges close to the performance of the proposed BM-TEQ. Moreover, the performance of both these proposed BM-PTEQ algorithms is improved as compared with the BM-TEQ.

Copyright © 2009 S. Sitjongsataporn and P. Yuvapoositanon. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

Discrete multitone (DMT) is a digital implementation technique widely used for high speed wired multicarrier transmission such as asymmetric digital subscriber lines (ADSLs) [1]. The cyclic prefix (CP) is inserted among DMT-symbols to arrange subchannels separately in order to eliminate intercarrier interference (ICI) and intersymbol interference (ISI). Conventional equalisation of DMT-based system consists of an adaptive (real) time-domain equaliser (TEQ) which shortens the convolutional result of TEQ and channel impulse response (CIR). So that ISI can be effectively handled by CP, and ICI can also be mitigated. A (complex) one-tap frequency-domain equaliser (FEQ) is applied subsequently to compensate for amplitude and phase of distortion [1, 2]. However, TEQs are not designed to achieve the maximum bit rate performance [3]. The so-called per-tone equalisation which is a frequency-domain equalisation scheme for each tone has been introduced in [4]. It is shown to give comparable bit

rate maximising characteristics with existing equalisation schemes.

In the literature, a few update algorithms for T -tap per-tone equalisers (PTEQs) are proposed in [4–7]. The per-tone equalisation scheme using a technique based on transferring the (real) TEQ-operations to the frequency-domain is done per tone after the fast Fourier transform (FFT) demodulation as suggested in [4]. This enables us to accomplish the signal-to-noise ratio (SNR) optimisation per tone, because the equalisation of each tone is independent of other tones. This PTEQ performance has been presented to be better than any TEQ-based receiver. In [4], the authors conclude that the result of complexity of TEQ (including one-tap FEQ) is comparable to PTEQ. To reduce complexity during initialisation of PTEQ, the tone grouping PTEQ approach is presented in [7–9] by combining tones. The idea of tone grouping is to compute the PTEQ for the center tone of each group, then to reuse it for the whole group. Another method to decrease the complexity of PTEQ is to consider a suitable length of the equaliser for every tone. A resource

allocation technique is presented for the variable-length equaliser in order to optimise the length distribution of PTEQ over tones with a relatively low complexity, as given in [10].

Based on the recursive least squares (RLS) algorithm, the adaptive PTEQs with inverse updating have been presented in [5, 7]. An RLS-based algorithm requires the second-order information as the autocorrelation matrix of the sliding discrete Fourier transform (DFT) of the received signal. In [5], it is shown that a significant part of RLS-based computations for storing and updating can be shared among the different tones, leading to sufficiently low initialisation complexity. A combined recursive least squares-least mean square (RLS-LMS) initialisation algorithm for PTEQs [7] is presented to exploit the advantages of both the fast convergence and low complexity. In [6], an adaptive recursive Levenberg-Marquardt (RLM) algorithm for PTEQs is proposed with no TEQ concerned.

In [11], the authors present a TEQ design as its optimal solution of the truly bit rate maximising time-domain equalisation (BM-TEQ) cost function. It is based on an exact formulation of the subchannel SNR as a function of the taps of TEQ. Its bit rate is smooth as a function of synchronisation delay, so it is shown to approach as well as the PTEQ performance. An adaptive RLM-based BM-TEQ design [12] is derived from the nonlinear and nonconvex cost criterion. This adaptive BM-TEQ has the same second-order statistics as that of the RLS-based adaptive PTEQ in [5]. Furthermore, many algorithms have been presented to adaptively initialise the TEQ and PTEQ schemes, but none of them truly maximises the bit rate of PTEQs framework in DMT-based systems.

The purpose of this paper is twofold. First, we introduce the bit rate maximising criterion of PTEQ. The PTEQ which attains this bit rate maximising capability is called a bit rate maximising per-tone equalisation (BM-PTEQ). Second, we apply an adaptive implementation to show how the solution of BM-PTEQ can be achieved in practice. We also show that the BM-PTEQ solution can be expressed in the form of the BM-TEQ of [11]. This leads us to the proposition that, given the proven superior performance of PTEQ over TEQ [13], the BM-PTEQ will continue to do better than the BM-TEQ of [11] in the sense of bit rate maximising performance.

We describe an overview of system model and notation in Section 2. The solution of the PTEQ design criterion is reviewed in Section 3. The derivation of proposed BM-PTEQ criterion is developed in Section 4. Section 5 shows that the proposed adaptive BM-PTEQ can be designed recursively using the nonlinear cost criterion. The simulation results are presented in Section 6. Finally, Section 7 concludes the paper.

2. System Model and Notation

In this section, we describe that the data model and notation based on an FIR model of the DMT transmission channel is presented as [4]

$$\mathbf{y} = \mathbf{H} \cdot \mathbf{X} + \mathbf{n},$$

$$\begin{aligned} & \underbrace{\begin{bmatrix} y_{k,l+\Delta} \\ \vdots \\ y_{k,N-l+\Delta} \end{bmatrix}}_{\mathbf{y}_{k,l+\Delta:N-l+\Delta}} \\ &= \underbrace{\begin{bmatrix} \mathbf{0}_{(1)} & \left[\begin{array}{c|ccc} \overline{\mathbf{h}}^{-T} & 0 & \cdots & \\ \vdots & \ddots & \ddots & \\ \cdots & 0 & \overline{\mathbf{h}}^{-T} & \end{array} \right] & \mathbf{0}_{(2)} \end{bmatrix}}_{\mathbf{H}} \cdot \underbrace{\begin{bmatrix} \mathcal{P}_\nu & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathcal{P}_\nu & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathcal{P}_\nu \end{bmatrix}}_{\mathbf{H}} \cdot \underbrace{\begin{bmatrix} \mathcal{I}_N & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathcal{I}_N & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathcal{I}_N \end{bmatrix}}_{\mathbf{H}} \\ & \cdot \underbrace{\begin{bmatrix} \mathbf{x}_{k-1,N} \\ \mathbf{x}_{k,N} \\ \mathbf{x}_{k+1,N} \end{bmatrix}}_{\mathbf{x}_{k-1:k+1,N}} + \underbrace{\begin{bmatrix} \eta_{k,l+\Delta} \\ \vdots \\ \eta_{k,N-l+\Delta} \end{bmatrix}}_{\boldsymbol{\eta}_{k,l+\Delta:N-l+\Delta}}, \end{aligned} \quad (1)$$

where l denotes the first considered sample of the k th received DMT-symbol. This depends on the number of tap of equaliser (T) and the synchronisation delay (Δ). The vector $\mathbf{y}_{k,i:j}$ of received samples i to j of k th DMT-symbol is $\mathbf{y}_{k,i:j} = [y_{k,i} \cdots y_{k,j}]^T$. A sequence of the $N \times 1 \mathbf{x}_{k,N}$ transmitted symbol vector is $\mathbf{x}_{k,N} = [x_{k,0} \cdots x_{k,N-1}]^T$. The size N is of inverse discrete Fourier transform (IDFT) and DFT. The parameter ν denotes the length of cyclic prefix. The matrices $\mathbf{0}_{(1)}$ and $\mathbf{0}_{(2)}$ are also the zero matrices of size $(N-l) \times (N-L+2\nu+\Delta+l)$ and $(N-l) \times (N+\nu-\Delta)$. The vector $\overline{\mathbf{h}}$ is the \mathbf{h} channel impulse response (CIR) vector in reverse order. The $(N+\nu) \times N$ matrix \mathcal{P}_ν is denoted by

$$\mathcal{P}_\nu = \left[\begin{array}{c|c} \mathbf{0}_{\nu \times (N-\nu)} & \mathbf{I}_\nu \\ \hline & \mathbf{I}_N \end{array} \right], \quad (2)$$

which adds the cyclic prefix. The \mathcal{I}_N is $N \times N$ IDFT matrix and modulates the input symbols. The $\boldsymbol{\eta}_{k,l+\Delta:N-l+\Delta}$ is a vector with additive white Gaussian noise (AWGN) and near-end cross-talk (NEXT).

Some notation will be used throughout this paper as follows: $E\{\cdot\}$ is the expectation operator and $\text{diag}(\cdot)$ is a diagonal matrix operator. The operators $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^*$ denote the transpose, Hermitian, and complex conjugate operator, respectively. The k is the DMT symbol index and \mathbf{I}_a is an $a \times a$ identity matrix. A tilde over the variable indicates the frequency domain. The vectors are in bold lowercase and matrices are in bold uppercase.

3. Per-Tone Equalisation

In this section, we show the concept of per-tone equaliser (PTEQ). We refer the readers to [4] for more details. The per-tone equalisation structure is based on transferring the TEQ-operations into the frequency-domain after DFT

