Decision–Feedback Differential Detection Based on Linear Prediction for MDPSK Signals Transmitted Over Ricean Fading Channels

Robert Schober and Wolfgang H. Gerstacker

Lehrstuhl für Nachrichtentechnik II,
Universität Erlangen-Nürnberg,
Cauerstraße 7/NT, D-91058 Erlangen, Germany
Tel.: +49 9131 8527668
Fax: +49 9131 8528919
E-mail: schober@nt.e-technik.uni-erlangen.de

Abstract

In this paper, linear prediction based decision–feedback differential detection (DF–DD) for M–ary differential phase–shift keying (MDPSK) signals transmitted over Ricean fading channels is proposed. This scheme can improve conventional differential detection (DD) significantly for a multitude of frequency–nonselective channels as will be shown analytically and by computer simulations. Prediction–based DF–DD is particularly well suited for application in mobile communications since the predictor coefficients may be updated regularly using the recursive least–squares (RLS) algorithm. Here, adaptation can start blind, i.e., no training sequence and no a priori knowledge about the channel statistics are required. A further important characteristic of the proposed detection scheme is that no degradation occurs under frequency offset. The bit error rate (BER) performance of QDPSK with genie–aided prediction–based DF–DD is analyzed and it is shown under which conditions the irreducible error floor of conventional DD can be removed entirely. In addition, the influence of Doppler shift is discussed. Finally, the proposed scheme is compared with a second DF–DD scheme which is based on multiple–symbol detection (MSD).

1 Introduction

In this paper, we propose and analyze a novel decision–feedback differential detection (DF–DD) scheme for M–ary differential phase–shift keying (MDPSK) signals transmitted over flat Ricean fading channels[1]. This scheme is based on linear prediction and outperforms conventional differential detection (DD) for AWGN, Rayleigh fading, and Ricean fading conditions. Moreover, it is less complex than multiple–symbol detection (MSD) [2, 3, 4, 5, 6] or schemes based on sequence estimation (e.g. [7, 8, 9, 10, 11]). Especially, it is well suited for application in mobile communications since the predictor coefficients may be adapted to the current statistics of a (nonstationary) channel using the recursive least–squares (RLS) algorithm [12]. Unlike previously proposed DF–DD schemes [13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26] adaptive prediction–based DF–DD has the important advantage that it does not degrade under frequency offset.

Adaptive DF–DD schemes are also proposed by Adachi and Hamamoto in [19] and [20], respectively. However, these schemes use an IIR filter in order to estimate a reference symbol. Since commonly encountered fading processes are better approximated by an autoregressive (AR) model than by a moving average (MA) model, an FIR predictor is better suited [27].

DF–DD based on linear prediction of a reference symbol has been first proposed by Svensson [17]. However, in [17] only Rayleigh fading is considered, and the predictor coefficients are fixed. Bin et al. [21, 22] also propose a prediction–based DF–DD scheme, which uses a linear programming algorithm for calculation of the predictor coefficients. For this, it is assumed that all signal amplitudes are constant within one observation interval. Clearly, this assumption causes a degradation for fast fading processes. Recently, Adachi [23] reported an adaptive DF–DD scheme which applies a linear FIR predictor. However, there the observation interval length is restricted to $N = 3$ (i.e., a second order predictor). Furthermore, it is assumed that both predictor coefficients are real and that their sum is equal to one. These assumptions are not always justified and may lead to degradations. For example, for frequency offset, complex–valued predictor coefficients result.

In contrast to the contributions cited above, we provide a closed–form bit error rate (BER) analysis, valid for QDPSK transmission with Ricean fading and frequency offset, and genie–aided prediction–based DF–DD at the receiver (i.e., it is assumed that all feedback symbols are correct). Moreover, we investigate under which conditions the error floor [28] of conventional DD can be removed entirely.

Finally, we compare prediction–based DF–DD with MSD–based DF–DD. For this, we give a novel MSD–based DF–DD decision rule for Ricean fading which may be considered as a generalization of known decision rules for the AWGN channel and the Rayleigh fading channel reported in [13, 15, 16] and [24], respectively. It will be shown that MSD–based and prediction–based DF–DD are identical for pure AWGN and Rayleigh fading conditions. Although, in general, both schemes are different, their performance is similar as will be

[1]Note, that Ricean fading also comprises Rayleigh fading and pure additive white Gaussian noise (AWGN) as special cases [1].

shown by computer simulations.

This paper is organized as follows. In Section 2, the transmission model is introduced. Prediction–based DF–DD is derived and analyzed in Sections 3 and 4, respectively. The adaptive implementation of prediction–based DF–DD is discussed in Section 5. In Section 6, the proposed prediction–based DF–DD scheme is compared with MSD–based DF–DD. Simulation results are given in Section 7 and some conclusions are drawn in Section 8.

2 Transmission Model

Fig. 1 shows a block diagram of the transmission model under consideration. All signals are represented by their complex–valued baseband equivalents and ideal symbol synchronization is assumed. The MDPSK symbols are denoted by $a[\cdot] \in \mathcal{A}_\phi = \{\mathrm{e}^{j(2\pi\frac{\nu}{M}+\phi)} | \nu \in \{0, 1, \ldots, M-1\}\}$, $\phi \in \{0, \frac{\pi}{M}\}$, and the corresponding differentially encoded MPSK symbols $b[k]$ are given by

$$b[k] = a[k]b[k-1], \quad k \in \mathbb{Z}. \tag{1}$$

Both, transmitter filter $H_t(f)$ and receiver filter $H_r(f)$, have square–root Nyquist characteristic. Hence, no intersymbol interference occurs as long as the continuous–time fading process $f_c(t)$ with equivalent single–sided bandwidth B_f and the factor $\mathrm{e}^{j2\pi\Delta f t}$, where Δf is the frequency offset between modulator and demodulator, do not change significantly during one symbol interval T. In this paper, we assume that these conditions are fulfilled for $B_f T \leq 0.05$ and $\Delta f T \leq 0.05$, respectively. Then, the received signal sample $r[k]$ may be written as

$$r[k] = r(kT) = \mathrm{e}^{j\Theta}\mathrm{e}^{j2\pi\Delta f T k} f[k]b[k] + n[k], \tag{2}$$

where the Ricean fading process $f[\cdot]$ and the noise process $n[\cdot]$ are correlated and uncorrelated complex Gaussian random processes, respectively. Furthermore, $f[\cdot]$ and $n[\cdot]$ are mutually uncorrelated. Θ denotes an unknown, constant, uniformly distributed phase shift. Due to an appropriate normalization, $f[k]$ has power $q_f^2 = \mathcal{E}\{|f[k]|^2\} = 1$ and $n[k]$ has variance $\sigma_n^2 = \mathcal{E}\{|n[k]|^2\} = \frac{N_0}{E_S}$. Here, $\mathcal{E}\{\cdot\}$ denotes expectation and E_S is the mean received energy per symbol, whereas N_0 is the single–sided power spectral density of the underlying passband noise process. The direct and the scattered (Rayleigh) component of the fading process are denoted by $f_d[k] = \mathcal{E}\{f[k]\} \triangleq \mathrm{e}^{j2\pi f_D T k} f_m$ (f_m is the magnitude of the mean of the fading process and f_D, $|f_D| < B_f$, is the Doppler shift) and $f_s[k] = f[k] - f_d[k]$, respectively. Finally, the Ricean factor K is defined as [1]

$$K \triangleq \frac{|f_m|^2}{\sigma_s^2}, \tag{3}$$

where $\sigma_s^2 = \mathcal{E}\{|f_s[k]|^2\}$ is the variance of the scattered component $f_s[k]$. At the receiver, the estimated symbol $\hat{a}[k] \in \mathcal{A}_\phi$ is determined by prediction–based DF–DD (cf. Fig. 1).

3 Prediction–Based DF–DD

In this section, the prediction–based DF–DD decision rule is derived. Using Eq. (1), Eq. (2) may be rewritten to

$$r[k] = a[k]\mathrm{e}^{j\Theta}\mathrm{e}^{j2\pi\Delta fTk}f[k]b[k-1] + n[k]. \tag{4}$$

For optimum coherent detection (CD) of $a[k]$, the reference symbol $\mathrm{e}^{j\Theta}\mathrm{e}^{j2\pi\Delta fTk}f[k]b[k-1]$ has to be known. For conventional DD, $r[k-1]$ is used as estimate for $\mathrm{e}^{j\Theta}\mathrm{e}^{j2\pi\Delta fTk}f[k]b[k-1]$ [29] and because of the noisy character of $r[k-1]$, a performance degradation compared to CD is inevitable. Here, we propose to estimate $\mathrm{e}^{j\Theta}\mathrm{e}^{j2\pi\Delta fTk}f[k]b[k-1]$ not only from $r[k-1]$ but from the last $N-1$ observed signal samples $r[k-\nu]$, $1 \le \nu \le N-1$. The coefficients p_ν^a, $1 \le \nu \le N-1$, of the proposed linear estimator are obtained by minimizing the mean–squared error (MSE) between $\mathrm{e}^{j\Theta}\mathrm{e}^{j2\pi\Delta fTk}f[k]b[k-1]$ and the estimated reference symbol

$$r_e[k-1] = \sum_{\nu=1}^{N-1} p_\nu^a r[k-\nu]. \tag{5}$$

Since $n[\cdot]$ is an uncorrelated Gaussian random process, this estimator is equivalent to a linear estimator for $(\mathrm{e}^{j\Theta}\mathrm{e}^{j2\pi\Delta fTk}f[k]b[k] + n[k])/a[k] = r[k]/a[k]$. The MSE variance σ^2_{MSE} for estimation of $r[k]/a[k]$ is given by

$$\begin{aligned}
\sigma^2_{\mathrm{MSE}} &= \mathcal{E}\left\{\left|\frac{r[k]}{a[k]} - r_e[k-1]\right|^2\right\} = \mathcal{E}\left\{\left|r[k] - a[k]r_e[k-1]\right|^2\right\} \\
&= \mathcal{E}\Bigg\{\Bigg|\mathrm{e}^{j\Theta}\mathrm{e}^{j2\pi\Delta fTk}f[k]b[k] + n[k] \\
&\qquad -a[k]\sum_{\nu=1}^{N-1} p_\nu^a\left(\mathrm{e}^{j\Theta}\mathrm{e}^{j2\pi\Delta fT(k-\nu)}f[k-\nu]b[k-\nu] + n[k-\nu]\right)\Bigg|^2\Bigg\}.
\end{aligned} \tag{6}$$

Using the substitution

$$p_\nu^a = \frac{b[k-1]}{b[k-\nu]}p_\nu = \prod_{\mu=1}^{\nu-1} a[k-\mu]p_\nu, \qquad 1 \le \nu \le N-1, \tag{7}$$

