Hindawi Publishing Corporation
EURASIP Journal on Wireless Communications and Networking
Volume 2007, Article ID 14683, 12 pages
doi:10.1155/2007/14683

Research Article

Minimum Probability of Error-Based Equalization Algorithms for Fading Channels

Janos Levendovszky,[1] Lorant Kovacs,[1] and Edward C. van der Meulen[2]

[1] *Department of Telecommunications, Budapest University of Technology and Economics, Magyar tudósok körútja 2, 1117 Budapest, Hungary*
[2] *Department of Mathematics, Katholieke Universiteit Leuven, Celestijnenlaan 200B, 3001 Leuven (Heverlee), Belgium*

Received 12 December 2006; Revised 14 March 2007; Accepted 29 April 2007

Recommended by George K. Karagiannidis

Novel channel equalizer algorithms are introduced for wireless communication systems to combat channel distortions resulting from multipath propagation. The novel algorithms are based on newly derived bounds on the probability of error (PE) and guarantee better performance than the traditional zero forcing (ZF) or minimum mean square error (MMSE) algorithms. The new equalization methods require channel state information which is obtained by a fast adaptive channel identification algorithm. As a result, the combined convergence time needed for channel identification and PE minimization still remains smaller than the convergence time of traditional adaptive algorithms, yielding real-time equalization. The performance of the new algorithms is tested by extensive simulations on standard mobile channels.

Copyright © 2007 Janos Levendovszky et al. 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

Since radio spectrum became scarce and expensive, one of the major concerns of wireless communication is to maximize spectral efficiency (SE). This implies that broadband services are implemented over narrowband radio channels which makes them susceptible to selective fading due to multipath propagation which may yield severe performance degradation [1]. As a result, efficient channel equalization techniques prove to be instrumental to combat intersymbol interference (ISI) in order to avoid large scale drops in system performance.

The effect of interferences is especially crucial in mobile communication systems which have two evolutionary paths: (i) 3G systems are launched based on WCDMA [2]; and (ii) the current 2G systems (GSM and IS-136) are updated to provide broadband services [1]. The latter strategy introduces a novel common physical layer "enhanced data rates for GSM evolution" (EDGE) for both TDMA schemes. EDGE improves spectral efficiency by applying 8PSK modulation format instead of binary Gaussian minimum-shift keying (GMSK). For the sake of seamless GSM-EDGE transfer, most of the system parameters remain unchanged (e.g., symbol time and symbol duration). However, in the case of 8PSK modulation the maximum likelihood sequence estimator (involving the Viterbi algorithm) can no longer be implemented on the current DSP technology due to its complexity [2]. As a result, fast channel equalizer algorithms have to be developed which are simple enough to run on the currently available hardware architectures even in the case of multilevel PSK modulation schemes.

This paper aims at developing small complexity channel equalizer algorithms by directly minimizing the PE instead of minimizing the mean square error (MSE) or the peak distortion (PD) [3]. Unfortunately, the direct minimization of PE with respect to the equalizer coefficients is of exponential complexity. Thus, we develop new bounds on which basis the equalizer coefficients can be optimized by fast algorithms. For the sake of simplicity, the novel algorithms presented in the paper are treated assuming a two-state modulation scheme (however the analysis can be easily extended to many-state schemes by introducing complex variables).

The first attempt to derive an equalizer based on the minimum PE strategy can be found in the work of Shimbo and Celebiler [4]. The optimal equalizer coefficients were only sought by exhaustive search, thus real-time adaptivity was not guaranteed. In recent years, some new results have been developed for minimum PE equalization. In [5] a low-complexity adaptive algorithm is proposed for 2 or 4-state modulation systems but the convergence is rather slow,

while in [6, 7] near minimum PE equalization is carried out by radial basis function neural networks which considerably increases the equalizer complexity. Minimum BER equalization for decision feedback equalizers can be found in [8, 9]. On the other hand, very complex equalizer schemes have been proposed for DS-CDMA systems in [10–12].

Recently other equalization strategies have also been developed, such as iterative (turbo) equalization algorithms which jointly optimize the equalization and the detection yielding similar performance than the maximum likelihood sequence estimation. For some recent results see [13–16]. Other solutions are based on the negentropy minimization principle [17] or the nearest neighbor classifier [18] but they yield complex algorithms.

The results are given in the following structure.

(i) In Section 2, the communication model will be outlined.
(ii) In Section 3, PE is expressed as a function of the equalizer coefficients and a gradient-based algorithm is introduced for minimization. Then new bounds are derived on PE to develop new equalizer algorithms with low complexity.
(iii) In Section 4, the performance and convergence properties of the new equalizer algorithms are analyzed numerically.

2. THE MODEL

To describe single-user communication over fading channel, we use the so-called equivalent discrete time white noise filter model (for further details see [3]).

The corresponding quantities are defined as follows:

(i) $y_k \in \{-1, 1\}$ denotes the transmitted information bit at time instant k being a sequence of identically distributed independent Bernoulli random variables with $P(y_k = 1) = P(y_k = -1) = 0.5$;
(ii) the discrete impulse response of the channel is denoted by h_k, $k = 0, \ldots, M$, where M denotes the span of ISI;
(iii) the noise is denoted by ν_k and is assumed to be a stationary zero mean white Gaussian random sequence with constant spectral density N_0;
(iv) the received sequence is denoted by x_k, which is linearly distorted and noisy version of the transmitted sequence given as

$$x_k = \sum_{j=0}^{M} h_j y_{k-j} + \nu_k, \tag{1}$$

and with the assumption of BPSK modulation and coherent demodulation, x_k is real;
(i) the equalizer is a linear FIR filter, the output of which is denoted by $\tilde{y}_k$

$$\tilde{y}_k = \sum_{i=0}^{J} w_i x_{k-i}, \tag{2}$$

where w_i, $i = 0, \ldots, J$, denote the free parameters of the equalizer which are subject to further optimization;
(ii) the decision is carried out by threshold detection in a symbol-by-symbol fashion:

$$\hat{y}_k = \operatorname{sgn}\{\tilde{y}_k\} = \operatorname{sgn}\left\{\sum_{i=0}^{J} w_i x_{k-i}\right\}; \tag{3}$$

(iii) the overall channel impulse response function is determined by the cascade of the channel and the equalizer

$$q_k = \sum_{i=0}^{L} h_i w_{k-i}, \quad k = 0, \ldots, L, \tag{4}$$

where $L = M + J$ denotes the support of the overall impulse response.

Traditional equalization algorithms aimed at minimizing the PD defined as

$$\mathbf{w}_{\text{opt}} : \min_{\mathbf{w}} \sum_{i=1}^{L} |q_i| \tag{5}$$

or MSE defined as

$$\mathbf{w}_{\text{opt}} : \min_{\mathbf{w}} \mathrm{E}\left[\left(\mathbf{y}_k - \sum_{j=0}^{J} w_j x_{k-j}\right)^2\right]. \tag{6}$$

The corresponding adaptive equalizer algorithm that minimizes the PD is called zero-forcing:

$$w_l(k+1) = w_l(k) - \gamma\left(y_k - \sum_{j=0}^{J} w_j x_{k-j}\right) y_{k-l} \tag{7}$$

and that which minimizes the MSE (often referred to as LMS) is

$$w_l(k+1) = w_l(k) - \gamma\left(y_k - \sum_{j=0}^{J} w_j x_{k-j}\right) x_{k-l}, \tag{8}$$

where γ is a sufficiently small step-size which governs the convergence. Both approaches involved linear stochastic approximation schemes but they fell short of efficient estimation as the goal functions did not have direct relationship with PE.

3. NOVEL CHANNEL EQUALIZATION METHODS

In this section, we express PE as a function of the equalizer coefficients $\mathbf{w}$ and we also demonstrate that equalization with respect to direct PE minimization is of exponential complexity. In order to circumvent this difficulty, we develop new bounds on PE and the equalization coefficients can be optimized by minimizing these bounds in real time.

