# Optimized Delay Diversity for Frequency-Selective Fading Channels

Thorsten Hehn, *Student Member, IEEE*, Robert Schober, *Member, IEEE*, and Wolfgang H. Gerstacker, *Member, IEEE*

***Abstract*—This paper proposes an optimized delay diversity (ODD) scheme for frequency-selective fading channels. The novel scheme requires knowledge of the channel impulse response (CIR) autocorrelation matrix at the transmitter, but the CIRs themselves have to be available only at the receiver side. A cost function for optimization of the ODD transmit filters is derived and a steepest descent algorithm for iterative calculation of the filter coefficients is provided. In addition, an upper bound on the cost function is derived and employed to prove the asymptotic optimality of the generalized DD (GDD) scheme in [Proc. IEEE Int. Conf. Communications (ICC), (2002) p. 1949] for very high signal–to-noise ratios (SNRs) and transmit filters of maximum length. However, for SNRs of practical interest and reasonable filter lengths, the novel ODD scheme significantly outperforms GDD for both optimum and suboptimum equalizations. It is also shown that, in contrast to the frequency-nonselective case, for frequency-selective channels, transmit diversity schemes designed under the high SNR assumption may perform poorly for practically relevant SNRs.**

***Index Terms*—Delay diversity (DD), equalization, frequency-selective channels, transmit diversity (TD).**

## I. Introduction

Although most of the initial research on transmit diversity (TD) assumed flat-fading channels [2]–[4], more recently it has been shown that TD can also lead to significant performance improvements for frequency-selective fading channels. Frequency-selective fading channels are encountered in, e.g., the Global System for Mobile Communications (GSM), the Enhanced Data rates for GSM Evolution (EDGE) system, and future broadband wireless systems.

TD schemes for frequency-selective fading channels that require processing of an entire burst have been proposed, e.g., in [5]–[9]. The scheme in [7] is a generalization of Alamouti's space–time block code (STBC), whereas the schemes in [5], [6] and [8], [9] are based on orthogonal frequency division multiplexing (OFDM) and single-carrier transmission combined with frequency-domain equalization, respectively. These burst-based TD schemes achieve good performance and are amenable to low-complexity receiver design. However, the channel is required to be constant over the entire data burst, and it is unclear how burst-based TD schemes can be adapted to time-variant channels. In addition, if these schemes are used to upgrade existing systems from single-antenna transmission to TD, the burst structure has to be modified and the receiver processing (channel estimation, equalization, etc.) has to be changed. These disadvantages may make burst-based TD schemes less attractive for the upgrade of existing systems. The application of Alamouti's STBC [4] in frequency-selective fading channels was proposed in [10] and [11]. Although tracking of time-variant channels is possible in this case, the resulting equalizer complexity is relatively high. A space–time trellis code (STTC) [3] designed for flat fading with two transmit antennas was adopted in [12] and [13] for frequency-selective fading. Since this particular STTC can be interpreted as a data-dependent delay diversity (DD) scheme [12], the overall channel impulse response (CIR) is similar to the single transmit antenna case, but data dependent. This data dependence of the overall CIR makes the design of low-complexity equalizers difficult. In addition, since the STTC in [3] was designed for the flat-fading case, optimality in frequency-selective channels cannot be guaranteed. Assuming high signal-to-noise ratios (SNRs), STTCs and STBCs for frequency-selective channels were designed in [14] and [15], respectively. However, in this paper, we will show that for frequency-selective channels TD schemes optimized for high SNRs may be strongly suboptimum for low-to-moderate SNRs.

In this paper, we optimize simple DD for frequency-selective fading channels. DD was first proposed in [2] for flat-fading channels and has recently been generalized to frequency-selective channels in [1]. DD has the advantage that the overall channel can be modeled as a single-input channel. Therefore, the same channel estimation, channel tracking, and equalization techniques as in the single transmit antenna case can be used. Also, existing mobile communication systems can be upgraded easily with DD since the burst structure does not have to be modified. In the generalized DD (GDD) scheme in [1], transmit antenna $n_T, 1 \leq n_T \leq N_T$, delays the transmitted data stream by $(n_T - 1)L$ symbols, where $N_T$ and $L$ are the number of transmit antennas and the CIR length, respectively. In that way, full TD is achieved; however, the resulting overall CIR may be excessively long, which implies a high equalizer complexity.

