Adaptive Noncoherent Linear Minimum ISI Equalization for MDPSK and MDAPSK Signals

Robert Schober[1], Wolfgang H. Gerstacker, and Johannes B. Huber

Lehrstuhl für Nachrichtentechnik II,
Universität Erlangen-Nürnberg,
Cauerstraße 7/NT, D-91058 Erlangen, Germany
Phone: +49 9131 8527668
Fax: +49 9131 8528919
E-mail: schober@LNT.de

Abstract

In this paper, a novel noncoherent linear equalization scheme is introduced and analyzed. In contrast to previously proposed noncoherent equalization schemes, the proposed scheme is not only applicable for M–ary differential phase–shift keying (MDPSK) but also for M–ary differential amplitude/phase–shift keying (MDAPSK). The novel scheme minimizes the variance of intersymbol interference (ISI) in the equalizer output signal. The optimum equalizer coefficients may be calculated directly from an eigenvalue problem. For an efficient recursive adaptation of the equalizer coefficients, a modified least–mean–square (LMS) and a modified recursive least–squares (RLS) algorithm are proposed. It is shown that the corresponding cost function has no spurious local minima which ensures global convergence of the adaptive algorithms. Simulations confirm the good performance of the proposed noncoherent equalization scheme and its robustness against frequency offset.

EDICS Number: 3–CEQU (Channel Modeling, Estimation, and Equalization)

[1]Corresponding Author

1 Introduction

The combination of linear or nonlinear equalization and coherent detection (CD) has been studied extensively in literature (see e.g. [1, 2, 3, 4] and references therein) and is applied in many existing communication systems. However, only few results are available for *noncoherent* equalization schemes, i.e., for the combination of linear or nonlinear equalization and noncoherent detection. Such noncoherent receivers have the important advantage of being more robust against phase noise and frequency offset than coherent equalization schemes.

A noncoherent linear minimum mean–squared error (MMSE) equalizer has been proposed first by Sehier et al. [5]. Recently, a novel noncoherent linear MMSE equalizer which can approach the power efficiency of a coherent linear MMSE equalizer has been proposed in [6]. Noncoherent decision–feedback equalization (DFE) schemes have been reported in [7, 8, 9, 10].

All above mentioned noncoherent linear and nonlinear equalizers have in common that they are only designed for M–ary differential phase–shift keying (MDPSK) signals. However, recently M–ary differential amplitude/phase–shift keying (MDAPSK) constellations have become very popular because of their higher spectral efficiency compared to MDPSK [11, 12, 13, 14, 15, 16, 17, 18, 19]. Therefore, it is desirable to derive a robust noncoherent equalization scheme for MDAPSK signals.

In this paper, we propose a novel noncoherent linear equalization (NLE) scheme which works for both MDPSK and MDAPSK signals equally well. Similar to the scheme given in [6], this NLE scheme consists of a linear equalizer combined with a decision–feedback differential detector (DF–DD). Using the equalizer output signal, the DF–DD generates a reference symbol either nonrecursively [20, 21, 14, 15] or recursively [22, 23, 15]. In contrast to the NLE scheme reported in [6], the NLE scheme proposed here minimizes the variance of intersymbol interference (ISI) in the equalizer output signal. This leads to a constrained minimization problem. The essential difference to all previously proposed noncoherent equalization schemes [5, 6, 7, 8, 9, 10] is that the absence of spurious local minima in the resulting cost function can be shown analytically. If an infinite–length filter is employed, a zero–forcing (ZF) equalizer results.

For adaptation of the equalizer coefficients a modified least–mean–square (LMS) and a modified recursive least–squares (RLS) algorithm are presented. Although these adaptive algorithms find the exact minimum of the above mentioned cost function for finite–length equalizers only in the MDPSK case, a very good solution close to the optimum one is also obtained for MDAPSK. Furthermore, simple conditions on the step size parameter for stability of the modified LMS algorithm are provided. In our computer simulations,

the proposed NLE scheme is compared with coherent ZF equalization and the noncoherent MMSE equalizer reported in [6]. Simulations also confirm the robustness of the novel NLE scheme against frequency offset.

This paper is organized as follows. In Section 2, the transmission model and the receiver structure are provided. The noncoherent minimum ISI equalizer (NMIE) is introduced and analyzed in Section 3, whereas modified LMS and RLS algorithms are presented in Section 4. Finally, some simulation results are given in Section 5 and some conclusions are drawn in Section 6.

2 Transmission Model and Receiver Structure

Fig. 1 shows a block diagram of the discrete–time transmission model. All signals are represented by their complex–baseband equivalents. For simplicity, only T–spaced equalizers are considered here. However, our results can be extended easily to the fractionally–spaced case. The transmitted MPSK/MAPSK symbol $s[k]$ is given by

$$s[k] \triangleq R[k]b[k], \qquad k \in \mathbb{Z}, \tag{1}$$

with absolute amplitude symbol $R[k]$, $R[k] \in \{R_1, \ldots, R_Z\}$ ($Z = 1$, i.e., $R[k] \equiv 1$, $\forall k$, for MDPSK) and absolute phase symbol $b[k] \in \mathcal{A}_b = \{\mathrm{e}^{j2\pi\nu/(M/Z)} | \nu \in \{0, 1, \ldots, M/Z-1\}\}$. For convenience, $\mathcal{E}\{|s[k]|^2\}$ ($\mathcal{E}\{\cdot\}$ denotes expectation) is normalized to unity. $s[k]$ is obtained from the MDPSK/MDAPSK symbols

$$\Delta s[k] \triangleq \Delta R[k]a[k], \tag{2}$$

via differential encoding:

$$s[k] = \Delta s[k]s[k-1]. \tag{3}$$

$\Delta R[k] = R[k]/R[k-1]$ ($\Delta R[k] \equiv 1$, $\forall k$, for MDPSK) and $a[k] = b[k]/b[k-1] \in \mathcal{A}_a = \{\mathrm{e}^{j2\pi\nu/(M/Z)} | \nu \in \{0, 1, \ldots, M/Z-1\}\}$, denote the amplitude and the phase difference symbol, respectively. The most popular MDAPSK signaling format is 16DAPSK (16–star QAM) [11]. It will be used for the simulations presented in Section 5. Here, $Z = 2$ is valid, i.e., there are two different amplitude levels R_1 and R_2, $R_1 < R_2$, and the amplitude difference is given by $\Delta R = R_2/R_1$ (cf. Fig. 2). One information bit is mapped to the amplitude difference symbol, while three information bits are Gray–mapped to the phase difference symbol. $\Delta R[k] = 1$ and $\Delta R[k] = \Delta R$ (if $R[k-1] = R_1$) or $\Delta R[k] = 1/\Delta R$ (if $R[k-1] = R_2$) is valid if the information bit is equal to zero and one, respectively.

The MPSK/MAPSK symbols $s[k]$ are transmitted over an ISI producing channel with unknown, constant phase shift Θ. The discrete–time received signal, sampled at times kT at the output of the receiver input filter, can be expressed as

$$r[k] = \mathrm{e}^{j\Theta} \sum_{\nu=0}^{L_h-1} h_\nu s[k-\nu] + n[k], \tag{4}$$

where h_ν, $0 \leq \nu \leq L_h - 1$, are the coefficients of the combined discrete–time impulse response of the cascade of transmit filter, channel, and receiver input filter; its length is denoted by L_h. For the receiver input filter, we assume a square–root Nyquist frequency response[2]. Thus, the zero mean complex Gaussian noise $n[\cdot]$ is white. Due to an appropriate normalization, the noise variance is $\sigma_n^2 = \mathcal{E}\{|n[k]|^2\} = N_0/E_S$. E_S and N_0 are the mean received energy per symbol and the single–sided power spectral density of the underlying passband noise process, respectively. The equalizer output symbol $q[k]$ may be written as

$$\begin{aligned} q[k] &= \sum_{\nu=0}^{L_c-1} c_\nu r[k-\nu] \\ &= \mathrm{e}^{j\Theta} g_{k_0} s[k-k_0] + \mathrm{e}^{j\Theta} \sum_{\substack{\nu=0 \\ \nu \neq k_0}}^{L_h+L_c-2} g_\nu s[k-\nu] + \sum_{\nu=0}^{L_c-1} c_\nu n[k-\nu], \end{aligned} \tag{5}$$

where c_ν are the equalizer coefficients and

$$g_\nu = \sum_{\mu=0}^{L_c-1} c_\mu h_{\nu-\mu} \tag{6}$$

are the coefficients of the combined impulse response of overall channel and equalizer; L_c is the equalizer length. The decision delay k_0 should be optimized since it can significantly affect performance (cf. e.g. [25]).

The next stage of the proposed receiver is a DF–DD, which determines an estimate $\Delta\hat{s}[k-k_0]$ for the transmitted symbol $\Delta s[k-k_0]$ based on a reference symbol $q_{\mathrm{ref}}[k-1]$. This reference symbol may be generated either nonrecursively [14, 15]

$$q_{\mathrm{ref}}[k-1] = \frac{\sum_{\nu=1}^{N-1} q[k-\nu] \prod_{\mu=1}^{\nu-1} \frac{1}{\Delta\hat{s}^*[k-k_0-\mu]}}{\sum_{\nu=1}^{N-1} \prod_{\mu=1}^{\nu-1} \frac{1}{|\Delta\hat{s}[k-k_0-\mu]|^2}}, \tag{7}$$

where N, $N \geq 2$, is the number of equalizer output symbols used for determination of $\Delta\hat{s}[k-k_0]$ (cf. Eq. (11)), or recursively [15]

$$q_{\mathrm{ref}}[k-1] = \frac{q[k-1] + W \frac{1}{\Delta\hat{s}^*[k-k_0-1]} q_{\mathrm{ref}}[k-2]}{1 + W \frac{1}{|\Delta\hat{s}[k-k_0-1]|^2}}, \tag{8}$$

[2]Note that this includes the whitened matched filter [24] as a special case.

where W, $W \geq 0$, is a design parameter (cf. Section 5). Here, $(\cdot)^*$ denotes complex conjugation. In order to clarify the role of $q_{\text{ref}}[k-1]$, we may first consider the nonrecursive reference symbol with $N = 2$. In this case, $q_{\text{ref}}[k-1] = q[k-1]$ results, i.e., the previous equalizer output symbol is employed to provide a phase and amplitude reference for the current equalizer output symbol $q[k]$. This reference is used for conventional differential detection (cf. Eq. (11)) [3, 11]. Since $q[k-1]$ is a noisy reference the performance of the resulting noncoherent receiver is worse than that of a coherent receiver which assumes perfect knowledge of the amplitude and phase reference, of course. In order to overcome this problem $N-1$, $N > 2$, equalizer output symbols may be used to improve the phase and amplitude reference. Due to the factors $\prod_{\mu=1}^{\nu-1} 1/\Delta\hat{s}^*[k-k_0-\mu]$ in Eq. (7), the desired signal components $\mathrm{e}^{j\Theta} g_{k_0} s[k-k_0-\nu]$ (cf. Eq. (5)) of $q[k-\nu]$ are added constructively ($\Delta\hat{s}[k-k_0-\mu] = \Delta s[k-k_0-\mu]$ is assumed), whereas the ISI and noise components add non–constructively. Hence, the signal–to–noise ratio of the reference symbol $q_{\text{ref}}[k-1]$ increases as N increases. For $N \to \infty$ a perfect reference is achieved, i.e., the DF–DD approaches the performance of a coherent detector. The denominator in Eq. (7) is real–valued and is necessary for normalization of the amplitude of the reference symbol. The recursive reference symbol has the same purpose as the nonrecursive reference symbol, however, in this case the degree of averaging for calculation of the reference symbol is controlled by W. In fact it can be shown that the recursive reference symbol can be derived from the nonrecursive one [15, 26].