Eq. (6) may be simplified to

$$\sigma^2_{\mathrm{MSE}} = \mathcal{E}\left\{\left|c[k] - \sum_{\nu=1}^{N-1} p_\nu c[k-\nu]\right|^2\right\}, \tag{8}$$

with the definitions

$$c[k] \triangleq \mathrm{e}^{j2\pi\Delta fTk}f[k] + n'[k], \tag{9}$$

$$n'[k] \triangleq \frac{\mathrm{e}^{-j\Theta}n[k]}{b[k]}. \tag{10}$$

Note, that $n'[\cdot]$ is uncorrelated white Gaussian noise with variance σ_n^2 like $n[\cdot]$. From Eq. (8) it can be observed that p_ν, $1 \le \nu \le N-1$, can be interpreted as coefficients of a linear

predictor of order $N-1$ for process $c[\cdot]$. Hence, p_ν, $1 \leq \nu \leq N-1$, can be determined from the *Yule–Walker equations* [12]

$$\boldsymbol{\Phi}_c \boldsymbol{p}^* = \boldsymbol{\varphi}_c, \tag{11}$$

with the $(N-1) \times (N-1)$ autocorrelation matrix (ACM) [12] $\boldsymbol{\Phi}_c$ of $c[\cdot]$ and

$$\boldsymbol{\varphi}_c \triangleq [\varphi_c[-1]\ \varphi_c[-2]\ \ldots\ \varphi_c[-N+1]]^T, \tag{12}$$

$$\varphi_c[\lambda] \triangleq \mathcal{E}\{c[k+\lambda]c^*[k]\} = \mathrm{e}^{j2\pi\Delta f T\lambda}\varphi_f[\lambda] + \sigma_n^2\delta[\lambda], \tag{13}$$

$$\boldsymbol{p} \triangleq [p_1\ p_2\ \ldots\ p_{N-1}]^T \tag{14}$$

($(\cdot)^*$ and $[\cdot]^T$ denote complex conjugation and transposition, respectively). Here, $\varphi_f[\cdot]$ is the autocorrelation function (ACF) of the fading process $f[\cdot]$ and $\delta[\cdot]$ denotes the unit pulse sequence, i.e., $\delta[0] = 1$, $\delta[\lambda] = 0$, $\lambda \neq 0$.

So far, we have assumed that $a[k-\nu]$, $1 \leq \nu \leq N-2$, is perfectly known (cf. Eq. (7)). Since this is not the case in a realizable receiver, we replace the unknown transmitted symbols $a[k-\nu]$, $1 \leq \nu \leq N-2$, by the decision–feedback symbols $\hat{a}[k-\nu]$, $1 \leq \nu \leq N-2$, in Eq. (7). The resulting reference symbol is denoted by $\hat{r}_e[k-1]$ and given by

$$\hat{r}_e[k-1] = \sum_{\nu=1}^{N-1} p_\nu \prod_{\mu=1}^{\nu-1} \hat{a}[k-\mu] r[k-\nu]. \tag{15}$$

The decision variable for prediction–based DF–DD is

$$g[k] = r[k]\hat{r}_e^*[k-1]. \tag{16}$$

The complex plane is divided into M sectors corresponding to the M possible values of $a[k]$ and $\hat{a}[k]$ is uniquely determined by the sector into which the complex number $g[k]$ falls. Fig. 2 shows the structure of the prediction–based DF–DD receiver. The described decision rule is equivalent to

$$\hat{a}[k] = \underset{\tilde{a}[k]}{\operatorname{argmax}} \left\{\mathrm{Re}\left\{\tilde{a}[k] r^*[k] \hat{r}_e[k-1]\right\}\right\}. \tag{17}$$

If the channel statistics are known a priori, Eq. (11) may be used for calculation of the predictor coefficients. In mobile communications, where the channel may be nonstationary, the adaptive algorithm described in Section 5 can be applied.

It should be mentioned that for $N = 2$, prediction–based DF–DD is identical with conventional DD as long as the predictor coefficient p_1 is a real number. In the following, we assume that the fading spectrum of the scattered component $f_s[\cdot]$ is symmetric and thus, p_1 is real if $\Delta f = 0$ and $f_D = 0$.

4 Performance Analysis for Genie–Aided Prediction–Based DF–DD and $M = 4$

In this section, the performance of genie–aided prediction–based DF–DD is analyzed. We concentrate on QDPSK ($M = 4$), since in this case a closed–form result for the BER can be

obtained. Moreover, QDPSK is used in many existing systems, like the United States Digital Cellular (IS–54, IS–136) and the Japanese Digital Cellular (PDC) [1]. The result of the BER analysis is used to investigate the limiting performance of prediction–based DF–DD.

4.1 BER Analysis

We restrict ourselves to genie–aided DF–DD since it is difficult to take into account the effect of error propagation. In general, for DF–DD schemes error propagation increases BER by approximately a factor of two (cf. [15, 16, 30]). However, this holds only if the errors are not burst–like. For error bursts this factor is smaller than two as will be shown by simulations in Section 7. For derivation of BER, $\phi = \pi/4$ (i.e., $\pi/4$–QDPSK) is assumed[2]. Our analysis follows essentially [15], however, here a Ricean fading channel with frequency offset is considered instead of an AWGN channel.

Since the BER does not depend on the transmitted symbol, without loss of generality the transmission of $a[k] = \mathrm{e}^{j\pi/4} \in \mathcal{A}_{\pi/4}$ is assumed. In this case, no symbol error occurs if the decision variable $g[k]$ (cf. Eq. (16)) falls into the 1st quadrant. The probability that $g[k]$ is located in the 2nd, 3rd, and 4th quadrant is denoted by P_2, P_3, and P_4, respectively. For Gray mapping only one bit error occurs if $g[k]$ falls into the 2nd or 4th quadrant and, in addition, $P_2 = P_4$ because of symmetry. If $g[k]$ falls into the 3rd quadrant, two bit errors occur. Thus, the BER is given by

$$P_b = \frac{1}{2}(P_2 + P_4) + P_3 = P_2 + P_3. \tag{18}$$

This corresponds to

$$P_b = \Pr\{\mathrm{Re}\{g[k]\} < 0\} = \Pr\{d[k] < 0\}, \tag{19}$$

where

$$d[k] \stackrel{\triangle}{=} g[k] + g^*[k] \stackrel{\triangle}{=} Cx[k]y^*[k] + C^*x^*[k]y[k] \tag{20}$$

is a special case of the quadratic form of Eq. (4B.1) of [29]. A closed–form result for P_b is given in the Appendix. As can be observed from Eqs. (52)–(64), P_b depends on the magnitude of the mean f_m of the fading process, the Doppler shift f_D, the first N samples of the ACF $\varphi_f[\cdot]$, the frequency offset Δf, and the channel noise variance σ_n^2. From this we can conclude that all fading spectra $S_f(\mathrm{e}^{j2\pi fT})$ [3] whose corresponding ACFs coincide in the first N samples cause the same BER. For $N = 2$, Eq. (52) gives the BER for conventional DD if $\Delta f = 0$ and $f_D = 0$.

To become more specific, we assume $\Delta f = 0$ and use Jakes fading model [31], i.e., $\varphi_f[\lambda]$ is given by

$$\varphi_f[\lambda] = \mathrm{e}^{j2\pi f_D T\lambda}|f_m|^2 + \sigma_s^2 \mathrm{J}_0(2\pi B_f T\lambda), \qquad \forall\lambda, \tag{21}$$

[2]Note, that the BERs for $\phi = \pi/4$ and $\phi = 0$ are identical since amplifier nonlinearities are not taken into account.

[3]$S_f(\mathrm{e}^{j2\pi fT})$ is the Fourier transform of the ACF $\varphi_f[\cdot]$.

where $\mathrm{J}_0(\cdot)$ denotes the zeroth order Bessel function of the first kind. Fig. 3 shows BER vs. $10\log_{10}(E_b/N_0)$ ($E_b = E_S/2$) for coherent QPSK with perfect channel state information (denoted by CD) and QDPSK with genie–aided prediction–based DF–DD for a Ricean fading channel with $f_D = 0$ and $10\log_{10}(K) = 10$ dB. Coherent QPSK is a lower bound for genie–aided prediction–based DF–DD as will be shown in the next section. For the DF–DD scheme, p_ν, $1 \le \nu \le N-1$, was calculated from Eq. (11). As N increases performance improves and BER is slightly lower for $B_fT = 0.0075$ than for $B_fT = 0.03$. However, in both cases, even for $N = 100$ BER of genie–aided DF–DD is higher than that of coherent QPSK.

In Fig. 4, $B_fT = 0.05$ and $10\log_{10}(K) = 0$ dB are assumed. Here, $f_DT = 0$ and $f_DT = 0.04$ are considered. Although the Ricean process with $f_DT = 0.04$ causes a higher BER for $N = 2$, for $N = 100$ the same performance as for $f_DT = 0$ is obtained[4]. The reason for this behaviour is explained in Section 4.2. Moreover, for $10\log_{10}(E_b/N_0) < 60$ dB no error floor can be observed for $N = 5$ and $N = 100$. Under which conditions it is, in principle, possible to avoid the error floor and in which cases the lower bound given by coherent QPSK can be approached is also discussed in Section 4.2.

4.2 Limiting Performance for $N \to \infty$

Using similar relations as in Section V of [24], P_b for $N \to \infty$ (cf. Eq. (52)) may be expressed as

$$P_b = \mathrm{Q_M}(a_1,\, b_1) - \frac{1}{2}\frac{s+1}{s}\mathrm{I}_0(a_1b_1)\exp\left(-\frac{a_1^2+b_1^2}{2}\right), \tag{22}$$

with

$$a_1 = \frac{1}{2}\left(1-\frac{1}{s}\right)\sqrt{\frac{|f_m|^2}{\sigma^2_{e,\min}}(s^2-1)}, \tag{23}$$

$$b_1 = \frac{1}{2}\left(1+\frac{1}{s}\right)\sqrt{\frac{|f_m|^2}{\sigma^2_{e,\min}}(s^2-1)}, \tag{24}$$

$$s = \sqrt{\frac{\sigma_s^2+\sigma_n^2+\sigma^2_{e,\min}}{\sigma_s^2+\sigma_n^2-\sigma^2_{e,\min}}}, \tag{25}$$

$$\sigma^2_{e,\min} = \exp\left(T\int\limits_{-1/(2T)}^{1/(2T)}\log\left(S_c(\mathrm{e}^{j2\pi fT})\right)\,\mathrm{d}f\right), \tag{26}$$

$$S_c(\mathrm{e}^{j2\pi fT}) = S_f(\mathrm{e}^{j2\pi fT}) + \sigma_n^2. \tag{27}$$

Here, $S_c(\mathrm{e}^{j2\pi fT})$ is the power spectral density of process $c[\cdot]$ and $\sigma^2_{e,\min}$ [32, 33] is the minimum error variance for prediction of $c[\cdot]$.