Manuscript received April 24, 2003; revised December 16, 2003; accepted June 9, 2004. The editor coordinating the review of this paper and approving it for publication is Y.-C. Liang. This work was supported in part by the Canada Research Chair Program and in part by a grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada. This paper was presented in part at the IEEE International Conference on Communications (ICC), Paris, France, June 2004.

T. Hehn and W. H. Gerstacker are with the University of Erlangen-Nuernberg, Erlangen D-91058, Germany (e-mail: hehn@LNT.de; gersta@LNT.de).

R. Schober is with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC V6T 1Z4, Canada (e-mail: rschober@ece.ubc.ca).

Digital Object Identifier 10.1109/TWC.2005.853818

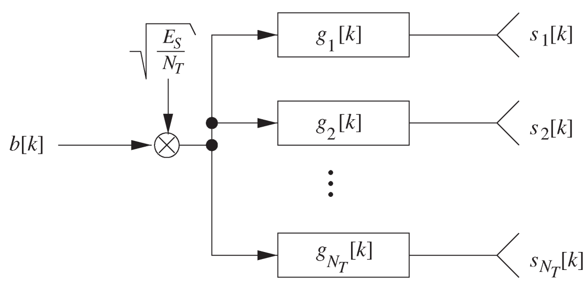


Fig. 1. Block diagram of transmitter for the general delay diversity scheme.

We combine the ideas of [1] and [2] and further generalize DD for frequency-selective channels by employing a finite-impulse response (FIR) filter of length $N$ at each transmit antenna. Based on a Chernoff bound on the pairwise error probability (PEP) for maximum-likelihood sequence estimation (MLSE), a cost function that is suitable for optimization of the FIR filter coefficients is derived. A closed-form optimization does not seem to be feasible, and we devise a steepest descent (SD) type of algorithm for filter search. Since it cannot be guaranteed that the SD algorithm finds the global maximum of the cost function, an upper bound for the maximum of the cost function is also derived. The bound shows that for very high SNRs, the GDD scheme in [1] is optimum in the sense that it maximizes the proposed cost function. However, it is also shown that for the realistic case of low-to-moderate SNRs, the optimized DD (ODD) significantly outperforms the GDD scheme in [1]. The simulations show that the performance advantage of ODD over GDD is preserved if suboptimum equalization strategies such as delayed decision-feedback sequence estimation (DFSE) [16] or decision-feedback equalization (DFE) [17] are used.

This paper is organized as follows. In Section II, DD and the adopted channel model are discussed. In Section III, we derive the cost function for optimization of the FIR transmit filters and propose a corresponding SD algorithm for filter search. In Section IV, an upper bound on the cost function is given and ODD and GDD are compared. Some simulations for GSM/EDGE channels are given in Section V, and Section VI concludes this paper.

## II. Preliminaries

This section discusses the proposed TD scheme and the adopted channel model. It is also explained in detail why DD is an attractive approach for the upgrade of existing systems such as GSM and EDGE. All signals are represented by their complex-baseband equivalents.

### A. DD

A block diagram of the transmitter of a general DD scheme is shown in Fig. 1. The independent identically distributed (i.i.d.) transmit symbols $b[k]$ belong to some scalar symbol alphabet $\mathcal{A}$ such as phase-shift keying (PSK) or quadratic amplitude modulation (QAM). We assume the normalization $\mathcal{E}\{|b[k]|^2\} = 1$ ($\mathcal{E}\{\cdot\}$: statistical expectation). The transmitted signal at antenna $n_T$, $1 \le n_T \le N_T$, in time interval $k$, $k \in \mathbb{Z}$, is given by

$$s_{n_T}[k] = \sqrt{\frac{E_S}{N_T}} \sum_{\nu=0}^{N-1} g_{n_T}[\nu] b[k-\nu] = \sqrt{\frac{E_S}{N_T}} g_{n_T}[k] * b[k] \quad (1)$$