It should be mentioned that for MDPSK, Eqs. (7) and (8) may be simplified. In this special case,

$$q_{\text{ref}}[k-1] = \frac{1}{N-1} \sum_{\nu=1}^{N-1} q[k-\nu] \prod_{\mu=1}^{\nu-1} \hat{a}[k-k_0-\mu], \tag{9}$$

and

$$\begin{aligned} q_{\text{ref}}[k-1] &= \frac{1}{1+W}\, q[k-1] + \frac{W}{1+W}\, \hat{a}[k-k_0-1] q_{\text{ref}}[k-2] \\ &\triangleq (1-\alpha)\, q[k-1] + \alpha\, \hat{a}[k-k_0-1] q_{\text{ref}}[k-2] \end{aligned} \tag{10}$$

are obtained for the nonrecursively and the recursively generated reference symbol, respectively. $\hat{a}[\cdot] \in \mathcal{A}_a$ are the estimated phase difference symbols and $\alpha \triangleq W/(1+W)$, $0 \leq \alpha < 1$, may be interpreted as forgetting factor. Note that the reference symbols according to Eq. (9) and (10) are – up to a normalization constant, which has no influence on the decision – identical with the reference symbols reported in [20, 21] and [22, 23], respectively. It can be shown that the nonrecursively and the recursively generated reference symbols are identical for $N = 2$, $W = 0$ ($\alpha = 0$) and for $N \to \infty$, $W \to \infty$ ($\alpha \to 1$).

The decision variable for estimation of $\Delta\hat{s}[k-k_0]$ is given by

$$d[k] = \frac{q[k]}{q_{\text{ref}}[k-1]}. \tag{11}$$

For the phase decision, i.e., for estimation of $\hat{a}[k-k_0]$, only the phase of $d[k]$ is of interest. Thus, the complex plane is divided into M/Z sectors corresponding to the M/Z possible values of $a[k-k_0]$ and the estimated phase difference symbol is determined by the sector into which $d[k]$ falls.

The estimated amplitude difference symbol $\Delta\hat{R}[k-k_0]$ is determined by the magnitude of $d[k]$. For simplicity, we only give the decision rule for 16DAPSK which is used exclusively in our simulations. $\Delta\hat{R}[k-k_0]$ is given by

$$\Delta\hat{R}[k-k_0] = \begin{cases} 1, & \text{if} \quad |d[k]| \geq \frac{1}{2}(1+1/\Delta R) \\ 1/\Delta R, & \text{if} \quad |d[k]| < \frac{1}{2}(1+1/\Delta R) \end{cases}, \tag{12}$$

if $\hat{R}[k-k_0-1] = \Delta\hat{R}[k-k_0-1]\hat{R}[k-k_0-2] = R_2$ and by

$$\Delta\hat{R}[k-k_0] = \begin{cases} \Delta R, & \text{if} \quad |d[k]| \geq \frac{1}{2}(1+\Delta R) \\ 1, & \text{if} \quad |d[k]| < \frac{1}{2}(1+\Delta R) \end{cases}, \tag{13}$$

if $\hat{R}[k-k_0-1] = R_1$. Note that the amplitude decision variable used in this paper is different from that proposed in [15], however, the performance of the resulting DF–DD is almost identical.

3 Noncoherent Minimum ISI Equalization

In this section, the NMIE is derived and analyzed. In order to simplify our analysis, we assume that the reference symbol $q_{\text{ref}}[k-1]$ is generated nonrecursively (Eqs. (7), (9)) in this section. However, the adaptive algorithms described in Section 4 work for recursively generated reference symbols, too.

3.1 Noncoherent Cost Function

First of all, the cost function for noncoherent minimum ISI equalization is derived. As usual, it is assumed that all decision–feedback symbols $\Delta\hat{s}[\cdot]$ are correct, i.e, $\Delta\hat{s}[k-k_0-\nu] = \Delta s[k-k_0-\nu]$, $\nu \geq 0$. Hence, the reference symbol according to Eq. (7) can be rewritten to

$$q_{\text{ref}}[k-1] = s[k-k_0-1]\frac{\sum_{\nu=1}^{N-1} q[k-\nu]s^*[k-k_0-\nu]}{\sum_{\nu=1}^{N-1} |s[k-k_0-\nu]|^2}, \tag{14}$$

where Eq. (3) is used. Now, Eq. (5) may be applied in Eq. (14). This yields

$$\begin{aligned} q_{\mathrm{ref}}[k-1] &= \mathrm{e}^{j\Theta} g_{k_0}\, s[k-k_0-1] \\ &\quad + \mathrm{e}^{j\Theta} \frac{\sum\limits_{\substack{\mu=0 \\ \mu\neq k_0}}^{L_h+L_c-2} g_\mu \sum\limits_{\nu=1}^{N-1} s[k-\mu-\nu]s^*[k-k_0-\nu]}{\sum\limits_{\nu=1}^{N-1} |s[k-k_0-\nu]|^2} s[k-k_0-1] \\ &\quad + \frac{\sum\limits_{\mu=0}^{L_c-1} c_\mu \sum\limits_{\nu=1}^{N-1} n[k-\mu-\nu]s^*[k-k_0-\nu]}{\sum\limits_{\nu=1}^{N-1} |s[k-k_0-\nu]|^2} s[k-k_0-1]. \end{aligned} \tag{15}$$

The NMIE minimizes the variance of the error signal

$$e[k] = \Delta \hat{s}[k-k_0] q_{\mathrm{ref}}[k-1] - q[k], \tag{16}$$

cf. Fig. 1. Note that in the absence of noise and ISI $e[k] = 0, \forall\, k$, for $c_{k_0} = 1$, $c_\nu = 0$, $\nu \neq k_0$. Using Eqs. (3), (5), and (15), Eq. (16) can be rewritten to

$$\begin{aligned} e[k] &= \mathrm{e}^{j\Theta} \sum_{\substack{\mu=0 \\ \mu\neq k_0}}^{L_h+L_c-2} g_\mu \left(\frac{\sum\limits_{\nu=1}^{N-1} s[k-\mu-\nu]s^*[k-k_0-\nu]}{\sum\limits_{\nu=1}^{N-1} |s[k-k_0-\nu]|^2} s[k-k_0] - s[k-\mu] \right) \\ &\quad + \sum_{\mu=0}^{L_c-1} c_\mu \left(\frac{\sum\limits_{\nu=1}^{N-1} n[k-\mu-\nu]s^*[k-k_0-\nu]}{\sum\limits_{\nu=1}^{N-1} |s[k-k_0-\nu]|^2} s[k-k_0] - n[k-\mu] \right). \end{aligned} \tag{17}$$

Taking into account the uncorrelatedness of $s[\cdot]$ and $n[\cdot]$, the variance $\sigma_e^2(\boldsymbol{c})$ ($\boldsymbol{c} \triangleq [c_0\ c_1\ \ldots\ c_{L_c-1}]^H$, $[\cdot]^H$ denotes Hermitian transposition) of the error signal $e[k]$ can be calculated to

$$\sigma_e^2(\boldsymbol{c}) = \mathcal{E}\{|e[k]|^2\} = \sum_{\substack{\mu=0 \\ \mu\neq k_0}}^{L_h+L_c-2} x_\mu |g_\mu|^2 + \gamma \sigma_n^2 \sum_{\mu=0}^{L_c-1} |c_\mu|^2, \tag{18}$$

where the definitions

$$x_\mu \triangleq 1 + \mathcal{E}\left\{ |s[k-k_0]|^2 \frac{\sum\limits_{\nu=1}^{N-1} \sum\limits_{\xi=1}^{N-1} s[k-\mu-\nu]s^*[k-\mu-\xi]s^*[k-k_0-\nu]s[k-k_0-\xi]}{\left(\sum\limits_{\nu=1}^{N-1} |s[k-k_0-\nu]|^2 \right)^2} \right\} \tag{19}$$

and

$$\gamma \triangleq 1 + \mathcal{E}\left\{ \frac{1}{\sum\limits_{\nu=1}^{N-1} |s[k-k_0-\nu]|^2} \right\} \tag{20}$$

are used. Because of the symmetry of the employed signal constellations, Eq. (19) may be rewritten to

$$x_\mu = 1 + \mathcal{E}\left\{ |s[k-k_0]|^2 \frac{\sum_{\nu=1}^{N-1} |s[k-\mu-\nu]|^2 |s[k-k_0-\nu]|^2}{\left(\sum_{\nu=1}^{N-1} |s[k-k_0-\nu]|^2\right)^2} \right\} \tag{21}$$

In Appendix A, it is shown that, in some special cases, the expressions for x_μ and γ can be simplified considerably.

As can be seen from Eq. (18), the error variance $\sigma_e^2(\boldsymbol{c})$ consists of an ISI (first sum) and a noise (second sum) component. x_μ and γ in Eq. (18) depend exclusively on the underlying signal constellation and on N. γ, $\gamma \geq 1$, may be interpreted as a noise amplification factor. This is especially obvious for the MDPSK case. Here, $\gamma = N/(N-1)$ is valid (cf. Eq. (79)), i.e., for $N = 2$, $\gamma = 2$ results whereas for $N \to \infty$, $\gamma = 1$ is obtained. The noise reduction for large N is a direct consequence of the averaging over different equalizer output symbols $q[k-\nu]$, $1 \leq \nu \leq N-1$, in the reference symbol (cf. Eq. (7)). Similarly x_μ, $0 \leq \mu \leq L_h + L_c - 2$, $\mu \neq k_0$, may be interpreted as weighting factors of the ISI coefficients g_μ. For MDPSK, $x_\mu = N/(N-1)$, $0 \leq \mu \leq L_h + L_c - 2$, $\mu \neq k_0$, holds, i.e., all ISI coefficients are weighted with the same factor. For MDAPSK, in general, the weighting factors may have different values which is a direct consequence of the fact that different *effective* symbols $\sum_{\nu=1}^{N-1} s[k-\mu-\nu]s^*[k-k_0-\nu]s[k-k_0-1]/(\sum_{\nu=1}^{N-1}|s[k-k_0-\nu]|^2)$, $0 \leq \mu \leq L_h + L_c - 2$, $\mu \neq k_0$, belonging to different g_μ in $q_{\mathrm{ref}}[k-1]$ (cf. Eq. (15)) may have different variances. Consequently, effective symbols with large variance cause large weighting factors and the corresponding ISI coefficients g_μ have to be more suppressed by the optimum equalizer coefficient vector $\boldsymbol{c}$.

In the following, the definition

$$\boldsymbol{h}_\mu \triangleq [h_\mu \; h_{\mu-1} \; \ldots \; h_{\mu-L_c+1}]^T, \tag{22}$$

where $[\cdot]^T$ denotes transposition, is used. Note that $h_\mu = 0$ for $\mu < 0$ and $\mu > L_h - 1$ is valid.

