NONCOHERENT RECEIVERS FOR DIFFERENTIAL SPACE–TIME MODULATION

Robert Schober and Lutz H.–J. Lampe

Laboratorium für Nachrichtentechnik, Universität Erlangen–Nürnberg
Cauerstraße 7, D–91058 Erlangen, Germany
e-mail: {*schober,llampe*}*@LNT.de*, WWW: *http://www.LNT.de/LNT2*

***Abstract* — In this paper, noncoherent receivers for differential space–time modulation (DSTM) are investigated. It is shown that the performance of the previously proposed conventional differential detection (DD) receiver is satisfactory only for very slow flat fading channels. However, conventional DD suffers from a considerable loss in performance even for moderately fast fading. In order to overcome this problem, multiple–symbol detection (MSD) and low–complexity decision–feedback differential detection (DF–DD) receivers are derived.**

1. Introduction

Recently, differential space–time modulation (DSTM) has received high attention [1, 2, 3]. DSTM provides transmit diversity while the coefficients of the flat fading channel have not to be known for the detection process. This may be crucial if fading conditions change too rapidly for a reliable channel estimation necessary for coherent detection, e.g. [4, 5].

For DSTM it is distinguished between constellations based on group designs and non–group designs [6]. Constellations whose elements form a group have the advantage that the matrix multiplication necessary for differential encoding has not to be performed explicitly [3]. Also the search for good codes is facilitated by the group property [6]. Among the group designs *diagonal signals* play a prominent role [3, 2]. Diagonal signals achieve full transmit diversity and are especially attractive for low data rates $R \leq 2\,\mathrm{bits}/(\mathrm{channel\,use})$. In addition, for diagonal signals and conventional differential detection (DD) a low–complex fast decoding algorithm exists [7], which makes them very attractive for implementation.

Although a lot of work has been done to design efficient DSTM schemes, the receiver side has received little attention so far. In the above cited references on DSTM only conventional DD is considered. In addition, it is assumed that the fading rate is very low, i.e., the channel coefficients are assumed to be constant over $2N_T$ modulation intervals where N_T is the number of transmit antennas. We will show that DSTM with conventional DD suffers from a severe performance degradation if the slow fading condition is not fulfilled.

Motivated by the fact that multiple–symbol detection (MSD) [8, 9] and decision–feedback differential detection (DF–DD) [10, 11] are very effective for single–antenna transmission schemes we derive corresponding decision rules for the multiple–antenna case. Thereby, we concentrate on diagonal signals because in this case relatively simple decision rules result. Analytical and simulation results show in good agreement the significantly enhanced performance of the proposed receivers for DSTM compared to conventional DD.

2. System Model

2.1. Flat Rayleigh Fading Channel Model

We consider a transmission scheme using N_T transmit antennas and N_R receive antennas. The transmitted symbols are grouped in blocks of size N_T and at time $N_Tk+\kappa$, $k \in \mathbb{Z}$, $0 \leq \kappa \leq N_T-1$, the symbol $s_\mu[N_Tk+\kappa]$ is transmitted by the μth, $0 \leq \mu \leq N_T-1$, antenna.

We model the equivalent complex–baseband representation of the received signal $r_\nu[N_Tk+\kappa]$ at the νth, $0 \leq \nu \leq N_R-1$, receive antenna as

$$r_\nu[N_Tk+\kappa] = \sum_{\mu=0}^{N_T-1} h_{\mu\nu}[N_Tk+\kappa]s_\mu[N_Tk+\kappa]+n_\nu[N_Tk+\kappa], \tag{1}$$