[4]Note, that although no genie is needed for $N = 2$, DF–DD is not identical with conventional DD for $f_DT = 0.04$.

Interestingly, in the limit $N \to \infty$, P_b depends only on the noise variance σ_n^2, the magnitude of the mean f_m of the fading process and the fading spectrum $S_f(\mathrm{e}^{j2\pi fT})$ (cf. Eqs. (22)–(27)). Since the Dirac delta impulse at Doppler shift $f = f_D$ corresponding to the direct component of the Ricean process does not contribute to the integral in Eq. (26) (cf. [24], Appendix A), the Doppler shift f_D has no influence on P_b for $N \to \infty$. This explains why for $N = 100$ the Ricean processes with $f_D T = 0.0$ and $f_D T = 0.04$ considered in Fig. 4 yield (approximately) the same performance. Since frequency offset (i.e., shift of fading spectrum) has also no influence on P_b for $N \to \infty$ (our simulations in Section 7 show that this is also true for finite N), without loss of generality, we assume $\Delta f = 0$ and $f_D = 0$ in the following.

Now, three special cases of Eqs. (22) and (26) are discussed.

A. $\sigma_n^2 \to 0$
From $\sigma_n^2 \to 0$, $S_c(\mathrm{e}^{j2\pi fT}) \to S_f(\mathrm{e}^{j2\pi fT})$ (cf. Eq. (27)) follows, i.e., $\sigma_{e,\min}^2$ depends exclusively on the fading spectrum. If $S_f(\mathrm{e}^{j2\pi fT})$ is bandlimited, $\sigma_{e,\min}^2$ is equal to zero [32, 33]. From Eq. (25) $s = 1$ is obtained for $\sigma_{e,\min}^2 = 0$ and Eqs. (23) and (24) yield $a_1 = 0$ and $b_1 = 0$, respectively. Hence, Eq. (22) gives $P_b = 0$, i.e., there is no irreducible error floor for bandlimited Ricean fading processes if prediction–based DF–DD is used. On the other hand, for realizable fading processes which cannot be bandlimited, $\sigma_{e,\min}^2 > 0$ and thus, $s > 1$ is obtained. This means $P_b > 0$ even in the absence of noise and there is always an irreducible error floor.

B. $B_f T \to 0$ (Slow Ricean Fading)
Here, the fading spectrum is given by

$$S_f(\mathrm{e}^{j2\pi fT}) = |f_m|^2 \delta_0(f) + \sigma_s^2 \delta_0(f), \qquad |f| \le 1/(2T), \tag{28}$$

where $\delta_0(\cdot)$ is the Dirac delta function [34]. Using similar methods as proposed in [24], it can be shown that in this special case

$$\mu_{xy} = \mu_{yy} = q_f^2 - |f_m|^2 = \sigma_s^2 \tag{29}$$

and

$$\bar{y} = f_m \tag{30}$$

follow from Eqs. (61), (62) and Eq. (64), respectively.

On the other hand, for coherent QPSK with perfect channel state information the decision variable $g_{CD}[k]$ is given by [29]

$$g_{CD}[k] = r[k]f^*[k] = b[k](f[k] + n[k]b^*[k]) \cdot f^*[k]. \tag{31}$$

Without loss of generality, $b[k] = \mathrm{e}^{j\pi/4}$ is assumed. If we define $C_{CD} \triangleq b[k] = \mathrm{e}^{j\pi/4} = C$, $x_{CD}[k] \triangleq f[k] + n[k]b^*[k]$ and $y_{CD}[k] \triangleq f[k]$, the first and second order moments of $x_{CD}[k]$ and $y_{CD}[k]$ are calculated to $\bar{x}_{CD} = \bar{y}_{CD} = f_m$, $\mu_{x_{CD}x_{CD}} = \sigma_s^2 + \sigma_n^2$, $\mu_{x_{CD}y_{CD}} =$

$\mu_{y_{CD}y_{CD}} = \sigma_s^2$, i.e., they are identical with those of genie–aided prediction–based DF–DD for $B_fT \to 0$ and $N \to \infty$ (cf. Eqs. (29), (30), (60), and (63)).

This means, for slow Ricean fading, genie–aided prediction–based DF–DD of QDPSK yields the same BER as coherent QPSK with perfect channel state information. Note, that for the AWGN channel this property has already been proved in [15] (for the AWGN channel prediction–based DF–DD and MSD–based DF–DD are identical (cf. Section 6)). It should be mentioned that realizable prediction–based DF–DD (i.e., without genie) is lower–bounded by optimum coherent Q**D**PSK [35], of course. The BER of coherent QDPSK is approximately by a factor of two higher than that of coherent QPSK [29]. On the other hand, the difference between genie–aided DF–DD and realizable DF–DD is also approximately a factor of two at high E_b/N_0 ratios. From this it can be concluded that realizable prediction–based DF–DD for QDPSK can approach its lower bound for $B_fT \to 0$.

C. Jakes Model

If Jakes fading model is assumed, $S_f(\mathrm{e}^{j2\pi fT})$ is given by

$$S_f(\mathrm{e}^{j2\pi fT}) = \begin{cases} |f_m|^2\delta_0(f) + \frac{\sigma_s^2}{\pi T\sqrt{B_f^2 - f^2}} & |f| < B_f \\ 0 & B_f \le |f| \le \frac{1}{2T} \end{cases}. \tag{32}$$

Using the methods proposed in Appendix A and B of [24], it can be shown that in this case $\sigma_{e,\min}^2$ may be expressed by

$$\sigma_{e,\min}^2 = \sigma_n^2 \left(\frac{\mathrm{e}\sigma_s^2}{2\pi B_f T \sigma_n^2} \right)^{2B_fT} \exp(2TC_0), \tag{33}$$

where C_0 is defined as

$$C_0 \stackrel{\triangle}{=} B_f \int_0^{\pi/2} \log(1 + q\sin\varphi)\sin\varphi \,\mathrm{d}\varphi, \tag{34}$$

with $q \stackrel{\triangle}{=} \frac{\pi B_f T \sigma_n^2}{\sigma_s^2}$. For small fading bandwidths B_f and high σ_s^2/σ_n^2 ratios, $q \ll 1$ and thus, $C_0 \approx 0$ holds. However, in general, C_0 has to be evaluated numerically.

Fig. 5 shows BER calculated from Eqs. (22)–(25) and (33) vs. $10\log_{10}(E_b/N_0)$ for QDPSK with genie–aided prediction–based DF–DD ($N \to \infty$). Note, that the BERs for $B_fT = 0$ are identical with those of coherent QPSK as has been shown in Case B. For $B_fT > 0$, however, genie–aided DF–DD performs worse than coherent QPSK even for $N \to \infty$. Thus, it has to be expected, that realizable DF–DD also cannot approach its lower bound (coherent Q**D**PSK) for $B_fT > 0$. Performance bounds for realizable DF–DD at high E_b/N_0 ratios can be obtained by multiplying the bounds of genie–aided DF–DD by a factor of two. Note, that in Fig. 5, no error floor can be observed since the assumed fading spectrum is bandlimited (cf. Case A). For finite K, BER increases with increasing normalized fading bandwidth B_fT.

5 Adaptation of the Predictor Coefficients Using the RLS Algorithm

In this section, the adaptation algorithm for the predictor coefficients is described. As has been already mentioned, p_ν, $1 \leq \nu \leq N-1$, could be calculated from Eq. (11), of course. However, for this it would be necessary to know the first N samples of the ACF $\varphi_c[\cdot]$. Thus, $\varphi_c[\lambda]$, $0 \leq \lambda \leq N-1$, would have to be estimated a priori. Since the channel statistics in mobile communications are typically nonstationary, this estimation would have to be repeated regularly.

In order to avoid the above mentioned problems, we propose the employment of an adaptation algorithm which computes the predictor coefficients recursively. Since the eigenvalue spread of the autocorrelation matrix $\boldsymbol{\Phi}_c$ can become quite large, the convergence speed of the simple least–mean–square (LMS) algorithm is quite slow [12] and the RLS algorithm is preferable.

The RLS algorithm does not minimize the MSE variance σ^2_{MSE} (cf. Eq. (6)) but the related cost function [12]

$$J[k] = \sum_{\mu=1}^{k} w^{k-\mu} \left| r[\mu] - \sum_{\nu=1}^{N-1} p_\nu[k] \hat{r}[\mu,\, \nu] \right|^2, \tag{35}$$

where $p_\nu[k]$, $1 \leq \nu \leq N-1$, are the predictor coefficients at time k and $\hat{r}[k,\, \nu]$ is defined as

$$\hat{r}[k,\, \nu] \stackrel{\triangle}{=} r[k-\nu] \prod_{j=0}^{\nu-1} \hat{a}[k-j], \qquad 1 \leq \nu \leq N-1. \tag{36}$$

w, $0 < w \leq 1$, is the forgetting factor of the RLS algorithm. For a forgetting factor $w = 1$ and $k \to \infty$, $p_\nu[k] \to p_\nu$[5], $1 \leq \nu \leq N-1$, is obtained if all random processes involved are ergodic. For the problem at hand, the RLS algorithm consists of the following equations [12]

$$\boldsymbol{k}_k = \frac{\boldsymbol{P}_{k-1} \hat{\boldsymbol{r}}_k}{w + \hat{\boldsymbol{r}}_k^H \boldsymbol{P}_{k-1} \hat{\boldsymbol{r}}_k}, \tag{37}$$

$$\xi[k] = r[k] - \boldsymbol{p}_{k-1}^T \hat{\boldsymbol{r}}_k, \tag{38}$$

$$\boldsymbol{P}_k = w^{-1} \boldsymbol{P}_{k-1} - w^{-1} \boldsymbol{k}_k \hat{\boldsymbol{r}}_k^H \boldsymbol{P}_{k-1}, \tag{39}$$

$$\boldsymbol{p}_k = \boldsymbol{p}_{k-1} + \boldsymbol{k}_k^* \xi[k], \tag{40}$$

where $[\cdot]^H$ denotes Hermitian transposition and the definitions

$$\hat{\boldsymbol{r}}_k \stackrel{\triangle}{=} [\hat{r}[k,\, 1]\ \hat{r}[k,\, 2]\ \ldots\ \hat{r}[k,\, N-1]]^T, \tag{41}$$

$$\boldsymbol{p}_k \stackrel{\triangle}{=} [p_1[k]\ p_2[k]\ \ldots\ p_{N-1}[k]]^T, \tag{42}$$

are used. For initialization of the algorithm we propose

$$\boldsymbol{P}_0 = \delta^{-1} \boldsymbol{I}, \tag{43}$$

$$\boldsymbol{p}_0 = [1\, 0\, 0\, \ldots\, 0]^T. \tag{44}$$

[5] p_ν, $1 \leq \nu \leq N-1$, are the coefficients calculated from Eq. (11)

Here, δ is a small positive constant [12]. The initialization of $\boldsymbol{p}_0$ according to Eq. (44) has the effect that the DF–DD scheme starts as a conventional differential detector. Thus, the adaptive prediction–based DF–DD algorithm can start blind without training sequence and without prior knowledge of the channel statistics. As k increases, $\boldsymbol{p}_k$ is adapted to the channel statistics. In a stationary environment $w = 1.0$ is optimum [12]. In mobile communications, however, the channel statistics are typically nonstationary and therefore, $w < 1.0$ may be preferred, since this improves the tracking capability of the RLS algorithm [36].