Using this and Eq. (6), Eq. (18) can be rewritten to

$$\begin{aligned} \sigma_e^2(\boldsymbol{c}) &= \boldsymbol{c}^H \left(\sum_{\substack{\mu=0 \\ \mu \neq k_0}}^{L_h+L_c-2} x_\mu \boldsymbol{h}_\mu \boldsymbol{h}_\mu^H \right) \boldsymbol{c} + \gamma \sigma_n^2 \boldsymbol{c}^H \boldsymbol{c} \\ &= \boldsymbol{c}^H \boldsymbol{H} \boldsymbol{c} + \gamma \sigma_n^2 \boldsymbol{c}^H \boldsymbol{c}, \end{aligned} \tag{23}$$

where the positive semi–definite Hermitian matrix $\boldsymbol{H}$ is defined as

$$\boldsymbol{H} \stackrel{\triangle}{=} \sum_{\substack{\mu=0 \\ \mu \neq k_0}}^{L_h+L_c-2} x_\mu \boldsymbol{h}_\mu \boldsymbol{h}_\mu^H. \tag{24}$$

The NMIE should minimize the error variance $\sigma_e^2(\boldsymbol{c})$, however, an unconstrained minimization of $\sigma_e^2(\boldsymbol{c})$ yields the all zero vector with L_c rows $\boldsymbol{c} = \boldsymbol{0}_{L_c}$. Therefore, a constraint has to be introduced. Here, the unit–energy constraint for the equalizer coefficients is employed[3]:

$$\boldsymbol{c}^H \boldsymbol{c} = 1. \tag{25}$$

Similar constraints are well known from impulse response truncation [27, 28] and blind equalization [29, 30, 31, 32].

Using Eqs. (23) and (25), the Lagrange cost function

$$J_L(\boldsymbol{c}) \stackrel{\triangle}{=} \boldsymbol{c}^H \boldsymbol{H} \boldsymbol{c} + \gamma \sigma_n^2 \, \boldsymbol{c}^H \boldsymbol{c} + \eta(1 - \boldsymbol{c}^H \boldsymbol{c}) \tag{26}$$

can be defined, where η is the Lagrange multiplier.

Eq. (26) shows that the cost function $J_L(\boldsymbol{c})$ does not depend on the channel phase Θ. Hence, it may be considered as *noncoherent* cost function. Our simulations show that the resulting NMIE is also robust against phase variations (e.g. frequency offset).

3.2 Stationary Points of the Cost Function

The stationary points of the cost function can be obtained by using the method for complex differentiation described in [4], Appendix B:

$$\frac{\partial J_L(\boldsymbol{c})}{\partial \boldsymbol{c}^*} = \boldsymbol{H} \boldsymbol{c} + (\gamma \sigma_n^2 - \eta) \, \boldsymbol{c}, \tag{27}$$

and setting the result equal to zero. This yields the eigenvalue problem

$$\boldsymbol{H} \boldsymbol{c} = (\eta - \gamma \sigma_n^2) \, \boldsymbol{c} \stackrel{\triangle}{=} \lambda \, \boldsymbol{c}, \tag{28}$$

where λ denotes the eigenvalue. Since $\boldsymbol{H}$ is a Hermitian positive semi–definite matrix, there are L_c vectors $\boldsymbol{c}_\nu$, $0 \leq \nu \leq L_c - 1$, of unit length[4], which solve Eq. (28), corresponding to L_c real, non–negative eigenvalues λ_ν, $0 \leq \nu \leq L_c - 1$ [4]. In Appendix B, it is shown, that the eigenvector $\boldsymbol{c}_{\text{opt}}$ corresponding to the minimum eigenvalue $\lambda_{\min}$ minimizes the cost function

[3] Note that in principle $\boldsymbol{c}^H \boldsymbol{c}$ could be normalized to an arbitrary positive value because the magnitude of g_{k_0} has no influence on the decision variable $d[k]$ (cf. Eq. (11)).

[4] Note that $\boldsymbol{c}_\nu$ is only unique up to a complex factor of magnitude one.

$J_L(\boldsymbol{c})$ and the error variance $\sigma_e^2(\boldsymbol{c})$, whereas the eigenvector corresponding to the maximum eigenvalue maximizes $J_L(\boldsymbol{c})$ and $\sigma_e^2(\boldsymbol{c})$. The remaining eigenvectors correspond to saddle points.

The minimum error variance is given by

$$\sigma_e^2(\boldsymbol{c}_{\text{opt}}) = J_L(\boldsymbol{c}_{\text{opt}}) = \lambda_{\min} + \gamma\sigma_n^2. \tag{29}$$

Eq. (28) shows that the noise variance σ_n^2 has no influence on the optimum equalizer coefficients $\boldsymbol{c}_{\text{opt}}$; it only influences the minimum error variance (cf. Eq. (29)). This implies that the resulting equalizer is related to a ZF equalizer. However, in contrast to a ZF equalizer, in general, the NMIE forces the coefficients g_ν, $\nu \neq k_0$, of the overall impulse response not to zero, but it minimizes them in the mean–square sense (cf. Eqs. (18), (26)). Hence, we refer to the novel equalizer as noncoherent minimum ISI equalizer. To the best of the authors' knowledge, there has not been reported a corresponding coherent equalizer, yet. In the next section, it will be shown that for an infinite number of taps the proposed NMIE is up to a constant identical with a ZF equalizer. It should be mentioned that the independence of the optimum equalizer coefficients of the noise variance is not a consequence of the error criterion (cf. Eq. (16)) but of the employed constraint (cf. Eq. (25)), i.e., different constraints might lead to a noise–dependent solution.

3.3 Infinite–Length Equalizer

The error variance corresponding to $\boldsymbol{c}_{\text{opt}}$ becomes minimum for $\lambda_{\min} = 0$, i.e., if $\boldsymbol{H}$ is a singular matrix. In this case

$$\sigma_e^2(\boldsymbol{c}_{\text{opt}}) = J_L(\boldsymbol{c}_{\text{opt}}) = \gamma\sigma_n^2 \tag{30}$$

results from Eq. (29). From Eq. (18), it can be seen that Eq. (30) is obtained, if and only if $g_\mu = 0$, $\mu \neq k_0$. This corresponds to a transfer function of the resulting equalizer given by

$$C(\mathrm{e}^{j2\pi fT}) = \frac{\mathrm{e}^{-j2\pi fTk_0}\, p}{H(\mathrm{e}^{j2\pi fT})}, \tag{31}$$

where $H(\mathrm{e}^{j2\pi fT})$ and p denote the transfer function[5] of the discrete–time channel and a complex constant, respectively. Up to the constant p, the infinite–tap NMIE given by Eq. (31) is identical with a ZF equalizer [1, 3].

From Eq. (25),

$$T \int_{-1/(2T)}^{1/(2T)} |C(\mathrm{e}^{j2\pi fT})|^2 \mathrm{d}f = 1 \tag{32}$$

[5]We assume that $H(\mathrm{e}^{j2\pi fT})$ has no spectral nulls.

follows. Using this and Eq. (31), $|p|$ can be calculated to

$$|p| = \frac{1}{\sqrt{T \int\limits_{-1/(2T)}^{1/(2T)} \frac{1}{|H(\mathrm{e}^{j2\pi fT})|^2} \mathrm{d}f}}. \tag{33}$$

Note that the phase of p is arbitrary. The transfer function $G(\mathrm{e}^{j2\pi fT})$ of the combination of equalizer and overall channel is given by

$$G(\mathrm{e}^{j2\pi fT}) = H(\mathrm{e}^{j2\pi fT}) C(\mathrm{e}^{j2\pi fT}) = \mathrm{e}^{-j2\pi fTk_0}\, p. \tag{34}$$

Since $G(\mathrm{e}^{j2\pi fT})$ is the Fourier transform of g_k,

$$g_{k_0} = p \tag{35}$$

follows.

3.4 Performance for $N \to \infty$

In the following, we derive the limiting performance of the infinite–length equalizer for $N \to \infty$. In this case,

$$q_{\mathrm{ref}}[k-1] = \mathrm{e}^{j\Theta} g_{k_0} s[k-k_0-1] \tag{36}$$

holds in the mean–square sense (cf. Eq. (15)) and the decision variable of the infinite–length equalizer is

$$d[k] = \frac{s[k-k_0]}{s[k-k_0-1]} + \frac{\sum\limits_{\nu=0}^{L_c-1} c_\nu n[k-\nu]}{\mathrm{e}^{j\Theta} g_{k_0} s[k-k_0-1]}. \tag{37}$$

From Eq. (37) it can be seen that the signal–to–noise ratio (SNR) of the NMIE can be expressed as

$$\begin{aligned} \mathrm{SNR} &= \frac{\mathcal{E}\left\{\left|\frac{s[k-k_0]}{s[k-k_0-1]}\right|^2\right\}}{\mathcal{E}\left\{\left|\frac{\sum_{\nu=0}^{L_c-1} c_\nu n[k-\nu]}{g_{k_0} s[k-k_0-1]}\right|^2\right\}} \\ &= \frac{|g_{k_0}|^2}{\sigma_n^2}, \end{aligned} \tag{38}$$

where Eq. (25) is used. Applying Eqs. (33) and (35) in Eq. (38) yields

$$\mathrm{SNR} = \frac{1}{\sigma_n^2 T \int\limits_{-1/(2T)}^{1/(2T)} \frac{1}{|H(\mathrm{e}^{j2\pi fT})|^2} \mathrm{d}f}. \tag{39}$$

The same expression can be obtained for a coherent infinite–length ZF equalizer for MPSK/MAPSK, i.e., if no differential encoding is employed [3]. Note that we assumed for our analysis $\Delta\hat{s}[k-k_0-\nu] = \Delta s[k-k_0-\nu]$, $\nu > 0$, i.e., all feedback symbols are correct. This explains, why the resulting equalizer (which is not implementable, of course) is lower–bounded by a coherent infinite–length ZF equalizer for MPSK/MAPSK. The simulations in Section 5 show that a realizable NMIE (i.e., without genie) is lower–bounded by a coherent infinite–length ZF equalizer for M**D**PSK/M**D**APSK since erroneous feedback symbols cause an increase of BER by a factor of two, which is typical for noncoherent receivers using DF–DDs [14, 21, 6].

4 Adaptive Algorithms for Adjustment of the NMIE Coefficients

In this section, we present a modified LMS and a modified RLS algorithm for adjustment of the NMIE coefficients. For these adaptive algorithms the reference symbol $q_{\text{ref}}[k-1]$ may be calculated nonrecursively (Eqs. (7), (9)) or recursively (Eqs. (8), (10)). However, for simplicity, the analysis of these algorithms (cf. Sections 4.3, 4.4) is again restricted to the case of a nonrecursively generated reference symbol.