where T–spaced sampling (T is the modulation interval) and perfect symbol synchronization are assumed. $h_{\mu\nu}[N_Tk+\kappa]$ denotes the zero–mean complex fading gain between the μth transmit and the νth receive antenna at time $N_Tk+\kappa$, whereas $n_\nu[N_Tk+\kappa]$ refers to the complex additive white Gaussian noise at receive antenna ν. As usual, it is assumed that all fading processes have identical statistical properties and are uncorrelated in space, i.e., $\mathcal{E}\{h_{\mu\nu}[\cdot]h^*_{\mu'\nu'}[\cdot]\} = 0$ for $\mu \neq \mu'$ or/and $\nu \neq \nu'$ ($\mathcal{E}\{\cdot\}$ and $(\cdot)^*$ denote expectation and complex conjugation, respectively), but correlated in time. Here, we use Clarkes model [12] for the fading autocorrelation function $\varphi_{hh}[\lambda] = \mathcal{E}\{h^*_{\mu\nu}[N_Tk+\kappa]h_{\mu\nu}[N_Tk+\kappa+\lambda]\} = \sigma_h^2 \cdot \mathrm{J}_0(2\pi B_f T\lambda)$, $0 \leq \mu \leq N_T-1$, $0 \leq \nu \leq N_R-1$, where σ_h^2, $\mathrm{J}_0(\cdot)$, and B_f refer to the variance of the fading processes, the zeroth order Bessel function of the first kind, and the one–sided bandwidth of the underlying continuous–time fading processes, respectively. It is worth mentioning that the received signal can only be modeled as in (1) as long as the continuous–time fading process is approximately constant during the symbol duration T. Here, we assume that this condition is fulfilled for $B_fT \leq 0.03$. The white Gaussian noise processes $n_\nu[\cdot]$ at different receive antennas are mutually uncorrelated and have equal variance $\sigma_n^2 = \mathcal{E}\{|n_\nu[N_Tk+\kappa]|^2\}$, $0 \leq \nu \leq N_R-1$. The transmitted symbols are normalized to $\mathcal{E}\{\sum_{\mu=0}^{N_T-1}|s_\mu[N_Tk+\kappa]|^2\} = 1$, i.e., the (mean) signal–to–noise ratio (SNR) per receive antenna is $\mathrm{SNR} = \sigma_h^2/\sigma_n^2$.

2.2. *Differential Space–Time Modulation*

At the transmitter, the signals are organized in $N_T \times N_T$ matrices $\boldsymbol{S}[k]$ with elements $s_\mu[N_T k + \kappa]$ in row κ and column μ. Matrix $\boldsymbol{S}[k]$ is obtained by differential encoding from $\boldsymbol{S}[k-1]$ and $\boldsymbol{V}[k]$

$$\boldsymbol{S}[k] = \boldsymbol{V}[k]\,\boldsymbol{S}[k-1]. \tag{2}$$

Here, we restrict our attention to constellations whose elements form a group under matrix multiplication, i.e., the possible values for $\boldsymbol{V}[k]$ and $\boldsymbol{S}[k]$ belong to a finite set with $L = 2^{N_T R}$ elements [6], where R denotes the data rate. In particular, we are interested in diagonal constellations where the unitary matrices $\boldsymbol{V}[k] = \boldsymbol{V}_{l[k]}$ are taken from the set $\mathcal{A} = \{\boldsymbol{V}_l = \mathrm{diag}\{\exp(j2\pi u_0 l/L), \exp(j2\pi u_1 l/L), \ldots, \exp(j2\pi u_{N_T-1} l/L)\} | l \in \{0, 1 \ldots, L-1\}\}$ ($\sqrt{-1} \triangleq j$ and $\mathrm{diag}\{\cdot\}$ denote the imaginary unit and a diagonal matrix, respectively). For given N_T and data rate R the coefficients u_μ, $0 \le \mu \le N_T - 1$, are optimized to achieve maximum diversity product [3]. For all examples considered in this paper, we adopt the values given in [3, Table I]. For almost all constellations given in [3, Table I] a *Gray labeling* for the bits may be constructed (cf. [13]). The constellation with $N_T = 4$, $R = 1\,\mathrm{bit}/(\mathrm{channel\,use})$ is an exception since each symbol has 8 nearest neighbor symbols. In this case, a natural labeling for $l = 0, 1 \ldots, 15$ may be used.