3.1. Weight optimization subject to minimizing the PE

Since our approach to equalization is based on minimizing the probability of error, first we express PE as a function of the equalizer coefficients as given in [4]:

$$P_E(\mathbf{w}) = \frac{1}{2^L}\sum_{\mathbf{y}\in\mathcal{Y}}\Phi\left(\frac{-q_0+\sum_{l=1}^{L} q_l y_l}{\sigma}\right) = \frac{1}{2^L}\sum_{\mathbf{y}\in\mathcal{Y}}\Phi\left(\frac{\sum_{l=0}^{L} q_l y_l}{\sigma}\right), \tag{9}$$

where $\Phi(\cdot)$ denotes the standard normal cumulative distribution function (cdf) defined as $\Phi(x) = (1/\sqrt{2\pi})\int_{-\infty}^{x}\exp(-t^2/2)\mathrm{d}t$, $\sigma^2 = N_0\sum_{j=0}^{J} w_j^2$, and $\mathcal{Y} = \{\mathbf{y} = (y_0, y_1, \ldots, y_L) \mid y_0 = -1;\ y_i \in \{-1, 1\},\ i = 1, \ldots, L\}$. Substituting (4) into (9), we obtain

$$P_E(\mathbf{w}) = \frac{1}{2^L}\sum_{\mathbf{y}\in\mathcal{Y}}\Phi\left(\frac{\sum_{n=0}^{J} w_n \sum_{l=n}^{M+n} h_{l-n} y_l}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}}\right). \tag{10}$$

To find the optimal weights of the equalizer which minimize this error probability, we have to solve the following equation:

$$\mathbf{w}_{\mathrm{opt}} : \mathrm{grad}_{\mathbf{w}}\, P_E(\mathbf{w}) = \mathbf{0}, \tag{11}$$

where the ith component of the gradient is

$$\frac{\partial P_E(\mathbf{w})}{\partial w_i} = \frac{1}{2^L\sqrt{2\pi N_0\left(\sum_{n=0}^{J} w_n^2\right)^3}}\sum_{\mathbf{y}\in\mathcal{Y}}\exp\left(\frac{-\left(\sum_{n=0}^{J} w_n \sum_{l=n}^{M+n} h_{l-n} y_l\right)^2}{2N_0\sum_{n=0}^{J} w_n^2}\right) \cdot\left[\left(\sum_{n=0}^{J} w_n^2\right)\cdot\left(\sum_{l=i}^{M+i} h_{l-i} y_l\right) - w_i\left(\sum_{n=0}^{J} w_n \sum_{l=n}^{M+n} h_{l-n} y_l\right)\right]. \tag{12}$$

The weights can be minimized by gradient descent, which yields the following equalization algorithm:

$$w_i(k+1) = w_i(k) - \gamma\frac{\partial P_E(\mathbf{w}(k))}{\partial w_i}. \tag{13}$$

Here $\mathbf{w}(k)$ is the value of the weight vector at the kth iteration and γ is a sufficiently small step-size.

In the forthcoming discussion, this procedure is termed as true gradient search (TGS). Unfortunately, performing TGS is computationally prohibitive because of the summation over an exponentially growing number of vectors in expression (12). This summation must be calculated in each step of algorithm (13). Thus, TGS can only be applied in practice if the support of the overall impulse response defined in (4) is very limited. Otherwise, near-optimal algorithms must be sought which lend themselves to real-time implementations. To ease this complexity, new bounds are derived on PE.

3.2. New bounds on PE

In this section, we derive new upperbounds on PE which can be used for channel equalization. For the sake of performance analysis, lower bounds are also derived which can help to evaluate the accuracy of the bounds.

To develop these upperbounds, first we introduce the concept of "diagonal dominancy" (or "eye-openness") which is often used in the literature related to digital communication theory [3, 4].

Definition 1. A sequence $\{a_k\}$ is called an "eye-opened" if $|a_0| > \sum_{k,k\neq 0}|a_k|$.

It should be noted that if a sequence a_k is "eye-opened," then the associated Toeplitz matrix $\mathbf{A}$, defined by $A_{ij} = a_{i-j}$, is diagonally dominant (i.e., $|A_{ii}| > \sum_{j,j\neq i}|A_{ij}|$).

Definition 2. The peak distortion (PD) of a linear filter with impulse response function a_k is defined as

$$\mathrm{PD} = \sum_{k,k\neq 0}|a_k|. \tag{14}$$

In the forthcoming discussion, the PD is related to the overall impulse response function q_k of the communication system given in (4), which is calculated as follows:

$$\mathrm{PD}(\mathbf{w}) = \sum_{k,k\neq 0}|q_k| = \sum_{k,k\neq 0}\left|\sum_{j=0}^{J} w_j h_{k-j}\right|. \tag{15}$$

The appearance of $\mathbf{w}$ in (15) in the notation for PD is due to the dependence of PD on the weights of the equalizer. Note that if q_k is eye-opened, then $\mathrm{PD}(\mathbf{w}) < |q_0|$.

Theorem 1. *The following bounds on $P_E(\mathbf{w})$ can be derived, where the inequalities provided with a star hold under the assumption of the "eye-openness" of the overall channel response function q_k (when $|q_0| > \sum_{k=1}^{L}|q_k|$):*

$$\Phi\left(\frac{-q_0}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}}\right) \overset{*}{\leq} P_E(\mathbf{w}) \leq \Phi\left(\frac{-q_0+\mathrm{PD}(\mathbf{w})}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}}\right), \tag{16a}$$

$$\frac{1}{2^{L+1}}\left[\Phi\left(\frac{-q_0+\mathrm{PD}(\mathbf{w})}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}}\right)+\Phi\left(\frac{-q_0-\mathrm{PD}(\mathbf{w})}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}}\right)\right] + \left(1-\frac{1}{2^L}\right)\Phi\left(\frac{-q_0}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}}\right) \overset{*}{\leq} P_E(\mathbf{w}) \overset{*}{\leq} \frac{1}{2}\left[\Phi\left(\frac{-q_0+\mathrm{PD}(\mathbf{w})}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}}\right)+\Phi\left(\frac{-q_0-\mathrm{PD}(\mathbf{w})}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}}\right)\right], \tag{16b}$$

$$\Phi\left(\frac{-h_0}{\sqrt{N_0}}\right) \overset{*}{\leq} P_E(\mathbf{w}). \tag{16c}$$

The proof of this theorem can be found in Appendix A.

It should be noted that the upperbound in (16b) is tighter than the upperbound in (16a) due to the relation

$$\Phi\left(\frac{-q_0 - \mathrm{PD}(\mathbf{w})}{\sqrt{N_0 \sum_{n=0}^{J} w_n^2}}\right) \le \Phi\left(\frac{-q_0 + \mathrm{PD}(\mathbf{w})}{\sqrt{N_0 \sum_{n=0}^{J} w_n^2}}\right) \tag{17}$$

which implies

$$\begin{aligned}\frac{1}{2}&\left[\Phi\left(\frac{-q_0 + \mathrm{PD}(\mathbf{w})}{\sqrt{N_0 \sum_{n=0}^{J} w_n^2}}\right) + \Phi\left(\frac{-q_0 - \mathrm{PD}(\mathbf{w})}{\sqrt{N_0 \sum_{n=0}^{J} w_n^2}}\right)\right] \\ &\le \Phi\left(\frac{-q_0 + \mathrm{PD}(\mathbf{w})}{\sqrt{N_0 \sum_{n=0}^{J} w_n^2}}\right).\end{aligned} \tag{18}$$

One must not forget, however, that upperbound (16b) is obtained under a stronger condition (eye-openness) than upperbound (16a), thus the latter one can be used in more general circumstances.

Unfortunately, due to the nondifferentiability of $\mathrm{PD}(\mathbf{w})$ it is still difficult to minimize the newly obtained bounds with respect to the weight vector. By using the Cauchy-Schwartz inequality to upperbound $\mathrm{PD}(\mathbf{w})$, differentiable bounds can be derived on $P_E(\mathbf{w})$.