4.1 Modified LMS Algorithm

A gradient algorithm is obtained from

$$\hat{\boldsymbol{c}}[k+1] = \boldsymbol{c}[k] - \delta_{\text{LMS}} \frac{\partial}{\partial \boldsymbol{c}^*[k]} |e[k]|^2 \tag{40}$$

and

$$\boldsymbol{c}[k+1] = \frac{\hat{\boldsymbol{c}}[k+1]}{||\hat{\boldsymbol{c}}[k+1]||_2}, \tag{41}$$

where δ_{LMS} is the adaptation step size parameter ($||\cdot||_2$ denotes the L_2–norm of a vector). $\boldsymbol{c}[k] \triangleq [c_0[k]\ c_1[k]\ \ldots\ c_{L_c-1}[k]]^H$ is the equalizer coefficient vector which is now time–variant. Eq. (41) ensures that the constraint given by Eq. (25) is fulfilled [27, 31, 32]. Here, $e[k]$ is given by

$$e[k] = \Delta\hat{s}[k-k_0] q_{\text{ref}}[k-1] - \boldsymbol{c}^H[k]\boldsymbol{r}[k], \tag{42}$$

with

$$\boldsymbol{r}[k] = [r[k]\ r[k-1]\ \ldots\ r[k-L_c+1]]^T. \tag{43}$$

Note that $q_{\mathrm{ref}}[k-1]$ depends only on $\boldsymbol{c}[k-\nu]$, $\nu > 0$, but not on $\boldsymbol{c}[k]$. Therefore, it has to be treated like a constant for differentiation with respect to $\boldsymbol{c}[k]$ (cf. [5, 7]). Hence, Eq. (40) may be rewritten to

$$\hat{\boldsymbol{c}}[k+1] = \boldsymbol{c}[k] + \delta_{\mathrm{LMS}} e^*[k]\boldsymbol{r}[k]. \tag{44}$$

The resulting modified LMS algorithm consists of Eqs. (41), (42), and (44). Note that $e^*[k]\boldsymbol{r}[k]$ does not depend on the channel phase Θ. Thus, $\boldsymbol{c}[k]$ is not influenced by Θ and the adaptive algorithm is noncoherent. Our simulations show that the proposed modified LMS algorithm is also very robust against phase changes (frequency offset).

4.2 Modified RLS Algorithm

For the modified RLS algorithm, we define the cost function

$$J_{\mathrm{RLS}}[k] \triangleq \sum_{\mu=1}^{k} w^{k-\mu} |e[k,\,\mu]|^2, \tag{45}$$

with

$$e[k,\,\mu] \triangleq \Delta\hat{s}[\mu - k_0] q_{\mathrm{ref}}[\mu - 1] - \boldsymbol{c}^H[k]\boldsymbol{r}[\mu]. \tag{46}$$

w, $0 < w \leq 1$, is the forgetting factor of the resulting RLS algorithm. The minimum of this cost function can be found by differentiating $J_{\mathrm{RLS}}[k]$ with respect to $\boldsymbol{c}[k]$ and setting the result equal to zero. This yields the equation

$$\boldsymbol{\Phi}[k]\boldsymbol{c}[k] = \boldsymbol{\varphi}[k], \tag{47}$$

with the definitions

$$\boldsymbol{\Phi}[k] \triangleq \sum_{\mu=1}^{k} w^{k-\mu}\boldsymbol{r}[\mu]\boldsymbol{r}^H[\mu] = \boldsymbol{r}[k]\boldsymbol{r}^H[k] + w\boldsymbol{\Phi}[k-1] \tag{48}$$

$$\begin{aligned}\boldsymbol{\varphi}[k] &\triangleq \sum_{\mu=1}^{k} w^{k-\mu}\boldsymbol{r}[\mu]\Delta\hat{s}^*[\mu - k_0] q_{\mathrm{ref}}^*[\mu - 1] \\ &= \boldsymbol{r}[k]\Delta\hat{s}^*[k - k_0] q_{\mathrm{ref}}^*[k-1] + w\boldsymbol{\varphi}[k-1].\end{aligned} \tag{49}$$

Here, $\Delta\hat{s}[k-k_0]q_{\mathrm{ref}}[k-1]$ may be considered as *desired* response and thus, the resulting modified RLS algorithm consists of the following equations [4]:

$$\boldsymbol{l}[k] = \frac{\boldsymbol{P}[k-1]\boldsymbol{r}[k]}{w + \boldsymbol{r}^H[k]\boldsymbol{P}[k-1]\boldsymbol{r}[k]}, \tag{50}$$

$$\xi[k] = \Delta\hat{s}[k-k_0]q_{\mathrm{ref}}[k-1] - \boldsymbol{c}^H[k-1]\boldsymbol{r}[k], \tag{51}$$

$$\hat{\boldsymbol{c}}[k] = \boldsymbol{c}[k-1] + \boldsymbol{l}[k]\xi^*[k], \tag{52}$$

$$\boldsymbol{P}[k] = w^{-1}\boldsymbol{P}[k-1] - w^{-1}\boldsymbol{l}[k]\boldsymbol{r}^H[k]\boldsymbol{P}[k-1], \tag{53}$$

$$\boldsymbol{c}[k] = \frac{\hat{\boldsymbol{c}}[k]}{||\hat{\boldsymbol{c}}[k]||_2}. \tag{54}$$

$\boldsymbol{P}[k]$ is initialized by

$$\boldsymbol{P}[0] = \delta_{\mathrm{RLS}}^{-1} \boldsymbol{I}_{L_c \times L_c}, \tag{55}$$

where δ_{RLS} is a small positive constant (typical value: 0.004) and $\boldsymbol{I}_{L_c \times L_c}$ is the $L_c \times L_c$ identity matrix. Note that $\boldsymbol{c}[k]$ is again independent of the channel phase Θ and the modified RLS algorithm is also robust against phase variations. Moreover, the equalizer setting delivered by both modified algorithms is only unique up to a complex factor with magnitude one and arbitrary phase which is mandatory for noncoherent equalizers [5, 6].

4.3 Stationary Points of the Adaptive Algorithms

For noncoherent equalizers the error signal $e[k]$ does not only depend on the current equalizer output symbol $q[k]$ but also on previously received symbols $q[k-\nu]$, $\nu > 0$. This is an important difference between conventional coherent equalizers and noncoherent equalizers. For calculation of the noncoherent cost function in Section 3, it was taken into account that $q[k-\nu]$, $\nu > 0$, also depends on $\boldsymbol{c}$. On the other hand, in order to obtain efficient, fast converging adaptive algorithms, in this section, we had to treat $\boldsymbol{c}[k-\nu]$, $\nu > 0$, as constants for derivation of $e[k]$ with respect to $\boldsymbol{c}[k]$ as it is also customary in literature [5, 7, 6, 10]. For this reason, it is interesting to investigate if the equalizer tap settings delivered by the adaptive algorithms minimize the cost function discussed in Section 3.1. Here, for convenience, again we restrict our analysis to nonrecursively generated reference symbols $q_{\mathrm{ref}}[k-1]$. Nevertheless, the results can be extended to the scheme with recursively generated reference symbol.

The modified LMS algorithm described in Section 4.1 has reached a stationary point, if $\mathcal{E}\{\boldsymbol{c}[k+1]\} = \mathcal{E}\{\boldsymbol{c}[k]\} = \boldsymbol{c}$ is valid. In this case, from Eqs. (41) and (44)

$$\mathcal{E}\{e^*[k]\boldsymbol{r}[k]\} = \frac{1}{\delta_{\mathrm{LMS}}} \mathcal{E}\{\hat{\boldsymbol{c}}[k+1] - \boldsymbol{c}[k]\} = -\eta' \boldsymbol{c} \tag{56}$$

can be obtained, where η' is a real constant. Using the assumption $\Delta\hat{s}[k-k_0-\nu] = \Delta s[k-k_0-\nu]$, $\nu \geq 0$, and Eqs. (4), (5), (15), and (42) (cf. also Eq. (17)), the ρth row of Eq. (56) may be rewritten to

$$\begin{aligned}
\mathcal{E}\{e^*[k]r[k-\rho]\} &= \mathcal{E}\left\{e^*[k]\left(\mathrm{e}^{j\Theta}\sum_{\nu=0}^{L_h-1} h_\nu s[k-\rho-\nu] + n[k-\rho]\right)\right\} \\
&= -\sum_{\substack{\mu=0 \\ \mu\neq k_0}}^{L_h+L_c-2} g_\mu^* h_{\mu-\rho} - \sigma_n^2 c_\rho^* \\
&= -\sum_{\substack{\mu=0 \\ \mu\neq k_0}}^{L_h+L_c-2} \boldsymbol{h}_\mu^H \boldsymbol{c} h_{\mu-\rho} - \sigma_n^2 c_\rho^* = -\eta' c_\rho^*.
\end{aligned} \tag{57}$$

Using this result, Eq. (56) can be transformed in the eigenvalue equation

$$\sum_{\substack{\mu=0 \\ \mu \neq k_0}}^{L_h+L_c-2} \boldsymbol{h}_\mu \boldsymbol{h}_\mu^H \boldsymbol{c} = \boldsymbol{H}' \boldsymbol{c} = (\eta' - \sigma_n^2)\boldsymbol{c} \stackrel{\triangle}{=} \lambda' \boldsymbol{c}, \tag{58}$$

where λ' denotes the eigenvalue. In order to show that the modified LMS algorithm enjoys global convergence, we may define a corresponding LMS cost function $J_L^{\mathrm{LMS}}(\boldsymbol{c})$

$$J_L^{\mathrm{LMS}}(\boldsymbol{c}) = \boldsymbol{c}^H \boldsymbol{H}' \boldsymbol{c} + \sigma_n^2 \boldsymbol{c}^H \boldsymbol{c} - \eta' \boldsymbol{c}^H \boldsymbol{c} + \tau \tag{59}$$

which can be obtained from Eq. (58) by integration. Here, τ is an arbitrary constant, which has no influence on the convergence properties of the modified LMS algorithm.

A comparison of Eq. (59) and Eq. (26) shows that both cost functions are very similar. Since $\boldsymbol{H}'$ is also a positive semi–definite Hermitian matrix and the unit–energy constraint (cf. Eq. (25)) is still valid, all statements about the stationary points of $J_L(\boldsymbol{c})$ (cf. Section 3.2) also apply to those of $J_L^{\mathrm{LMS}}(\boldsymbol{c})$. This means, $J_L^{\mathrm{LMS}}(\boldsymbol{c})$ has no spurious local minima and the modified LMS algorithm enjoys global convergence.

From a comparison of Eq. (58) and Eq. (28) it can be seen, that the solution found by the adaptive algorithm minimizes the cost function $J_L(\boldsymbol{c})$, if the eigenvector corresponding to the minimum eigenvalue of $\boldsymbol{H}'$ coincides with the eigenvector corresponding to the minimum eigenvalue of $\boldsymbol{H}$. This is the case for MDPSK (for arbitrary N) and for MDAPSK with $N \to \infty$ since in these cases $x_\mu = x$, $0 \leq \mu \leq L_c + L_h - 2$, $\mu \neq k_0$, is valid (cf. Appendix A). For the infinite–tap equalizer, $g_\mu = 0$, $\mu \neq k_0$, results as has been shown in Section 3. In this case, the weighting factors x_μ, $0 \leq \mu \leq L_c + L_h - 2$, $\mu \neq k_0$, have no influence on the optimum equalizer coefficient vector $\boldsymbol{c}_{\mathrm{opt}}$, since $\boldsymbol{c}_{\mathrm{opt}}$ is orthogonal to $\boldsymbol{h}_\mu$, $\mu \neq k_0$. Hence, the modified LMS algorithm also converges to the optimum infinite–length equalizer.

For MDAPSK with a finite–length equalizer and $N < \infty$ the tap settings delivered by the modified LMS algorithm do not minimize the cost function. However, our simulations show, that the modified LMS algorithm also delivers a solution close to the optimum one for MDAPSK with finite–length equalizers and $N < \infty$. The reason for this behaviour is that even for MDAPSK with small N, the weighting factors x_μ, $\mu \neq k_0$, differ not significantly, e.g., for 16DAPSK with $\Delta R = 2.0$ and $N = 2$, $x_\mu = 2.56$, $0 \leq \mu \leq L_c + L_h - 2$, $\mu \notin \{k_0 - 1, k_0\}$, and $x_{k_0-1} = 3.13$, result.