where $*$ refers to discrete-time convolution and $E_S$ is the average transmit energy. The FIR transmit filters have length $N$ and the energy of their coefficients $g_{n_T}[\nu]$, $0 \le \nu \le N-1$, is normalized to $\sum_{n_T=1}^{N_T} \sum_{\nu=0}^{N-1} |g_{n_T}[\nu]|^2 = N_T$. The optimization of these filters will be discussed in Section III. The GDD scheme in [1] is obtained as a special case for $N = (N_T - 1)L + 1$, and $g_{n_T}[\nu] = 1$ for $\nu = (n_T - 1)L$ and $g_{n_T}[\nu] = 0$ for $\nu \ne (n_T - 1)L$, $1 \le n_T \le N_T$. However, in the following, we will use the term GDD for any DD scheme with $g_{n_T}[\nu] = 1$ for $\nu = (n_T - 1)L_0$ and $g_{n_T}[\nu] = 0$ for $\nu \ne (n_T - 1)L_0$, where $0 \le L_0 \le L$, $1 \le n_T \le N_T$. Note that $L_0 < L$ may be desirable in a practical system for complexity reasons (cf. Section II-B).

### B. Channel Model

We assume $N_R$ receive antennas and that the coefficients $h_{n_T n_R}[l]$, $0 \le l \le L-1$, of the discrete-time overall CIR between transmit antenna $n_T$ and receive antenna $n_R$ are mutually correlated zero-mean Gaussian random variables. Furthermore, in this paper, it was assumed that the CIRs $h_{n_T n_R}[l]$ are constant within one data burst but change independently from burst to burst (block fading model). Note that the coefficients of the discrete-time overall CIRs contain the combined effects of transmit pulse shaping, multipath Rayleigh fading, receive filtering, and sampling. In this paper, for all numerical and simulation results presented, the GSM/EDGE system parameters were adopted, i.e., we assume a linearized Gaussian minimum-shift keying (GMSK) transmit pulse [18], the typical urban area (TU) and equalizer test (EQ) power delay profiles [19], and a square-root-raised cosine receive filter with a roll-off factor of 0.3. The CIR coefficients are spatially and temporally correlated, and their $N_T N_R L \times N_T N_R L$ autocorrelation matrix (ACM) is defined as

$$\mathbf{\Phi} \triangleq \mathcal{E}\{\mathbf{h}\mathbf{h}^{\mathrm{H}}\} \quad (2)$$

with $\mathbf{h} \triangleq [h_{11}[0] h_{11}[1], \ldots, h_{11}[L-1] h_{21}[0], \ldots, h_{N_T N_R}[L-1]]^{\mathrm{T}}$, where $[\cdot]^{\mathrm{T}}$ and $[\cdot]^{\mathrm{H}}$ denote transposition and Hermitian transposition, respectively. The spatial correlation can be attributed to insufficient antenna spacing, whereas the temporal correlation is caused by transmit pulse shaping and receive filtering. Although all the derivations are valid for the general case $N_R \ge 1$, for all numerical and simulation results presented in this paper, the practically most important case of one receive antenna was considered. In addition, all CIRs corresponding to different antenna pairs have the same average energy and identical statistical properties. For simplicity, it was assumed that the spatial correlation between two CIRs is identical for all CIR coefficients and defines the spatial correlation parameter $\rho_{\mu\nu} \triangleq \mathcal{E}\{h_{\mu 1}[l] h^*_{\nu 1}[l]\} / \sqrt{\sigma^2_{\mu 1}[l] \sigma^2_{\nu 1}[l]}$,

$0 \leq l \leq L-1$, where $\sigma^2_{n_T n_R}[l] \triangleq \mathcal{E}\{|h_{n_T n_R}[l]|^2\}$ and $(\cdot)^*$ denotes complex conjugation.