Figs. 6a) and 7a) show $p_\nu[k]$, $1 \leq \nu \leq N-1$, vs. k for an AWGN channel ($10\log_{10}(E_b/N_0) = 8$ dB) and a Rayleigh fading channel ($10\log_{10}(E_b/N_0) = 20$ dB, $B_fT = 0.03$, Jakes model), respectively. The corresponding learning curves can be seen in Figs. 6b) and 7b), where the mean–squared error $J'[k] = \mathcal{E}\{|\xi[k]|^2\}$ was approximated by averaging over 1000 adaptation processes. In both cases, w was chosen to 0.99. In Figs. 6a) and 7a), the lines denoted by 'optimum' correspond to the solution of Eq. (11). Obviously, in all cases the adapted coefficients converge to the optimum setting. Figs. 6b) and 7b) show that $J'[k]$ approaches the minimum MSE variance σ^2_{MMSE} (σ^2_{MMSE} is obtained form σ^2_{MSE} by inserting the predictor coefficients calculated from Eq. (11) in Eq. (6)) as k increases. It should be mentioned that the DF–DD receiver outperforms conventional DD before steady state is reached. Assuming $N = 2$ (i.e., conventional DD), for the AWGN channel and the Rayleigh fading channel σ^2_{MMSE} is equal to 0.15 and 0.026, respectively. For all predictor orders considered in Figs. 6b) and 7b) $J'[k]$, $k > 10$, is smaller than σ^2_{MMSE} for $N = 2$. Since the mean–squared error directly influences the BER, it can be expected that the proposed adaptive DF–DD receiver outperforms conventional DD for $k > 10$.

6 Comparison with MSD–Based DF–DD

In this section, we investigate the relation between the proposed prediction–based DF–DD scheme and MSD–based DF–DD. Since MSD–based DF–DD is derived directly from maximum–likelihood sequence detection, it may be considered as the optimum DF–DD scheme and will be used as a benchmark for prediction–based DF–DD in our simulations (cf. Section 7).

In the following, we compare MSD–based DF–DD and prediction–based DF–DD for three special cases. For simplicity, we assume $\Delta f = 0$ and $f_D = 0$ in this section.

A. $K = 0$ (Rayleigh Fading Channel)
MSD–based DF–DD for the Rayleigh fading channel was considered in [24]. It can be shown that the decision rule obtained in [24] is equivalent to Eq. (17). Hence, prediction–based DF–DD and MSD–based DF–DD are identical for Rayleigh fading.

B. $K \to \infty$ (AWGN Channel)
For the AWGN channel, $\varphi_c[0] = q_f^2 + \sigma_n^2$ and $\varphi_c[\lambda] = q_f^2$, $\lambda \neq 0$, results. Thus, for $\sigma_n^2 > 0$ from Eq. (11), $p_\nu = p > 0$, $1 \leq \nu \leq N-1$, is obtained. This is used in Eq. (15)

and then, Eq. (15) is inserted into Eq. (17). If it is taken into account that the positive multiplicative constant p does not influence the decision,

$$\hat{a}[k] = \underset{\tilde{a}[k]}{\operatorname{argmax}} \left\{ \operatorname{Re} \left\{ \tilde{a}[k] r^*[k] \sum_{\nu=1}^{N-1} r[k-\nu] \prod_{j=1}^{\nu-1} \hat{a}[k-j] \right\} \right\} \tag{45}$$

results as decision rule for prediction–based DF–DD. This decision rule was derived in a different way by Leib et al. [13] and Edbauer [15]. Adachi et al. showed in [16], that it also may be obtained from MSD. We conclude, that prediction–based and MSD–based DF–DD are identical for the AWGN channel.

C. $0 < K < \infty$

To the best of the authors knowledge, the following MSD–based DF–DD decision rule for $0 < K < \infty$ has not been published in the open technical literature, yet. It can be derived from the MSD scheme proposed for Ricean fading in [5]. Using a similar derivation as in [16] and [24] (i.e., introducing decision–feedback symbols in the MSD metric according to Eq. (5) of [5]), the following MSD–based DF–DD rule is obtained[6]:

$$\begin{aligned} \hat{a}[k] = \underset{\tilde{a}[k]}{\operatorname{argmax}} \Bigg\{ & 2\operatorname{Re} \left\{ \tilde{a}[k] r^*[k] \sum_{\mu=1}^{N-1} t_{0\mu} r[k-\mu] \prod_{j=1}^{\mu-1} \hat{a}[k-j] \right\} \\ & + \ln \left(\mathrm{I}_0 \left(2|f_m| \left| \tilde{a}[k] r^*[k] \sum_{\mu=0}^{N-1} t_{0\mu} \prod_{j=1}^{N-2} \hat{a}[k-j] \right. \right. \right. \\ & \left. \left. \left. + \sum_{\nu=1}^{N-1} r^*[k-\nu] \sum_{\mu=0}^{N-1} t_{\nu\mu} \prod_{j=\nu}^{N-2} \hat{a}[k-j] \right| \right) \right) \Bigg\}, \end{aligned} \tag{46}$$

where $\tilde{a}[k] \in \mathcal{A}_\phi$ is a trial symbol. Here, $t_{\nu\mu}$, $0 \le \nu, \mu \le N-1$, are the elements of matrix

$$\boldsymbol{T} = -\boldsymbol{M}_c^{-1}, \tag{47}$$

where $\boldsymbol{M}_c$ denotes the $N \times N$ covariance matrix of process $c[\cdot]$ (cf. Eq. (9), $\Delta f = 0$, $f_D T = 0$).

Further analysis of Eq. (46) shows that prediction–based DF–DD and MSD–based DF–DD are not only equivalent for $K = 0$ and $K \to \infty$, but also for the special cases $B_f T = 0$ (slow Ricean fading) and $N = 2$. For $N = 2$, both schemes are also equivalent to conventional DD ($\Delta f = 0$, $f_D = 0$ is assumed).

On the other hand, a comparison of Eq. (17) and Eq. (46) indicates that in general the decision rules for prediction–based DF–DD and MSD–based DF–DD are different. However, our simulations in the next section show that for $\Delta f = 0$ the performance of both schemes is always almost identical.

[6]Note, that Eq. (46) is also valid for the special cases $K = 0$ and $K \to \infty$, of course.

7 Simulation Results

In the following, prediction–based DF–DD will be referred to as *adaptive* prediction–based DF–DD or simply adaptive DF–DD if the predictor coefficients are calculated using the adaptive algorithm discussed in Section 5.

For all simulations presented here, continuous transmission, i.e., no burst structure, is assumed. Furthermore, we restrict ourselves to the case $f_D = 0$. First the influence of the forgetting factor w of the RLS algorithm on the BER of adaptive prediction–based DF–DD is examined.

Influence of w on BER of Adaptive Prediction–Based DF–DD

Fig. 8 shows BER vs. w for the AWGN channel. For $0.99 \leq w \leq 1$, BER is almost independent of w for $10\log_{10}(E_b/N_0) = 5$ dB and $10\log_{10}(E_b/N_0) = 8$ dB. For $w < 0.99$ the RLS algorithm becomes instable in some cases and BER increases. It should be mentioned, that this instability typically arises after transmission of several tens of thousands of symbols and can be avoided by using a more stable implementation of the RLS algorithm (e.g. QR–RLS algorithm [12]). In mobile communication applications, typically bursts of several hundreds of symbols are transmitted. Thus, the stability problem is avoided if the RLS algorithm is reinitialized after reception of some bursts. Note, that the RLS algorithm is more stable at the higher E_b/N_0 ratio and for larger N.

In Fig. 9, it can be observed, that for Rayleigh fading ($B_fT = 0.03$) the stability of the RLS algorithm is not a major problem. For $10\log_{10}(E_b/N_0) = 30$ dB and $10\log_{10}(E_b/N_0) = 60$ dB the algorithm is stable for $0.95 \leq w \leq 1$ and BER is almost constant.

AWGN Channel

Fig. 10 shows BER vs. $10\log_{10}(E_b/N_0)$ for QDPSK transmitted over the AWGN channel and detected with adaptive (prediction–based) DF–DD ($w = 1$). For comparison, the BER curves for QDPSK with conventional DD, adaptive genie–aided (prediction–based) DF–DD ($w = 1$), MSD–based DF–DD (equivalent to prediction–based DF–DD with fixed predictor coefficients according to Eq. (11)), and CD are also included in Fig. 10. The loss of the adaptive scheme when compared to the MSD–based DF–DD is very small (less than 0.1 dB). At BER $= 10^{-3}$ the gain of adaptive DF–DD over conventional DD is 0.7 dB and 1.2 dB for $N = 3$ and $N = 5$, respectively. As N increases, CD is approached since it is a lower bound for realizable noncoherent detectors. For the calculated BER (denoted by 'Theory') Eq. (11) was used for determination of the predictor coefficients, of course. The simulated and calculated BERs of genie–aided DF–DD are in good agreement and it can be observed, that error propagation increases BER by approximately a factor of two. For conventional DD (equivalent to MSD–based and prediction–based DF–DD with $N = 2$) no feedback is required and hence, there is no error propagation.

Frequency Offset Δf

Since practical receivers often have to cope with frequency offset, we consider an AWGN channel with $\Delta f > 0$. Fig. 11 shows BER vs. normalized frequency offset $\Delta f T$ for adaptive (prediction–based) DF–DD ($w = 1$) and MSD–based DF–DD optimized for AWGN. For adaptive (prediction–based) DF–DD BER is hardly influenced by a normalized frequency offset of up to 0.04 (corresponding to $\Delta f = 971$ Hz for IS–136), whereas BER increases for MSD–based DF–DD with increasing frequency offset, especially for large N. It can be shown analytically that the predictor coefficients perfectly account for the phase rotation introduced by frequency offset and hence, prediction–based DF–DD is insensitive to frequency offset. Note, that MSD–based DF–DD with $N = 2$ and optimized for AWGN is identical with conventional DD. However, here adaptive DF–DD with $N = 2$ is *not* identical with conventional DD since for $\Delta f \neq 0$ the predictor coefficient p_1 is not a real number. Fig. 11 clearly illustrates the robustness of adaptive DF–DD against frequency offset.