It is also interesting to note that the parameter N has no influence on the steady–state equalizer tap setting obtained by the modified LMS algorithm (cf. Eq. (58)).

For simplicity, we focused on the modified LMS algorithm in this section, however, all results presented are also valid for the modified RLS algorithm presented in Section 4.2.

4.4 Stability Analysis of the Modified LMS Algorithm

A general stability analysis for the proposed algorithms is very difficult. Therefore, in this section, we investigate only the stability of the mean $\mathcal{E}\{\boldsymbol{c}[k]\}$ [4, 3]. Although the resulting bounds on the step size parameter δ_{LMS} do not ensure convergence of the modified LMS algorithm, they are very useful hints for a proper selection of δ_{LMS}. Here, a similar procedure like that for the conventional LMS algorithm is used (cf. [3]), i.e., the instantaneous gradient $e^*[k]\boldsymbol{r}[k]$ in Eq. (44) is replaced by the mean gradient $\mathcal{E}\{e^*[k]\boldsymbol{r}[k]\}$:

$$\hat{\boldsymbol{c}}[k+1] = \boldsymbol{c}[k] + \delta_{\mathrm{LMS}}\mathcal{E}\{e^*[k]\boldsymbol{r}[k]\}. \tag{60}$$

The resulting algorithm may be considered as a modified *steepest descent algorithm* [4].

If the equalizer coefficients are time–variant, the resulting error signal $e[k]$ is given by

$$\begin{aligned} e[k] \;=\; & \mathrm{e}^{j\Theta}\left(\frac{\sum_{\nu=1}^{N-1} g_{k_0}[k-\nu]|s[k-k_0-\nu]|^2}{\sum_{\nu=1}^{N-1}|s[k-k_0-\nu]|^2} - g_{k_0}[k]\right)s[k-k_0] \\ & + \mathrm{e}^{j\Theta}\sum_{\substack{\mu=0\\ \mu\neq k_0}}^{L_h+L_c-2}\left(\frac{\sum_{\nu=1}^{N-1} g_{\mu}[k-\nu]s[k-\mu-\nu]s^*[k-k_0-\nu]}{\sum_{\nu=1}^{N-1}|s[k-k_0-\nu]|^2}s[k-k_0] - g_{\mu}[k]s[k-\mu]\right) \\ & + \sum_{\mu=0}^{L_c-1}\left(\frac{\sum_{\nu=1}^{N-1} c_{\mu}[k-\nu]n[k-\mu-\nu]s^*[k-k_0-\nu]}{\sum_{\nu=1}^{N-1}|s[k-k_0-\nu]|^2}s[k-k_0] - c_{\mu}[k]n[k-\mu]\right), \end{aligned} \tag{61}$$

(cf. Eqs. (5), (14),(42)) with

$$g_{\mu}[k] = \boldsymbol{c}^H[k]\boldsymbol{h}_{\mu}. \tag{62}$$

In contrast to the time–invariant case (cf. Eq. (17)), in the time–variant case $g_{k_0}[k]$ is not constant and thus, $g_{k_0}[k]$ has an influence on $e[k]$. This makes a theoretical stability analysis of the modified steepest descent algorithm very difficult if not impossible. Therefore, for the sake of mathematical tractability, in the following, we make the simplifying assumption $g_{k_0}[k] \approx g_{k_0}[k-\nu]$, $1 \leq \nu \leq N-1$, i.e., $\boldsymbol{c}[k] \approx \boldsymbol{c}[k-\nu]$, $1 \leq \nu \leq N-1$. This simplification is justified only if N is not too large and for small values of δ_{LMS} or if the algorithm operates close to steady state. The advantage of this approximation is that the first term in Eq. (61) may be neglected for calculation of $\mathcal{E}\{e^*[k]\boldsymbol{r}[k]\}$. Hence, $\mathcal{E}\{e^*[k]\boldsymbol{r}[k]\}$ may be calculated in a similar way like the gradient in Eqs. (57), (58), and Eq. (60) can be rewritten to

$$\hat{\boldsymbol{c}}[k+1] = \boldsymbol{c}[k] - \delta_{\mathrm{LMS}}\left(\boldsymbol{H}' + \sigma_n^2\,\boldsymbol{I}_{L_c\times L_c}\right)\boldsymbol{c}[k]. \tag{63}$$

Since $\boldsymbol{H}'$ is a Hermitian matrix, it can be expressed as [4]

$$\boldsymbol{H}' \triangleq \boldsymbol{C}\boldsymbol{\Lambda}'\boldsymbol{C}^H, \tag{64}$$

with the definitions

$$\boldsymbol{C} \triangleq [\boldsymbol{c}_0\ \boldsymbol{c}_1\ \ldots\ \boldsymbol{c}_{L_c-1}], \tag{65}$$

$$\boldsymbol{\Lambda}' \triangleq \mathrm{diag}\{\lambda'_0,\ \lambda'_1,\ \ldots,\ \lambda'_{L_c-1}\}. \tag{66}$$

Here, $\mathrm{diag}\{y_0,\ y_1,\ \ldots,\ y_{L_y-1}\}$ denotes an $L_y \times L_y$ diagonal matrix and without loss of generality, $\lambda'_0 < \lambda'_1 \leq \ldots \leq \lambda'_{L_c-1}$ is assumed. $\boldsymbol{c}_\nu$ ($\boldsymbol{c}_\nu^H \boldsymbol{c}_\nu = 1$) is the eigenvector of matrix $\boldsymbol{H}'$ which corresponds to eigenvalue λ'_ν, $0 \leq \nu \leq L_c - 1$. For simplicity, we assume again that the minimum eigenvalue λ'_0 has multiplicity one. Now, we introduce the transformed coefficient vectors

$$\boldsymbol{u}[k] \triangleq \boldsymbol{C}^H \boldsymbol{c}[k], \tag{67}$$

$$\hat{\boldsymbol{u}}[k] \triangleq \boldsymbol{C}^H \hat{\boldsymbol{c}}[k]. \tag{68}$$

Using this, Eqs. (63) and (41) may be transformed into

$$\hat{\boldsymbol{u}}[k+1] = \left(\boldsymbol{I}_{L_c \times L_c} - \delta_{\mathrm{LMS}}\left(\boldsymbol{\Lambda}' + \sigma_n^2 \boldsymbol{I}_{L_c \times L_c}\right)\right) \boldsymbol{u}[k], \tag{69}$$

and

$$\boldsymbol{u}[k+1] = \frac{\hat{\boldsymbol{u}}[k+1]}{||\hat{\boldsymbol{u}}[k+1]||_2} \tag{70}$$

The μth element $\hat{u}_\mu[k+1]$ of $\hat{\boldsymbol{u}}[k+1]$ can be obtained from

$$\hat{u}_\mu[k+1] = \left(1 - \delta_{\mathrm{LMS}}\left(\lambda'_\mu + \sigma_n^2\right)\right) u_\mu[k],\ \ 0 \leq \mu \leq L_c - 1, \tag{71}$$

where $u_\mu[k]$ denotes the μth element of vector $\boldsymbol{u}[k]$. For a stable optimum solution, Eqs. (69) and (70) have to converge to $\boldsymbol{u}[k] = [1\ 0\ \ldots\ 0]^T$ for $k \to \infty$ since only in this case $\boldsymbol{c}[k] = \boldsymbol{c}_0$ results[6]. Hence, from Eqs. (71) and (70), the conditions

$$1 - \delta_{\mathrm{LMS}}\left(\lambda'_0 + \sigma_n^2\right) > 0 \tag{72}$$

and

$$1 - \delta_{\mathrm{LMS}}\left(\lambda'_0 + \sigma_n^2\right) > \left|1 - \delta_{\mathrm{LMS}}\left(\lambda'_\mu + \sigma_n^2\right)\right|,\ \ \mu \neq 0 \tag{73}$$

can be derived. Eq. (72) has to be fulfilled for a stable solution and Eq. (73) is necessary for dominance of the *natural mode* [4] corresponding to the minimum eigenvalue λ'_0. Hence,

[6]Note that $\{\boldsymbol{c}_\nu | \nu \in \{0,\ 1,\ \ldots,\ L_c - 1\}\}$ is an (orthonormal) basis for $\mathbb{C}^{L_c}$, cf. Eq. (67).

the modified steepest descent algorithm is stable if the step size parameter δ_{LMS} satisfies the following condition:

$$0 \leq \delta_{\text{LMS}} < \frac{2}{\lambda'_0 + \lambda'_{L_c-1} + 2\sigma_n^2} \tag{74}$$

(for $\delta_{\text{LMS}} = 0$ the stability of the algorithm is obvious). In order to avoid the calculation of the eigenvalues of $\boldsymbol{H}'$ in a practical application, a simple lower bound for $\frac{2}{\lambda'_0+\lambda'_{L_c-1}+2\sigma_n^2}$ may be used. Since the trace $\text{tr}\{\boldsymbol{H}'\}$ of $\boldsymbol{H}'$ is equal to $\text{tr}\{\boldsymbol{\Lambda}'\} = \sum_{\nu=0}^{L_c-1} \lambda_\nu \geq \lambda'_0 + \lambda'_{L_c-1}$, it can be expected that the steepest descent algorithm is stable for

$$0 \leq \delta_{\text{LMS}} < \frac{2}{\text{tr}\{\boldsymbol{H}'\} + 2\sigma_n^2}. \tag{75}$$

Although some simplifications were necessary for derivation of Eq. (75), which is valid for the modified steepest descent algorithm, it may also be considered as a rule of thumb for a proper choice of δ_{LMS} for the modified LMS algorithm. For a further discussion of the modified steepest descent algorithm we refer to [26].

4.5 Speed of Convergence of the Modified Adaptive Algorithms

Figs. 3a) and b) show learning curves for the proposed modified LMS ($\delta_{\text{LMS}} = 0.01$) and RLS ($w = 1.0$) algorithms, respectively, for a QDPSK constellation ($M = 4$). The impulse response of the channel (referred to as Channel A) used is $h_0 = 0.304$, $h_1 = 0.903$, $h_2 = 0.304$ ($L_h = 3$), and $10\log_{10}(E_b/N_0) = 10$ dB ($E_b = E_S/\log_2(M) = E_S/2$) is valid. The equalizer length is chosen to $L_c = 7$ and the decision delay is $k_0 = 4$. In order to demonstrate that the convergence speed of the proposed algorithms is similar to that of the corresponding conventional adaptive algorithms, we also included the learning curves for a conventional LMS ($\delta_{\text{LMS}} = 0.01$) and a conventional RLS ($w = 1.0$) algorithm [3, 4]. However, it has to be mentioned that the error signals of the modified and the conventional algorithms are completely different. Therefore, a direct comparison of the steady–state error is **not** possible. For the modified LMS algorithm, $J'[k] = \mathcal{E}\{|e[k]|^2\}$ is valid, whereas $J'[k] = \mathcal{E}\{|\xi[k]|^2\}$ is used for the modified RLS algorithm. For the conventional algorithms, the definitions proposed in [4] are used. In all cases, averaging was done over 1000 adaptation processes; $\boldsymbol{c}[0]$ was initialized with $c_{k_0}[0] = 1.0$, $c_\mu[0] = 0$, $\mu \neq k_0$, and a training sequence was used. For the modified algorithms, the reference symbol was calculated nonrecursively. It can be seen from Figs. 3a) and b), that the steady–state error of the modified LMS and RLS algorithm decreases as N increases. The dashed lines correspond to the theoretical steady–state error variance of an infinite–length NMIE calculated from Eqs. (30) and (79). There is a good agreement between theory and simulation. The simulated steady–state error is slightly higher

since a finite–length equalizer is used and because of gradient noise. The convergence speed of the modified RLS algorithm is considerably higher for $N = 5, 10$ than for $N = 2, 3$, whereas the convergence speed of the modified LMS algorithm is hardly influenced by N. This behaviour of the modified LMS algorithm could also be expected from the analysis in Section 4.4. From Eq. (71) it can be seen that the transient behaviour of the corresponing steepest descent algorithm does not depend on N. Since the steepest descent and the LMS algorithm are closely related, it is clear that the convergence speed of the latter is also independent of N. Note that for Channel A and the equalizer used in this example, Eq. (75) suggests $0 \leq \delta_{\text{LMS}} < 0.3$, i.e., our choice ($\delta_{\text{LMS}} = 0.01$) fulfills this condition. In order to keep the steady–state error as small as possible, i.e., close to the theoretical lower limit, we have chosen δ_{LMS} considerably smaller than the upper limit proposed by Eq. (75).