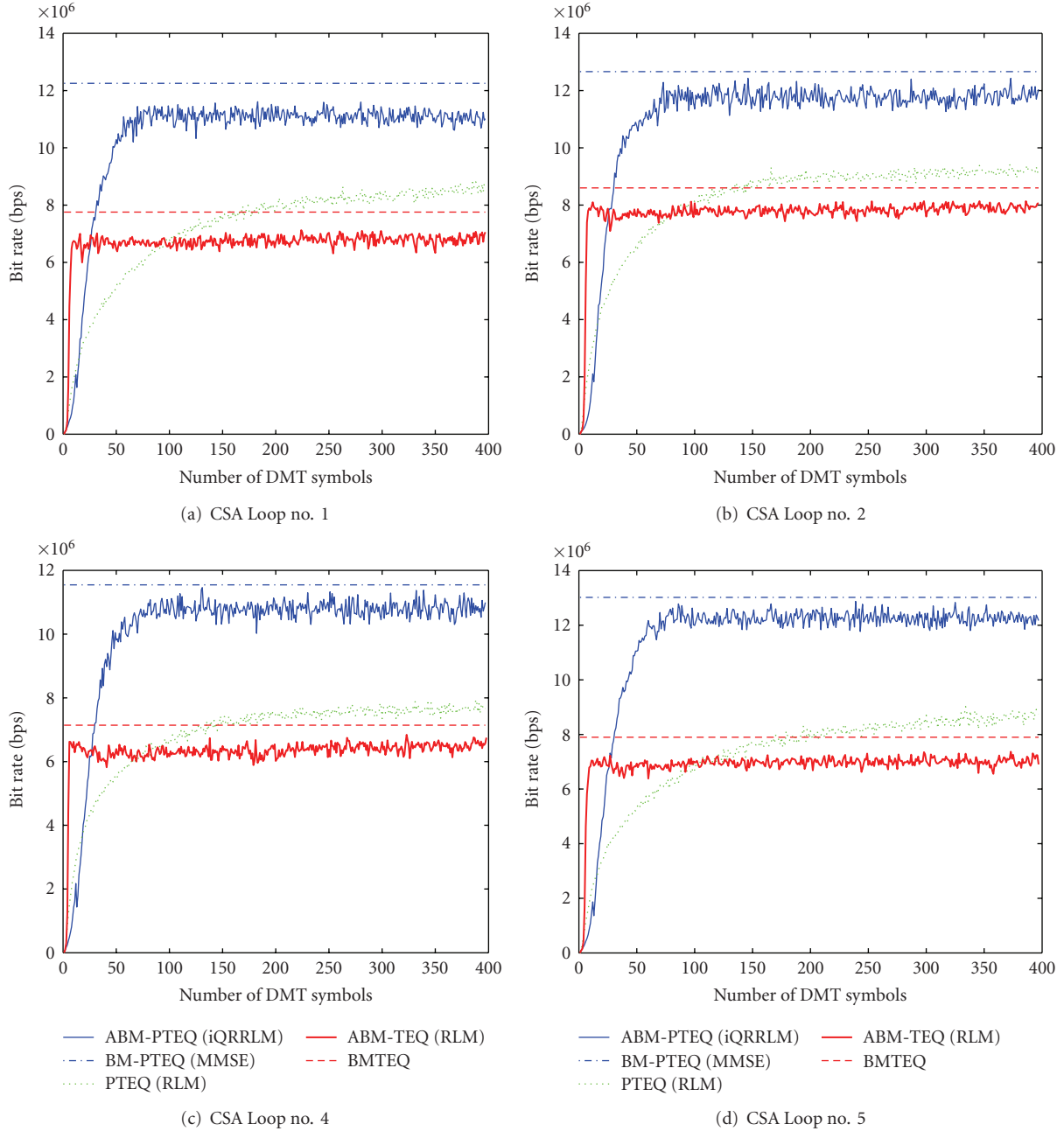


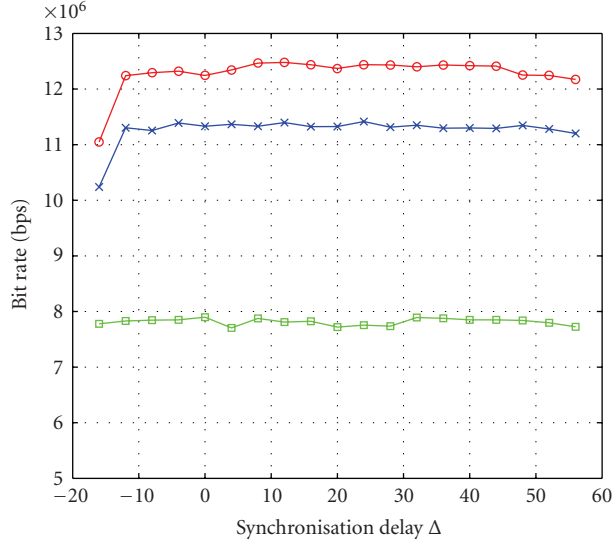
FIGURE 1: Learning curves of bit rate convergence of proposed adaptive iQRRLM-based BM-PTEQ (ABM-PTEQ), adaptive RLM-based PTEQ [6], and adaptive RLM-based BM-TEQ (ABM-TEQ) [12] which compared with BM-TEQ [11] and proposed MMSE-based BM-PTEQ for ADSL downstream starting at tones 38 to 255, when the samples of CSA loop are (a) CSA Loop no. 1, (b) CSA Loop no. 2, (c) CSA Loop no. 4, and (d) CSA Loop no. 5.

demodulation, which results in a T -tap PTEQ for each tone separately. For each tone i ($i = 1, \dots, n$), the TEQ-operations are shown as follows [4]:

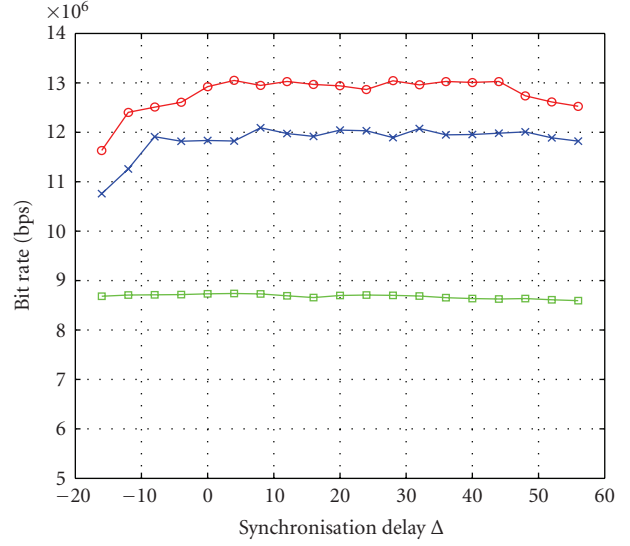
$$\tilde{d}_n = \underbrace{\tilde{z}_n}_{1\text{-tap FEQ}} \cdot \underbrace{\text{row}_n(\mathcal{F}_N \cdot \mathbf{Y} \cdot \mathbf{w})}_{1 \text{ DFT}}, \quad (3)$$

$$= \underbrace{\text{row}_n(\mathcal{F}_N \cdot \mathbf{Y})}_{T \text{ DFTs}} \cdot \underbrace{\mathbf{w} \cdot \tilde{z}_n}_{T\text{-tap FEQ } \mathbf{v}_n}, \quad (4)$$

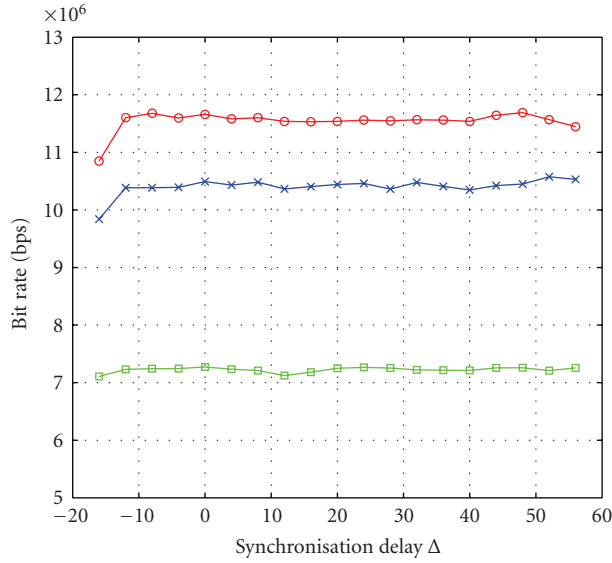
where \tilde{d}_n is the output after frequency-domain equalisation for tone n . The \tilde{z}_n is the (complex) one-tap FEQ for tone n . The parameter \mathbf{w} is of (real) T -tap TEQ and \mathcal{F}_N is an $N \times N$ DFT matrix [4]. Note that \mathbf{Y} is an $N \times T$ Toeplitz matrix of received signal samples as vector \mathbf{y} in (1). From (4), the T DFT-operations are cheaply calculated by means of a sliding DFT. It is demonstrated in [4] that every T -tap FEQ \mathbf{v}_n exists a T -tap PTEQ $\tilde{\mathbf{p}}_n$ which consists of only one DFT and $T - 1$ real difference terms as its input.



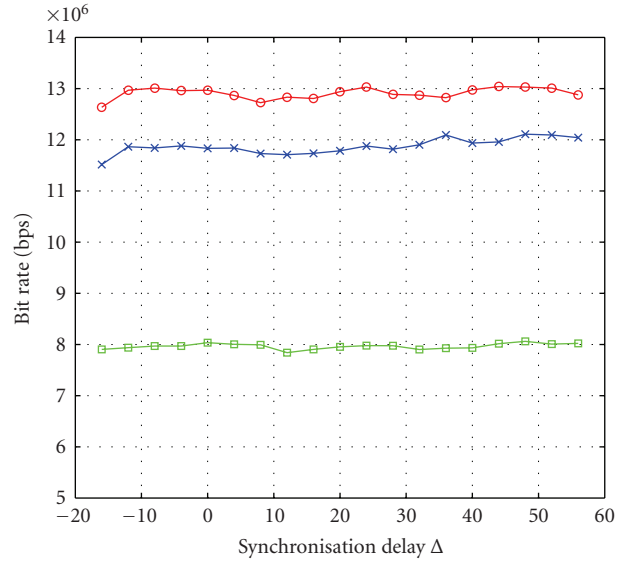
(a) CSA Loop no. 1



(b) CSA Loop no. 2



(c) CSA Loop no. 4



(d) CSA Loop no. 5

× ABM-PTEQ (iQRRLM)
 × BM-PTEQ (MMSE)
 × BM-TEQ

× ABM-PTEQ (iQRRLM)
 × BM-PTEQ (MMSE)
 × BM-TEQ

FIGURE 2: Bit rate as a function of the synchronisation delay Δ for ADSL downstream starting at tones 38 to 255, when the samples of CSA loop are (a) CSA Loop no. 1, (b) CSA Loop no. 2, (c) CSA Loop no. 4, and (d) CSA Loop no. 5.

The PTEQ output $\hat{x}_{k,n}$ can be specified as follows:

$$\hat{x}_{k,n} = \tilde{\mathbf{p}}_n^H \cdot \underbrace{\begin{bmatrix} \mathbf{I}_{T-1} & \mathbf{0} & -\mathbf{I}_{T-1} \\ \mathbf{0} & \mathcal{F}_N(n,:) \end{bmatrix}}_{\mathbf{F}_n} \cdot \mathbf{y}, \quad (5)$$

$$= \tilde{\mathbf{p}}_n^H \cdot \tilde{\mathbf{y}}_{k,n}, \quad (6)$$

where $\tilde{\mathbf{p}}_n$ is the T -tap complex-valued PTEQ vector for tone n . The \mathbf{F}_n is a $(T-1) \times (N+T-1)$ matrix [4]. The $\mathcal{F}_N(n,:)$ is the n th row of \mathcal{F}_N . By using the sliding DFT, the first

block row of matrix \mathbf{F}_n in (5) extracts the difference terms, while the last row corresponds to the usual DFT operation as detailed in [4, 10]. The vector \mathbf{y} is of channel output samples as described in (1). The $\tilde{\mathbf{y}}_{k,n}$ is the sliding DFT output for tone n at symbol k .