The received signal at antenna $n_R$ is given by

$$r_{n_R}[k] = \sqrt{\frac{E_S}{N_T}} \sum_{n_T=1}^{N_T} h_{n_T n_R}[k] * g_{n_T}[k] * b[k] + n_{n_R}[k]$$
$$= \sqrt{\frac{E_S}{N_T}} \sum_{l=0}^{L+N-2} h^{\text{eq}}_{n_R}[l] b[k-l] + n_{n_R}[k] \quad (3)$$

where the equivalent CIR corresponding to the receive antenna $n_R$ is defined as

$$h^{\text{eq}}_{n_R}[k] \triangleq \sum_{n_T=1}^{N_T} h_{n_T n_R}[k] * g_{n_T}[k]. \quad (4)$$

$n_{n_R}[k]$ is the additive white Gaussian noise (AWGN) process at receive antenna $n_R$ and has variance $\sigma_n^2 = N_0$, where $N_0$ denotes the single-sided power spectral density of the underlying continuous-time passband noise process.

From (3), it can be observed that for delay TD, an equivalent single-input multiple-output (SIMO) system with CIRs $h^{\text{eq}}_{n_R}[k]$, $1 \leq n_R \leq N_R$, exists. Therefore, the same channel estimation, channel tracking, and channel equalization techniques as in the single transmit antenna case can be used. However, the equivalent CIRs have length $L_{\text{eq}} = L + N - 1$, and in order to minimize receiver complexity, $N$ should be chosen as small as possible. If $N$ is chosen small enough so that $L_{\text{eq}}$ does not exceed the maximum CIR length that can be tolerated by currently used receivers, existing single transmit antenna wireless systems can be upgraded to DD without changing the receiver at all.

## III. ODD

Based on a Chernoff upper bound on the PEP for optimum MLSE, a cost function that is suitable for optimization of the FIR transmit filter coefficients is first derived. Then, we pose the optimization problem to be solved and show that a closed-form optimization is not possible. Therefore, an SD type of algorithm that can be used for filter search is derived.

### A. Cost Function

Assume the transmission of a block of $K$ symbols $b[k]$, $0 \leq k \leq K-1$. For antenna $n_R$, the vector $\mathbf{r}_{n_R} \triangleq [r_{n_R}[0] r_{n_R}[1], \ldots, r_{n_R}[K+L_{\text{eq}}-1]]^{\mathrm{T}}$ of received signal samples is given by

$$\mathbf{r}_{n_R} = \sqrt{\frac{E_S}{N_T}} \mathbf{B}\mathbf{G}\mathbf{h}_{n_R} + \mathbf{n}_{n_R} \quad (5)$$

where the definitions $\mathbf{h}_{n_R} \triangleq [h_{1n_R}[0] h_{1n_R}[1], \ldots, h_{N_T n_R}[L-1]]^{\mathrm{T}}$ and $\mathbf{n}_{n_R} \triangleq [n_{n_R}[0] n_{n_R}[1], \ldots, n_{n_R}[K+L_{\text{eq}}-1]]^{\mathrm{T}}$ are used. The $(K+L_{\text{eq}}-1) \times L_{\text{eq}}$ data matrix $\mathbf{B}$ contains the vector $[b[k]\ b[k-1], \ldots, b[k-(L_{\text{eq}}-1)]$ in row $k$, $0 \leq k \leq K-1$ ($b[k]=0$ for $k<0$ and $k \geq K$). The $L_{\text{eq}} \times LN_T$ filter matrix $\mathbf{G}$ is defined as

$$\mathbf{G} \triangleq [\mathbf{G}_1\ \mathbf{G}_2, \ldots, \mathbf{G}_{N_T}] \quad (6)$$

$$\mathbf{G}_{n_T} \triangleq \begin{pmatrix} g_{n_T}[0] & 0 & \cdots & 0 \\ g_{n_T}[1] & g_{n_T}[0] & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ g_{n_T}[N-1] & \ddots & \ddots & g_{n_T}[0] \\ 0 & g_{n_T}[N-1] & \ddots & g_{n_T}[1] \\ \vdots & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \vdots & g_{n_T}[N-1] \end{pmatrix}. \quad (7)$$

Now, the received vectors of different antennas are stacked into one vector and yield

$$\mathbf{r} = \sqrt{\frac{E_S}{N_T}} \left(\mathbf{I}_{N_R} \otimes \mathbf{B}\mathbf{G}\right) \mathbf{h} + \mathbf{n} \quad (8)$$

where the definitions $\mathbf{r} \triangleq [\mathbf{r}_1^{\mathrm{T}} \mathbf{r}_2^{\mathrm{T}}, \ldots, \mathbf{r}_{N_R}^{\mathrm{T}}]^{\mathrm{T}}$ and $\mathbf{n} \triangleq [\mathbf{n}_1^{\mathrm{T}} \mathbf{n}_2^{\mathrm{T}}, \ldots, \mathbf{n}_{N_R}^{\mathrm{T}}]^{\mathrm{T}}$ are used. $\mathbf{I}_X$ denotes the $X \times X$ identity matrix and $\otimes$ is the Kronecker product [20].