For comparison Figs. 4a), and b) show the learning curves of the proposed modified LMS ($\delta_{\text{LMS}} = 0.01$, upper limit according to Eq. (75): $\delta_{\text{LMS}} = 0.3$) and RLS ($w = 1.0$) algorithms, respectively, for a 16DAPSK constellation ($\Delta R = 2.0$). The same equalizer length, decision delay, and equalizer initialization are used as for Fig. 3. Also, Channel A and a training sequence are used. However, now $10 \log_{10}(E_b/N_0) = 16$ dB ($E_b = E_S/\log_2(M) = E_S/4$) is valid. Again the dashed lines correspond to the theoretical steady–state error variance of an infinite–length NMIE calculated from Eqs. (30) and (80). As for QDPSK, the agreement between simulation and theory is very good. It is also confirmed that the modified LMS algorithm is hardly influenced by N, whereas the convergence speed of the modified RLS algorithm increases with N. It should be mentioned that if N is chosen too large and steady state is not reached yet, $\boldsymbol{c}[k]$ changes considerably in the interval $[k_1,\, k_1 + N - 1]$ which has a negative influence on the reference symbol $q_{\text{ref}}[k-1]$, and the convergence speed decreases again.

In conclusion, we can state that the proposed modified adaptive algorithms have similar convergence speeds as their conventional counterparts (cf. Fig. 3) and the steady–state error is close to the theoretical minimum.

5 Simulation Results

In this section, the performance of the proposed NMIE is evaluated for QDPSK and 16DAPSK by computer simulations. The equalizer coefficients are adapted by the modified LMS algorithm. $\delta_{\text{LMS}} = 0.001$ is chosen much lower than the upper limit suggested by Eq. (75) since here only the steady–state behaviour is of interest. A training sequence is transmitted until steady state is reached; then the equalizer operates in the decision–directed mode. In

the simulations, the novel NMIE scheme is compared with coherent ZF equalization [3] for MDPSK and MDAPSK. The coherent scheme determines an estimate $\hat{s}[k-k_0]$ for the transmitted MPSK/MAPSK symbol $s[k-k_0]$, and subsequently differential encoding is inverted ($\Delta\hat{s}[k-k_0] = \hat{s}[k-k_0]/\hat{s}[k-k_0-1]$).

First, 16DAPSK signals ($\Delta R = 2.0$) are transmitted over Channel A ($h_0 = 0.304$, $h_1 = 0.903$, $h_2 = 0.304$). Figs. 5) and 6) show the bit error rate (BER) vs. $10\log_{10}(E_b/N_0)$ for nonrecursively and recursively generated reference symbols $q_{\text{ref}}[k-1]$, respectively. $L_c = 7$ and $k_0 = 4$ are used for NMIE and ZF equalization. As N and W increase the power efficiency of the noncoherent scheme improves. The performance of coherent ZF equalization is approached for $N \gg 1$ and $W \gg 1$. Note that for $N = 2$ and $W = 0$ the same performance is obtained since in this case the nonrecursively and recursively generated reference symbols are identical (cf. Section 2).

Figs. 7 and 8 show again BER vs. $10\log_{10}(E_b/N_0)$ for nonrecursively and recursively generated reference symbols $q_{\text{ref}}[k-1]$, respectively. Here, QDPSK signals are transmitted over Channel B ($L_h = 5$, $h_0 = 0.18$, $h_1 = -0.25$, $h_2 = 0.9$, $h_3 = -0.25$, $h_4 = 0.18$). The equalizer length is chosen to $L_c = 11$, whereas a decision delay of $k_0 = 6$ is used. Besides the BERs for the proposed NMIE and the coherent ZF equalizer, for comparison the BER of the noncoherent MMSE equalizer (NMMSEE) proposed in [6] is also included in Figs. 7 and 8. It can be observed that the NMMSEE performs only slightly better than the NMIE. For the scheme with recursively generated reference symbol, α is used as parameter instead of W since this is customary in the MDPSK literature [22, 23, 6]. The power efficiency of the proposed NMIE improves as N and α increase. For $N \gg 1$ and $\alpha \to 1$, coherent ZF equalization is approached.

So far, zero frequency offset has been assumed. However, practical receivers often have to cope with carrier phase variations. Therefore, Figs. 9a) and b) show BER vs. normalized frequency offset ΔfT (T is the symbol duration) for QDPSK transmitted over Channel A. $10\log_{10}(E_b/N_0) = 12$ dB, $L_c = 7$, and $k_0 = 4$ are valid. It can be seen that the sensitivity to frequency offset increases with increasing N and α. On the other hand, for zero frequency offset power efficiency improves for higher values of N and α. Hence, there is a trade–off between performance under pure AWGN condition and robustness against frequency offset. Note that a coherent ZF or MMSE equalizer degrades severely (BER = 0.5) even for very low frequency offsets ($\Delta fT < 0.0001$) since the corresponding conventional adaptive gradient algorithms fail to follow the phase changes. Note also that the modified LMS algorithm applied for the NMIE does not follow the frequency offset since it is a noncoherent adaptive algorithm.

6 Conclusions

In this paper, a novel noncoherent equalizer is presented and analyzed. It consists of a combination of a linear equalizer and a DF–DD. The proposed linear equalizer minimizes the variance of ISI in the equalizer output signal and the DF–DD, whose reference symbol may be generated either nonrecursively or recursively, removes the dependence on the channel phase. The optimum equalizer coefficients may be calculated from an eigenvalue problem which is obtained from a constrained optimization task.

It is shown that for infinite–length equalizers, the resulting equalization scheme is equivalent to a coherent ZF equalizer. For an efficient adaptation of the equalizer coefficients novel modified LMS and RLS algorithms are proposed. The convergence speed of these algorithms is comparable to that of the corresponding conventional (coherent) adaptive algorithms. It is also shown that the modified algorithms enjoy global convergence. In addition, the stability of the modified LMS algorithm is investigated. Simulation results demonstrate the good performance of the proposed equalization scheme for MDPSK and MDAPSK constellations. Moreover, it is shown that the novel scheme is robust against frequency offsets and that there is a trade–off between performance under pure AWGN conditions and insensitivity against frequency offset.

Acknowledgement

The authors would like to thank TCMC (Technology Centre for Mobile Communication), Philips Semiconductors, Germany, for supporting this work.

Appendix A

In this appendix, x_μ (cf. Eq. (21)), $0 \leq \mu \leq L_c + L_h - 2$, $\mu \neq k_0$, and γ (cf. Eq. (20)) are calculated for some special cases:

A. $N \to \infty$

For $N \to \infty$, Eqs. (21) and (20) yield

$$x \stackrel{\triangle}{=} x_\mu = 1, \tag{76}$$

and

$$\gamma = 1, \tag{77}$$

respectively. Note that these results are valid for arbitrary MDPSK and MDAPSK modulation formats.

B. MDPSK

Since $|s[k]| = 1, \forall\, k$, holds for MDPSK,

$$x \;\triangleq\; x_\mu = 1 + \frac{1}{N-1} = \frac{N}{N-1}, \tag{78}$$

$$\gamma \;=\; 1 + \frac{1}{N-1} = \frac{N}{N-1}, \tag{79}$$

is obtained.

C. 16DAPSK

For 16DAPSK, γ can be calculated to

$$\gamma = 1 + \frac{1}{2^{N-1}} \sum_{\nu=0}^{N-1} \binom{N-1}{\nu} \frac{1}{(N-1-\nu)R_2^2 + \nu R_1^2}. \tag{80}$$

Appendix B

In this appendix, the stationary points of $J_L(\boldsymbol{c})$ obtained from Eq. (28) are examined. We assume $0 \leq \lambda_0 \triangleq \lambda_{\min} \leq \lambda_1 \leq \ldots \leq \lambda_{L_c-2} \leq \lambda_{L_c-1} \triangleq \lambda_{\max}$ and denote the corresponding eigenvectors by $\boldsymbol{c}_\nu$, $0 \leq \nu \leq L_c - 1$. First, it is assumed that the eigenvalues λ_0 and λ_{L_c-1} have multiplicity one.

Now, Eq. (26) may be rewritten to

$$J_L(\boldsymbol{c}_\nu) = \sigma_e^2(\boldsymbol{c}_\nu) = \lambda_\nu + \gamma\sigma^2. \tag{81}$$

Since $\boldsymbol{c}_\nu$, $0 \leq \nu \leq L_c - 1$, are the only stationary points, $J_L(\boldsymbol{c})$ attains its absolute maximum for $\boldsymbol{c} = \boldsymbol{c}_{L_c-1}$ corresponding to the maximum eigenvalue $\lambda_{L_c-1} = \lambda_{\max}$. On the other hand, $J_L(\boldsymbol{c})$ has an unique absolute minimum for $\boldsymbol{c} = \boldsymbol{c}_0 \triangleq \boldsymbol{c}_{\text{opt}}$ corresponding to the minimum eigenvalue $\lambda_0 = \lambda_{\min}$.