4. A Bit Rate Maximising Per-Tone Equalisation

In this section, we introduce the BM-PTEQ criterion with an exact subchannel SNR model. In the derivation of the cost

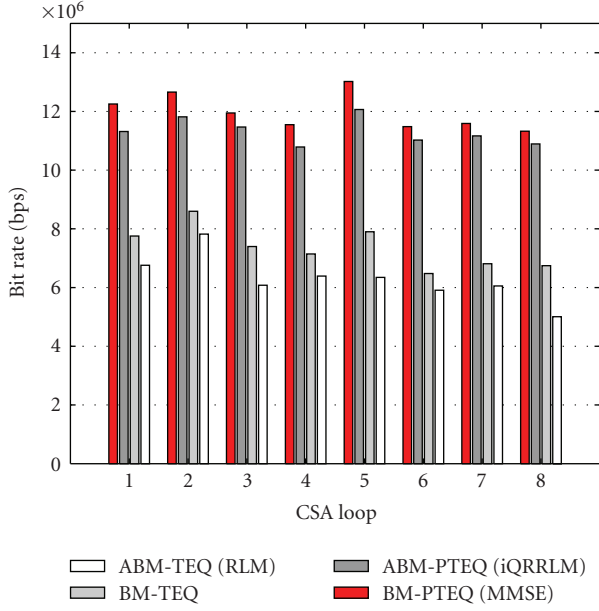


FIGURE 3: The bit rate performance of the BM-TEQ [11], adaptive RLM-based BM-TEQ (ABM-TEQ) [12], proposed MMSE-based BM-PTEQ, and proposed adaptive iQRRLM-based BM-PTEQ (ABM-PTEQ) for all CSA loop nos. 1–8 at starting tones 38 to 255 downstream ADSL when fixed $\Delta = 45$.

function of BM-PTEQ, we start from the bit rate expression as given in [14]. The total number of bits transmitted in one DMT-symbol is defined by

$$b_{\text{PTEQ}} = \sum_{n \in N_d} \log_2 \left(1 + \frac{\text{SNR}_n}{\Gamma_n} \right), \quad (7)$$

where N_d is the range of active tones, and SNR_n denotes the SNR on tone n . The constant Γ_n is a function of the desired probability of error, coding gain, and system margin. We notice that an integer number of bits is allocated to optimise the transmit power per tone after equalisation.

4.1. An Exact Subchannel SNR Model. For the BM-PTEQ criterion to be derived, it is important to define the dependence of the subchannel SNR on PTEQs. The SNR on tone n can be written as

$$\text{SNR}_n = \frac{\varepsilon_{s,n}}{\varepsilon_{e,n}}, \quad (8)$$

where $\varepsilon_{s,n}$ is the desired received signal energy on tone n , and $\varepsilon_{e,n}$ is the energy in the error signal on tone n at the FEQ output. The signal energy portions $\varepsilon_{s,n}$ and $\varepsilon_{e,n}$ in the subchannel SNR model (8) are determined at the FFT outputs, as assumed in [14, 15].

Following [16, 17], the PTEQ output on tone n can be written as

$$\tilde{\mathbf{p}}_n^H \tilde{\mathbf{y}}_{k,n} = \beta_n \tilde{\mathbf{x}}_{k,n} + \underbrace{\tilde{i}_{c,n} + \tilde{i}_{\eta,n}}_{\tilde{\eta}_n}, \quad (9)$$

where the $\tilde{\mathbf{p}}_n$ is the complex PTEQ vector on tone n , and $\tilde{\mathbf{y}}_{k,n}$ is the n th sliding DFT output vector for tone n at symbol k . The $\beta_n \tilde{\mathbf{x}}_{k,n}$ is a scaled version of the transmitted frequency-domain DMT-symbol $\tilde{\mathbf{x}}_{k,n}$. The error $\tilde{\eta}_n$ is the sum of residual ISI/ICI $\tilde{i}_{c,n}$ and noise $\tilde{i}_{\eta,n}$ at the n th PTEQ output. When the scalar β_n in (9) is equal to 1, the desired signal component at the PTEQ output is unbiased, in case of unconstrained MMSE PTEQ $\tilde{\mathbf{p}}_{*,n}$ as

$$\tilde{\mathbf{p}}_{*,n} = \frac{E\{\tilde{\mathbf{y}}_{k,n}^H \tilde{\mathbf{x}}_{k,n}\}}{E\{\tilde{\mathbf{y}}_{k,n}^H \tilde{\mathbf{y}}_{k,n}\}}. \quad (10)$$

With MMSE PTEQ, the desired signal energy $\varepsilon_{s,n} = E\{|\tilde{\mathbf{x}}_{k,n}|^2\}$ is equal to $\sigma_{\tilde{\mathbf{x}}_n}^2$. The error energy $\varepsilon_{e,n}$ in (8) is the mean square error $E\{|\tilde{\eta}_n|^2\}$ at the PTEQ output. It takes residual ISI/ICI $\tilde{i}_{c,n}$ and all external noise $\tilde{i}_{\eta,n}$ sources into account. The ratio of signal energy $E\{|\tilde{\mathbf{x}}_{k,n}|^2\}$ over the estimated error energy $E\{|\tilde{\eta}_n|^2\}$ yields an estimated SNR on tone n needed in the bit rate calculation. So the SNR in (8) is suitable to calculate the transmitted power allocation scheme.

Therefore, the exact subchannel SNR model (8) can be rewritten as

$$\text{SNR}_n^{\max} = \frac{\varepsilon_{s,n}}{\varepsilon_{e,n}} = \frac{E\{|\tilde{\mathbf{x}}_{k,n}|^2\}}{E\{|\tilde{\eta}_n|^2\}} = \frac{\sigma_{\tilde{\mathbf{x}}_n}^2}{E\{|\tilde{\mathbf{x}}_{k,n} - \tilde{\mathbf{p}}_{*,n}^H \tilde{\mathbf{y}}_{k,n}|^2\}}. \quad (11)$$

Introducing the compact notation for the $1 \times T$ correlation vectors $\sum_{\tilde{\mathbf{x}}_n}$ and $T \times T$ matrix $\sum_{\tilde{\mathbf{y}}_n}^2$ as

$$\sum_{\tilde{\mathbf{x}}_n} = E\{\tilde{\mathbf{x}}_{k,n}^* \tilde{\mathbf{y}}_{k,n}\}, \quad (12)$$

$$\sum_{\tilde{\mathbf{y}}_n}^H = E\{\tilde{\mathbf{y}}_{k,n}^H \tilde{\mathbf{y}}_{k,n}\}, \quad (13)$$

$$\sum_{\tilde{\mathbf{y}}_n}^2 = E\{\tilde{\mathbf{y}}_{k,n}^H \tilde{\mathbf{y}}_{k,n}\}, \quad (14)$$

and expanding the denominator of (11) gives

$$\begin{aligned} E\{|\tilde{\eta}_n|^2\} &= E\{|\tilde{\mathbf{x}}_{k,n} - \tilde{\mathbf{p}}_{*,n}^H \tilde{\mathbf{y}}_{k,n}|^2\} \\ &= \sigma_{\tilde{\mathbf{x}}_n}^2 - \tilde{\mathbf{p}}_{*,n} \sum_{\tilde{\mathbf{y}}_n} - \tilde{\mathbf{p}}_{*,n}^H \sum_{\tilde{\mathbf{y}}_n}^H + |\tilde{\mathbf{p}}_{*,n}|^2 \sum_{\tilde{\mathbf{y}}_n}^2 \\ &= \sigma_{\tilde{\mathbf{x}}_n}^2 \left(\frac{\sigma_{\tilde{\mathbf{x}}_n}^2 \sum_{\tilde{\mathbf{y}}_n}^2}{|\sum_{\tilde{\mathbf{y}}_n}|^2} - 1 \right), \end{aligned} \quad (15)$$

where $\tilde{\mathbf{p}}_{*,n}$ is the unconstrained MMSE PTEQ as defined in (10).

We obtain a compact maximum SNR model SNR_n^{\max} by replacing (15) in (11) as

$$\begin{aligned} \text{SNR}_n^{\max} &= \frac{\sigma_{\tilde{x}_n}^2}{\sigma_{\tilde{x}_n}^2 \left(\sigma_{\tilde{x}_n}^2 \sum_{\tilde{y}_n}^2 / \left| \sum_{\tilde{x}\tilde{y}_n} \right|^2 - 1 \right)} \\ &= \frac{\left| \sum_{\tilde{x}\tilde{y}_n} \right|^2}{\sigma_{\tilde{x}_n}^2 \sum_{\tilde{y}_n}^2 - \left| \sum_{\tilde{x}\tilde{y}_n} \right|^2} \\ &= \frac{\left| \sum_{\tilde{x}\tilde{y}_n} \right|^2}{\left(\sigma_{\tilde{x}_n}^2 \sum_{\tilde{y}_n}^2 \right) \left\{ 1 - \left| \sum_{\tilde{x}\tilde{y}_n} \right|^2 / \sigma_{\tilde{x}_n}^2 \sum_{\tilde{y}_n}^2 \right\}} \\ &= \frac{\rho_n^2}{1 - \rho_n^2}, \end{aligned} \quad (16)$$

with

$$\rho_n^2 = \frac{\left| \sum_{\tilde{x}\tilde{y}_n} \right|^2}{\sigma_{\tilde{x}_n}^2 \sum_{\tilde{y}_n}^2}, \quad (17)$$

where ρ_n^2 is a squared normalised correlation function of FFT output $\tilde{y}_{k,n}$ and $\tilde{x}_{k,n}$ at the PTEQ output. We note that the SNR_n^{\max} in (16) is an exact (maximum) subchannel SNR model per tone at the PTEQs outputs, which is achieved by using the MMSE PTEQ $\tilde{\mathbf{p}}_{*,n}$ in (10) as described in [4]. So this BM-PTEQ design criterion will be defined by means of the unconstrained MMSE PTEQ $\tilde{\mathbf{p}}_{*,n}$ as given in (10). This will be used to maximise the bit rate capacity with regard to an integer number of bits allocation as given in (7).