In the following, it was assumed that a sequence $b_\alpha[\cdot] \in \mathcal{A}^K$ is transmitted but a different sequence $b_\beta[\cdot] \in \mathcal{A}^K$ is detected. The corresponding data matrices are denoted by $\mathbf{B}_\alpha$ and $\mathbf{B}_\beta$ ($\mathbf{B}_\beta \neq \mathbf{B}_\alpha$), respectively. Using (8) and similar techniques as in [3], it can be shown that the PEP for MLSE can be upper bounded as

$$P_e = \Pr\left\{ \left\| \mathbf{r} - \sqrt{\frac{E_S}{N_T}} \left(\mathbf{I}_{N_R} \otimes \mathbf{B}_\alpha \mathbf{G}\right) \mathbf{h} \right\|^2 > \left\| \mathbf{r} - \sqrt{\frac{E_S}{N_T}} \left(\mathbf{I}_{N_R} \otimes \mathbf{B}_\beta \mathbf{G}\right) \mathbf{h} \right\|^2 \right\}$$
$$\leq \prod_{q=1}^{L_{\text{eq}} N_R} \frac{1}{1 + \frac{\lambda_q(\mathbf{Z}) E_S}{(4 N_T N_0)}} \quad (9)$$

where $\Pr\{\cdot\}$ and $\|\cdot\|^2$ denote the probability of the event in brackets and the $L_2$ norm, respectively. In this paper, $\lambda_q(\mathbf{X})$ denotes the eigenvalues of matrix $\mathbf{X}$, which are assumed to be real valued and ordered as $\lambda_1(\mathbf{X}) \geq \lambda_2(\mathbf{X}) \geq \ldots$. In (9), the $(K+L_{\text{eq}}-1)N_R \times (K+L_{\text{eq}}-1)N_R$ matrix $\mathbf{Z}$ is given by

$$\mathbf{Z} \triangleq \left(\mathbf{I}_{N_R} \otimes \mathbf{B}_{\beta|\alpha}\mathbf{G}\right) \mathbf{\Phi} \left(\mathbf{I}_{N_R} \otimes \mathbf{B}_{\beta|\alpha}\mathbf{G}\right)^{\mathrm{H}} \quad (10)$$

where the difference sequence matrix is defined as $\mathbf{B}_{\beta|\alpha} \triangleq \mathbf{B}_\beta - \mathbf{B}_\alpha$. Note that $\mathbf{Z}$ has at most $L_{\text{eq}} N_R$ nonzero eigenvalues. From (9) and (10), it can be observed that the maximum possible diversity order $LN_TN_R$ can be achieved only if the CIR ACM $\mathbf{\Phi}$ has full rank and if the FIR filter length is chosen as proposed in [1], i.e., $N = (N_T - 1)L + 1$. On the other hand, in practice, $N < (N_T - 1)L + 1$ might be desirable for complexity reasons, and it is not clear if the specific choice of $\mathbf{G}$ in [1] minimizes the upper bound on the PEP.

In order to minimize the upper bound in (9), we have to maximize the cost function

$$d(\mathbf{g}, \mathbf{B}_{\beta|\alpha}) \stackrel{\triangle}{=} \prod_{q=1}^{L_{\mathrm{eq}}N_R} \left(1 + \frac{E_S}{4N_TN_0}\lambda_q(\mathbf{Z})\right) \quad (11)$$

where the FIR coefficient vector of length $NN_T$ is defined as $\mathbf{g} \stackrel{\triangle}{=} [g_1[0]g_1[1], \ldots, g_1[N-1]g_2[0], \ldots, g_{N_T}[N-1]]^{\mathrm{T}}$. In order to minimize the maximum PEP, we first have to find the optimum $\mathbf{g}_{\mathrm{opt}}^{\beta|\alpha}$ for all possible difference matrices $\mathbf{B}_{\beta|\alpha}$. The overall optimum vector $\mathbf{g}_{\mathrm{opt}}$ is the one that corresponds to the difference matrix $\mathbf{B}_{\beta|\alpha}$ that yields the minimum $d(\mathbf{g}_{\mathrm{opt}}^{\beta|\alpha}, \mathbf{B}_{\beta|\alpha})$.