3. Noncoherent Receivers

3.1. *Multiple–Symbol Detection*

Using (1), the signals received at antennas 0 to $N_R - 1$ can be collected in a vector

$$\boldsymbol{r}[N_T k+\kappa] = \boldsymbol{s}[N_T k+\kappa]\boldsymbol{H}[N_T k+\kappa] + \boldsymbol{n}[N_T k+\kappa], \tag{3}$$

with the definitions $\boldsymbol{r}[N_T k+\kappa] \triangleq [r_0[N_T k+\kappa]\ \ldots\ r_{N_R-1}[N_T k+\kappa]]$, $\boldsymbol{s}[N_T k+\kappa] \triangleq [s_0[N_T k+\kappa]\ \ldots\ s_{N_T-1}[N_T k+\kappa]]$, $\boldsymbol{n}[N_T k+\kappa] \triangleq [n_0[N_T k+\kappa]\ \ldots\ n_{N_R-1}[N_T k+\kappa]]$, and the element of matrix $\boldsymbol{H}[N_T k+\kappa]$ in column μ and row ν is $h_{\mu\nu}[N_T k+\kappa]$, $0 \le \mu \le N_T - 1$, $0 \le \mu \le N_R - 1$.

Now, the vectors $\boldsymbol{r}[\cdot]$ received in the observation interval $[N_T k + N_T - 1,\ N_T(k-(N-1))]$ of N matrix symbols are collected in a matrix $\bar{\boldsymbol{R}}[k] \triangleq [\boldsymbol{r}^T[N_T k + N_T - 1]\ \ldots\ \boldsymbol{r}^T[N_T(k-(N-1))]]^T$ ($[\cdot]^T$ denotes transposition). Using (3), $\bar{\boldsymbol{R}}[k]$ can be expressed as

$$\bar{\boldsymbol{R}}[k] = \bar{\boldsymbol{S}}[k]\bar{\boldsymbol{H}}[k] + \bar{\boldsymbol{N}}[k], \tag{4}$$

where the definitions $\bar{\boldsymbol{H}}[k] \triangleq [\boldsymbol{H}^T[N_T k + N_T - 1]\ \ldots\ \boldsymbol{H}^T[N_T(k-(N-1))]]^T$, $\bar{\boldsymbol{N}}[k] \triangleq [\boldsymbol{n}^T[N_T k + N_T - 1]\ \ldots\ \boldsymbol{n}^T[N_T(k-(N-1))]]^T$ and

$$\bar{\boldsymbol{S}}[k] \triangleq \begin{bmatrix} \boldsymbol{s}[N_T k + N_T - 1] & \boldsymbol{0}_{N_T} & \ldots & \boldsymbol{0}_{N_T} \\ \boldsymbol{0}_{N_T} & \vdots & \vdots & \boldsymbol{0}_{N_T} \\ \vdots & \ddots & \ddots & \vdots \\ \boldsymbol{0}_{N_T} & \ldots & \boldsymbol{0}_{N_T} & \boldsymbol{s}[N_T(k-(N-1))] \end{bmatrix}$$

($\boldsymbol{0}_{N_T}$ denotes the N_T dimensional all–zero row vector) are used.

Since $h_{\mu\nu}[\cdot]$ and $n_\nu[\cdot]$, $0 \le \mu \le N_T - 1$, $0 \le \nu \le N_R - 1$, are zero–mean complex Gaussian random processes, the probability density function (pdf) of $\bar{\boldsymbol{R}}[k]$ conditioned on $\bar{\boldsymbol{S}}[k]$ is given by

$$p(\bar{\boldsymbol{R}}[k]|\bar{\boldsymbol{S}}[k]) = \frac{1}{(\pi^{NN_T}\det\{\boldsymbol{C}_R[k]\})^{N_R}} \cdot \exp\left(-\mathrm{tr}\left(\bar{\boldsymbol{R}}^H[k]\boldsymbol{C}_R^{-1}[k]\bar{\boldsymbol{R}}[k]\right)\right) \tag{5}$$

($\det\{\cdot\}$ and $\mathrm{tr}(\cdot)$ denote the determinant and the trace of a matrix, respectively), where $\boldsymbol{C}_R[k]$ denotes the conditional covariance matrix

$$\boldsymbol{C}_R[k] \triangleq \mathcal{E}\{\bar{\boldsymbol{R}}[k]\bar{\boldsymbol{R}}^H[k]|\bar{\boldsymbol{S}}[k]\}. \tag{6}$$