Theorem 2. *The following additional bounds on $P_E(\mathbf{w})$ can be derived, where the inequality* (19b) *(provided with a star) holds under the assumption of "eye-openness" of the overall channel response function:*

$$\begin{aligned}&P_E(\mathbf{w}) \le \Phi(G(\mathbf{w})), \\ &\textit{where } G(\mathbf{w}) = \frac{-q_0 + \sqrt{L \sum_{l=1}^{L} \left(\sum_{j=0}^{M} w_j h_{l-j}\right)^2}}{\sqrt{N_0 \sum_{n=0}^{J} w_n^2}},\end{aligned} \tag{19a}$$

$$\begin{aligned}&P_E(\mathbf{w}) \overset{*}{\le} Q(\mathbf{w}) \\ &\textit{where } Q(\mathbf{w}) = \frac{1}{2}\left[\Phi\left(\frac{-q_0 + \sqrt{L \sum_{l=1}^{L} \left(\sum_{j=0}^{M} w_j h_{l-j}\right)^2}}{\sqrt{N_0 \sum_{n=0}^{J} w_n^2}}\right)\right. \\ &\qquad \left. + \Phi\left(\frac{-q_0 - \sqrt{L \sum_{l=1}^{L} \left(\sum_{j=0}^{M} w_j h_{l-j}\right)^2}}{\sqrt{N_0 \sum_{n=0}^{J} w_n^2}}\right)\right].\end{aligned} \tag{19b}$$

Note that bounds (16a) and (16b) are *tighter* than (19b) and (19a) (due to the application of the Cauchy-Schwartz inequality in the latter ones). On the other hand, the advantage of (19b) and (19a) is that both $G(\mathbf{w})$ and $Q(\mathbf{w})$ are *differentiable* functions with respect to the weights, which can give rise to gradient-based equalization algorithms.

The proof of this theorem can be found in Appendix B.

3.3. Channel equalization by minimizing the bounds on PE

Channel equalization is performed by recursively minimizing the new bounds with respect to weights. Since upperbounds (16a) and (16b) are *nondifferentiable* with respect to the weights (because of the absolute values occurring in the corresponding formulas due to $\mathrm{PD}(\mathbf{w})$), the bounds derived in Theorem 2 are used for equalization. The new equalizer algorithms are obtained by minimizing bounds (19a) and (19b) by gradient search.

Bound-based equalization algorithm related to bound (19a) *(BBEAd)*

$$w_i(k+1) = w_i(k) - \gamma \frac{\partial G(\mathbf{w})}{\partial w_i}, \tag{20}$$

where γ is a sufficiently small step-size and

$$\begin{aligned}\frac{\partial G(\mathbf{w})}{\partial w_i} &= \frac{-\delta_{i0} h_0 + 2\sqrt{L} \dfrac{\left(\sum_{k=0}^{M} h_k^2 - \delta_{i0} h_0\right) w_i + \sum_{l=0}^{M} h_l \sum_{j=0,\, j\ne i}^{J} w_j h_{i+l-j}}{\sqrt{\sum_{l=1}^{M} \left(\sum_{n=0}^{J} w_n h_{l-n}\right)^2}}}{\sqrt{N_0 \sum_{n=0}^{J} w_n^2}} \\ &\quad - 2 w_i \frac{-w_0 h_0 + \sqrt{L}\sqrt{\sum_{l=1}^{L} \left(\sum_{n=0}^{J} w_n h_{l-n}\right)^2}}{\sqrt{N_0} \left(\sum_{n=0}^{J} w_n^2\right)^{3/2}}.\end{aligned} \tag{21}$$

This procedure is obtained from minimizing $G(\mathbf{w})$ by gradient search where $G(\mathbf{w})$ is defined in Theorem 2. Since $\Phi(\cdot)$ is monotone, it is enough to minimize $G(\mathbf{w})$.

Bound-based equalization algorithm related to bound (19b) *(BBEAe)*

$$w_i(k+1) = w_i(k) - \gamma \frac{\partial Q(\mathbf{w})}{\partial w_i}, \tag{22}$$

where γ is a sufficiently small step-size and

$$\begin{aligned}\frac{\partial Q(\mathbf{w})}{\partial w_i} &= \frac{1}{2\sqrt{2\pi}} \exp\left(\frac{-q_0 + \sqrt{L\left(\sum_{l=1}^{L} \sum_{j=0}^{J} w_j h_{l-j}\right)^2}}{\sqrt{N_0 \sum_{n=0}^{J} w_n^2}}\right) \\ &\quad \cdot \left\{\frac{-\delta_{i0} h_0}{\sqrt{N_0 \sum_{n=0}^{J} w_n^2}}\right. \\ &\qquad + \frac{2\sqrt{L} \dfrac{\left(\sum_{k=0}^{M} h_k^2 - \delta_{i0} h_0\right) w_i + \sum_{l=0}^{M} h_l \sum_{j=0,\, j\ne i}^{J} w_j h_{i+l-j}}{\sqrt{\sum_{l=1}^{M} \left(\sum_{n=0}^{J} w_n h_{l-n}\right)^2}}}{\sqrt{N_0 \sum_{n=0}^{J} w_n^2}} \\ &\qquad \left. - 2 w_i \frac{-w_0 h_0 + \sqrt{L}\sqrt{\sum_{l=1}^{L} \left(\sum_{n=0}^{J} w_n h_{l-n}\right)^2}}{\sqrt{N_0} \left(\sum_{n=0}^{J} w_n^2\right)^{3/2}}\right\}\end{aligned}$$

$$+\frac{1}{2\sqrt{2\pi}}\exp\left(\frac{-q_0-\sqrt{L\left(\sum_{l=1}^{L}\sum_{j=0}^{J}w_jh_{l-j}\right)^2}}{\sqrt{N_0\sum_{n=0}^{J}w_n^2}}\right)$$

$$\cdot\left\{\frac{-\delta_{i0}h_0}{\sqrt{N_0\sum_{n=0}^{J}w_n^2}}\right.$$

$$-\frac{2\sqrt{L}\dfrac{\left(\sum_{k=0}^{M}h_k^2-\delta_{i0}h_0\right)w_i+\sum_{l=0}^{M}h_l\sum_{j=0,\,j\neq i}^{J}w_jh_{i+l-j}}{\sqrt{\sum_{l=1}^{M}\left(\sum_{n=0}^{J}w_nh_{l-n}\right)^2}}}{\sqrt{N_0\sum_{n=0}^{J}w_n^2}}$$

$$\left.-2w_i\frac{-w_0h_0-\sqrt{L}\sqrt{\sum_{l=1}^{L}\left(\sum_{n=0}^{J}w_nh_{l-n}\right)^2}}{\sqrt{N_0}\left(\sum_{n=0}^{J}w_n^2\right)^{3/2}}\right\}. \tag{23}$$

One must note that the advantage of minimizing bounds (19a) and (19b) is that in the gradients of $G(\mathbf{w})$ and $Q(\mathbf{w})$ there is no summation over an exponentially growing set. In this way a much faster equalization can be obtained by applying algorithm (20) or (22) than (13).

3.4. Obtaining channel-state information

In order to run the proposed algorithms, channel-state information is needed (the channel impulse response function h_k appears in expressions (13) and (20)). There are plenty of real-time adaptive channel identification algorithms [1] which provide fast and simple channel-state information. In this paper, we identify the channel with an adaptive FIR filter, the coefficients of which are updated as follows:

$$g_n(k+1)=g_n(k)-\Delta\left(x_k-\sum_{i=0}^{M}g_iy_{k-i}\right)y_{k-n}, \tag{24}$$

where symbols y_k come from a sufficiently large training sequence: $\tau^K=(y_k,x_k)$, $k=1,2,\ldots,K$, where y_k denotes the transmitted sequence (known at the receiver side prior to start of the real communication), while x_k is the observed input at the receiver. Algorithm (24) minimizes the MSE between the unknown channel impulse response function h_i, $i=1,\ldots,M$, and the FIR filter coefficients g_i, $i=0,1,2,\ldots,M$. Here x_k denotes the received sequence at the output of the channel. In stable state, algorithm (24) provides weights for which $g_i=h_i$ in mean square if the degree of the FIR filter is larger than the channel impulse response (overmodeling).

It is noteworthy that the adaptive channel identifier (24) converges rather fast to the true channel-state because of the narrow eigenvalue-spectrum of the underlying matrices (for further details see [3]). Hence, the combination of identification and equalization can provide real-time solutions for low PE reception of digital information.

4. NUMERICAL RESULTS

In this section, a detailed performance analysis is given to evaluate the PE achieved by the different equalization methods and comparing their convergence speed and algorithmic complexities.