4.2. The BM-PTEQ Cost Function. With the use of (7) and (16), the BM-PTEQ cost function criterion is the solution of

$$\begin{aligned} \arg \max_{\tilde{\mathbf{p}}_{*,n}} b_{\text{PTEQ}}^{\max} &= \arg \max_{\tilde{\mathbf{p}}_{*,n}} \sum_{n \in N_d} \log_2 \left(1 + \frac{\text{SNR}_n^{\max}}{\Gamma_n} \right) \\ &= \arg \max_{\tilde{\mathbf{p}}_{*,n}} \sum_{n \in N_d} \log_2 \left(1 + \frac{\rho_n^2}{\Gamma_n (1 - \rho_n^2)} \right) \\ &= \arg \max_{\tilde{\mathbf{p}}_{*,n}} \sum_{n \in N_d} \log_2 \left(\frac{\Gamma_n (1 - \rho_n^2) + \rho_n^2}{\Gamma_n (1 - \rho_n^2)} \right) \\ &= \arg \max_{\tilde{\mathbf{p}}_{*,n}} \sum_{n \in N_d} \log_2 \left(\frac{\Gamma_n + (1 - \Gamma_n) \rho_n^2}{\Gamma_n (1 - \rho_n^2)} \right). \end{aligned} \quad (18)$$

By rearranging (10) in terms of compact notation in (13) and (14), the unconstrained MMSE PTEQ $\tilde{\mathbf{p}}_{*,n}$ is given as

$$\tilde{\mathbf{p}}_{*,n} = \frac{\sum_{\tilde{x}\tilde{y}_n}^H}{\sum_{\tilde{y}_n}^2}, \quad (19)$$

and the squared normalised correlation parameter ρ_n^2 in (17) is rewritten as

$$\rho_n^2 = \frac{\sum_{\tilde{x}\tilde{y}_n} \sum_{\tilde{x}\tilde{y}_n}^H}{\sigma_{\tilde{x}_n}^2 \sum_{\tilde{y}_n}^2}. \quad (20)$$

Therefore, the BM-PTEQ cost function using the unconstrained MMSE PTEQs $\tilde{\mathbf{p}}_{*,n}$ in (19) when considering the maximum subchannel SNR at FEQs outputs in (16) is introduced as

$$\begin{aligned} &\arg \max_{\tilde{\mathbf{p}}_{*,n}} b_{\text{PTEQ}}^{\max} \\ &= \arg \max_{\tilde{\mathbf{p}}_{*,n}} \sum_{n \in N_d} \log_2 \frac{\Gamma_n + (1 - \Gamma_n) \left(\sum_{\tilde{x}\tilde{y}_n} \sum_{\tilde{x}\tilde{y}_n}^H / \sigma_{\tilde{x}_n}^2 \sum_{\tilde{y}_n}^2 \right)}{\Gamma_n \left\{ 1 - \sum_{\tilde{x}\tilde{y}_n} \sum_{\tilde{x}\tilde{y}_n}^H / \sigma_{\tilde{x}_n}^2 \sum_{\tilde{y}_n}^2 \right\}} \\ &= \arg \max_{\tilde{\mathbf{p}}_{*,n}} \sum_{n \in N_d} \log_2 \frac{\Gamma_n \sigma_{\tilde{x}_n}^2 + \tilde{\mathbf{p}}_{*,n} \sum_{\tilde{x}\tilde{y}_n} - \tilde{\mathbf{p}}_{*,n}^H \Gamma_n \sum_{\tilde{x}\tilde{y}_n}}{\Gamma_n \sigma_{\tilde{x}_n}^2 - \tilde{\mathbf{p}}_{*,n} \Gamma_n \sum_{\tilde{x}\tilde{y}_n}} \\ &= \arg \max_{\tilde{\mathbf{p}}_{*,n}} \sum_{n \in N_d} \log_2 \frac{\Gamma_n \left\{ \sigma_{\tilde{x}_n}^2 \right\} + (1 - \Gamma_n) \left\{ \tilde{\mathbf{p}}_{*,n}^H \sum_{\tilde{y}_n}^2 \tilde{\mathbf{p}}_{*,n} \right\}}{\Gamma_n \left\{ \sigma_{\tilde{x}_n}^2 - \tilde{\mathbf{p}}_{*,n}^H \sum_{\tilde{y}_n}^2 \tilde{\mathbf{p}}_{*,n} \right\}} \\ &= \arg \max_{\tilde{\mathbf{p}}_{*,n}} \sum_{n \in N_d} \log_2 \frac{\tilde{\mathbf{p}}_{*,n} \Gamma_n \tilde{\mathbf{p}}_{*,n}^H + \tilde{\mathbf{p}}_{*,n} \rho_n^2 \tilde{\mathbf{p}}_{*,n}^H - \tilde{\mathbf{p}}_{*,n} \Gamma_n \rho_n^2 \tilde{\mathbf{p}}_{*,n}^H}{\tilde{\mathbf{p}}_{*,n} \Gamma_n \tilde{\mathbf{p}}_{*,n}^H - \tilde{\mathbf{p}}_{*,n} \Gamma_n \rho_n^2 \tilde{\mathbf{p}}_{*,n}^H} \\ &= \arg \max_{\tilde{\mathbf{p}}_{*,n}} \sum_{n \in N_d} \log_2 \frac{\tilde{\mathbf{p}}_{*,n} \left\{ \Gamma_n (1 - \rho_n^2) + \rho_n^2 \right\} \tilde{\mathbf{p}}_{*,n}^H}{\tilde{\mathbf{p}}_{*,n} \left\{ \Gamma_n (1 - \rho_n^2) \right\} \tilde{\mathbf{p}}_{*,n}^H} \\ &= \arg \max_{\tilde{\mathbf{p}}_{*,n}} \sum_{n \in N_d} \log_2 \frac{\tilde{\mathbf{p}}_{*,n} \left\{ \Gamma_n \left(\sigma_{\tilde{x}_n}^2 \sum_{\tilde{y}_n}^2 - g \right) + g \right\} \tilde{\mathbf{p}}_{*,n}^H}{\tilde{\mathbf{p}}_{*,n} \left\{ \Gamma_n \left(\sigma_{\tilde{x}_n}^2 \sum_{\tilde{y}_n}^2 - g \right) \right\} \tilde{\mathbf{p}}_{*,n}^H} \\ &= \arg \max_{\tilde{\mathbf{p}}_{*,n}} \sum_{n \in N_d} \log_2 \frac{\tilde{\mathbf{p}}_{*,n} \mathbf{A}_n \tilde{\mathbf{p}}_{*,n}^H}{\tilde{\mathbf{p}}_{*,n} \mathbf{B}_n \tilde{\mathbf{p}}_{*,n}^H}, \end{aligned} \quad (21)$$

where g represents $\sum_{\tilde{x}\tilde{y}_n} \sum_{\tilde{x}\tilde{y}_n}^H$ and \mathbf{A}_n and \mathbf{B}_n depend on the second order statistics information $\sigma_{\tilde{x}_n}^2$, $\sum_{\tilde{y}_n}^2$ and $\sum_{\tilde{x}\tilde{y}_n}$

$$\begin{aligned} \mathbf{A}_n &= \Gamma_n \left(\sigma_{\tilde{x}_n}^2 \sum_{\tilde{y}_n}^2 - \sum_{\tilde{x}\tilde{y}_n} \sum_{\tilde{x}\tilde{y}_n}^H \right) + \sum_{\tilde{x}\tilde{y}_n} \sum_{\tilde{x}\tilde{y}_n}^H, \\ \mathbf{B}_n &= \Gamma_n \left(\sigma_{\tilde{x}_n}^2 \sum_{\tilde{y}_n}^2 - \sum_{\tilde{x}\tilde{y}_n} \sum_{\tilde{x}\tilde{y}_n}^H \right) \end{aligned} \quad (22)$$

Clearly, (21) has the exact form for the BM-TEQ solution of [11] with only a trivial interchange of the maximisation and minimisation operations for the argument. Therefore, the solution to achieve BM-PTEQ $\tilde{\mathbf{p}}_{*,n}$ can be also achieved with the same methodology for the bit rate maximising TEQ of [11]. This leads us to the crucial point that, given the proven superior performance of PTEQ over TEQ [13], the BM-PTEQ will always continue to do better than the BM-TEQ of [11] in the sense of bit rate maximising performance.

Proposition 1. *The bit rate performance of the BM-PTEQ is greater than or equal to that of the BM-TEQ,*

$$b_{\text{PTEQ}}^{\max} \geq b_{\text{TEQ}}^{\max} \quad (23)$$

where b_{TEQ}^{\max} represents the maximum bit rate achievable from the BM-TEQ of [11].

5. An Adaptive Bit Rate Maximising Per-Tone Equalisation

In Section 5.1, we introduce the constrained nonlinear exponentially weighted cost function for the complex-valued PTEQ. This criterion is translated with the deterministic approach to accomplish the maximum number of bits per DMT-symbol. With this nonlinear criterion in Section 5.1, we introduce an adaptive BM-PTEQ algorithm based on RLM algorithm in Section 5.2.

5.1. The Constrained Nonlinear BM-PTEQ Cost Function. This criterion follows from the constrained nonlinear optimisation problem as described in [12], which is modified for the complex-valued PTEQs criterion as

$$\max_{\tilde{\mathbf{p}}_{*,n}} \sum_{n \in N_d} \log_2 \left(1 + \frac{\text{SNR}_n}{\Gamma_n} \right), \quad (24)$$

with

$$\text{SNR}_n = \frac{\sigma_{\tilde{x}_n}^2}{E \left\{ \left| \tilde{x}_{k,n} - \tilde{\mathbf{p}}_{*,n}^H \tilde{\mathbf{y}}_{k,n} \right|^2 \right\}}, \quad (25)$$

subject to

$$\tilde{\mathbf{p}}_{*,n} = \frac{E \left\{ \tilde{\mathbf{y}}_{k,n}^H \tilde{x}_{k,n} \right\}}{E \left\{ \tilde{\mathbf{y}}_{k,n}^H \tilde{\mathbf{y}}_{k,n} \right\}} = \frac{\sum_{\tilde{x}}^H \tilde{y}_n}{\sum_{\tilde{y}}^2}, \quad \forall n \in N_d, \quad (26)$$

where $\tilde{x}_{k,n}$ is the k th transmitted DMT-symbol on tone n . The $\sigma_{\tilde{x}_n}^2 = E \{ |\tilde{x}_{k,n}|^2 \}$ is a variance and $\tilde{\mathbf{y}}_{k,n}$ is the k th unequalised $T \times 1$ symbol vector after sliding DFT at tone n . We aim to maximise the number of bits per DMT-symbol in (24) subject to the unconstrained MMSE PTEQ $\tilde{\mathbf{p}}_{*,n}$ in (26) with the subchannel SNR on n tone in (25).

A constrained optimisation criterion is typically restated as a cost minimisation

$$J(\tilde{\mathbf{p}}_{*,n}) = \sum_{n \in N_d} \log_2 \left(1 + \frac{\text{SNR}_n}{\Gamma_n} \right). \quad (27)$$

By means of the least squares criterion, the gradient of (27) with respect to PTEQs $\tilde{\mathbf{p}}_{*,n}$ can be rewritten compactly with an exponentially weighted over K DMT-symbols as (see also in the appendix)

$$\nabla_{\tilde{\mathbf{p}}_{*,n}} J = \sum_{n \in N_d} \sum_{k=1}^K \lambda^{K-k} \tilde{\mathbf{y}}_{k,n}^H \tilde{\mathbf{y}}_{k,n} e_{k,n}^*, \quad (28)$$

with

$$\tilde{\gamma}_{k,n} = \frac{\text{SNR}_n^2}{\sigma_{\tilde{x}_{k,n}}^2 (\Gamma_n + \text{SNR}_n)}, \quad (29)$$

$$e_{k,n} = E \left\{ \left| \tilde{x}_{k,n} - \tilde{\mathbf{p}}_{*,n}^H \tilde{\mathbf{y}}_{k,n} \right|^2 \right\},$$

where $\tilde{\gamma}_{k,n}$ is a tone-dependent weight and $e_{k,n}$ is the error on tone n at symbol k .