The MSD decision rule for general DSTM can be obtained by maximizing (5) with respect to the vector of transmitted symbols $\boldsymbol{l}[k] \triangleq [l[k]\ l[k-1]\ \ldots\ l[k-(N-2)]]$. In general, this decision rule will be rather complicated and computationally complex. However, for *diagonal signals* [3] significant simplifications are possible. It can be shown that for this important special case the MSD decision rule for $\boldsymbol{l}[k]$ is given by

$$\hat{\boldsymbol{l}}[k] = \underset{\boldsymbol{l}[k]}{\mathrm{argmax}}\left\{\Re\left\{\sum_{\xi_1=0}^{N-1}\sum_{\xi_2=\xi_1+1}^{N-1} t_{\xi_1\xi_2} \sum_{\mu=0}^{N_T-1}\prod_{m=\xi_1}^{\xi_2-1}\exp\left(j\frac{2\pi u_\mu l[k-m]}{L}\right) \sum_{\nu=0}^{N_R-1} r_\nu^*[N_T(k-\xi_1)+\mu]\,r_\nu[N_T(k-\xi_2)+\mu]\right\}\right\} \tag{7}$$

where the decisions $\hat{l}[k-m]$, $0 \le m \le N-2$, are collected in vector $\hat{\boldsymbol{l}}[k]$ and $\Re\{\cdot\}$ denotes the real part of a complex number. The coefficients $t_{\xi_1\xi_2}$, $0 \le \xi_1, \xi_2 \le N-1$, are the elements of matrix

$$\boldsymbol{T} \triangleq -(\boldsymbol{C}_h^{N_T} + \sigma_n^2 \boldsymbol{I}_N)^{-1} \tag{8}$$

where $\boldsymbol{C}_h^{N_T}$ is the (temporal) autocorrelation matrix of process $h_{\mu\nu}[N_T k]$, which is independent from μ and ν. (7) may be viewed as a generalization of the MSD decision rule for M–ary differential phase–shift keying (MDPSK) transmission over a flat Rayleigh fading channel (cf. [8]) to diagonal DSTM and multiple receive antennas.

Unfortunately, the above MSD decision rule requires the calculation of $2^{N_T R(N-1)}/(N_T R(N-1))$ metrics per bit decision, i.e., complexity is exponential in number of transmit antennas N_T, data rate R, and observation window size N. However, it is well known from the single antenna case that the power efficiency of MSD increases with increasing N. Thus, a low–complexity scheme with a similar performance as MSD but whose complexity is almost independent of N will be derived in the next section.

3.2. Decision–Feedback Differential Detection

A. Decision Rule: A very simple way to reduce complexity is to introduce decision feedback, i.e., in (7) the $l[k-m]$ are replaced by previously decided symbols $\hat{l}[k-m]$, $1 \le m \le N-2$, and a decision is made only on $l[k]$. If all irrelevant terms are neglected, the resulting DF–DD decision rule is

$$\hat{l}[k] = \underset{l}{\operatorname{argmax}} \left\{ \Re\left\{ \sum_{\mu=0}^{N_T-1} \sum_{\nu=0}^{N_R-1} \exp\left(j\frac{2\pi u_\mu l}{L}\right) r_\nu^*[N_T k+\mu]\hat{r}_{\mathrm{ref},\nu}[N_T(k-1)+\mu] \right\}\right\} \quad (9)$$

with the reference signal

$$\hat{r}_{\mathrm{ref},\nu}[N_T(k-1)+\mu] = \sum_{\xi=1}^{N-1} t_{0\xi} \prod_{m=1}^{\xi-1} \exp\left(j\frac{2\pi u_\mu \hat{l}[k-m]}{L}\right) r_\nu[N_T(k-\xi)+\mu]. \quad (10)$$

(9) and (10) may be viewed as a generalization of the DF–DD decision rule for the single–antenna case proposed in [14]. Now, only $2^{N_T R}/(N_T R)$ metrics per bit decision have to be calculated, i.e., complexity is only exponential in N_T and R. For the special case $N=2$ both the MSD (7) and the DF–DD (9) decision rules are identical to the decision rule for conventional DD given e.g. in [3].