4.1. Channel characteristics and channel-state information

The simulations were made in the case of four different channel models representing multipath propagation in different practical scenarios. The corresponding channel characteristics are given by their impulse response as follows: $\mathbf{h}^{(1)}=[1;0.6;-0.3]^T$, $\mathbf{h}^{(2)}=[1;0.3;-0.2;0.3;-0.1;0.1]^T$, $\mathbf{h}^{(3)}=[1;0.6;-0.45]^T$, and $\mathbf{h}^{(4)}=[1.2;1.1;-0.2]^T$.

One must note that $\mathbf{h}^{(3)}$ and $\mathbf{h}^{(4)}$ are non-minimumphase channels. In this case, PE can be decreased by introducing a delay D into the equalization in the following way: (i) instead of (3) the decision is carried out by $\hat{y}_{k-D}=\operatorname{sgn}\{\tilde{y}_k\}=\operatorname{sgn}\{\sum_{i=0}^{J}w_ix_{k-i}\}$, (ii) and bound (19a) (see Theorem 2) must be modified by substituting

$$G(\mathbf{w})=\frac{-q_D+\sqrt{L\sum_{l=0;\,l\neq D}^{L}\left(\sum_{j=0}^{M}w_jh_{l-j}\right)^2}}{\sqrt{N_0\sum_{n=0}^{J}w_n^2}}. \tag{25}$$

In order to achieve the best performance, one should choose $D=0$ in the case of minimumphase channels, or $D=J$ in the case of non-minimumphase channels.

As far as the channel-state information is concerned, we investigated two scenarios:

(i) at first the exact channel-state information (the impulse response of the channel) was assumed to be known at the receiver side. Therefore the equalizer algorithms were run by using the corresponding $\mathbf{h}$ vector;

(ii) secondly, no channel-state information was assumed to be available at the receiver side, thus channel equalization was preceded by an adaptive channel identifier algorithm given in (24).

4.2. The PE versus SNR

In this section, we numerically investigate PE with respect to SNR. The performance was analyzed by having 2 up to 8 equalizer coefficients. In the case of the TGS, BBEAd, and BBEAe algorithms, the weight vector of the equalizer was normalized by setting $\mathbf{w}^T\mathbf{w}=1$, since PE is invariant to the normalization. The step-size of the gradient descent algorithms was not changed during the optimization and it falls into the interval of $10^{-4}\cdots10^{-1}$ for TGS, $5\cdot10^{-5}\cdots10^{-2}$ for BBEAd and BBEAe, and $0.01\cdots0.2$ for AMBER in the case of different channels and different SNRs, respectively. For the sake of comparison, the exact PE was calculated by formula (9) using the exact channel-state information.

The PE-SNR curves are plotted for the different channels by Figures 1, 2, 3, and 4, respectively. In the case of

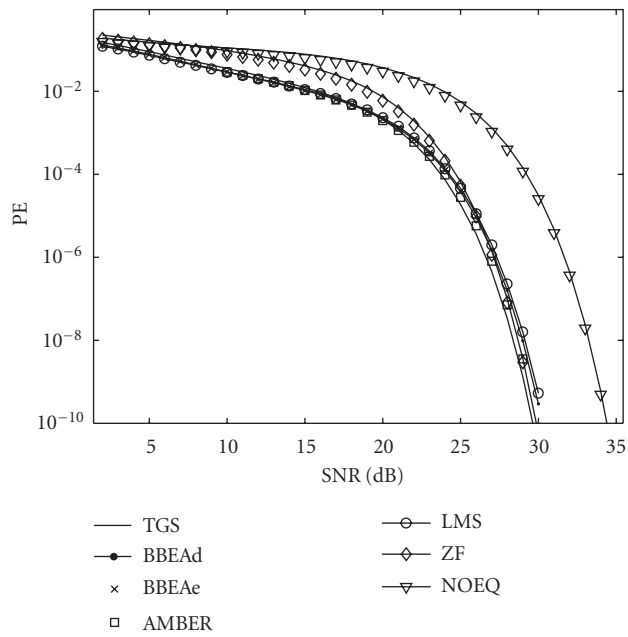

Figure 1: PE versus SNR performance of the different methods in the case of channel $\mathbf{h}^{(1)}$ and 3 equalizer coefficients.

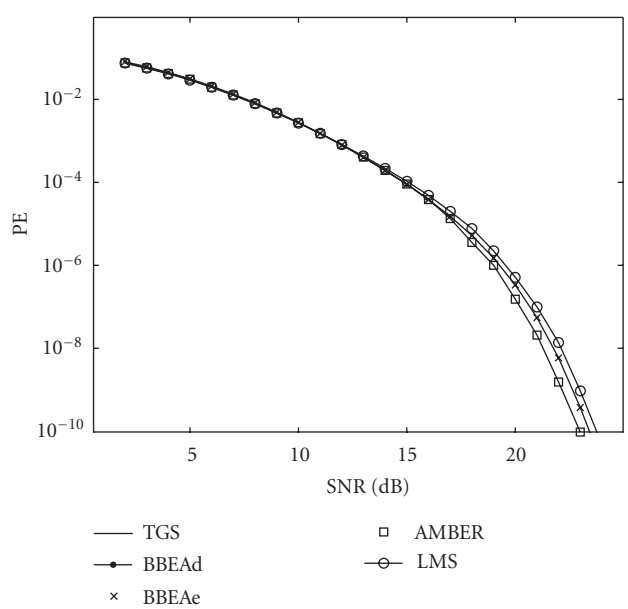

Figure 2: PE versus SNR performance of the different methods in the case of channel $\mathbf{h}^{(2)}$ and 6 equalizer coefficients.

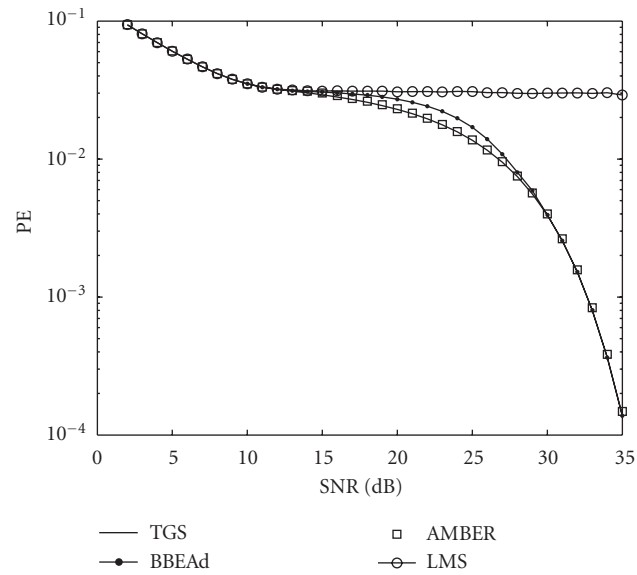

Figure 3: PE versus SNR performance of the different methods in the case of channel $\mathbf{h}^{(3)}$, 3 equalizer coefficients and $D = 2$.

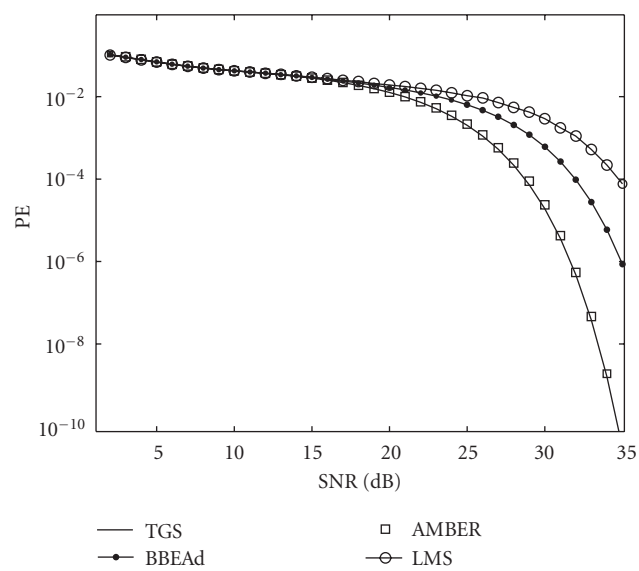

Figure 4: PE versus SNR performance of the different methods in the case of channel $\mathbf{h}^{(4)}$, 3 equalizer coefficients and $D = 2$.