Hence, $\tilde{\gamma}_{k,n}$ is replaced by an instantaneous a priori estimate based on the previous parameter tap-weight estimate vector $\hat{\mathbf{p}}_{k-1,n}$ on tone n at symbol $k - 1$. Consequently, the tone-dependent weight estimate $\hat{\gamma}_{k,n}$ at tone n for each symbol k is given as

$$\hat{\gamma}_{k,n} = \frac{\widehat{\text{SNR}}_{k,n}^2}{\sigma_{\tilde{x}_{k,n}}^2 (\Gamma_n + \widehat{\text{SNR}}_{k,n})}, \quad (30)$$

where

$$\widehat{\text{SNR}}_{k,n} = \frac{\sigma_{\tilde{x}_n}^2}{\left| \tilde{x}_{k,n} - \hat{\mathbf{p}}_{k-1,n}^H \tilde{\mathbf{y}}_{k,n} \right|^2}. \quad (31)$$

The gradient in (28) is also applied to the nonlinear weighted problem with varying weight estimate $\hat{\gamma}_{k,n}$ and the instantaneous estimate SNR at each symbol k for n tone $\widehat{\text{SNR}}_{k,n}$. We note that the denominator of $\widehat{\text{SNR}}_{k,n}$ in (31) is equal to the MSE with the previous tap-weight estimate vector $\hat{\mathbf{p}}_{k-1,n}$ at the PTEQ output.

Therefore, a constrained nonlinear exponentially weighted least squares cost function for the complex-valued PTEQ tap-weight estimate vector $\hat{\mathbf{p}}_{k,n}$ is defined as

$$J_{NL}(\hat{\mathbf{p}}_{k,n}) = \sum_{n \in N_d} \frac{1}{2} \sum_{k=1}^K \lambda^{K-k} \hat{\gamma}_{k,n} |\tilde{e}_{k,n}|^2, \quad (32)$$

$$\tilde{e}_{k,n} = \tilde{x}_{k,n} - \hat{\mathbf{p}}_{k-1,n}^H \tilde{\mathbf{y}}_{k,n}, \quad (33)$$

where $\tilde{e}_{k,n}$ is the a priori estimate error at each DMT-symbol. With the nonlinear cost function in (32), an adaptive algorithm introduced in Section 5.2 can achieve the same performance as the BM-PTEQ cost function in (21) with these approximations in (30) and (31).

5.2. An Adaptive BM-PTEQ Algorithm. In this section, we introduce the methodology in solving the nonlinear cost function in (32) recursively at each symbol k based on an adaptive recursive Levenberg-Marquardt (RLM) algorithm updating of $T \times 1$ PTEQ tap-weight vector $\hat{\mathbf{p}}_{k,n}$ at tone n for $n \in N_d$. The iterative Levenberg Marquadt (LM) method is classical and well-known strategies for solving nonlinear batch optimisation problems. The recursive LM is definitely modified for adaptively solving nonlinear problems by earlier algorithms as the recursive identification system presented in [18] and neural network for nonlinear adaptive filter training described in [19].

The constrained nonlinear exponentially least squares cost criterion in (32) for a complex-valued tap-weight estimate PTEQ $\hat{\mathbf{p}}_{k,n}$ at DMT-symbol k on tone n is defined as

$$J(\hat{\mathbf{p}}_{k,n}) = \frac{1}{2} \sum_{k=1}^K \lambda^{K-k} \hat{\gamma}_{k,n} |\tilde{e}_{k,n}|^2, \quad (34)$$

where $\hat{\gamma}_{k,n}$ is a scalar of tone-dependent weight estimate as given in (30) and $\tilde{e}_{k,n}$ is the a priori estimate error as described in (33).

Following [18], a tap-weight estimate PTEQ $\hat{\mathbf{p}}_{k,n}$ can be obtained at each DMT-symbol k as

$$\hat{\mathbf{p}}_{k,n} = \hat{\mathbf{p}}_{k-1,n} + \check{\mathbf{R}}_{k,n}^{-1} \check{\mathbf{g}}_{k,n}, \quad (35)$$

where the gradient estimate $\check{\mathbf{g}}_{k,n}$ is derived by differentiating the cost function in (34) with respect to $\hat{\mathbf{p}}_{k,n}$ in (35) as

$$\check{\mathbf{g}}_{k,n} = \nabla_{\hat{\mathbf{p}}_{k,n}} J = \hat{\gamma}_{k,n} \check{\mathbf{Y}}_{k,n}^H \check{\mathbf{e}}_{k,n}^*. \quad (36)$$

Based on LM method [20], the regularised approximation Hessian $\check{\mathbf{R}}_{k,n}$ is reformed as

$$\check{\mathbf{R}}_{k,n} = \sum_{k=1}^K \lambda^{K-k} \left(\left\{ \hat{\gamma}_{k,n} \check{\mathbf{Y}}_{k,n} \check{\mathbf{Y}}_{k,n}^H \right\} + \delta_{k,n} \text{diag} \{ \mathbf{R}_{k,n} \} \right), \quad (37)$$

$$\mathbf{R}_{k,n} = \sum_{k=1}^K \lambda^{K-k} \hat{\gamma}_{k,n} \check{\mathbf{Y}}_{k,n} \check{\mathbf{Y}}_{k,n}^H, \quad (38)$$

where $\mathbf{R}_{k,n}$ is the approximation Hessian for the complexed PTEQ. The $\delta_{k,n}$ is the regularisation parameter at symbol k [19], in which this algorithm ensures the stability by taking the changing of the approximation Hessian over symbol into account. Hence, the regularised approximation Hessian $\check{\mathbf{R}}_{k,n}$ is regularised for stability reason by the second term in (37).

With the recursion method, the tap-weight estimate PTEQ vector $\hat{\mathbf{p}}_{k,n}$ is updated as

$$\hat{\mathbf{p}}_{k,n} = \hat{\mathbf{p}}_{k-1,n} + (1 - \lambda) \hat{\mathbf{R}}_{k,n}^{-1} \check{\mathbf{g}}_{k,n}, \quad (39)$$

where

$$\hat{\mathbf{R}}_{k,n} = \lambda \hat{\mathbf{R}}_{k-1,n} + (1 - \lambda) \left\{ \left\{ \hat{\gamma}_{k,n} \check{\mathbf{Y}}_{k,n} \check{\mathbf{Y}}_{k,n}^H \right\} + \delta_{k,n} \text{diag} \left\{ \hat{\gamma}_{k,n} \check{\mathbf{Y}}_{k,n} \check{\mathbf{Y}}_{k,n}^H \right\} \right\}, \quad (40)$$

where λ is the forgetting-factor, $0 < \lambda < 1$. The regularised approximation Hessian $\check{\mathbf{R}}_{k,n}$ in (37) is replaced by an exponentially weighted estimate approximation Hessian $\hat{\mathbf{R}}_{k,n}$ in (40).

5.2.1. The Modified Inverse Regularised Approximation Hessian Matrix. Unfortunately, the matrix inversion lemma cannot be used directly on the updating approximation Hessian $\hat{\mathbf{R}}_{k,n}$ in (40). So, we rearrange $\hat{\mathbf{R}}_{k,n}$

$$\hat{\mathbf{R}}_{k,n} = \lambda \hat{\mathbf{R}}_{k-1,n} + (1 - \lambda) \hat{\gamma}_{k,n} \left\{ \left\{ \check{\mathbf{Y}}_{k,n} \check{\mathbf{Y}}_{k,n}^H \right\} + \delta_{k,n} \text{diag} \left\{ \check{\mathbf{Y}}_{k,n} \check{\mathbf{Y}}_{k,n}^H \right\} \right\}, \quad (41)$$

by adding the $\boldsymbol{\varphi}_{k,n}$ matrix and $\boldsymbol{\psi}_{k,n}$ matrix into (41) (*The matrix inversion lemma*. Let A and B be two positive definite M -by- M matrices related by $A = B^{-1} + C \cdot D^{-1} \cdot C^H$, where D is a positive definite N -by- M matrix and C is an M -by- N matrix. We may express the inverse of the matrix A by $A^{-1} = B - BC(D + C^H BC)^{-1} C^H B$).

We then introduce how to define $\tilde{\mathbf{R}}_{k,n}$ as

$$\tilde{\mathbf{R}}_{k,n} = \lambda \tilde{\mathbf{R}}_{k-1,n} + (1 - \lambda) \hat{\gamma}_{k,n} \left\{ \boldsymbol{\psi}_{k,n} \boldsymbol{\varphi}_{k,n} \boldsymbol{\psi}_{k,n}^H \right\}, \quad (42)$$

where

$$\boldsymbol{\psi}_{k,n} = \left[\begin{array}{c|ccc} \check{\mathbf{Y}}_{k,n}^{(1)} & 0 & 0 & \cdots & 0 \\ \check{\mathbf{Y}}_{k,n}^{(2)} & 1 & 0 & \cdots & 0 \\ \check{\mathbf{Y}}_{k,n}^{(3)} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \check{\mathbf{Y}}_{k,n}^{(T)} & 0 & 0 & \cdots & 1 \end{array} \right], \quad (43)$$

$$\mathbf{Y}_{k,n} = \delta_{k,n} \text{diag} \left\{ \check{\mathbf{Y}}_{k,n} \check{\mathbf{Y}}_{k,n}^H \right\} = \begin{bmatrix} \Upsilon_{11} & \mathbf{0}^T \\ \mathbf{0} & \Upsilon_{22} \end{bmatrix}, \quad (44)$$

$$\boldsymbol{\varphi}_{k,n} = \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \Upsilon_{22} \end{bmatrix}, \quad (45)$$

where $\boldsymbol{\psi}_{k,n}$ denotes the $T \times T$ matrix. The Υ_{22} is the $(T - 1) \times (T - 1)$ block diagonal matrix. The size of zero vector $\mathbf{0}$ is of $1 \times (T - 1)$, and the size of the identity matrix $\tilde{\mathbf{I}}$ is of $(T - 1) \times (T - 1)$. Notice that $\mathbf{Y}_{k,n}$ in (44) and $\boldsymbol{\varphi}_{k,n}$ in (45) are the $T \times T$ block diagonal matrices. Hence, the $\boldsymbol{\varphi}_{k,n}$ is nonsingular, if and only if its inverse exists [21]. With the approximation Hessian $\tilde{\mathbf{R}}_{k,n}$ assumed to be positive definite and therefore nonsingular, we can apply the matrix inversion lemma to the modified approximation Hessian $\tilde{\mathbf{R}}_{k,n}$ in (42) instead of $\hat{\mathbf{R}}_{k,n}$ in (41).