Rayleigh Fading

Fig. 12 shows BER for conventional DD, adaptive (prediction–based) DF–DD ($w = 0.99$), genie–aided adaptive (prediction–based) DF–DD ($w = 0.99$), MSD–based DF–DD, and CD with perfect channel state information vs. $10 \log_{10}(E_b/N_0)$ for Rayleigh fading with $B_f T = 0.03$. Note, that for Rayleigh fading MSD–based DF–DD and prediction–based DF–DD with fixed predictor coefficients according to Eq. (11) are identical. The increase of BER due to coefficient adaptation is negligible. It can be observed, that the error floor of conventional DD can be reduced considerably by DF–DD. For $N = 3$ and $N = 5$, it is, respectively, 34 and 80 times lower than for conventional DD. At low E_b/N_0 ratios, conventional DD performs almost as good as CD and thus, the gain of DF–DD is small. Note also, that a certain loss of prediction–based DF–DD when compared with CD is unavoidable because here, the normalized fading bandwidth $B_f T$ is larger than zero (cf. Section 4.2, Case B and C). Moreover, at low E_b/N_0 ratios burst errors dominate and hence, error propagation increases BER by a factor smaller than two. On the other hand, at high E_b/N_0 ratios single error events prevail and BER increases by approximately a factor of two. For determination of the theoretical BER of genie–aided DF–DD, the first N samples of the ACF were estimated from the simulated fading process and the predictor coefficients were calculated from Eq. (11). Although the simulated fading process is based on Jakes fading model [31], it cannot be bandlimited, of course. Thus, an irreducible error floor is unavoidable (cf. Section 4.2).

Ricean Fading

Fig. 13 shows BER vs. $10 \log_{10}(E_b/N_0)$ for QDPSK with conventional DD, adaptive (prediction–based) DF–DD ($w = 0.99$), genie–aided adaptive (prediction–based) DF–DD ($w = 0.99$), MSD–based DF–DD, and CD with perfect channel state information. $10 \log_{10}(K) = 10$ dB, $B_f T = 0.0075$, and $f_D T = 0$ are used for the simulation. It can be seen that adaptive DF–DD performs as good as MSD–based DF–DD. At BER $= 10^{-3}$, the gain of

these techniques over conventional DD is 0.5 dB and 0.75 dB for $N = 3$ and $N = 5$, respectively. At this BER, there is still a gap of 0.85 dB between adaptive DF–DD with $N = 5$ and CD which could be further narrowed by increasing N. The BER analysis presented in Section 4.1 is perfectly confirmed by the simulated BERs for genie–aided adaptive DF–DD and conventional DD. For calculation of the theoretical BER the predictor coefficients were determined in the same way like for Rayleigh fading.

The same detection schemes as in Fig. 13 are applied in Fig. 14, however, now $10 \log_{10}(K) = 3$ dB, $B_f T = 0.03$, and $f_D T = 0$ are valid. Again adaptive (prediction–based) DF–DD performs almost as good as MSD–based DF–DD. At low E_b/N_0 ratios, the gap between conventional DD and CD is small and DF–DD cannot improve conventional DD. Again a certain loss of prediction–based DF–DD when compared with CD is unavoidable because of $B_f T > 0$. On the other hand, the error floor caused by conventional DD is reduced significantly by both DF–DD schemes. Again, there is a good agreement between the simulated BERs for conventional DD and genie–aided adaptive DF–DD and the corresponding theoretical BERs. At high E_b/N_0 ratios, BER of adaptive DF–DD is approximately by a factor of two larger than BER of genie–aided adaptive DF–DD. At low E_b/N_0 ratios, where burst errors dominate, this factor is significantly smaller than two.

8 Conclusions

In this paper, prediction–based DF–DD for MDPSK signals transmitted over Ricean fading channels has been proposed and analyzed. The BER of QDPSK with genie–aided prediction–based DF–DD is calculated and the limiting performance for $N \to \infty$ is investigated. It is shown, that for large values of N the Doppler shift has no influence on the performance of prediction–based DF–DD and that in principle the error floor of conventional DD can be avoided for bandlimited Ricean fading processes. The decision rule for prediction–based DF–DD is also compared with that of MSD–based DF–DD. Simulations indicate that prediction–based DF–DD can improve conventional DD significantly for AWGN, Rayleigh fading, and Ricean fading channels. Furthermore, it has the very important advantage that it does not degrade under frequency offset. The proposed scheme is well suited for a practical implementation since the optimum predictor coefficients can be found efficiently by using an adaptive algorithm. This algorithm can start blind, i.e., no training sequence and no a priori knowledge about the channel statistics are required at the receiver.

Acknowledgement

The authors would like to thank TCMC (Technology Centre for Mobile Communication), Philips Semiconductors, Germany, for supporting this work. Furthermore, they wish to thank Prof. J. B. Huber for helpful discussions and his encouragement throughout the development of this paper.

Appendix

In this appendix, a closed–form result for P_b is calculated. By using Eqs. (4), (9), (10) and (15), Eq. (16) may be rewritten as

$$g[k] \triangleq a[k]c[k]\left(\sum_{\nu=1}^{N-1} p_\nu c[k-\nu]\right)^*, \tag{48}$$

where because of the assumption of genie–aiding, $\hat{a}[k-\nu] = a[k-\nu]$, $1 \leq \nu \leq N-2$, has been used. A comparison of Eqs. (20) and (48) shows that C, $x[k]$, and $y[k]$ can be defined as

$$C \triangleq a[k] = \mathrm{e}^{j\frac{\pi}{4}}, \tag{49}$$

$$x[k] \triangleq c[k], \tag{50}$$

$$y[k] \triangleq \sum_{\nu=1}^{N-1} p_\nu c[k-\nu]. \tag{51}$$

If the BER expression of Eq. (4B.21) of [29] is adapted to the problem at hand,

$$P_b = \mathrm{Q_M}(a_1, b_1) - \frac{v_2}{v_1+v_2}\mathrm{I_0}(a_1 b_1)\exp\left(-\frac{a_1^2+b_1^2}{2}\right) \tag{52}$$

results, where $\mathrm{Q_M}(\cdot,\cdot)$ and $\mathrm{I_0}(\cdot)$ are the Marcum Q–function and the zeroth order modified Bessel function of the first kind, respectively. Here, the following definitions are used:

$$a_1 \triangleq \frac{\sqrt{2v_1^2 v_2(\alpha_1 v_2 - \alpha_2)}}{v_1+v_2}, \tag{53}$$

$$b_1 \triangleq \frac{\sqrt{2v_1 v_2^2(\alpha_1 v_1 + \alpha_2)}}{v_1+v_2}, \tag{54}$$

$$\alpha_1 \triangleq |\bar{x}|^2\mu_{yy} + |\bar{y}|^2\mu_{xx} - \bar{x}^*\bar{y}\,\mu_{xy} - \bar{x}\bar{y}^*\mu_{xy}^*, \tag{55}$$

$$\alpha_2 \triangleq C\bar{x}^*\bar{y} + C^*\bar{x}\bar{y}^*, \tag{56}$$

$$v_1 \triangleq \sqrt{w^2 + \frac{1}{\mu_{xx}\mu_{yy} - |\mu_{xy}|^2}} - w, \tag{57}$$

$$v_2 \triangleq \sqrt{w^2 + \frac{1}{\mu_{xx}\mu_{yy} - |\mu_{xy}|^2}} + w, \tag{58}$$

$$w \triangleq \frac{C\mu_{xy}^* + C^*\mu_{xy}}{2(\mu_{xx}\mu_{yy} - |\mu_{xy}|^2)}, \tag{59}$$

$$\mu_{xx} \triangleq \mathcal{E}\{|x[k]-\bar{x}|^2\} = \varphi_c[0] - |\bar{x}|^2, \tag{60}$$

$$\mu_{xy} \triangleq \mathcal{E}\{(x[k]-\bar{x})(y[k]-\bar{y})^*\} = \sum_{\nu=1}^{N-1} p_\nu^*\varphi_c[\nu] - \bar{x}\bar{y}^*, \tag{61}$$

$$\mu_{yy} \triangleq \mathcal{E}\{|y[k]-\bar{y}|^2\} = \sum_{\nu=1}^{N-1}\sum_{\mu=1}^{N-1} p_\nu p_\mu^*\varphi_c[\mu-\nu] - |\bar{y}|^2, \tag{62}$$

$$\bar{x} \triangleq \mathcal{E}\{x[k]\} = \mathrm{e}^{j2\pi(\Delta f + f_D)Tk} f_m, \tag{63}$$

$$\bar{y} \triangleq \mathcal{E}\{y[k]\} = \mathrm{e}^{j2\pi(\Delta f + f_D)Tk} f_m \sum_{\nu=1}^{N-1} \mathrm{e}^{-j2\pi(\Delta f + f_D)T\nu} p_\nu. \tag{64}$$

References

[1] T.S. Rappaport. *Wireless Communications.* Prentice Hall, Upper Saddle River, NJ, 1996.

[2] D. Makrakis and P. T. Mathiopoulos. Optimal Decoding in Fading Channels: A Combined Envelope, Multiple Differential and Coherent Detection Approach. In *Proceedings of Globecom'89*, pages 1551–1557, 1989.

[3] D. Divsalar and M.K. Simon. Multiple–Symbol Differential Detection of MPSK. *IEEE Trans. on Commun.*, COM-38:300–308, March 1990.

[4] P. Ho and D. Fung. Error Performance of Multiple–Symbol Differential Detection of PSK Signals Transmitted Over Correlated Rayleigh Fading Channels. *IEEE Trans. on Commun.*, COM-40:25–29, October 1992.

[5] D. Divsalar and M.K. Simon. Maximum–Likelihood Differential Detection of Uncoded and Trellis Coded Amplitude Phase Modulation over AWGN and Fading Channels – Metrics and Performance. *IEEE Trans. on Commun.*, COM-94:76–89, January 1994.

[6] M. Visintin. Differential PSK Block Demodulation over a Flat Correlated Rayleigh–Fading Channel. *IEEE Trans. on Commun.*, COM-45:9–11, January 1997.

[7] D. Makrakis, P. T. Mathiopoulos, and D. P. Bouras. Optimal Decoding of Coded PSK and QAM Signals in Correlated Fast Fading Channels and AWGN: A Combined Envelope, Multiple Differential and Coherent Detection Approach. *IEEE Trans. on Commun.*, COM-42:63–74, January 1994.

[8] G. M. Vitetta and D. P. Taylor. Viterbi Decoding of Differentially Encoded PSK Signals Transmitted over Rayleigh Frequency–Flat Fading Channels. *IEEE Trans. on Commun.*, COM-43:1256–1259, Feb-Apr 1995.