B. Relation to Linear Prediction: First, observe that $\boldsymbol{C}_h^{N_T} + \sigma_n^2 \boldsymbol{I}_N$ (8) may be interpreted as autocorrelation matrix of the process

$$c_{\mu\nu}[N_T k] \triangleq h_{\mu\nu}[N_T k] + n_\nu[N_T k], \quad (11)$$

$0 \le \mu \le N_T-1$, $0 \le \nu \le N_R-1$. Therefore, using the same approach as in [14, Section V], it is straightforward to show that the coefficients $t_{0\xi}$, $1 \le \xi \le N-1$, can be expressed as

$$t_{0\xi} = \frac{p_\xi}{\sigma_e^2}, \qquad 1 \le \xi \le N-1, \quad (12)$$

where p_ξ are the coefficients of the $(N-1)$st order predictor for process $c_{\mu\nu}[N_T k]$ and σ_e^2 is the resulting prediction error variance. Since a multiplicative real constant does not influence the decision rule, in (10) the coefficients $t_{0\xi}$ may be replaced by the predictor coefficients p_ξ, $1 \le \xi \le N-1$. This new formulation of the decision rule emphasizes the *estimator–correlator* structure [15] of the DF–DD receiver and facilitates its implementation. In particular, the predictor coefficients can be adaptively calculated employing the recursive least squares (RLS) algorithm (cf. [13]).

C. Fast Decoding Algorithm: Although the complexity of the DF–DD decision rule (9) is almost independent of the observation window size N, it is still exponential in data rate R and number of transmit antennas N_T. In [7] a (suboptimum) fast decoding algorithm is given for conventional DD of DSTM with diagonal signals whose complexity is only polynomial in R and N_T. Fortunately, as will be shown in the following this decoding algorithm can also be applied to DF–DD.

For application of the fast decoding algorithm of [7] a decision rule of the form

$$\hat{l}[k] = \underset{l}{\operatorname{argmax}} \left\{ \sum_{\mu=0}^{N_T-1} \sum_{\nu=0}^{N_R-1} A_{\mu\nu}^2[k] \cos((u_\mu l - \varphi_{\mu\nu}[k])\frac{2\pi}{L}) \right\} \quad (13)$$

is necessary, where amplitude $A_{\mu\nu}[k]$ and phase difference $\varphi_{\mu\nu}[k]$ (in units of $2\pi/L$) have to be independent of the current trial symbol l. A comparison with (9) shows that a representation according to (13) is obtained by setting

$$A_{\mu\nu}[k] \triangleq \sqrt{\left| r_\nu[N_T k+\mu]\hat{r}_{\mathrm{ref},\nu}^*[N_T(k-1)+\mu]\right|} \quad (14)$$

$$\varphi_{\mu\nu}[k] \triangleq \arg\left\{ r_\nu[N_T k+\mu]\hat{r}_{\mathrm{ref},\nu}^*[N_T(k-1)+\mu]\right\} \frac{L}{2\pi}, \quad (15)$$

where $\arg\{\cdot\}$ refers to the phase of a complex number. Thus, the fast decoding algorithm of [7] can be applied directly to DF–DD to further reduce computational complexity. However, since in this paper we are primarily interested in the achievable power efficiency of DF–DD, in all results presented in Section 5 the exact and more complex decision rule according to (9) is used.

4. Performance Analysis for DF–DD

In this section, the achievable performance of DF–DD for diagonal DSTM is analyzed. Since it is difficult to take the effect of decision errors in the feedback symbols into account, we first analyze genie–aided DF–DD, i.e., it is assumed that all feedback symbols are correct. This is a standard approach for the analysis of DF–DD (cf. e.g. [16, 14]).

4.1. Pairwise Error Probability

First, we analyze the pairwise error probability $P_e(l_1, l_2)$, i.e., the probability of detecting $\hat{l}[k] = l_2$, when $l[k] = l_1$ ($l_1, l_2 \in \{0, 1, \ldots, L-1\}$, $l_1 \neq l_2$) is transmitted. Using a similar approach as the authors in [5] for conventional DD and slow fading, we obtain [13]

$$P_e(l_1, l_2) = \frac{1}{\pi} \int_0^{\pi/2} \prod_{\mu=0}^{N_T-1} \left(\frac{1}{1+\frac{\alpha_\mu(l_1,l_2)}{4\cos^2\Theta}} \right)^{N_R} \mathrm{d}\Theta, \quad (16)$$

where

$$\alpha_\mu(l_1, l_2) = 4\,\frac{\sigma_h^2 + \sigma_n^2 - \sigma_e^2}{\sigma_e^2}\,(1 - d_\mu^2(l_1, l_2)) \quad (17)$$

is valid. $\sigma_e^2 = \varphi_{hh}[0] + \sigma_n^2 - \sum_{\xi=1}^{N-1} p_\xi^* \varphi_{hh}[N_T \xi]$ and $d_\mu(l_1, l_2) = |\cos(\pi u_\mu (l_1 - l_2)/L)|$ denote the prediction error variance for process $c_{\mu\nu}[N_T k]$ and the μth singular value of the matrix $\frac{1}{2}(\boldsymbol{I}_{N_T} + \boldsymbol{V}_{l_1}^H \boldsymbol{V}_{l_2})$ [3], respectively.