In the following, it is shown that the remaining stationary points $\boldsymbol{c}_\nu$, $1 \leq \nu \leq L_c - 2$, are saddle points. For this, we introduce the vector

$$\boldsymbol{c}'_\nu = \frac{\boldsymbol{c}_\nu + \varepsilon \boldsymbol{z}}{||\boldsymbol{c}_\nu + \varepsilon \boldsymbol{z}||_2}, \tag{82}$$

where $||\cdot||_2$ denotes the L_2–norm of a vector and $0 < \varepsilon \ll 1$ is a small positive constant. Vector $\boldsymbol{z}$ is orthogonal to $\boldsymbol{c}_\nu$ and normalized to unit norm, i.e.,

$$\boldsymbol{c}_\nu^H \boldsymbol{z} \;=\; 0, \tag{83}$$

$$\boldsymbol{z}^H \boldsymbol{z} \;=\; 1. \tag{84}$$

Vector $\boldsymbol{c}'_\nu$ is close to $\boldsymbol{c}_\nu$ and meets also the unit–energy constraint, i.e., ${\boldsymbol{c}'_\nu}^H \boldsymbol{c}'_\nu = 1$. Now, we make use of the relation $||\boldsymbol{c}_\nu + \varepsilon \boldsymbol{z}||_2^2 = 1 + \varepsilon^2$ and insert $\boldsymbol{c}'_\nu$ in Eq. (26). This yields

$$J_L(\boldsymbol{c}'_\nu) = \frac{1}{1+\varepsilon^2}(\boldsymbol{c}_\nu^H \boldsymbol{H} \boldsymbol{c}_\nu + \varepsilon \boldsymbol{c}_\nu^H \boldsymbol{H} \boldsymbol{z} + \varepsilon \boldsymbol{z}^H \boldsymbol{H} \boldsymbol{c}_\nu + \varepsilon^2 \boldsymbol{z}^H \boldsymbol{H} \boldsymbol{z}) + \gamma\sigma_n^2. \tag{85}$$

$\boldsymbol{c}_\nu^H \boldsymbol{H} \boldsymbol{z} = 0$ and $\boldsymbol{z}^H \boldsymbol{H} \boldsymbol{c}_\nu = 0$ result from Eqs. (28) and (83). Therefore, Eq. (85) can be simplified to

$$J_L(\boldsymbol{c}'_\nu) = \frac{1}{1+\varepsilon^2}(\lambda_\nu + \varepsilon^2 \boldsymbol{z}^H \boldsymbol{H} \boldsymbol{z}) + \gamma\sigma_n^2, \qquad 1 \le \nu \le L_c - 2. \tag{86}$$

Since distinct eigenvectors of a Hermitian matrix are orthogonal [4], we may set $\boldsymbol{z} = \boldsymbol{c}_0$. In this case, Eq. (86) can be rewritten to

$$J_L(\boldsymbol{c}'_\nu) = \frac{1}{1+\varepsilon^2}(\lambda_\nu + \varepsilon^2 \lambda_{\min}) + \gamma\sigma_n^2 < J_L(\boldsymbol{c}_\nu), \qquad 1 \le \nu \le L_c - 2. \tag{87}$$

On the other hand, for $\boldsymbol{z} = \boldsymbol{c}_{L_c-1}$,

$$J_L(\boldsymbol{c}'_\nu) = \frac{1}{1+\varepsilon^2}(\lambda_\nu + \varepsilon^2 \lambda_{\max}) + \gamma\sigma_n^2 > J_L(\boldsymbol{c}_\nu), \qquad 1 \le \nu \le L_c - 2, \tag{88}$$

results. This shows that depending on the direction of $\boldsymbol{z}$, $J_L(\boldsymbol{c}'_\nu)$ is smaller or larger than $J_L(\boldsymbol{c}_\nu)$. Hence, $\boldsymbol{c}_\nu$, $1 \le \nu \le L_c - 2$, can only correspond to saddle points.

So far, we assumed that $\lambda_{\min}$ and $\lambda_{\max}$ have multiplicity one. If the maximum eigenvalue has a multiplicity larger than one, this has no influence on the uniqueness of the global minimum. If the minimum eigenvalue has multiplicity $m > 1$, there are m different eigenvectors which yield the same error variance $\sigma_e^2(\boldsymbol{c})$. Thus, there are m optimum equalizer settings which yield the same performance. In practice, it is very unlikely that $\lambda_{\min}$ or $\lambda_{\max}$ have a multiplicity larger than one, and thus, for simplicity, we do not focus on this case in our analysis.

References

[1] R. W. Lucky, J. Salz, and E. J. Jr. Weldon. *Principles of Data Communication.* McGraw–Hill, New York, 1968.

[2] S.U.H. Qureshi. Adaptive Equalization. *Proceedings of the IEEE*, 73:1349–1387, September 1985.

[3] J.G. Proakis. *Digital Communications.* McGraw–Hill, New York, Third Edition, 1995.

[4] S. Haykin. *Adaptive Filter Theory.* Prentice-Hall, Upper Saddle River, New Jersey, Third Edition, 1996.

[5] P. Sehier and G. Kawas Kaleh. Adaptive equaliser for differentially coherent receiver. *IEE Proceedings-I*, 137:9–12, February 1990.

[6] R. Schober, W. H. Gerstacker, and J. B. Huber. Adaptive Linear Equalization Combined With Noncoherent Detection of MDPSK Signals. *IEEE Trans. on Commun.*, COM-48:733–738, May 2000.

[7] A. Masoomzadeh-Fard and S. Pasupathy. Nonlinear Equalization of Multipath Fading Channels with Noncoherent Demodulation. *IEEE Journal on Selected Areas in Communications*, SAC-14:512–520, April 1996.

[8] N. Benvenuto and L. Tomba. Performance Comparison of Space Diversity and Equalization Techniques for Indoor Radio Systems. *IEEE Transactions on Vehicular Technology*, 46:358–368, May 1997.

[9] A. E. Jones and R. Perry. Frequency–Offset Invariant Equalizer for Broadband Wireless Networks. In *Proceedings of the IEEE Int. Symposium on Personal, Indoor and Mobile Radio Communications (PIMRIC)*, Boston, 1998.

[10] R. Schober and W. H. Gerstacker. Adaptive Noncoherent DFE for MDPSK Signals Transmitted Over ISI Channels. *To be published IEEE Trans. on Commun.*, July 2000.

[11] W. T. Webb, L. Hanzo, and R. Steele. Bandwidth efficient QAM schemes for Rayleigh fading channels. *Proceedings of the IEEE*, 138:169–175, June 1991.

[12] W. T. Webb and L. Hanzo. *Modern Quadrature Amplitude Modulation.* Pentech Press, London, 1994.

[13] N. A. Svensson. On Differentially Encoded Star 16QAM with Differential Detection and Diversity. *IEEE Transactions on Vehicular Technology*, 44:586–593, August 1995.

[14] F. Adachi and M. Sawahashi. Decision Feedback Differential Detection of Differentially Encoded 16APSK Signals. *IEEE Trans. on Commun.*, COM-44:416–418, August 1996.

[15] R.-Y. Wei and M.-C. Lin. Modified decision feedback differential detection for differentially encoded 16 APSK signals. *Electronics Letters*, 34(4):336–337, 1998.

[16] X. Dong, T. T. Tjhung, and F. Adachi. Error Probability Analysis for 16 STAR–QAM in Frequency–Selective Rician Fading with Diversity Reception. *IEEE Transactions on Vehicular Technology*, 47:924–935, August 1998.

[17] R. Schober, W. H. Gerstacker, and J. B. Huber. Decision–feedback differential detection scheme for 16–DAPSK. *Electronics Letters*, 34(19):1812–1813, September 1998.

[18] M. Machida, S. Handa, and S. Oshita. Multiple–Symbol Differential Detection of APSK Based on MAP–Criterion. In *Proceedings of the IEEE Global Telecommunication Conference (GLOBECOM'98)*, pages 2740–2744, Sydney, November 1998.

[19] H. Kong and E. Shwedyk. Adaptive MLSE of QAM signals over frequency nonselective Rayleigh fading channels. *IEE Proc.–Commun*, 145:426–430, December 1998.

[20] H. Leib and S. Pasupathy. The Phase of a Vector Perturbed by Gaussian Noise and Differentially Coherent Receivers. *IEEE Transactions on Information Theory*, IT-34:1491–1501, November 1988.

[21] F. Edbauer. Bit Error Rate of Binary and Quaternary DPSK Signals with Multiple Differential Feedback Detection. *IEEE Trans. on Commun.*, COM-40:457–460, March 1992.

[22] H. Leib. Data–Aided Noncoherent Demodulation of DPSK. *IEEE Trans. on Commun.*, COM-43:722–725, Feb-Apr 1995.

[23] N. Hamamoto. Differential Detection with IIR Filter for Improving DPSK Detection Performance. *IEEE Trans. on Commun.*, COM-44:959–966, August 1996.

[24] G.D. Forney, Jr. Maximum–Likelihood Sequence Estimation of Digital Sequences in the Presence of Intersymbol Interference. *IEEE Transactions on Information Theory*, IT-18:363–378, 1972.

[25] P.A. Voois, I. Lee, and J.M. Cioffi. The Effect of Decision Delay in Finite-Length Decision Feedback Equalization. *IEEE Transactions on Information Theory*, IT-42:618–621, March 1996.

[26] R. Schober. *Noncoherent Detection and Equalization for MDPSK and MDAPSK Signals*. PhD thesis, University Erlangen–Nürnberg, Erlangen, 2000.

[27] D. D. Falconer and F. R. Magee. Adaptive Channel Memory Truncation for Maximum Likelihood Sequence Estimation. *Bell Technical Journal*, 52:1541–1562, 1973.

[28] N. Al-Dhahir and J.M. Cioffi. Efficiently Computed Reduced–Parameter Input–Aided MMSE Equalizers for ML Detection: A Unified Approach. *IEEE Transactions on Information Theory*, IT-42:903–915, May 1996.

[29] A. Benveniste, M. Goursat, and G. Rouget. Robust Identification of a Nonminimum Phase System: Blind Adjustment of a Linear Equalizer in Data Communications. *IEEE Trans. on Automatic Control*, AC-25:385–399, June 1980.

[30] A. Benveniste, M. Goursat, and G. Rouget. Analysis of Stochastic Approximation Schemes with Discontinuous and Dependent Forcing Terms with Applications to Data Communication Algorithms. *IEEE Trans. on Automatic Control*, AC-25:1042–1058, December 1980.

[31] O. Shalvi and E. Weinstein. New Criteria for Blind Deconvolution of Nonminimum Phase Systems (Channels). *IEEE Transactions on Information Theory*, IT-36:312–321, March 1990.

[32] Z. Xu and M. K. Tsatsanis. Adaptive Minimum Variance Methods for Direct Blind Multichannel Equalization. In *Proceedings of ICASSP'99*, pages 2105–2108, 1998.

Figures:

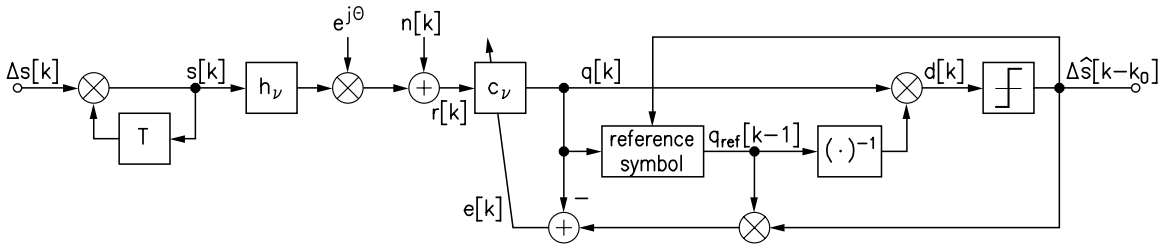

Figure 1: Block diagram of the discrete–time transmission model under consideration with noncoherent linear minimum ISI equalization.

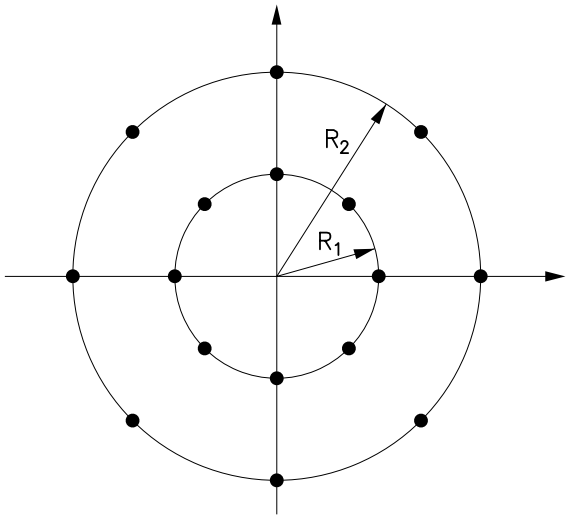

Figure 2: 16DAPSK signal constellation.