[9] G. M. Vitetta and D. P. Taylor. Maximum Likelihood Decoding of Uncoded and Coded PSK Signal Sequences Transmitted Over Rayleigh Flat–Fading Channels. *IEEE Trans. on Commun.*, COM-43:2750–2758, November 1995.

[10] X. Yu and S. Pasupathy. Innovations–Based MLSE for Rayleigh Fading Channels. *IEEE Trans. on Commun.*, COM-43:1534–1544, Feb-Apr 1995.

[11] G. Colavolpe, P. Castoldi, and R. Raheli. Linear predictive receivers for fading channels. *Electronics Letters*, 34:1289–1290, June 1998.

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

[13] 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.

[14] D. Makrakis and K. Feher. Optimal Noncoherent Detection of PSK Signals. *Electr. Letters*, 26:398–400, March 1990.

[15] 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.

[16] F. Adachi and M. Sawahashi. Decision Feedback Multiple–Symbol Differential Detection for M–ary DPSK. *Electronics Letters*, 29(15):1385–1387, 1993.

[17] A. Svensson. Coherent detection based on linear prediction and decision feedback for QDPSK. *Electronics Letters*, 30(20):1642–1643, 1994.

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

[19] F. Adachi. Adaptive differential detection for M-ary DPSK. *Proceedings of the IEE-Commun.*, 143:21–28, February 1996.

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

[21] L. Bin and P. Ho. A New Decision–Feedback DPSK Receiver with Blind Phase Prediction. In *Proceedings of Globecom'97*, pages 337–341, Phoenix, Arizona, USA, November 1997.

[22] L. Bin and P. Ho. Data–Aided Linear Prediction Receiver for Coherent DPSK and CPM Transmitted over Rayleigh Flat–Fading Channels. *to appear in IEEE Transactions on Vehicular Technology*, 1999.

[23] F. Adachi. Adaptive Differential Detection Using Linear Prediction for M-ary DPSK. *IEEE Transactions on Vehicular Technology*, 47:909–918, October 1998.

[24] R. Schober, W. H. Gerstacker, and J. B. Huber. Decision–Feedback Differential Detection of MDPSK for Flat Rayleigh Fading Channels. *IEEE Trans. on Commun.*, COM-47:1025–1035, July 1999.

[25] R. Schober, W. Gerstacker, and J. Huber. Improving Differential Detection of MDPSK by Noise Prediction. In *Proceedings of Globecom'98*, Sydney, November 1998.

[26] R. Schober, W. H. Gerstacker, and J. B. Huber. Improving Differential Detection of MDPSK by Nonlinear Noise Prediction and Sequence Estimation. *IEEE Trans. on Commun.*, COM-47:1161–1172, August 1999.

[27] S. M. Kay. *Modern Spectral Estimation: Theory and Application.* Prentice–Hall, New Jersey, 1988.

[28] B. D. Nelin P. A. Bello. The Influence of Fading Spectrum on the Binary Error Probabilities of Incoherent and Differentially Coherent Matched Filter Receivers. *IRE Transactions on Communication Systems*, CS-10:160–168, June 1962.

[29] J.G. Proakis. *Digital Communications.* McGraw–Hill, New York, Second Edition, 1989.

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

[31] Ed. W. C. Jakes, Jr. *Microwave Mobile Communications.* Prentice–Hall, New Jersey, 1974.

[32] J. Makhoul. Linear Prediction: A Tutorial Review. *Proceedings of the IEEE*, 63:561–580, April 1975.

[33] M. B. Priestley. *Spectral Analysis and Time Series, Volume 2: Multivariate Series, Prediction and Control.* Academic Press, Inc., Orlando, Florida, 1981.

[34] E. J. Borowski and J. M.Borwein. *The Harper Collins Dictionary of Mathematics.* Harper Perennial, New York, 1991.

[35] M. K. Simon and D. Divsalar. On the Optimality of Classical Coherent Receivers of Differential Encoded M–PSK. *IEEE Communication Letters*, 1:67–70, May 1997.

[36] E. Eleftheriou and D. D. Falconer. Tracking Properties and Steady–State Performance of RLS Adaptive Fitler Algorithms. *IEEE Trans. on Acoustics, Speech, and Signal Processing*, 34:1097–1110, October 1986.

Figures

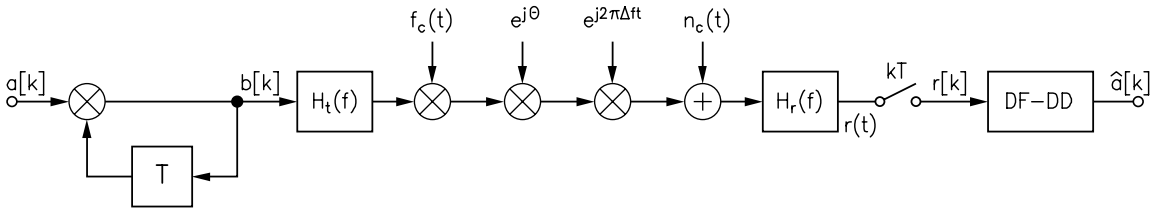

Figure 1: Block diagram of the transmission model.

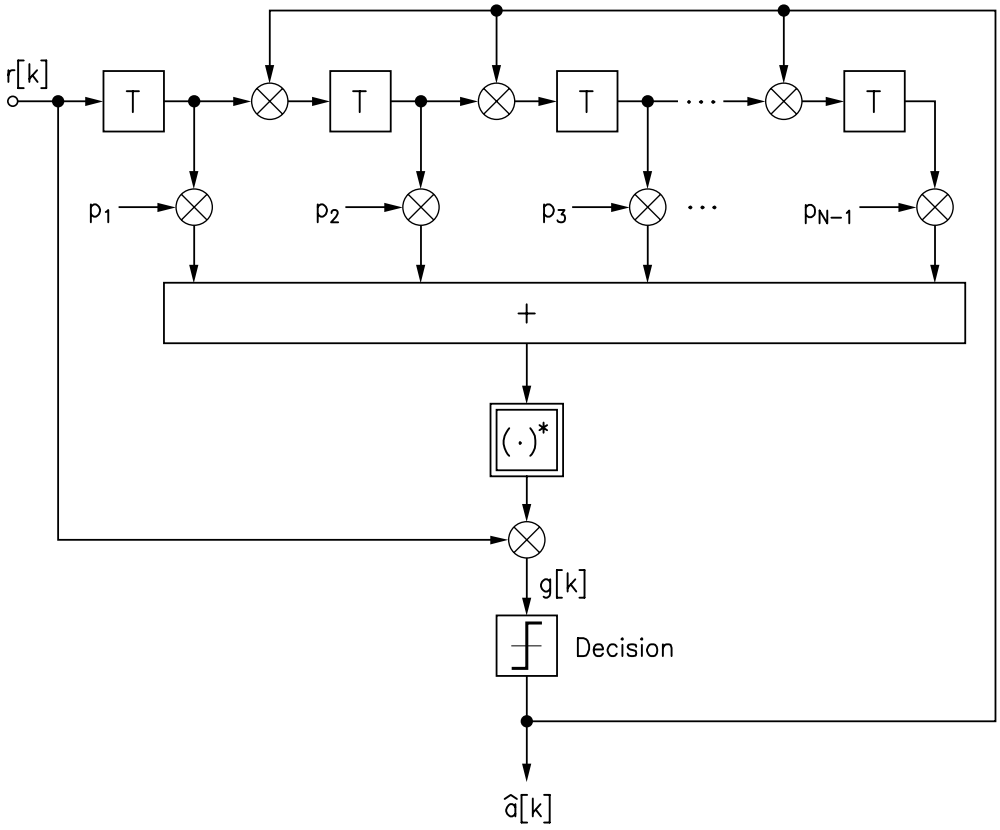

Figure 2: Structure of the prediction–based DF–DD receiver.

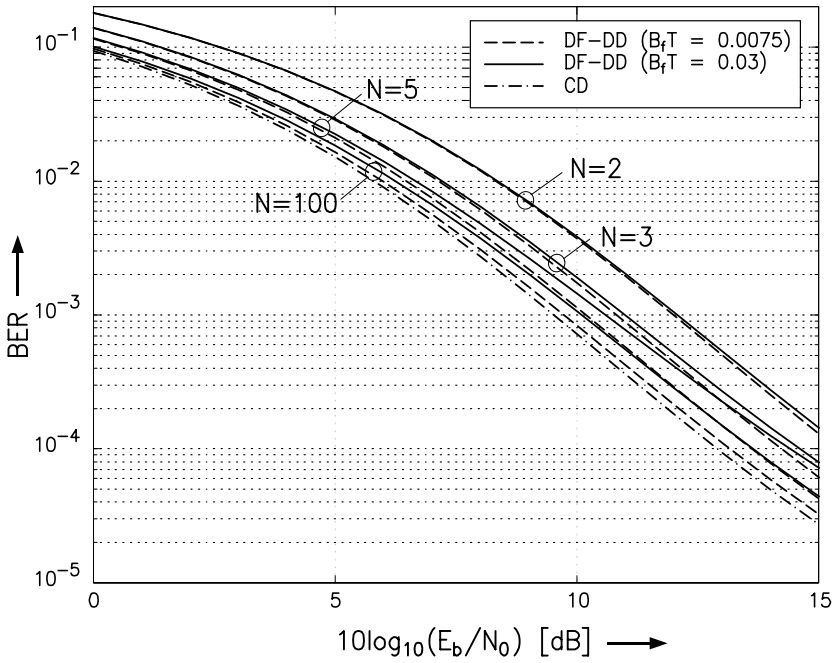

Figure 3: BER of coherent QPSK (denoted by 'CD') and QDPSK vs. $10\log_{10}(E_b/N_0)$ for conventional DD ($N = 2$) and genie–aided prediction–based DF–DD ($N = 3, 5, 100$) for Jakes fading model ($10\log(K) = 10$ dB, $f_DT = 0$).