non-minimumphase channels (Figures 3 and 4) the new methods far outperform the classical ones, while in the case of minimumphase channels the benefit is not so large. The best results were obtained by the TGS method which yields a 1–6 dB gain in SNR related to the traditional solutions. Furthermore, Figure 3 depicts such an example when traditional algorithms cannot provide better performance even though with increasing SNR, while the new methods are capable to further decrease PE. The BBEAd algorithm gets very close to the performance of TGS, but it runs much faster (due to the newly derived bounds on PE). It is noteworthy that TGS needs exponential complexity in each step to calculate the gradient of PE by the exact summation according to formula (10). Therefore, TGS can only be applied in practice if the support of the overall impulse response (channel impulse response convolved with the equalizer impulse response, defined in (4)) is very limited. This, in turn, puts severe limitations on the number of equalizer coefficients when using TGS. This argument also prompts the use of

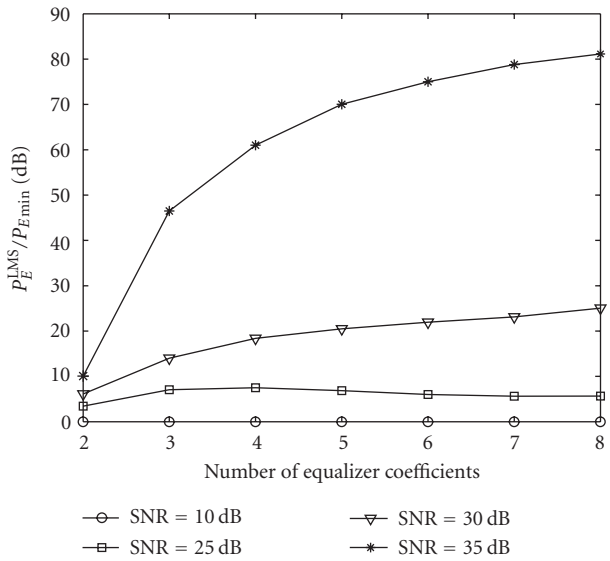

FIGURE 5: The ratio $P_E^{\mathrm{LMS}}/P_{E\min}$ versus the number of equalizer coefficients for channel $\mathbf{h}^{(3)}$.

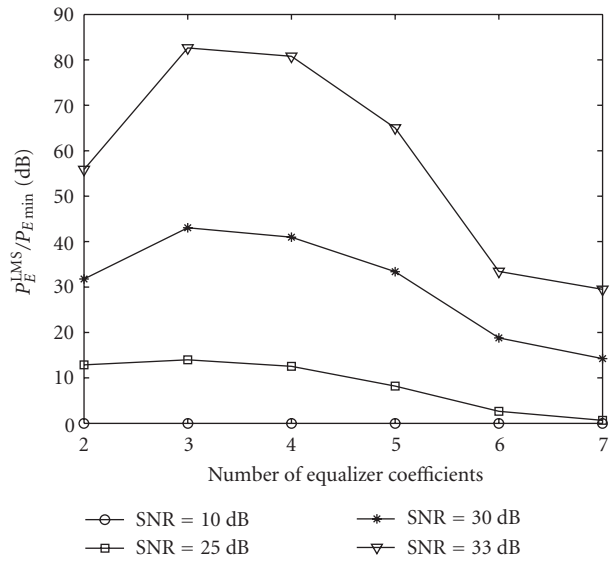

FIGURE 6: The ratio $P_E^{\mathrm{LMS}}/P_{E\min}$ versus the number of equalizer coefficients for channel $\mathbf{h}^{(4)}$.

the new algorithms where the complexity is not exponential with respect to the support of the overall impulse response. The performance of BBEAe is very close to BBEAd for minimumphase channels, which is explained by the fact that among the two terms in bound (19b) one term dominates the other, in the case of small noise. As a result, bound (19b) converges to bound (19a). When the noise is large, then all algorithms will yield similar performance, which are demonstrated by Figures 1–4. Furthermore BBEAe does not converge in the case of non-minimumphase ones. For the sake of comparison we also plotted the PE-SNR curve of the AMBER algorithm (for details see [5]), which exhibits almost the same performance, as TGS. On the other hand, the traditional equalizer methods (ZF and LMS) yield significantly worse performance than the new bounds. It is noteworthy, however, that in the case of minimumphase channels the performance of the LMS method can come close to the minimum PE solution.

In Figures 5 and 6, the $P_E^{\mathrm{LMS}}/P_{E\min}$ ratio is depicted with respect to the number of equalizer coefficients in the case of different SNR values, for two different channels. One can see that, on the one hand (in the case of $\mathbf{h}^{(3)}$), the difference between the performance increases in favor of the minimum PE solution as the number of equalizer coefficients grow. On the other hand, in the case of $\mathbf{h}^{(4)}$, the performance of LMS and the minimum PE solution converges to each other as the number of equalizer coefficients grow. Hence, the gain obtained by increasing the number of equalizer coefficients depends on the type of channel to be equalized.

4.3. Convergence time and numerical complexity

In this section, the convergence properties of the obtained algorithms are analyzed in comparison with their numerical complexity.

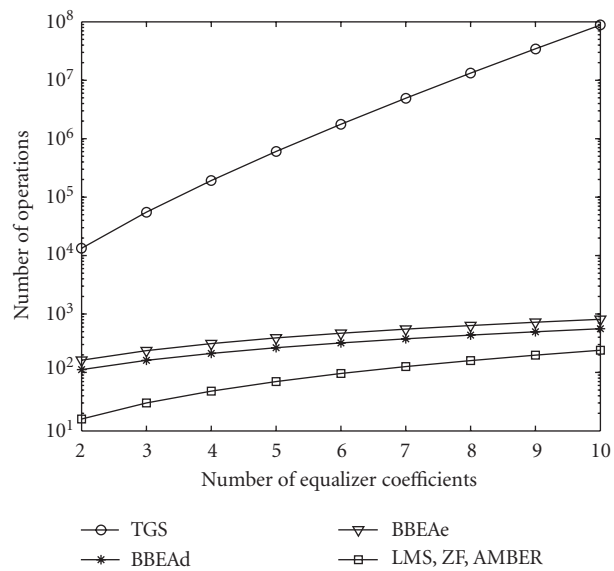

FIGURE 7: Number of operations required for a single update of the equalizer coefficients in the case of 5 channel coefficients.

The numerical complexity is measured by the number of additions and multiplications required for a single update of the equalizer coefficients. The complexity of different algorithms are depicted by Figure 7 in the case of 5 channel coefficients. Note that TGS needs exponential number of summations, while the other methods are much simpler providing real-time equalization.

The convergence properties of the different algorithms in the case of SNR = 30 dB are compared in Figure 8, where the convergence time is averaged over channel characteristics $\mathbf{h}^{(1)}$, $\mathbf{h}^{(2)}$, $\mathbf{h}^{(3)}$, and shown in the case of two and six equalizer

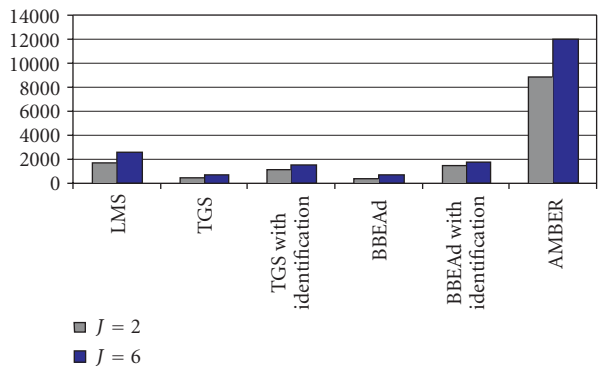

Figure 8: Convergence time of the different algorithms in the case of 2 and 6 equalizer coefficients.