In order to get a more intuitive insight, the Chernoff upper bound [17] on $P_e(l_1, l_2)$ may be considered. Using the same technique as proposed in [5, Appendix B] for conventional DD, we obtain

$$P_e(l_1, l_2) \leq \frac{1}{2} \prod_{\mu=0}^{N_T-1} \left(\frac{1}{1 + \frac{\alpha_\mu(l_1, l_2)}{4}} \right)^{N_R} . \quad (18)$$

Observe that (16) and (18) are generalizations of the corresponding equations given in [5] which are only valid for the special case $N = 2$ and $B_f T = 0$.

(17) and (18) show that $P_e(l_1, l_2)$ decreases if the prediction error variance σ_e^2 decreases. On the other hand, σ_e^2 increases for a given observation window size N if N_T is increased as long as the fading bandwidth B_f is larger than zero since the process $c_{\mu\nu}[N_T k]$ becomes less correlated. This means, there may be situations when the error rate increases if the number of transmit antennas is increased.

The limiting performance of DF–DD is obtained for $N \rightarrow \infty$. Borrowing again from [14, Section VI], it is straightforward to show that for Clarkes fading model [12] and $N \rightarrow \infty$

$$\sigma_e^2 = \sigma_n^2 \left(\frac{\mathrm{e}\, \sigma_\mathrm{h}^2}{2\pi N_T B_f T \sigma_n^2} \right)^{2 N_T B_f T} \exp(2 N_T B_f T I) \quad (19)$$

is obtained for the prediction error variance, where e is the Euler number and I is defined as

$$I \triangleq \int_0^{\pi/2} \ln(1 + \pi N_T B_f T \sigma_n^2 / \sigma_h^2 \sin\Theta) \sin\Theta \, \mathrm{d}\Theta. \quad (20)$$

Note that we presume $N_T B_f T < 1/2$, i.e., process $h_{\mu\nu}[N_T k]$ is bandlimited. Therefore, there is no error floor for $\sigma_n^2 = 0$ since in this case $\sigma_e^2 = 0$ and consequently $P_e(l_1, l_2) = 0$ result [13]. For $B_f T \rightarrow 0$, $\sigma_e^2 = \sigma_n^2$ holds and the pairwise error probability for DF–DD is identical to that of coherent detection. For $B_f T > 0$ and $\sigma_n^2 > 0$, $\sigma_e^2 > \sigma_n^2$ holds which implies an inevitable loss of DF–DD compared to coherent detection with perfect channel state information (CSI).

4.2. *Approximation for BER*

If we use the union bound to approximate the symbol error rate for genie–aided DF–DD and assume that $R N_T$ bits are Gray mapped to one symbol, we obtain for the bit error rate (BER) the approximation

$$P_b^{\mathrm{genie}} \approx \frac{1}{R N_T} \sum_{l=1}^{L-1} P_e(l, 0). \quad (21)$$

For realizable DF–DD, the influence of decision errors has to be taken into account for $N > 2$. Our investigations have shown that for realizable DF–DD BER has to be approximately doubled since each decision error is very likely to cause a second decision error. This is in accordance with the literature for single–antenna DF–DD (e.g. [16, 14]). Hence, the BER of realizable DF–DD can be approximated as

$$P_b \approx \begin{cases} P_b^{\mathrm{genie}}, & N = 2 \\ 2 \cdot P_b^{\mathrm{genie}}, & N > 2 \end{cases} . \quad (22)$$

5. Results and Discussion

Since DF–DD and MSD yield almost the same error performance [13], in this paper we only show simulation and numerical results for DF–DD. We have simulated a system as described in Section 2, i.e., independent Rayleigh fading channels are assumed between any pair of transmit and receive antennas and the diagonal signals for DSTM are taken from [3, Table I]. For the numerical results, the BER approximation given by (22), (21), (16) is employed. For the limiting case $N \rightarrow \infty$, the prediction error variance necessary for calculation of BER is obtained from (19). An approximation for BER of DSTM with coherent detection is given by the lower part of (22) ($N > 2$) by setting $\sigma_e^2 \equiv \sigma_n^2$.