We make the following identifications as $A = \tilde{\mathbf{R}}_{k,n}$, $B^{-1} = \lambda \tilde{\mathbf{R}}_{k-1,n}$, $C = \boldsymbol{\psi}_{k,n}$, $D^{-1} = (1 - \lambda) \hat{\gamma}_{k,n} \boldsymbol{\varphi}_{k,n}$. By substituting these definitions in the matrix inversion lemma, we then obtain the following recursive equation for the inverse of the modified approximation Hessian $\tilde{\mathbf{R}}_{k,n}$ as

$$\tilde{\mathbf{R}}_{k,n}^{-1} = \lambda^{-1} \tilde{\mathbf{R}}_{k-1,n}^{-1} - \lambda^{-1} \tilde{\mathbf{R}}_{k-1,n}^{-1} \tilde{\mathbf{K}}_{k,n} \boldsymbol{\psi}_{k,n}^H, \quad (46)$$

$$\tilde{\mathbf{K}}_{k,n} = \frac{\lambda^{-1} \tilde{\mathbf{R}}_{k-1,n}^{-1} \boldsymbol{\psi}_{k,n}}{\left\{ (1 - \lambda)^{-1} \hat{\gamma}_{k,n}^{-1} \boldsymbol{\varphi}_{k,n}^{-1} \right\} + \left\{ \lambda^{-1} \boldsymbol{\psi}_{k,n}^H \tilde{\mathbf{R}}_{k-1,n}^{-1} \boldsymbol{\psi}_{k,n} \right\}}, \quad (47)$$

where $\hat{\gamma}_{k,n}$ is a scalar of tone-dependent weight estimate as given in (30).

Consequently, the tap-weight estimate PTEQ vector $\hat{\mathbf{p}}_{k,n}$ can be computed as

$$\hat{\mathbf{p}}_{k,n} = \hat{\mathbf{p}}_{k-1,n} + (1 - \lambda) \tilde{\mathbf{R}}_{k,n}^{-1} \check{\mathbf{g}}_{k,n}, \quad (48)$$

where $\tilde{\mathbf{R}}_{k,n}^{-1}$ is introduced above in (46) and $\check{\mathbf{g}}_{k,n}$ is the gradient estimate in (36).

5.2.2. An Adaptive Inverse Square-Root Recursive Levenberg-Marquardt (iQR-RLM) Algorithm. We consider the Givens rotation-based adaptive inverse square-root (QR) algorithm. An adaptive inverse QR algorithm is a QR decomposition-based recursive least squares (QR-RLS) algorithm that is designed to obtain explicit weight extraction by working directly with the incoming data matrix via the QR decomposition [22]. Accordingly, the QR-RLS algorithm is numerically more stable than the standard RLS algorithm [23].

Notice that the modified inverse approximation Hessian $\tilde{\mathbf{R}}_{k,n}^{-1}$ in (46) is also derived in a similar fashion with the inverse correlation $\Phi_{k,n}^{-1}$ of RLS algorithm as described in [23]. Hence, the form of $\tilde{\mathbf{R}}_{k,n}^{-1}$ in (46) of RLM algorithm is similar to the inverse correlation $\Phi_{k,n}^{-1}$ of RLS algorithm. We then introduce the Givens rotation-based adaptive inverse QR algorithm, which can be applied for $\tilde{\mathbf{R}}_{k,n}^{-1}$ of RLM algorithm for computing the PTEQ tap-weight estimate $\hat{\mathbf{p}}_{k,n}$ at symbol k for tone $n \in N_d$.

For convenience of computation, let

$$\begin{aligned} \mathbf{D}_{k,n} &\triangleq \tilde{\mathbf{R}}_{k,n}^{-1} \\ \mathbf{z}_{k,n} &= \left\{ (1-\lambda)^{-1} \hat{\gamma}_{k,n}^{-1} \boldsymbol{\varphi}_{k,n}^{-1} \right\} + \left\{ \lambda^{-1} \boldsymbol{\psi}_{k,n}^H \mathbf{D}_{k-1,n} \boldsymbol{\psi}_{k,n} \right\}. \end{aligned} \quad (49)$$

Using these definitions in (49), we may rewrite $\tilde{\mathbf{R}}_{k,n}^{-1}$ (46) as

$$\mathbf{D}_{k,n} = \lambda^{-1} \mathbf{D}_{k-1,n} - \lambda^{-1} \mathbf{D}_{k-1,n} \boldsymbol{\psi}_{k,n} \mathbf{z}_{k,n}^{-1} \boldsymbol{\psi}_{k,n}^H \lambda^{-1} \mathbf{D}_{k-1,n}. \quad (50)$$

There are 4-matrix terms that constitute the right-hand side of (50), we may introduce the 2×2 block matrix \mathbf{G} as

$$\mathbf{G} = \begin{bmatrix} \mathbf{z}_{k,n} & \lambda^{-1} \boldsymbol{\psi}_{k,n}^H \mathbf{D}_{k-1,n} \\ \lambda^{-1} \mathbf{D}_{k-1,n} \boldsymbol{\psi}_{k,n} & \lambda^{-1} \mathbf{D}_{k-1,n} \end{bmatrix}. \quad (51)$$

We then redefine the block matrix \mathbf{G} in (51) using the Cholesky factorisation as

$$\begin{aligned} \mathbf{G} &= \mathbb{A} \mathbb{A}^H, \\ \mathbb{A} &= \begin{bmatrix} (1-\lambda)^{-1/2} \hat{\gamma}_{k,n}^{-1/2} \boldsymbol{\varphi}_{k,n}^{-1/2} & \lambda^{-1/2} \boldsymbol{\psi}_{k,n}^H \mathbf{D}_{k-1,n}^{1/2} \\ \mathbf{0} & \lambda^{-1/2} \mathbf{D}_{k-1,n}^{1/2} \end{bmatrix}, \end{aligned} \quad (52)$$

where $\mathbf{0}$ is the null vector, the prearray \mathbb{A} is an upper triangular matrix and $\mathbf{D}_{k-1,n}$ indicates with its factor

$$\mathbf{D}_{k-1,n} = \mathbf{D}_{k-1,n}^{1/2} \mathbf{D}_{k-1,n}^{H/2}. \quad (53)$$

We may set the prearray \mathbb{A} to resulting the postarray \mathbb{B} transformation for iQR-RLM algorithm using the matrix factorisation lemma as

$$\begin{aligned} \mathbb{A} \Theta &= \mathbb{B}, \\ \begin{bmatrix} (1-\lambda)^{-1/2} \hat{\gamma}_{k,n}^{-1/2} \boldsymbol{\varphi}_{k,n}^{-1/2} & \lambda^{-1/2} \boldsymbol{\psi}_{k,n}^H \mathbf{D}_{k-1,n}^{1/2} \\ \mathbf{0} & \lambda^{-1/2} \mathbf{D}_{k-1,n}^{1/2} \end{bmatrix} \Theta & \\ &= \begin{bmatrix} \mathbf{z}_{k,n}^{1/2} & \mathbf{0}^T \\ \tilde{\mathbb{K}}_{k,n} \mathbf{z}_{k,n}^{1/2} & \mathbf{D}_{k,n}^{1/2} \end{bmatrix}, \end{aligned} \quad (54)$$

where Θ is a unitary rotation and $\tilde{\mathbb{K}}_{k,n}$ is described in (47) (*The matrix factorisation lemma*. Given any \mathbf{A} and \mathbf{B} $n \times m$ matrices with dimension $n \leq m$, this lemma states by following [23] as $\mathbf{A} \Theta \Theta^H \mathbf{A}^H = \mathbf{B} \mathbf{B}^H$, if and only if, there exists a unitary matrix Θ such that $\mathbf{A} \Theta = \mathbf{B}$ and $\Theta \Theta^H = \mathbf{I}$).

We note that $\mathbf{D}_{k,n}^{1/2}$ in the right-hand side of (54) is the lower triangular matrix. In virtue of the product of square-

root matrix its Hermitian transpose

$$\mathbf{D}_{k,n} = \mathbf{D}_{k,n}^{1/2} \mathbf{D}_{k,n}^{H/2} \quad (55)$$

is always nonnegative matrix as derived in [24].

Therefore, the tap-weight estimate PTEQ vector $\hat{\mathbf{p}}_{k,n}$ based on iQR-RLM algorithm can be performed

$$\hat{\mathbf{p}}_{k,n} = \hat{\mathbf{p}}_{k-1,n} + (1-\lambda) \mathbf{D}_{k,n} \tilde{\mathbf{g}}_{k,n}, \quad (56)$$

where $\mathbf{D}_{k,n}$ is defined in (55) and $\tilde{\mathbf{g}}_{k,n}$ is the gradient estimate in (36).

5.2.3. The Adaptive Regularisation Parameter. Both the convergence rate and stability are affected by a suitable choice of the regularisation parameter $\delta_{k,n}$ such that a small $\delta_{k,n}$ could cause the RLM algorithm to be unstable, while a large $\delta_{k,n}$ could deduce slow convergence [18]. So the parameter $\delta_{k,n}$ should be adapted during convergence. An adaptive regularisation parameter algorithm based on the instantaneous estimates of the predicted and actual cost criterion reduction is proposed in [19]. Hence, we apply this algorithm for an adaptive iQR-RLM algorithm as explained below.

Following [19], the predicted instantaneous cost reduction $\tilde{r}_{p,k,n}$ of the criterion in (34) for each update of iQRRLM-based algorithm (56) is computed as

$$\tilde{r}_{p,k,n} = (1-\lambda) \left[\hat{\gamma}_{k,n} \tilde{\mathbf{Y}}_{k,n}^H \tilde{\mathbf{e}}_{k,n}^* \right]^H \mathbf{D}_{k,n} \left[\hat{\gamma}_{k,n} \tilde{\mathbf{Y}}_{k,n}^H \tilde{\mathbf{e}}_{k,n}^* \right], \quad (57)$$

$$\tilde{\mathbf{e}}_{k,n} = \tilde{\mathbf{x}}_{k,n} - \hat{\mathbf{p}}_{k-1,n}^H \tilde{\mathbf{Y}}_{k,n}, \quad (58)$$

where $\hat{\gamma}_{k,n}$ is a scalar of tone-dependent weight estimate as given in (30). The error $\tilde{\mathbf{e}}_{k,n}$ is a priori estimate error, and $\mathbf{D}_{k,n}$ is the inverse of modified approximation Hessian in (55).