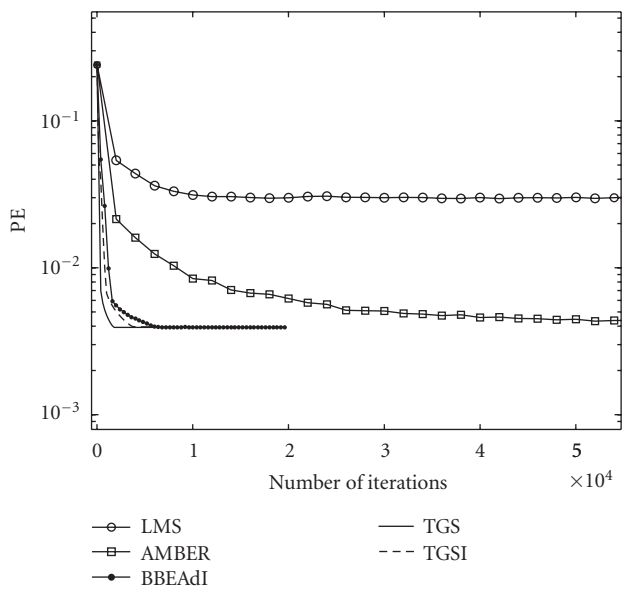

Figure 9: PE versus the number of iterations in the case of channel $\mathbf{h}^{(3)}$ and 30 dB SNR (10 runs are averaged); symbol "I" in the legend refers to *the case of applying* a plugged-in channel identifier.

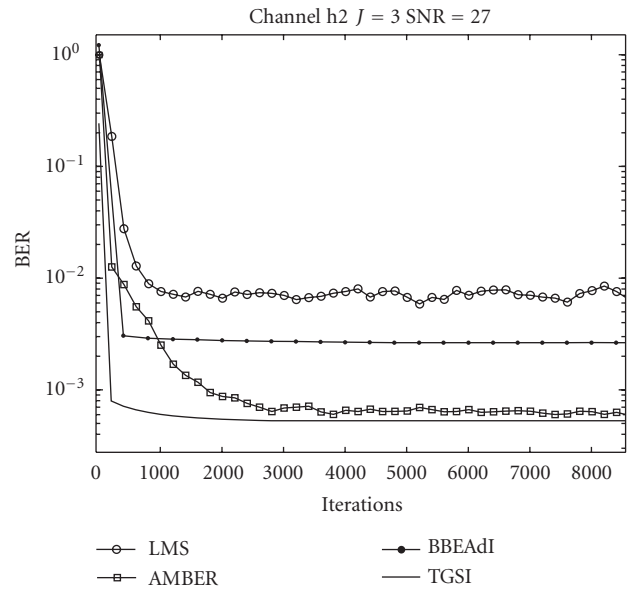

Figure 10: PE versus the number of iterations in the case of channel $\mathbf{h}^{(4)}$ and 27 dB SNR (10 runs are averaged); symbol "I" in the legend refers to *the case of applying* a plugged-in channel identifier.

coefficients. The convergence time is measured by the number of iterations from the initial state to the one, where the relative changes of PE will not exceed 5%. All the equalizer algorithms were started from the same initial state of ($\mathbf{w} = [1, 0, \ldots, 0]^T$) and the step-sizes of the algorithms were optimized empirically. In the case of Figures 9 and 10, we set $\gamma = 0.1$ for TGS and TGSI, $\gamma = 0.002$ for BBEAd, $\gamma = 0.01$ for LMS, and $\gamma = 0.2$ for AMBER, respectively. In the case of AMBER, we set the learning threshold to $\tau = 0.5$ providing increased convergence speed (for further details see [5]).

Fast convergence can still be maintained in the case of unknown channel-state, when an adaptive channel identification precedes the equalization algorithms. We started the identification and the equalizer algorithms at the same time instant and all updates of the equalizer were calculated applying the actual estimation of the channel. We used the simple traditional channel identification algorithm given in (24). The number of the channel coefficients was assumed to be known. Convergence curves for channels $\mathbf{h}^{(3)}$ and $\mathbf{h}^{(4)}$ for a given SNR with adaptive channel identification can be seen in Figures 9 and 10.

One must note that the convergence time of the proposed algorithms are about 10 times smaller than the traditional algorithms or the AMBER algorithm (which also minimizes PE but its convergence is apparently much slower). This justifies the use of this algorithms in real-time, high-data speed applications.

5. CONCLUSIONS

In this paper, novel channel equalizer algorithms have been developed based on newly derived bounds on PE. Due to the simplicity of the bounds, fast equalization algorithms can be obtained, the performance of which are close to optimum. Since these bounds need channel state information, the equalizer is preceded by an adaptive channel identifier. The combined convergence of channel identification and the new bound-based equalization is still much faster than other algorithms (e.g., AMBER, ZF, or LMS). The new methods yielded better performance than the traditional ZF and LMS equalizer algorithms. The operational complexity of the new bound-based algorithms is also low, requiring very simple calculations similarly to AMBER, ZF, or LMS. These benefits make the new algorithms suitable for real-time applications. In Table 1, the properties of the new and traditional algorithms are compared.

Table 1: Comparison of the new and traditional algorithms.

Algorithms	Performance (PE)	Adaptivity	Speed of convergence	Order of operational complexity/iterations (J: length of equalizer, M: length of channel)
TGS	Excellent	Adaptivity is ensured by additional channel identifier (short training sequence is needed)	Fast (even with adaptive channel identifier)	$O(3MJ^2 \cdot 2^{M+J})$
BBEAd, BBEAe	Good	Adaptivity is ensured by additional channel identifier (short training sequence is needed)	Fast (even with adaptive channel identifier)	$O(J^2 + 5MJ)$
AMBER	Excellent	Large training sequence required adaptation	Slow	$O(2J^2)$
ZF, LMS	Poor	Large training sequence required adaptation	Slow	$O(2J^2)$

APPENDICES

A. PROOF OF THEOREM 1

First we prove the right-hand side inequality of (16a), which can be deduced from (10). Since

$$\begin{aligned}\sum_{n=0}^{J} w_n \sum_{l=n}^{M+n} h_{l-n} y_l &= -q_0 + \sum_{n=1}^{J} w_n \sum_{l=n}^{M+n} h_{l-n} y_l \\ &= -q_0 + \sum_{l=n}^{M+n} \left(\sum_{n=1}^{J} w_n h_{l-n} \right) y_l \\ &\le -q_0 + \mathrm{PD}(\mathbf{w}),\end{aligned} \tag{A.1}$$

each term behind the summation sign in (10) can be upperbounded by

$$\Phi\left(\frac{\sum_{n=0}^{J} w_n \sum_{l=n}^{M+n} h_{l-n} y_l}{\sqrt{N_0 \sum_{n=0}^{J} w_n^2}} \right) \le \Phi\left(\frac{-q_0 + \mathrm{PD}(\mathbf{w})}{\sqrt{N_0 \sum_{n=0}^{J} w_n^2}} \right). \tag{A.2}$$

Now the overall expression (10) can be upperbounded in the following way:

$$\begin{aligned}P_E(\mathbf{w}) &= \frac{1}{2^L} \sum_{\mathbf{y} \in \mathcal{Y}} \Phi\left(\frac{\sum_{n=0}^{J} w_n \sum_{l=n}^{M+n} h_{l-n} y_l}{\sqrt{N_0 \sum_{n=0}^{J} w_n^2}} \right) \\ &\le \frac{1}{2^L} \sum_{\mathbf{y} \in \mathcal{Y}} \Phi\left(\frac{-q_0 + \mathrm{PD}(\mathbf{w})}{\sqrt{N_0 \sum_{n=0}^{J} w_n^2}} \right) \\ &= \Phi\left(\frac{-q_0 + \mathrm{PD}(\mathbf{w})}{\sqrt{N_0 \sum_{n=0}^{J} w_n^2}} \right)\end{aligned} \tag{A.3}$$

which proves the right-hand side of (16a).