The complexity of the optimization problem could be reduced significantly if it can be known in advance which $\mathbf{B}_{\beta|\alpha}$ minimizes $d(\mathbf{g}_{\mathrm{opt}}^{\beta|\alpha}, \mathbf{B}_{\beta|\alpha})$. On the other hand, since it is well known that, in general, for frequency-selective fading channels the performance of MLSE can be tightly approximated by the matched-filter bound [21], it seems to be reasonable to assume that for most channels $d(\mathbf{g}_{\mathrm{opt}}^{\beta|\alpha}, \mathbf{B}_{\beta|\alpha})$ is minimized by sequences that differ only by the minimum Euclidean distance. For those sequences, $b_\alpha[k] = b_\beta[k]$, $0 \le k \le K-1$, $k \ne k_0$, and $|b_\alpha[k_0] - b_\beta[k_0]|^2 = d_{\min}^2$, where $d_{\min}^2$ denotes the squared minimum Euclidean distance of the adopted signal constellation $\mathcal{A}$. With this assumption, (11) can be simplified to

$$\begin{aligned} d(\mathbf{g}) &= \prod_{q=1}^{L_{\mathrm{eq}}N_R} \left(1 + \gamma\lambda_q\left((\mathbf{I}_{N_R} \otimes \mathbf{G})\,\mathbf{\Phi}\,(\mathbf{I}_{N_R} \otimes \mathbf{G})^{\mathrm{H}}\right)\right) \\ &= \det\left(\mathbf{I}_{L_{\mathrm{eq}}N_R} + \gamma\,(\mathbf{I}_{N_R} \otimes \mathbf{G})\,\mathbf{\Phi}\,(\mathbf{I}_{N_R} \otimes \mathbf{G})^{\mathrm{H}}\right) \quad (12)\end{aligned}$$

where the effective SNR $\gamma$ is defined as $\gamma \stackrel{\triangle}{=} d_{\min}^2 E_S/(4N_TN_0)$ and $\det(\cdot)$ denotes the determinant of a matrix. Although the optimization procedure proposed in the following sections can be easily extended to the more general cost function $d(\mathbf{g}, \mathbf{B}_{\beta|\alpha})$, for complexity reasons and practical importance we restrict our attention to the simpler cost function $d(\mathbf{g})$ in (12), which does not depend on the specific signal constellation $\mathcal{A}$ used, but only on the minimum Euclidean distance $d_{\min}$ of that constellation. Note that unlike most other papers on TD design, cf., e.g., [3], [14], and [15], we do not make the high SNR assumption ($\gamma \gg 1$) in (12). It will be shown that ODD filters designed for $\gamma \gg 1$ do not perform well at SNRs of practical interest.

### B. Optimization Problem

The optimum filter coefficients are obtained by maximizing the cost function $d(\mathbf{g})$ subject to the energy constraint $\mathbf{g}^{\mathrm{H}}\mathbf{g} = N_T$. In an attempt to arrive at a closed-form solution, the Lagrange cost function [20] may be defined as

$$J(\mathbf{g}) \stackrel{\triangle}{=} d(\mathbf{g}) + \nu_0 \cdot (\mathbf{g}^{\mathrm{H}}\mathbf{g} - N_T) \quad (13)$$

where $\nu_0$ denotes the Lagrange multiplier. In order to find the optimum coefficient vector $\mathbf{g}_{\mathrm{opt}}$, we have to differentiate $J(\mathbf{g})$ with respect to $\mathbf{g}^*$. The gradient vector of $d(\mathbf{g})$ is given by [23]