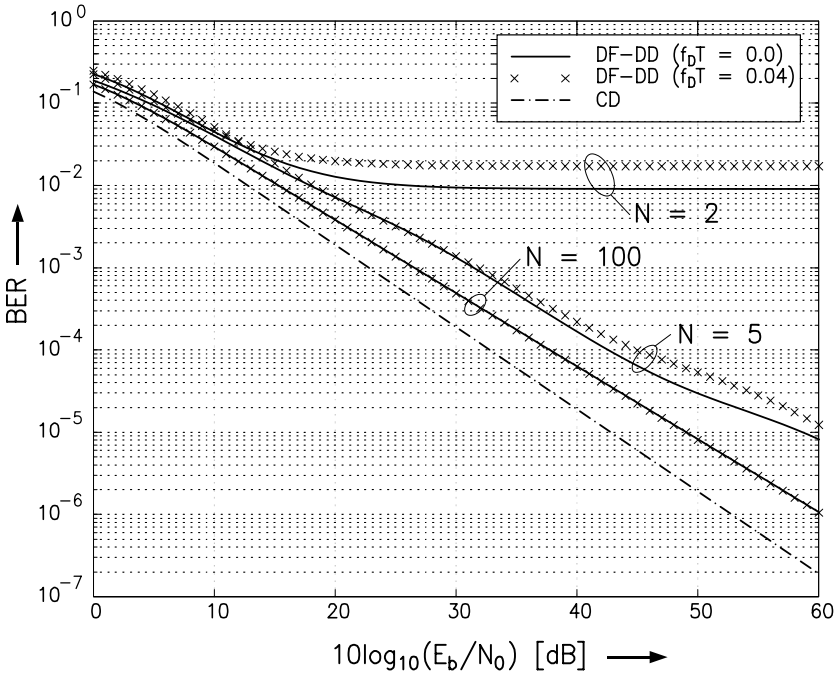

Figure 4: BER of coherent QPSK (denoted by 'CD') and QDPSK vs. $10\log_{10}(E_b/N_0)$ for genie–aided prediction–based DF–DD ($N = 2, 5, 100$) for Jakes fading model ($10\log(K) = 3$ dB, $B_fT = 0.05$).

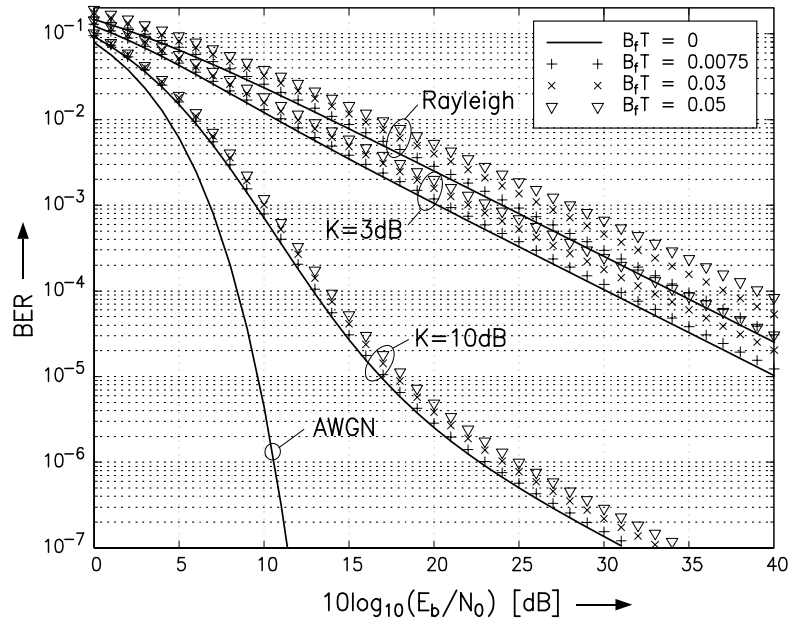

Figure 5: BER vs. $10\log_{10}(E_b/N_0)$ for QDPSK with genie–aided prediction–based DF–DD ($N \to \infty$) for different fading bandwidths B_f and Ricean factors K. Jakes fading model is assumed.

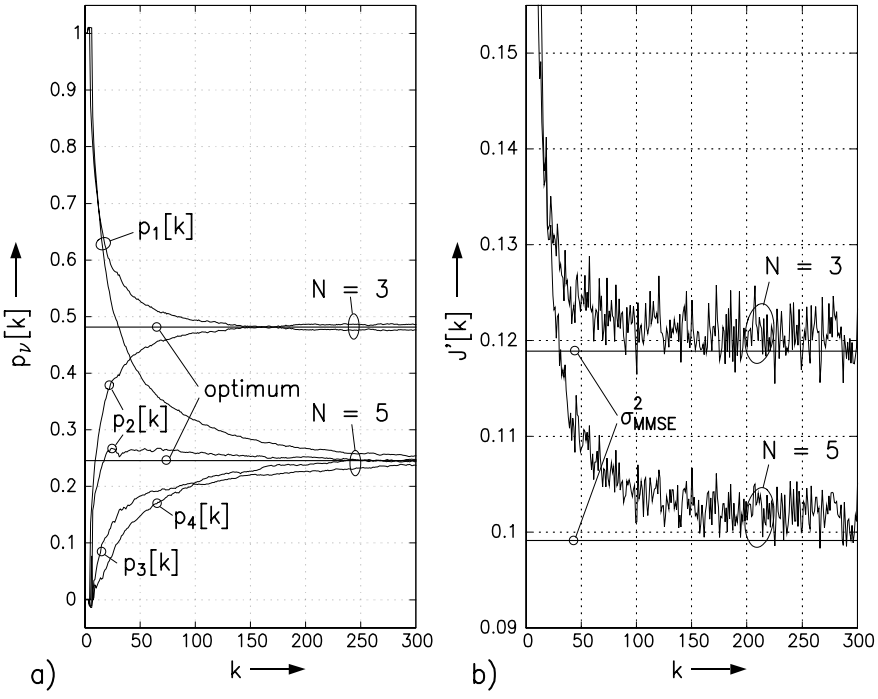

Figure 6: a) Predictor coefficients $p_\nu[k]$, $1 \leq \nu \leq N-1$, vs. k and b) learning curve (mean–squared error $J'[k]$ vs. k) of the adaptation algorithm for an AWGN channel ($10\log_{10}(E_b/N_0) = 8$ dB, $w = 0.99$). Predictors of order 2 ($N = 3$) and 4 ($N = 5$) are considered.

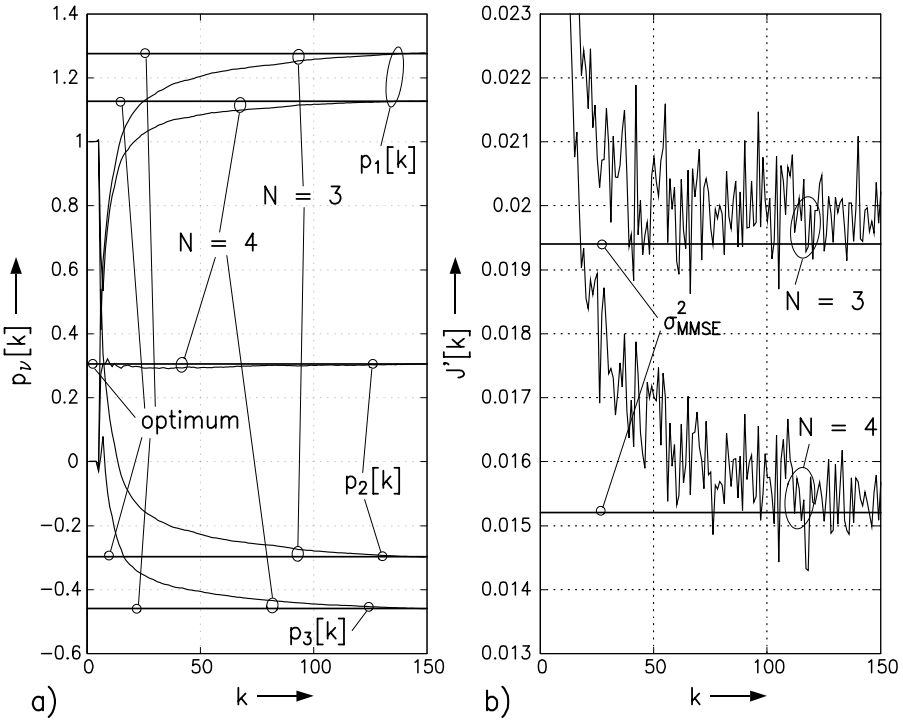

Figure 7: a) Predictor coefficients $p_\nu[k]$, $1 \leq \nu \leq N-1$, vs. k and b) learning curve (mean–squared error $J'[k]$ vs. k) of the adaptation algorithm for a Rayleigh fading channel ($10\log_{10}(E_b/N_0) = 20$ dB, $B_f T = 0.03$, $w = 0.99$). Predictors of order 2 ($N = 3$) and 3 ($N = 4$) are considered.

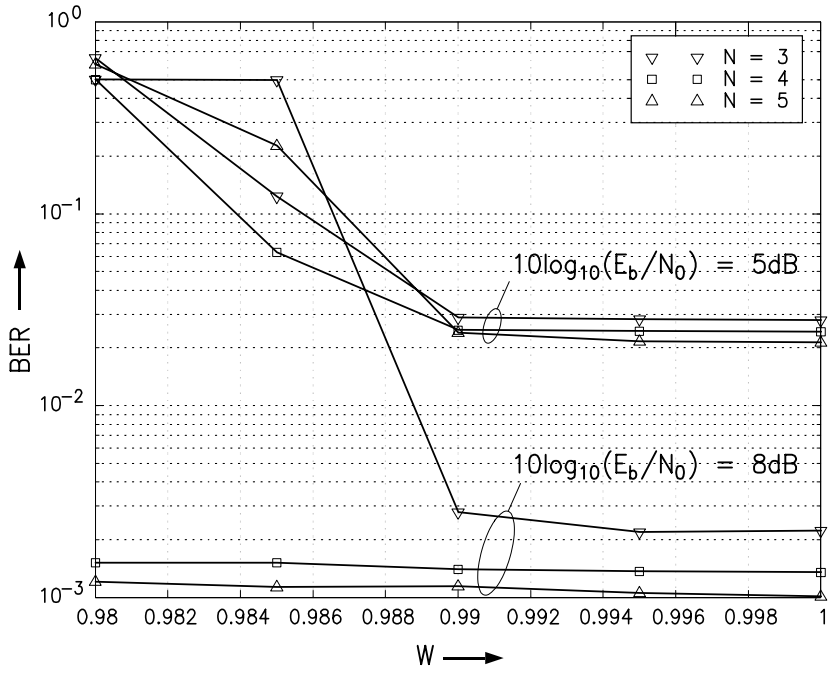

Figure 8: BER vs. forgetting factor w for the AWGN channel.