In the further discussion, we focus on the case where only one receive antenna is present ($N_R = 1$), since the influence of multiple receive antennas should be clear from (16), (18).

For our first example, $R = 1\,\mathrm{bit}/(\mathrm{channel\,use})$ and $B_f T = 0.03$ are valid. Figs. 1a), b), and c) show BER vs. SNR for $N_T = 1$, 2, and 3 transmit antennas, respectively. All figures show that DF–DD with $N > 2$ can yield significant gains over conventional DD ($N = 2$). Fortunately, the major part of the achievable gain can be realized with moderate observation window size N, i.e., the gap between the curves for $N = 5$ and $N \rightarrow \infty$ is relatively small. Observe that for conventional DD BER increases with increasing number of transmit antennas. In this case, the influence of the higher prediction error variance on BER is stronger than the improvement due to enhanced diversity. However, for $N = 5$ power efficiency improves with increasing number of transmit antennas, i.e., the receiver can take advantage of the higher diversity. The gap between coherent detection with perfect CSI and optimum DF–DD ($N \rightarrow \infty$) increases with increasing N_T since the minimum achievable prediction error variance (cf. (19)) increases. Nevertheless, for large values of N no error floor can be observed since the underlying fading processes are bandlimited. Remarkably, there is a very good agreement between the numerical approximation and the simulation results.

Fig. 2 shows BER vs. $B_f T$ for $R = 2\,\mathrm{bits}/(\mathrm{channel\,use})$ and SNR $\rightarrow \infty$, i.e., the error floor is investigated. Only numerical results are shown. We restrict ourselves to the interval $0.01 \leq B_f T \leq 0.03$ since BER becomes very small for smaller normalized fading bandwidths. It can be observed that the error floor for conventional DD ($N = 2$) can be reduced by orders of magnitude if the observation window size is increased to $N = 3$ or $N = 4$. This statement holds for $N_T = 1 - 4$. From the considerations in Section 4.1 it can be concluded that the error floor is removed completely for $N \rightarrow \infty$ since the underlying fading processes are strictly bandlimited. Again, depending on N and $B_f T$, there is a trade–off between enhanced diversity and increased prediction error variance when varying N_T from 1 to 4. Remarkably, in case of conventional DD performance deteriorates with increased number of transmit antennas for almost all $B_f T$ considered in Fig. 2.

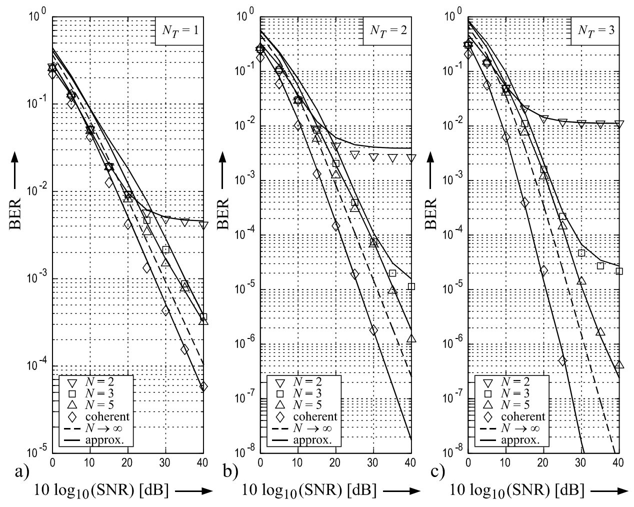

Figure 1: BER vs. $10 \log_{10}(\mathrm{SNR})$ for DF–DD with a) $N_T = 1$, b) $N_T = 2$, and c) $N_T = 3$ transmit antennas. $\bigtriangledown$, $\square$, $\triangle$, and $\diamond$ denote simulation points. $R = 1\,\mathrm{bit}/(\mathrm{channel\,use})$, $B_f T = 0.03$, and $N_R = 1$ are valid.