The actual instantaneous cost reduction $\tilde{r}_{a,k,n}$ is determined by using a priori estimate error $\tilde{\mathbf{e}}_{k,n}$ in (58) and a posteriori estimate error $\tilde{\xi}_{k,n}$ as

$$\tilde{r}_{a,k,n} = \hat{\gamma}_{k,n} \left\{ |\tilde{\mathbf{e}}_{k,n}|^2 - |\tilde{\xi}_{k,n}|^2 \right\}, \quad (59)$$

$$\tilde{\xi}_{k,n} = \tilde{\mathbf{x}}_{k,n} - \hat{\mathbf{p}}_{k,n}^H \tilde{\mathbf{Y}}_{k,n}.$$

Then, the values for $\delta_{k,n}$ can be adapted using the following criterion.

- (i) Increase $\delta_{k-1,n}$ by a factor of α if $\tilde{r}_{a,k,n}/\tilde{r}_{p,k,n}$ is smaller than a threshold ζ .
- (ii) Decrease $\delta_{k-1,n}$ by a factor of $1/\alpha$ if $\tilde{r}_{a,k,n}/\tilde{r}_{p,k,n}$ is larger than a threshold $1 - \zeta$.

The adaptive regularisation parameter $\delta_{k,n}$ method is summarised as

$$\delta_{k,n} = \begin{cases} \alpha \cdot \delta_{k-1,n}, & \text{if } \tilde{r}_{a,k,n} < \zeta \tilde{r}_{p,k,n}, \\ \frac{1}{\alpha} \cdot \delta_{k-1,n}, & \text{if } \tilde{r}_{a,k,n} > (1-\zeta) \tilde{r}_{p,k,n}, \\ \delta_{k-1,n}, & \text{otherwise,} \end{cases} \quad (60)$$

where $0 < \zeta < 0.5$ and a typical value is of 0.25.

Therefore, the iQR-RLM algorithm for BM-PTEQ using adaptive regularisation method is summarised as described in Algorithm 1.

Starting with the soft-constrained initialisation as: $\hat{\mathbf{p}}(0) = \mathbf{0}$

For $n \in N_d$, $n = 1, 2, \dots$, compute.

for $k = 1, 2, \dots, K$

(1) To arrange the block diagonal matrices $\Psi_{k,n}$, $\mathbf{Y}_{k,n}$ and $\Phi_{k,n}$ as:

$$\Psi_{k,n} = \begin{bmatrix} \tilde{\mathbf{y}}_{k,n}^{(1)} & 0 & 0 & \cdots & 0 \\ \tilde{\mathbf{y}}_{k,n}^{(2)} & 1 & 0 & \cdots & 0 \\ \tilde{\mathbf{y}}_{k,n}^{(3)} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \tilde{\mathbf{y}}_{k,n}^{(T)} & 0 & 0 & \cdots & 1 \end{bmatrix},$$

$$\mathbf{Y}_{k,n} = \delta_{k,n} \text{diag}\{\tilde{\mathbf{y}}_{k,n} \tilde{\mathbf{y}}_{k,n}^H\} = \begin{bmatrix} \Upsilon_{11} & \mathbf{0} \\ \mathbf{0} & \Upsilon_{22} \end{bmatrix},$$

$$\Phi_{k,n} = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \Upsilon_{22} \end{bmatrix},$$

$$\text{where } \tilde{\mathbf{y}}_{k,n} = [\tilde{\mathbf{y}}_{k,n}^{(1)} \tilde{\mathbf{y}}_{k,n}^{(2)} \cdots \tilde{\mathbf{y}}_{k,n}^{(T)}]^T.$$

(2) To compute $\widehat{\text{SNR}}_{k,n}$ and $\hat{\gamma}_{k,n}$ as:

$$\widehat{\text{SNR}}_{k,n} = \frac{\sigma_{\tilde{x}_n}^2}{|\tilde{\mathbf{x}}_{k,n} - \hat{\mathbf{p}}_{k-1,n}^H \tilde{\mathbf{y}}_{k,n}|^2},$$

$$\hat{\gamma}_{k,n} = \frac{\widehat{\text{SNR}}_{k,n}^2}{\sigma_{\tilde{x}_n}^2 (\Gamma_n + \widehat{\text{SNR}}_{k,n})}.$$

(3) To compute $\mathbf{D}_{k,n}$ as:

$$\mathbb{A} = \begin{bmatrix} (1-\lambda)^{-1/2} \hat{\gamma}_{k,n}^{-1/2} \Phi_{k,n}^{-1/2} & \lambda^{-1/2} \Psi_{k,n}^H \mathbf{D}_{k-1,n}^{1/2} \\ \mathbf{0} & \lambda^{-1/2} \mathbf{D}_{k-1,n}^{1/2} \end{bmatrix},$$

$$\mathbb{A} \Theta = \begin{bmatrix} \mathbb{B}_{11} & \mathbf{b}_{12} \\ \mathbf{b}_{21} & \mathbb{B}_{22} \end{bmatrix}, \quad \text{where } \Theta \text{ is a unitary rotation,}$$

$$\mathbf{D}_{k,n} = \mathbb{B}_{22} \mathbb{B}_{22}^H.$$

(4) To compute $\hat{\mathbf{p}}_{k,n}$ as:

$$\hat{\mathbf{p}}_{k,n} = \hat{\mathbf{p}}_{k-1,n} + (1-\lambda) \mathbf{D}_{k,n} \tilde{\mathbf{g}}_{k,n},$$

$$\text{where } \tilde{\mathbf{g}}_{k,n} = \hat{\gamma}_{k,n} \tilde{\mathbf{y}}_{k,n}^H \tilde{\mathbf{e}}_{k,n}^*,$$

$$\tilde{\mathbf{e}}_{k,n} = \tilde{\mathbf{x}}_{k,n} - \hat{\mathbf{p}}_{k-1,n}^H \tilde{\mathbf{y}}_{k,n}.$$

(5) To compute $\delta_{k,n}$ as:

$$\delta_{k,n} = \begin{cases} \alpha \cdot \delta_{k-1,n} & \text{if } \tilde{r}_{a_{k,n}} < \zeta \tilde{r}_{p_{k,n}}, \\ \frac{1}{\alpha} \cdot \delta_{k-1,n} & \text{if } \tilde{r}_{a_{k,n}} > (1-\zeta) \tilde{r}_{p_{k,n}}, \\ \delta_{k-1,n} & \text{otherwise,} \end{cases}$$

$$\text{where } \tilde{r}_{p_{k,n}} = (1-\lambda) [\hat{\gamma}_{k,n} \tilde{\mathbf{y}}_{k,n}^H \tilde{\mathbf{e}}_{k,n}^*]^H \mathbf{D}_{k,n} [\hat{\gamma}_{k,n} \tilde{\mathbf{y}}_{k,n}^H \tilde{\mathbf{e}}_{k,n}^*],$$

$$\tilde{r}_{a_{k,n}} = \hat{\gamma}_{k,n} \{ |\tilde{\mathbf{e}}_{k,n}|^2 - |\tilde{\xi}_{k,n}|^2 \},$$

$$\tilde{\xi}_{k,n} = \tilde{\mathbf{x}}_{k,n} - \hat{\mathbf{p}}_{k,n}^H \tilde{\mathbf{y}}_{k,n}.$$

end
end

ALGORITHM 1: Summary of the proposed adaptive iQRRLM-based BM-PTEQ.

6. Simulation Results

In this section, we performed transmission simulations for the ADSL downstream including AWGN and NEXT over the entire test channel. The used tones were starting at tones 38 to 255, and the unused tones were set to zero. The bit allocation calculation requires an estimate of SNR on tone $n \in N_d$, when the noise energy is estimated after per-tone equalisation. The carrier serving area (CSA) loop nos. 1–8 were used for the test channel, which comprises 512 coefficients of CIR. The length of CP (ν) was 32. The SNR gap of 9.8 dB, the coding gain of 4.2 dB, the noise margin of 6 dB, and the input signal power of -40 dBm/Hz were used for all active tones. The AWGN with a power of -140 dBm/Hz and NEXT coming from 12 ADSL disturbers were included. All simulations were done for $T = 32$, $f_s = 2.208$ MHz, and $N = 512$.

We compare the proposed MMSE-based BM-PTEQ with iterative method, the proposed adaptive BM-PTEQ with adaptive iQRRLM-based design, with the RLM-based PTEQ approach [6], with other BM-TEQ such as BM-TEQ with iterative scheme [11] and with the recursive method [12]. The BM-TEQ was initialised with $\mathbf{w} = [1 \ 0 \ \cdots \ 0]^T$, as presented in [11]. The proposed iQRRLM-based BM-PTEQ can be computed with the soft-constrained initialisation. The regularisation parameter δ_k of adaptive RLM-based PTEQ [6], adaptive RLM-based BM-TEQ [12], and proposed adaptive iQRRLM-based BM-PTEQ were initialised at $\delta_0 = 10^{-3}$ for all active tones. The forgetting-factor λ of the adaptive RLM-based PTEQ [6], the RLM-based adaptive BM-TEQ (ABM-TEQ) [12], and the proposed adaptive iQRRLM-based BM-PTEQ (ABM-PTEQ) were increased from $\lambda = 0.95$ during the first 150 update-symbols to $\lambda = 0.99$ for the remaining updated symbols. The adaptation parameter α of δ_k of the proposed iQRRLM-based adaptive BM-PTEQ was fixed at $\alpha = 2$.

Figure 1 depicts that the learning curves of bit rate convergence of all adaptive algorithms as a function of the number of updated DMT-symbols for the samples of CSA loop no. 1, no. 2, no. 4 and no. 5. The proposed iQR-RLM adaptive BM-PTEQ (ABM-PTEQ) is compared with the RLM-based adaptive BM-TEQ (ABM-TEQ) [12]. The bit rate of the RLM-based adaptive BM-TEQ [12] curves closely to reach the maximum bit rate of BM-TEQ [11]. Meanwhile, the learning curve of proposed adaptive BM-PTEQ with iQRRLM-based algorithm converges nearly to the truly MMSE-based BM-PTEQ. Approximately, 100 updated symbols are appeared to converge to steady-state condition for the proposed iQRRLM-based adaptive BM-PTEQ. The curve of proposed iQRRLM-based adaptive BM-PTEQ has slower convergence than the RLM-based adaptive BM-TEQ. The adaptive RLM-based PTEQ has the slowest convergence. In [11], the performance of BM-TEQ has shown closely to PTEQ and the learning curve of adaptive RLM-based BM-TEQ compared with adaptive RLM-based PTEQ [6] in both these figures reveal to confirm.