$$\frac{\partial d(\mathbf{g})}{\partial \mathbf{g}^*} = \left(\frac{\partial d(\mathbf{g})}{\partial g_1[0]}\ \frac{\partial d(\mathbf{g})}{\partial g_1[1]}, \ldots, \frac{\partial d(\mathbf{g})}{\partial g_{N_T}[N-1]}\right)^{\mathrm{H}} \quad (14)$$

$$\frac{\partial d(\mathbf{g})}{\partial g_{n_T}[\nu]} = \gamma d(\mathbf{g}) \cdot \mathrm{tr}\left(\left(\mathbf{I}_{L_{\mathrm{eq}}N_R} + \gamma\bar{\mathbf{G}}\mathbf{\Phi}\bar{\mathbf{G}}^{\mathrm{H}}\right)^{-1} \bar{\mathbf{E}}_{n_T\nu}\mathbf{\Phi}\bar{\mathbf{G}}^{\mathrm{H}}\right) \quad (15)$$

where $\mathrm{tr}(\cdot)$ denotes the trace of a matrix and the bar over a matrix $\mathbf{X}$ means $\bar{\mathbf{X}} \stackrel{\triangle}{=} \mathbf{I}_{N_R} \otimes \mathbf{X}$. The elements of the $L_{\mathrm{eq}} \times LN_T$ matrix $\mathbf{E}_{n_T\nu}$ are 1 at those positions where matrix $\mathbf{G}$ has entries $g_{n_T}[\nu]$ [cf. (7)] and zero otherwise.

Equations (14) and (15) show that the system of equations that has to be solved for direct maximization of $J(\mathbf{g})$ is highly nonlinear in $\mathbf{g}$. Consequently, it seems to be not possible to find a closed-form solution for the optimum vector $\mathbf{g}_{\mathrm{opt}}$. On the other hand, $J(\mathbf{g})$ could be maximized using the Optimization Toolbox of MATLAB. Alternatively, in the next section, we propose an SD algorithm [22] for recursive maximization of $J(\mathbf{g})$.

### C. SD Algorithm

Since we have a closed-form expression for the gradient vector $\partial d(\mathbf{g})/(\partial \mathbf{g}^*)$, an SD algorithm for maximizing $J(\mathbf{g})$ can be obtained. If the number of iterations is denoted by $i$, $I \in \{0, 1, \ldots\}$, the SD algorithm is given by

$$\mathbf{g}_0[i+1] = \mathbf{g}[i] + \mu_0 \frac{\partial d\left(\mathbf{g}[i]\right)}{\partial \mathbf{g}^*[i]} \quad (16)$$

$$\mathbf{g}[i+1] = \sqrt{\frac{N_T}{\mathbf{g}_0^{\mathrm{H}}[i+1]\mathbf{g}_0[i+1]}}\,\mathbf{g}_0[i+1] \quad (17)$$

where $\mu_0$ is a small positive adaptation constant. The ODD filter coefficients used in Sections IV and V were obtained with $\mu_0 = 0.01$. Equation (17) ensures that the constraint in (13) is fulfilled. Unfortunately, the cost function $J(\mathbf{g})$ has local maxima. Therefore, a careful initialization of the SD algorithm is important for finding the global maximum of $J(\mathbf{g})$ or at least a large local maximum. Since the simple GDD scheme is used as a benchmark for the proposed ODD scheme, for the sake of comparison we will restrict their attention to the case $N = (N_T - 1)L_0 + 1$, $0 \le L_0 \le L$, and initialize $\mathbf{g}[i]$ with the corresponding GDD filter coefficients (cf. Section II-A). Extensive simulations have shown that this is a good choice. Indeed, testing a variety of different channels, we could not find an example where another (randomly chosen) initialization would lead to a coefficient vector $\mathbf{g}$ that yielded a larger value of the respective cost function than the coefficient vector obtained with the GDD initialization. It is also worth mentioning that for all considered test cases convergence was achieved after 100 to 200 iterations. Note that for a given base station the CIR ACM is constant. Therefore, the adaptation of the ODD filter coefficients has to be performed only once and the speed of convergence of the SD algorithm is not important.

## IV. Performance Comparison

This section derives an upper bound for $d(\mathbf{g}_{\text{opt}})$ and subsequently the maximum of the cost function for the filter coefficients found by the proposed SD algorithm is compared with this bound.