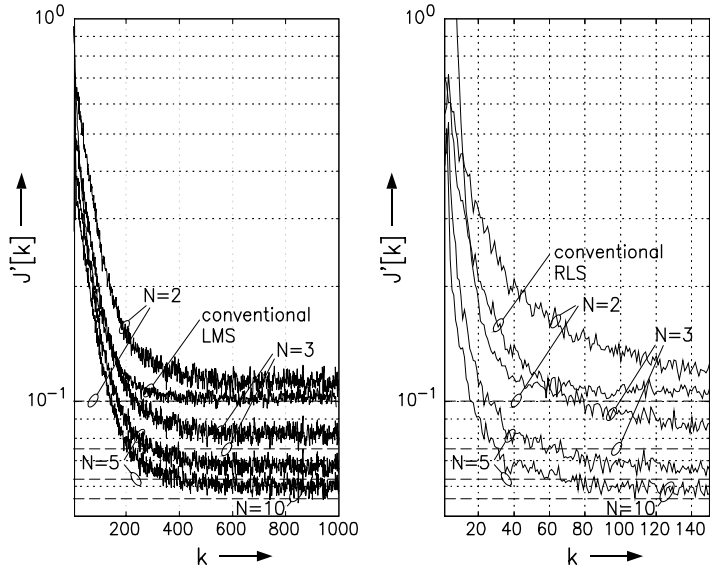

Figure 3: Learning curves for a) modified and conventional LMS algorithm; b) modified and conventional RLS algorithm (QDPSK constellation). Note that the error signal of the conventional algorithms is not the same as that of the modified algorithms. Thus, a direct comparison of the steady–state performance is **not** possible.

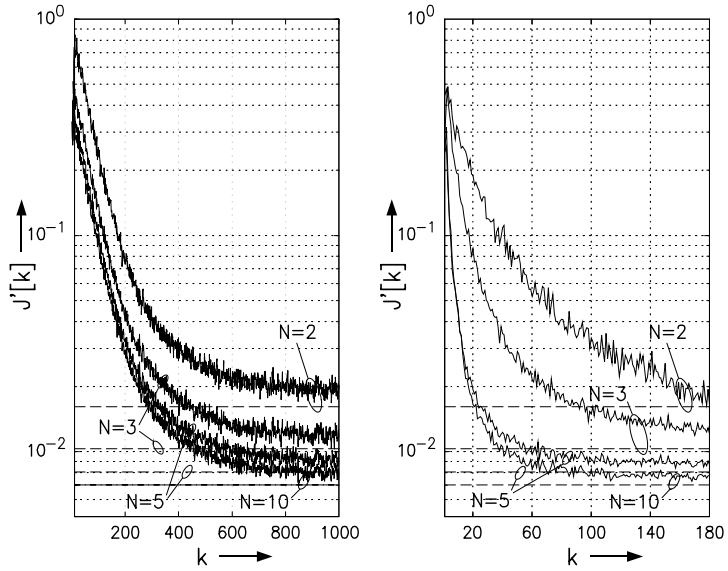

Figure 4: Learning curves for a) modified LMS algorithm; b) modified RLS algorithm (16DAPSK constellation).

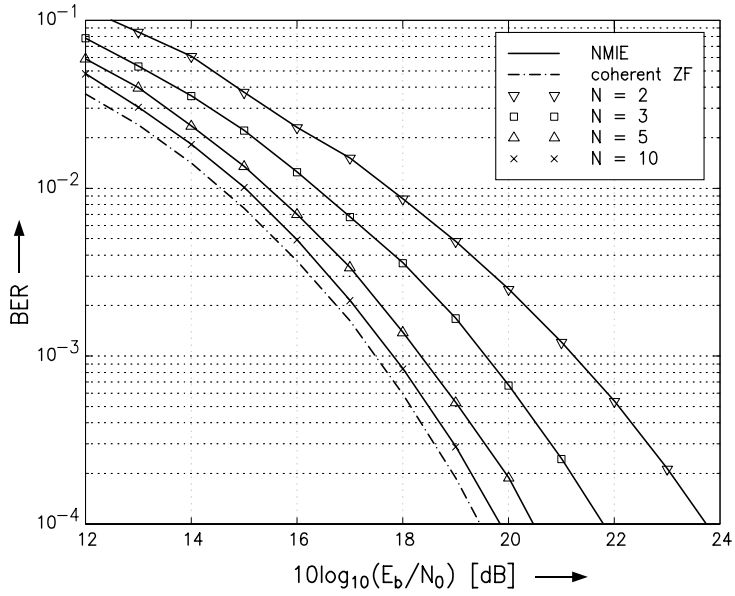

Figure 5: BER vs. $10\log_{10}(E_b/N_0)$ for the NMIE with nonrecursively generated reference symbol. 16DAPSK and Channel A are used. For comparison the BER for a coherent linear ZF equalizer is also shown.

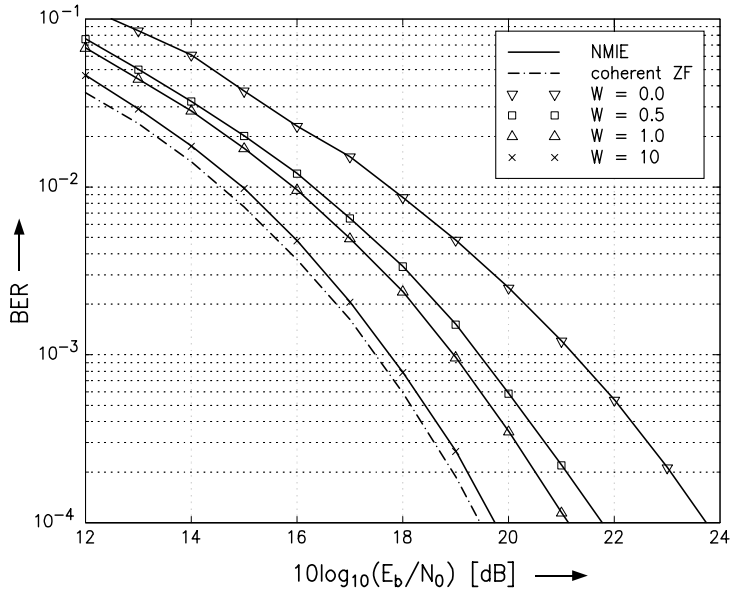

Figure 6: BER vs. $10\log_{10}(E_b/N_0)$ for the NMIE with recursively generated reference symbol. 16DAPSK and Channel A are used. For comparison the BER for a coherent linear ZF equalizer is also shown.

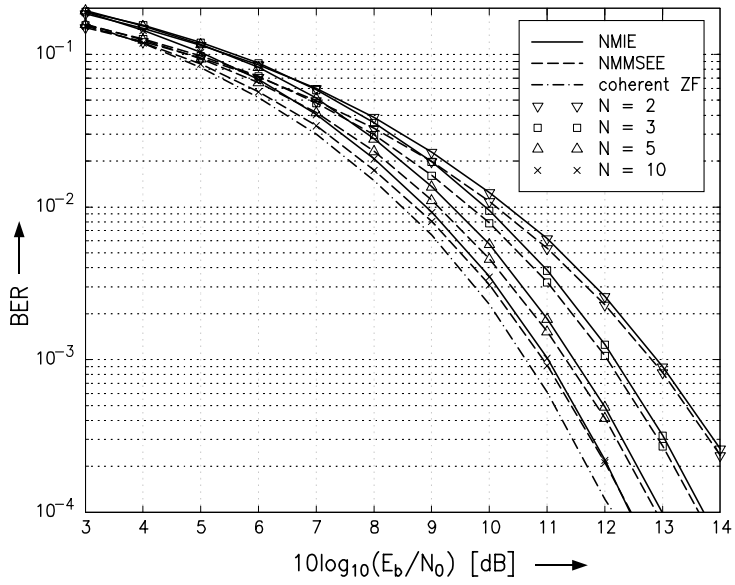

Figure 7: BER vs. $10\log_{10}(E_b/N_0)$ for the NMIE with nonrecursively generated reference symbol. QDPSK and Channel B are used. For comparison the BERs for the NMMSEE proposed in [6] and a coherent linear ZF equalizer are also shown.

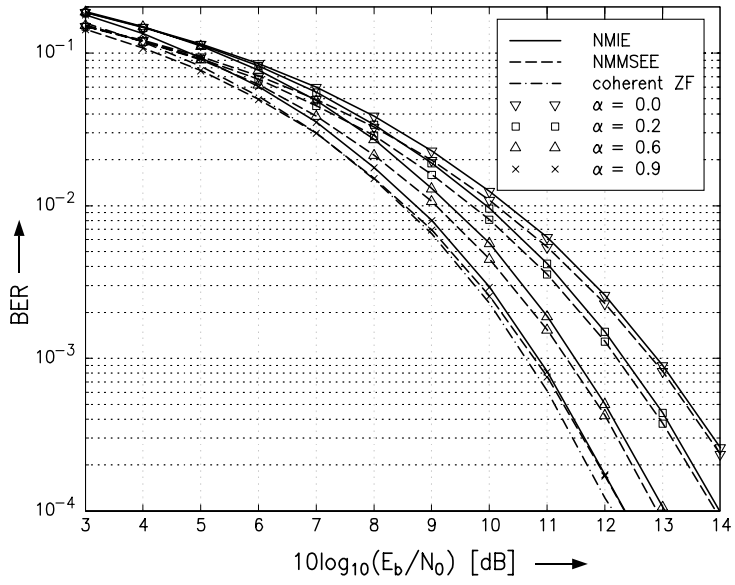

Figure 8: BER vs. $10\log_{10}(E_b/N_0)$ for the NMIE with recursively generated reference symbol. QDPSK and Channel B are used. For comparison the BERs for the NMMSEE proposed in [6] and a coherent linear ZF equalizer are also shown.

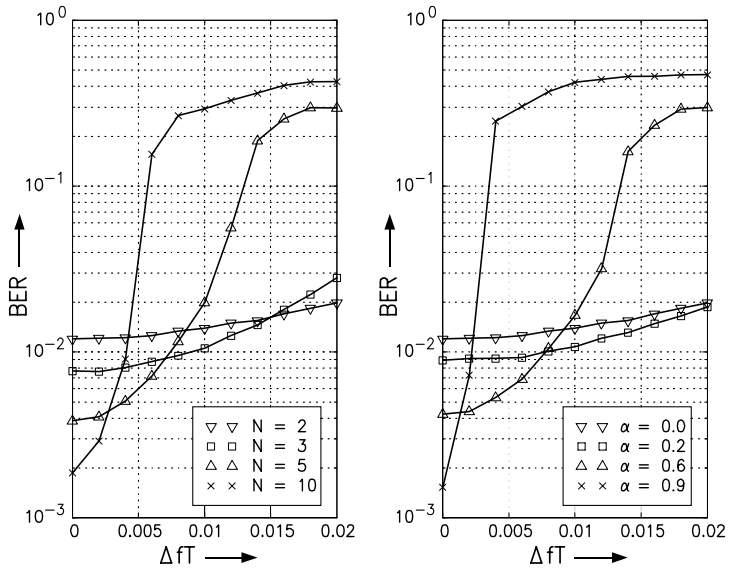

Figure 9: BER vs. ΔfT for NMIE with a) nonrecursively and b) recursively generated reference symbol. QDPSK and Channel A are used.