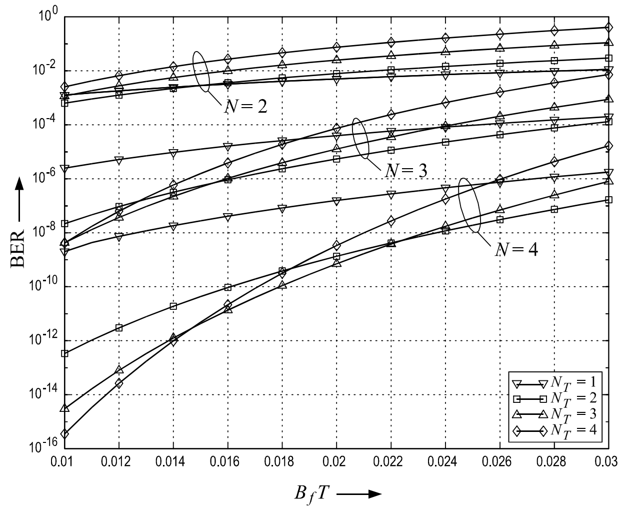

Figure 2: BER vs. $B_f T$ for DF–DD. $R = 2\mathrm{bits}/(\mathrm{channel\,use})$, $\mathrm{SNR} \to \infty$, and $N_R = 1$ are valid.

6. Conclusions

In this paper, it has been shown that conventional DD is suitable for DSTM only for very slow fading channels. In particular, increasing the number of transmit antennas is not beneficial for moderately fast or fast fading if conventional DD is used. Thus, we have derived MSD and DF–DD receivers for DSTM with diagonal signals. Both yield almost the same performance, however, the latter requires a lower computational complexity [13]. Simulation and analytical results confirm that large performance gains can be realized by DF–DD while complexity is increased only moderately.

References

[1] V. Tarokh and H. Jafarkhani. A Differential Detection Scheme for Transmit Diversity. *IEEE Journal on Selected Areas in Communications*, SAC-18:1168–1174, July 2000.

[2] B.L. Hughes. Differential Space–Time Modulation. *IEEE Trans. on Inf. Theory*, IT-46:2567–2578, November 2000.

[3] B.M. Hochwald and W. Sweldens. Differential Unitary Space–Time Modulation. *IEEE Trans. on Commun.*, COM-48:2041–2052, December 2000.

[4] S.M. Alamouti. A Simple Transmitter Diversity Scheme for Wireless Communications. *IEEE Journal on Selected Areas in Communications*, SAC-16:1451–1458, October 1998.

[5] B.M. Hochwald and T.L. Marzetta. Unitary Space–Time Modulation for Multiple–Antenna Communications in Rayleigh Flat Fading. *IEEE Trans. on Inf. Theory*, IT-46:543–564, March 2000.

[6] B. Hassibi, B.M. Hochwald, A. Shokrollahi, and W. Sweldens. Representation Theory for High–Rate Multiple–Antenna Code Design. *Preprint, available at: http://mars.bell-labs.com*, March 2000.

[7] K.L. Clarkson, W. Sweldens, and A. Zheng. Fast multiple antenna differential decoding. *IEEE Trans. on Commun.*, 49, February 2001.

[8] 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.

[9] 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.

[10] P.Y. Kam and C.H. Teh. Reception of PSK Signals Over Fading Channels Via Quadrature Amplitude Estimation. *IEEE Trans. on Commun.*, COM-31:1024–1027, August 1983.

[11] R. Schober and W. H. Gerstacker. Decision–Feedback Differential Detection Based on Linear Prediction for MDPSK Signals Transmitted Over Ricean Fading Channels. *IEEE Journal on Sel. Areas in Commun. (JSAC): Wireless Communications Series*, 18:391–402, March 2000.

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

[13] R. Schober and L.H.-J. Lampe. Noncoherent Receivers for Differential Space–Time Modulation. *Submitted to IEEE Trans. on Commun.*, January 2001.

[14] 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.

[15] T. Kailath. Correlation Detection of Signals Perturbed by a Random Channel. *IRE Trans. Information Theory*, IT-6:361–366, July 1960.

[16] 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.

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