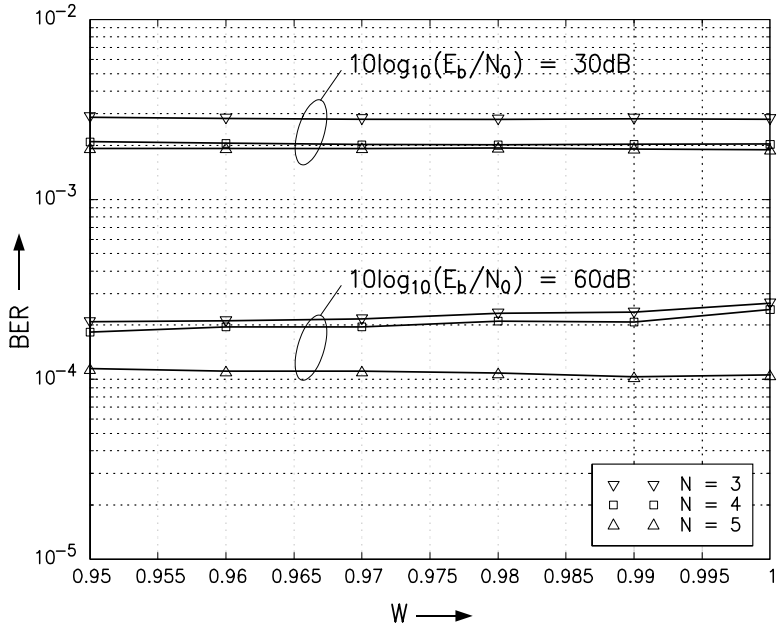

Figure 9: BER vs. forgetting factor w for the Rayleigh fading channel ($B_f T = 0.03$, Jakes model).

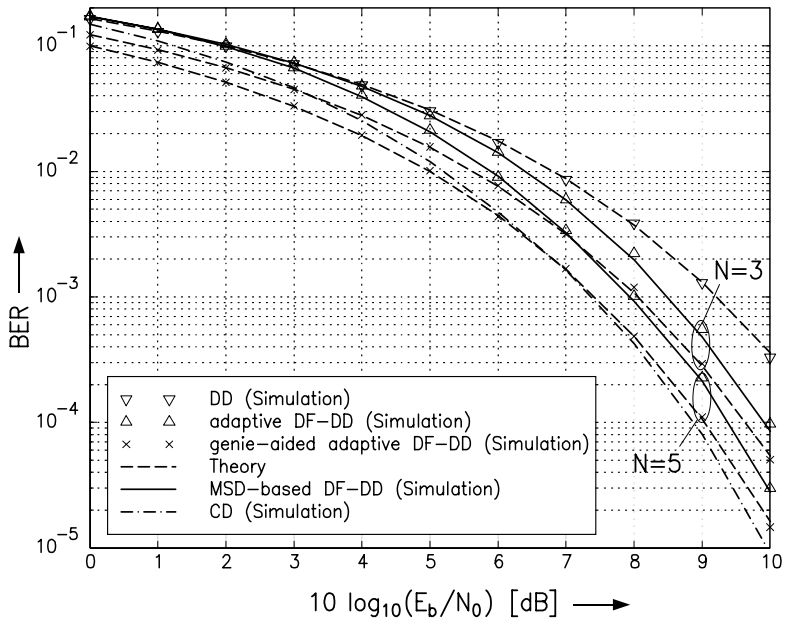

Figure 10: BER vs. $10\log_{10}(E_b/N_0)$ for QDPSK with conventional DD, adaptive prediction–based DF–DD ($w = 1$), genie–aided adaptive (prediction–based) DF–DD ($w = 1$), MSD–based DF–DD (here equivalent to prediction–based DF–DD with optimum fixed predictor coefficients), and CD for the AWGN channel. The theoretical BERs for conventional DD and genie–aided prediction–based DF–DD are also contained.

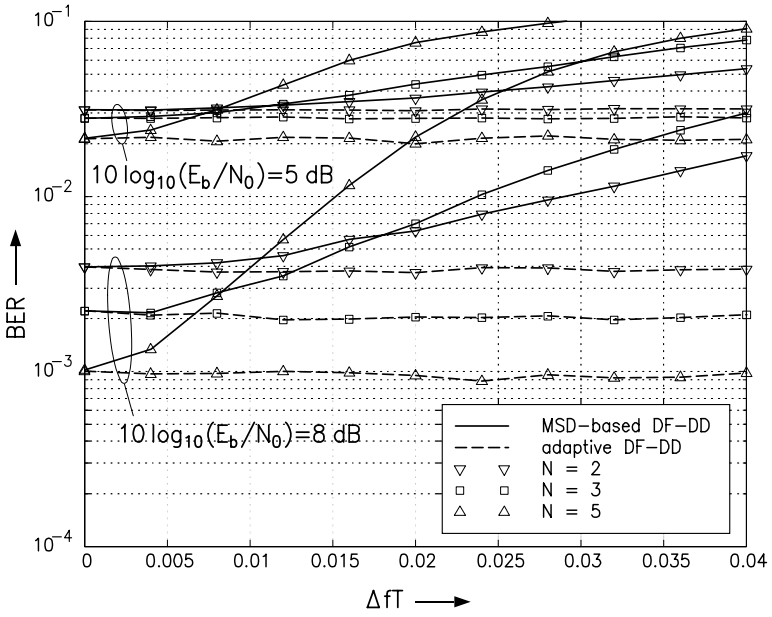

Figure 11: BER vs. normalized frequency offset ΔfT for QDPSK with MSD–based DF–DD optimized for AWGN and adaptive (prediction–based) DF–DD ($w = 1$).

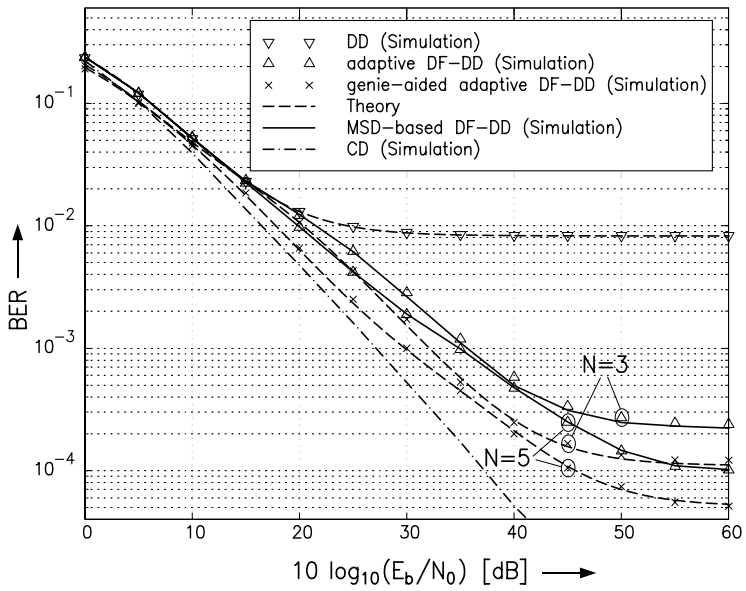

Figure 12: BER vs. $10\log_{10}(E_b/N_0)$ for QDPSK with conventional DD, adaptive prediction–based DF–DD ($w = 0.99$), genie–aided adaptive (prediction–based) DF–DD ($w = 0.99$), MSD–based DF–DD (here equivalent to prediction–based DF–DD with optimum fixed predictor coefficients), and CD for Rayleigh fading ($B_fT = 0.03$). The theoretical BERs for conventional DD and genie–aided prediction–based DF–DD are also contained.

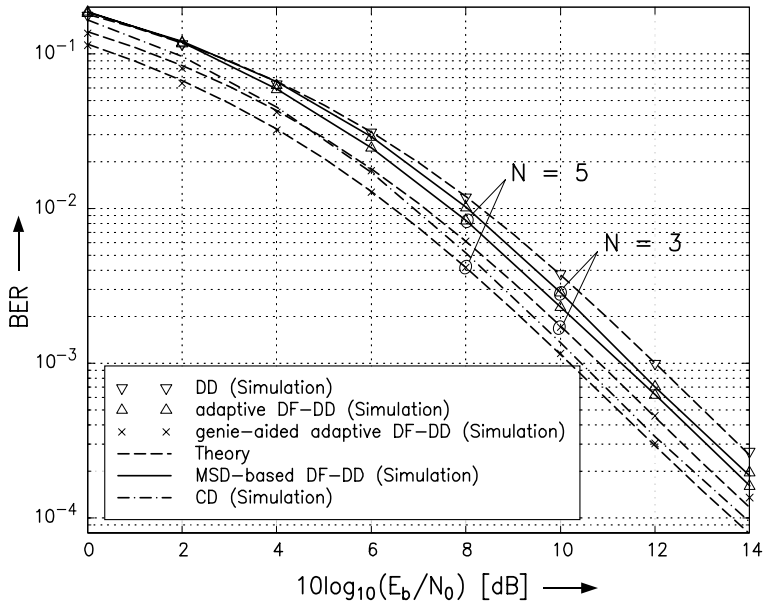

Figure 13: BER vs. $10\log_{10}(E_b/N_0)$ for QDPSK with conventional DD, adaptive (prediction–based) DF–DD ($w = 0.99$), genie–aided adaptive DF–DD, MSD–based DF–DD, and CD for Ricean fading ($10\log_{10}(K) = 10$ dB, $B_fT = 0.0075$, $f_DT = 0$). The theoretical BERs for conventional DD and genie–aided prediction–based DF–DD are also contained.

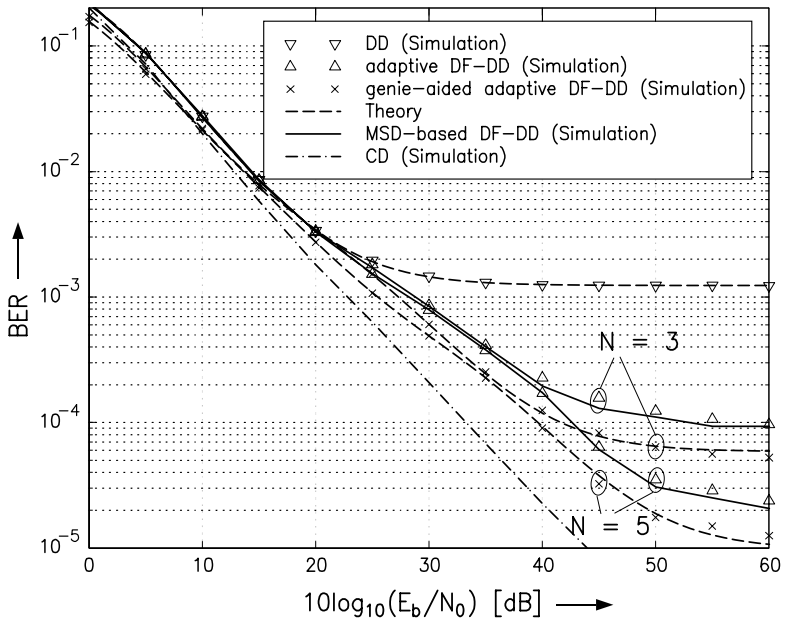

Figure 14: BER vs. $10\log_{10}(E_b/N_0)$ for QDPSK with conventional DD, adaptive (prediction–based) DF–DD ($w = 0.99$), genie–aided adaptive DF–DD, MSD–based DF–DD, and CD for Ricean fading ($10\log_{10}(K) = 3$ dB, $B_fT = 0.03$, $f_DT = 0$). The theoretical BERs for conventional DD and genie–aided prediction–based DF–DD are also contained.