The lowerbound of (16a) can be proven by recalling the bit error probability given in the form of (9), which can be rewritten as

$$\begin{aligned}&P_E(\mathbf{w}) \\ &= \frac{1}{2^L} \sum_{\mathbf{y} \in \{-1,1\}^L} \frac{1}{2} \left[\Phi\left(\frac{-q_0 + \sum_{l=1}^{L} q_l y_l}{\sigma} \right) + \Phi\left(\frac{-q_0 - \sum_{l=1}^{L} q_l y_l}{\sigma} \right) \right].\end{aligned} \tag{A.4}$$

If q_k is eye-opened, then, recalling (14) and the ensuing discussion, we obtain

$$\begin{aligned}-q_0 + \sum_{l=1}^{L} q_l y_l &< -q_0 + \mathrm{PD}(\mathbf{w}) < 0, \\ -q_0 - \mathrm{PD}(\mathbf{w}) &< -q_0 - \sum_{l=1}^{L} q_l y_l < 0.\end{aligned} \tag{A.5}$$

Since the inequality

$$\frac{1}{2} \left(\Phi(-a + \epsilon) + \Phi(-a - \epsilon) \right) \ge \Phi(-a) \tag{A.6}$$

is fulfilled for all $a > \epsilon > 0$ for the Gaussian cdf, we can apply this result, taking $a = -q_0/\sigma$ and $\epsilon = \sum_{l=1}^{L} q_l y_l / \sigma$, to obtain the following inequality:

$$\begin{aligned}&\frac{1}{2} \left[\Phi\left(\frac{-q_0 + \sum_{l=1}^{L} q_l y_l}{\sigma} \right) + \Phi\left(\frac{-q_0 - \sum_{l=1}^{L} q_l y_l}{\sigma} \right) \right] \\ &\quad \ge \Phi\left(\frac{-q_0}{\sqrt{N_0 \sum_{n=0}^{J} w_n^2}} \right).\end{aligned} \tag{A.7}$$

Thus the bit error probability can be lowerbounded as

$$\begin{aligned}P_E(\mathbf{w}) &\ge \frac{1}{2^L} \sum_{\mathbf{y} \in \{-1,1\}^L} \Phi\left(\frac{-q_0}{\sqrt{N_0 \sum_{n=0}^{J} w_n^2}} \right) \\ &= \Phi\left(\frac{-q_0}{\sqrt{N_0 \sum_{n=0}^{J} w_n^2}} \right),\end{aligned} \tag{A.8}$$

proving the left-hand side of (16a).

The proof of the upperbound in (16b) is based on the inequality

$$\Phi(-a+\epsilon)+\Phi(-a-\epsilon) \le \Phi(-a+b)+\Phi(-a-b), \quad \forall a>b>\epsilon>0, \tag{A.9}$$

which can be easily verified from the properties of the Gaussian cdf. Casting $a = q_0$, $\epsilon = \sum_{l=1}^{L} q_l y_l$, and $b = \mathrm{PD}(\mathbf{w})$, and making use of the eye-openness again, we obtain from representation (14) that

$$\begin{aligned} P_E(\mathbf{w}) &= \frac{1}{2^{L+1}} \sum_{\mathbf{y}\in\{-1,1\}^L} \left[\Phi\left(\frac{-q_0+\sum_{l=1}^{L} q_l y_l}{\sigma}\right) + \Phi\left(\frac{-q_0-\sum_{l=1}^{L} q_l y_l}{\sigma}\right)\right] \\ &\le \frac{1}{2^{L+1}} \sum_{\mathbf{y}\in\{-1,1\}^L} \left[\Phi\left(\frac{-q_0+\mathrm{PD}(\mathbf{w})}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}}\right) + \Phi\left(\frac{-q_0-\mathrm{PD}(\mathbf{w})}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}}\right)\right] \\ &= \frac{1}{2}\left[\Phi\left(\frac{-q_0+\mathrm{PD}(\mathbf{w})}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}}\right) + \Phi\left(\frac{-q_0-\mathrm{PD}(\mathbf{w})}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}}\right)\right]. \end{aligned} \tag{A.10}$$

For the lowerbound in (16b), we reshuffle the sum of the bit error probability in expression (A.4) and apply the definition of σ^2 and $\mathrm{PD}(\mathbf{w})$, yielding

$$\begin{aligned} P_E(\mathbf{w}) &= \frac{1}{2^{L+1}}\left[\Phi\left(\frac{-q_0+\mathrm{PD}(\mathbf{w})}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}}\right) + \Phi\left(\frac{-q_0-\mathrm{PD}(\mathbf{w})}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}}\right)\right] \\ &\quad + \frac{1}{2^{L+1}} \sum_{\mathbf{y}\in\{-1,1\}^L, \mathbf{y}\neq \mathrm{sgn}\{\mathbf{q}\}} \left[\Phi\left(\frac{-q_0+\sum_{l=1}^{L} q_l y_l}{\sigma}\right) + \Phi\left(\frac{-q_0-\sum_{l=1}^{L} q_l y_l}{\sigma}\right)\right] \\ &= \frac{1}{2^{L+1}}\left[\Phi\left(\frac{-q_0+\mathrm{PD}(\mathbf{w})}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}}\right) + \Phi\left(\frac{-q_0-\mathrm{PD}(\mathbf{w})}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}}\right)\right] \\ &\quad + \frac{1}{2^{L}} \sum_{\mathbf{y}\in\{-1,1\}^L, \mathbf{y}\neq \mathrm{sgn}\{\mathbf{q}\}} \frac{1}{2} \\ &\quad \times \left[\Phi\left(\frac{-q_0+\sum_{l=1}^{L} q_l y_l}{\sigma}\right) + \Phi\left(\frac{-q_0-\sum_{l=1}^{L} q_l y_l}{\sigma}\right)\right]. \end{aligned} \tag{A.11}$$

Due to expression (A.7), this can be lowerbounded by

$$\begin{aligned} &\frac{1}{2^{L+1}}\left[\Phi\left(\frac{-q_0+\mathrm{PD}(\mathbf{w})}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}}\right) + \Phi\left(\frac{-q_0-\mathrm{PD}(\mathbf{w})}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}}\right)\right] \\ &\quad + \frac{1}{2^{L}} \sum_{\mathbf{y}\in\{-1,1\}^L, \mathbf{y}\neq \mathrm{sgn}\{\mathbf{q}\}} \Phi\left(\frac{-q_0}{\sigma}\right) \\ &= \frac{1}{2^{L+1}}\left[\Phi\left(\frac{-q_0+\mathrm{PD}(\mathbf{w})}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}}\right) + \Phi\left(\frac{-q_0-\mathrm{PD}(\mathbf{w})}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}}\right)\right] \\ &\quad + \frac{1}{2^{L}}(2^L-1)\Phi\left(\frac{-q_0}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}}\right) \\ &= \frac{1}{2^{L+1}}\left[\Phi\left(\frac{-q_0+\mathrm{PD}(\mathbf{w})}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}}\right) + \Phi\left(\frac{-q_0-\mathrm{PD}(\mathbf{w})}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}}\right)\right] \\ &\quad + \left(1-\frac{1}{2^{L}}\right)\Phi\left(\frac{-q_0}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}}\right) \end{aligned} \tag{A.12}$$

which completes the proof of the lowerbound in (16b).

To prove the lower bound (16c), we first observe that since $q_0 = w_0 h_0$,

$$\frac{-q_0}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}} = \frac{-q_0/w_0}{\sqrt{N_0\left(1+\sum_{n=1}^{J}\left(w_n^2/w_0^2\right)\right)}} \ge \frac{-h_0}{\sqrt{N_0}}. \tag{A.13}$$

Applying these formulas to the lowerbound of (16a), we have

$$\begin{aligned} P_E(\mathbf{w}) &\ge \Phi\left(\frac{-q_0}{\sqrt{N_0\sum_{n=0}^{J} w_n^2}}\right) \\ &= \Phi\left(\frac{-q_0/w_0}{\sqrt{N_0\left(1+\sum_{n=1}^{J}\left(w_n^2/w_0^2\right)\right)}}\right) \ge \Phi\left(\frac{-h_0}{\sqrt{N_0}}\right), \end{aligned} \tag{A.14}$$

which concludes the derivation of the lowerbound (16c).

B. PROOF OF THEOREM 2

The proof of inequalities (19a) and (19b) follows from the application of the Cauchy-Schwarz inequality to $\mathrm{PD}(\mathbf{w})$ in the bounds given by the right-hand side of (16a) and (16b), respectively. Namely,

$$\begin{aligned} \mathrm{PD}(\mathbf{w}) &= \sum_{l=1}^{L} |q_l| = \left|\sum_{l=1}^{L} q_l \,\mathrm{sgn}\{q_l\}\right| = |\langle \mathbf{q}, \mathrm{sgn}\{\mathbf{q}\}\rangle| \\ &\le \|\mathrm{sgn}\{\mathbf{q}\}\|\|\mathbf{q}\| = \sqrt{L}\sqrt{\sum_{l=1}^{L} q_l^2} \\ &= \sqrt{L\sum_{l=1}^{L}\left(\sum_{j=0}^{M} w_j h_{l-j}\right)^2}. \end{aligned} \tag{B.1}$$

The quantity $\sqrt{L\sum_{l=1}^{L}(\sum_{j=0}^{M} w_j h_{l-j})^2}$ can now be substituted for $\mathrm{PD}(\mathbf{w})$ in the upperbounds of (16a) and (16b), which results in expressions (19a) and (19b), respectively.