Figure 2 illustrates the bit rate as a function of synchronisation delay Δ of T -tap complexed equalisers for the samples of CSA loop no. 1, no. 2, no. 4, and no. 5, when the numbers of taps of equalisers equal 32 ($T = 32$). The proposed BM-PTEQ and ABM-PTEQ are compared with the BM-TEQ [11] design. It is noticed that the proposed ABM-PTEQ performance has the same direction with the proposed BM-PTEQ design along the number of increasing delay for all samples of CSA loop. The performance of the BM-TEQ confirms that its bit rate has been smooth as a function of delay, as presented in [11]. The proposed ABM-PTEQ and BM-PTEQ appear to give higher bit rate than BM-TEQ design for a given range of synchronisation delay.

Figure 3 reveals the bit rate performance of the proposed MMSE-based BM-PTEQ and adaptive iQRRLM-based BM-PTEQ (ABM-PTEQ) for all CSA loop nos. 1–8 at starting tones 38 to 255 ADSL downstream when the fixed delay equals 45 ($\Delta = 45$). The performance of proposed ABM-PTEQ is compared with BM-TEQ [11] and adaptive RLM-based BM-TEQ (ABM-TEQ) [12]. It is shown that the proposed ABM-PTEQ is similar to the performance of BM-PTEQ approach. The bit rate of proposed ABM-PTEQ can be improved as compared to the BM-TEQ and the ABM-TEQ design.

7. Conclusion

In this paper, we present the BM-PTEQ design with the nonlinear bit rate maximising cost function. The proposed BM-PTEQ cost function is derived from the exact subchannel SNR model at the FEQ outputs. Since, the solution to achieve the BM-PTEQ criterion is exactly the same form of that of the BM-TEQ, we conclude that the BM-PTEQ can always perform better than or equal to the BM-TEQ in the sense of bit rate maximising performance. For achievable BM-PTEQ in practice, we then introduce the methodology of adaptive inverse-QR RLM-based BM-PTEQ design by the nonlinear bit rate maximising cost criterion. The proposed BM-PTEQ and iQRRLM-based ABM-PTEQ can ensure the performance of maximum bit rate. Simulation results with several ADSL parameters show that the proposed BM-PTEQ and iQRRLM-based ABM-PTEQ are able to improve superior bit rate performance as compared with BM-TEQ and ABM-TEQ design for all CSA loop.

Appendix

The constrained optimisation criterion is given as

$$J(\tilde{\mathbf{p}}_{*,n}) = \sum_{n \in N_d} \log_2 \left(1 + \frac{\text{SNR}_n}{\Gamma_n} \right). \quad (\text{A.1})$$

The derivation of the gradient of (A.1) with respect to PTEQ $\tilde{\mathbf{p}}_{*,n}$ is

$$\begin{aligned}
\nabla_{\tilde{\mathbf{p}}_{*,n}} J &= \sum_{n \in N_d} \frac{\partial}{\partial \tilde{\mathbf{p}}_{*,n}} \left\{ \log_2 \left(1 + \frac{\text{SNR}_n}{\Gamma_n} \right) \right\} \\
&= \sum_{n \in N_d} \left(\frac{\Gamma_n}{\Gamma_n + \text{SNR}_n} \right) \frac{\partial}{\partial \tilde{\mathbf{p}}_{*,n}} \left(1 + \frac{\text{SNR}_n}{\Gamma_n} \right) \\
&= \sum_{n \in N_d} \left(\frac{\sigma_{\tilde{x}_n}^2}{\Gamma_n + \text{SNR}_n} \right) \frac{\tilde{\mathbf{y}}_{k,n}^H e_{k,n}^*}{\left(E \left\{ \left| \tilde{x}_{k,n} - \tilde{\mathbf{p}}_{*,n}^H \tilde{\mathbf{y}}_{k,n} \right|^2 \right\} \right)^2} \\
&= \sum_{n \in N_d} \frac{\text{SNR}_n^2}{\sigma_{\tilde{x}_n}^2 (\Gamma_n + \text{SNR}_n)} \tilde{\mathbf{y}}_{k,n}^H e_{k,n}^* \\
&= \sum_{n \in N_d} \tilde{\gamma}_{k,n} \tilde{\mathbf{y}}_{k,n}^H e_{k,n}^*,
\end{aligned} \tag{A.2}$$

where SNR_n is the SNR on tone n , and $\tilde{\gamma}_{k,n}$ is a tone-dependent weight at symbol k on tone n

$$\begin{aligned}
\text{SNR}_n &= \frac{\sigma_{\tilde{x}_n}^2}{E \left\{ \left| \tilde{x}_{k,n} - \tilde{\mathbf{p}}_{*,n}^H \tilde{\mathbf{y}}_{k,n} \right|^2 \right\}}, \\
\tilde{\gamma}_{k,n} &= \frac{\text{SNR}_n^2}{\sigma_{\tilde{x}_n}^2 (\Gamma_n + \text{SNR}_n)}, \\
e_{k,n} &= E \left\{ \left| \tilde{x}_{k,n} - \tilde{\mathbf{p}}_{*,n}^H \tilde{\mathbf{y}}_{k,n} \right| \right\}.
\end{aligned} \tag{A.3}$$

The gradient in (A.2) of the constrained optimisation criterion in (A.1) with respect to PTEQ $\tilde{\mathbf{p}}_{*,n}$ can be expressed with the exponentially weighted over K DMT-symbols as

$$\nabla_{\tilde{\mathbf{p}}_{*,n}} J = \sum_{n \in N_d} \sum_{k=1}^K \lambda^{K-k} \tilde{\gamma}_{k,n} \tilde{\mathbf{y}}_{k,n}^H e_{k,n}^*, \tag{A.4}$$

where λ is an exponential weighting factor or forgetting factor.

Acknowledgment

This work was supported by the Shell Centennial Education Fund, Thailand.

References

- [1] R. Baldemair and P. Frenger, "A time-domain equalizer minimizing intersymbol and intercarrier interference in DMT systems," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '01)*, vol. 1, pp. 381–385, 2001.
- [2] S. Sitjongsatoporn and P. Yuvapoositanon, "An adaptive step-size order statistic time domain equaliser for discrete multi-tone systems," in *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS '07)*, pp. 1333–1336, New Orleans, LA, USA, May 2007.
- [3] T. Pollet, M. Peeters, M. Moonen, and L. Vandendorpe, "Equalization for DMT-based broadband modems," *IEEE Communications Magazine*, vol. 38, no. 5, pp. 106–113, 2000.
- [4] K. Van Acker, G. Leus, M. Moonen, O. van de Wiel, and T. Pollet, "Per tone equalization for DMT-based systems," *IEEE Transactions on Communications*, vol. 49, no. 1, pp. 109–119, 2001.
- [5] G. Ysebaert, K. Vanbleu, G. Cuyppers, M. Moonen, and T. Pollet, "Combined RLS-LMS initialization for per tone equalizers in DMT-receivers," *IEEE Transactions on Signal Processing*, vol. 51, no. 7, pp. 1916–1927, 2003.
- [6] S. Sitjongsatoporn and P. Yuvapoositanon, "Recursive Levenberg-Marquardt per-tone equalisation for discrete multi-tone systems," in *Proceedings of the 3rd International Symposium on Communications, Control, and Signal Processing (ISCCSP '08)*, pp. 1062–1066, St. Julians, Malta, March 2008.
- [7] K. Van Acker, G. Leus, M. Moonen, and T. Pollet, "RLS-based initialization for per-tone equalizers in DMT receivers," *IEEE Transactions on Communications*, vol. 51, no. 6, pp. 885–889, 2003.
- [8] K. Van Acker, G. Leus, M. Moonen, and T. Pollet, "Frequency domain equalization with tone grouping in DMT/ADSL-receivers," in *Proceedings of the Conference Record of the Asilomar Conference on Signals, Systems and Computers (Asilomar '99)*, vol. 2, pp. 1067–1070, 1999.
- [9] K. Vanbleu, G. Ysebaert, G. Cuyppers, and M. Moonen, "Bitrate maximizing per group equalization for DMT-based systems," *Signal Processing*, vol. 86, no. 10, pp. 2952–2965, 2006.
- [10] P. K. Pandey and M. Moonen, "Resource allocation in ADSL variable length per-tone equalizers," *IEEE Transactions on Signal Processing*, vol. 56, no. 5, pp. 2161–2164, 2008.
- [11] K. Vanbleu, G. Ysebaert, G. Cuyppers, M. Moonen, and K. Van Acker, "Bitrate-maximizing time-domain equalizer design for DMT-based systems," *IEEE Transactions on Communications*, vol. 52, no. 6, pp. 871–876, 2004.
- [12] K. Vanbleu, G. Ysebaert, G. Cuyppers, and G. Leus, "Adaptive bitrate maximizing TEQ design for DMT-based systems," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '04)*, vol. 4, pp. 1057–1060, May 2004.
- [13] K. Van Acker, *Equalization and echo cancellation for DMT-based DSL modems*, Ph.D. thesis, Katholieke Universiteit Leuven, Leuven, Belgium, 2001.
- [14] N. Al-Dhahir and J. M. Cioffi, "Optimum finite-length equalization for multicarrier transceivers," *IEEE Transactions on Communications*, vol. 44, no. 1, pp. 56–64, 1996.
- [15] W. Henkel and T. Kessler, "Maximizing the channel capacity of multicarrier transmission by suitable adaptation of the time-domain equalizer," *IEEE Transactions on Communications*, vol. 48, no. 12, pp. 2000–2004, 2000.
- [16] J. M. Cioffi, G. P. Dudevoir, M. V. Eyuboglu, and G. D. Forney Jr., "MMSE decision-feedback equalizers and coding—part I: equalization results," *IEEE Transactions on Communications*, vol. 43, no. 10, pp. 2582–2594, 1995.
- [17] E. de Carvalho and D. T. M. Slock, "Burst mode equalization: optimal approach and suboptimal continuous-processing approximation," *Signal Processing*, vol. 80, no. 10, pp. 1999–2015, 2000.
- [18] L. Ljung and T. Söderström, *Theory and Practice of Recursive Identification*, Cambridge, Mass, USA, MIT Press, 1983.
- [19] L. S. H. Ngia and J. Sjöberg, "Efficient training of neural nets for nonlinear adaptive filtering using a recursive Levenberg-Marquardt algorithm," *IEEE Transactions on Signal Processing*, vol. 48, no. 7, pp. 1915–1927, 2000.
- [20] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific*

Computing, Cambridge University Press, Cambridge, UK, 1993.

- [21] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1999.
- [22] J. Ma, K. K. Parhi, and E. F. Deprettere, "A unified algebraic transformation approach for parallel recursive and adaptive filtering and svd algorithms," *IEEE Transactions on Signal Processing*, vol. 49, no. 2, pp. 424–437, 2001.
- [23] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, Upper Saddle River, NJ, USA, 1996.
- [24] A. H. Sayed and T. Kailath, "A state-space approach to adaptive RLS filtering," *IEEE Signal Processing Magazine*, vol. 11, no. 3, pp. 18–60, 1994.