C. COMPARISON TO CLASSICAL EQUALIZATION ALGORITHMS

For the sake of performance analysis, one may want to compare the newly derived new bounds on PE with the classical ones related to the PD and MSE [3], given as follows:

$$P_E(\mathbf{w}) \le \Phi\left(\frac{-q_0}{\sqrt{N_0 w_0^2}} + \frac{\mathrm{PD}(\mathbf{w})}{\sqrt{N_0 w_0^2}}\right), \tag{C.1a}$$

$$P_E(\mathbf{w}) \le \mathrm{E}\left[\left(y_k - \sum_{n=0}^{J} w_n x_{k-n}\right)^2\right], \tag{C.1b}$$

where $\mathrm{E}[X]$ denotes the expected value of X. Inequality (C.1a) can be easily derived from the upperbound in expression (16a), taking into account the monotonicity of $\Phi(\cdot)$ and

$$\sqrt{N_0 \sum_{n=0}^{J} w_n^2} \ge \sqrt{N_0 w_0^2}, \tag{C.2}$$

where $w_0 = 1/h_0$.

It is noteworthy that this bound depends on $\mathbf{w}$ through $\mathrm{PD}(\mathbf{w})$ and this dependence is monotone. This fact can give ground to reducing the minimization of bound (C.1a) to the minimization of $\mathrm{PD}(\mathbf{w})$. It can be proven [3] that the optimal weight vector $\mathbf{w}^{(f)}$ which minimizes the function $\mathrm{PD}(\mathbf{w})$ under the condition that $w_0 = 1/h_0$ is the solution of the following set of linear equations:

$$\sum_{n=0}^{l} w_n h_{l-n} = \delta_{l0}, \quad l = 0, \ldots, N \tag{C.3}$$

under the assumption of eye-openness.

The MMSE equalization algorithm can be derived from the Chebyshev inequality

$$P(|y_k - \tilde{y}_k| \ge \epsilon) \le \frac{\mathrm{E}[(y_k - \tilde{y}_k)^2]}{\epsilon^2}. \tag{C.4}$$

The bit probability of a correct decision can be lower-bounded as follows:

$$P_C = P(y_k = \mathrm{sgn}\{\tilde{y}_k\}) \ge P(|y_k - \tilde{y}_k| < 1), \tag{C.5}$$

therefore

$$\begin{aligned} P_E &= 1 - P_C \le 1 - P(|y_k - \tilde{y}_k| < 1) = P(|y_k - \tilde{y}_k| \ge 1) \\ &\le \mathrm{E}[(y_k - \tilde{y}_k)^2] = \mathrm{E}\left[\left(y_k - \sum_{n=0}^{J} w_n v_{k-n}\right)^2\right]. \end{aligned} \tag{C.6}$$

This gives rise to the following goal function:

$$\mathbf{w}^{(g)} : \min_{\mathbf{w}} \mathrm{E}\left[\left(y_k - \sum_{n=0}^{J} w_n x_{k-n}\right)^2\right]. \tag{C.7}$$

The solution of this problem has been treated in numerous papers (see, e.g., [3]) and is given by the following set of linear equations:

$$\sum_{n=0}^{J} w_n \left(\sum_{m=0}^{M-|i-n|} h_m h_{m+|i-n|} + N_0 \delta_{in}\right) = h_0 \delta_{i0}, \quad \forall i = 0, \ldots, N. \tag{C.8}$$

One must note that these bounds yield the optimal equalizer coefficients as a solution of a set of linear equations, whereas the gradients of the newly derived bounds are nonlinear. As a result, there is a tradeoff between the performance (sharpness of the bound) and the complexity of the weight optimization. Namely, low complexity weight optimization (which reduces to solving a set of linear equations in the case of minimizing the PD or MSE) can result in poor performance. When, however, more sophisticated bounds, which yield better performance, are minimized, then a relatively time-consuming gradient descent must be performed. Finally, when one wants to minimize the error probability itself, then the complexity of the weight optimization algorithm becomes enormous, which prevents its practical implementation.

REFERENCES

[1] R. Steele and L. Hanzo, Eds., *Mobile Radio Communications*, John Wiley & Sons, New York, NY, USA, 1999.

[2] W. H. Gerstacker and R. Schober, "Equalization concepts for EDGE," *IEEE Transactions on Wireless Communications*, vol. 1, no. 1, pp. 190–199, 2002.

[3] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, NY, USA, 1995.

[4] O. Shimbo and M. I. Celebiler, "The probability of error due to intersymbol interference and Gaussian noise in digital communication systems," *IEEE Transactions on Communications*, vol. 19, no. 2, pp. 113–119, 1971.

[5] C.-C. Yeh and J. R. Barry, "Adaptive minimum bit-error rate equalization for binary signaling," *IEEE Transactions on Communications*, vol. 48, no. 7, pp. 1226–1235, 2000.

[6] S. Chen, B. Mulgrew, and L. Hanzo, "Least bit error rate adaptive nonlinear equalisers for binary signalling," *IEE Proceedings: Communications*, vol. 150, no. 1, pp. 29–36, 2003.

[7] P. Chandra Kumar, P. Saratchandran, and N. Sundararajan, "Minimal radial basis function neural networks for nonlinear channel equalisation," *IEE Proceedings: Vision, Image and Signal Processing*, vol. 147, no. 5, pp. 428–435, 2000.

[8] V. Bhatia, B. Mulgrew, and A. Th. Georgiadis, "Minimum BER DFE equalizer in alpha stable noise," in *Proceedings of the 12th European Signal Processing Conference (EUSIPCO '04)*, Vienna, Austria, September 2004.

[9] R. Y. Yen, M. T. Cha, and M.-H. Young, "A stochastic unbiased minimum mean error rate algorithm for decision feedback equalizers," in *Proceedings of the 24th IASTED International Conference on Signal Processing, Pattern Recognition, and Applications (SPPRA '06)*, pp. 160–164, Innsbruck, Austria, February 2006.

[10] A. Dua, U. B. Desai, and R. K. Mallik, "Minimum probability of error-based methods for adaptive multiuser detection in multipath DS-CDMA channels," *IEEE Transactions on Wireless Communications*, vol. 3, no. 3, pp. 939–948, 2004.

[11] S. Chen, A. K. Samingan, B. Mulgrev, and L. Hanzo, "Adaptive minimum-BER linear multiuser detection," in *Proceedings of*

IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '01), vol. 4, pp. 2253–2256, Salt Lake City, Utah, USA, May 2001.

[12] A. Hjørungnes and M. Debbah, "Minimum BER FIR receiver filters for DS-CDMA systems," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '05)*, vol. 4, pp. 2287–2291, St. Louis, Mo, USA, November-December 2005.

[13] B. Dong and X. Wang, "Sampling-based soft equalization for frequency-selective MIMO channels," *IEEE Transactions on Communications*, vol. 53, no. 2, pp. 278–288, 2005.

[14] J. X. Yu, Y. Li, Z.-Q. Luo, and S. Yoshida, "Split soft-decision equalization for wireless channels with large delay spread," *IEEE Transactions on Communications*, vol. 53, no. 2, pp. 269–277, 2005.

[15] S. X. Ng, M.-S. Yee, and L. Hanzo, "Coded modulation assisted radial basis function aided turbo equalization for dispersive Rayleigh-fading channels," *IEEE Transactions on Wireless Communications*, vol. 3, no. 6, pp. 2198–2206, 2004.

[16] C. Laot, R. Le Bidan, and D. Leroux, "Low-complexity MMSE turbo equalization: a possible solution for EDGE," *IEEE Transactions on Wireless Communications*, vol. 4, no. 3, pp. 965–974, 2005.

[17] S. Choi and T.-W. Lee, "A negentropy minimization approach to adaptive equalization for digital communication systems," *IEEE Transactions on Neural Networks*, vol. 15, no. 4, pp. 928–936, 2004.

[18] P. Savazzi, L. Favalli, E. Costamagna, and A. Mecocci, "A sub-optimal approach to channel equalization based on the nearest neighbor rule," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 9, pp. 1640–1648, 1998.