### A. Upper Bound on $d(\mathbf{g})$

The main difficulty in finding the maximum of $d(\mathbf{g})$ is the special structure of matrix $\bar{\mathbf{G}}$. Therefore, in the following, unstructured matrices $\bar{\mathbf{G}}$ for maximization of the cost function were considered. The resulting maximum is only an upper bound for the maximum achievable with structured matrices $\bar{\mathbf{G}}$, and the corresponding optimum matrix $\bar{\mathbf{G}}_{\text{opt}}$, in general, does not correspond to realizable FIR transmit filters. Using this approach, it is shown in the Appendix that $d(\mathbf{g})$ can be upper bounded as

$$d(\mathbf{g}_{\text{opt}}) \leq \left(\frac{LN_TN_R}{Q}\gamma\right)^Q \times \left(1+\frac{1}{\gamma LN_TN_R}\sum_{q=1}^{Q}\frac{1}{\lambda_q(\mathbf{\Phi})}\right)^Q \prod_{q=1}^{Q}\lambda_q(\mathbf{\Phi}) \quad (18)$$

where $Q$ is the maximum integer $1 \leq Q \leq L_{\text{eq}}N_R$ that yields $\sigma_Q^2(\bar{\mathbf{G}}_{\text{opt}}) > 0$ with $\sigma_q^2(\bar{\mathbf{G}}_{\text{opt}})$ as defined in (28).

Regarding the upper bound in (18), the following interesting observations were made.

1) If all eigenvalues of the CIR ACM $\mathbf{\Phi}$ are larger than zero, for high SNRs ($\gamma \gg 1$), $Q = L_{\text{eq}}N_R$ is valid [cf. (28)]. Therefore, the right hand side of (18) is approximately proportional to $\gamma^{L_{\text{eq}}N_R}$, which suggests that exploiting the maximum available diversity order $L_{\text{eq}}N_R$ is the optimum strategy for high SNRs.
2) If not all eigenvalues of matrix $\mathbf{\Phi}$ are identical, which will be the case for most practical frequency-selective channels, for small-to-medium SNRs, it can be inferred from (28) that $Q < L_{\text{eq}}N_R$ is optimum. This suggests that under these circumstances FIR transmit filters that do not exploit the maximum available diversity order may be preferable.

Considering the above discussion, the right hand side of (18) suggests that FIR transmit filters that are optimum for high SNRs may be suboptimum for SNRs of practical interest. This is supported by the numerical and simulation results presented in Sections IV-C and V, respectively.

### B. Special Case: $N=(N_T-1)L+1$ and $\gamma\lambda_{LN_TN_R}(\mathbf{\Phi})\gg 1$

If a GDD with $N=(N_T-1)L+1$ was assumed, $\mathbf{G} = \mathbf{I}_{LN_T}$ is obtained from (6). Consequently, (12) yields $d(\mathbf{g}) = \det(\mathbf{I}_{LN_TN_R}+\gamma\mathbf{\Phi})$. Thus, assuming $\lambda_q(\mathbf{\Phi}) > 0$, $1 \leq q \leq LN_TN_R$, and a large enough SNR to assure $\gamma\lambda_{LN_TN_R}(\mathbf{\Phi}) \gg 1$, we obtain

$$\begin{aligned} d(\mathbf{g}) &= \prod_{q=1}^{LN_TN_R}(1+\gamma\lambda_q(\mathbf{\Phi})) \\ &\approx \gamma^{LN_TN_R}\prod_{q=1}^{LN_TN_R}\lambda_q(\mathbf{\Phi}) \\ &= \gamma^{LN_TN_R}\det(\mathbf{\Phi}). \end{aligned} \quad (19)$$

On the other hand, making the same assumptions, from (28) $Q = LN_TN_R$ results. Therefore, we obtain for the upper bound in (18)

$$\begin{aligned} d(\mathbf{g}_{\text{opt}}) &\leq \gamma^{LN_TN_R}\left(1+\frac{1}{\gamma LN_TN_R}\text{tr}(\mathbf{\Phi}^{-1})\right)^{LN_TN_R}\det(\mathbf{\Phi}) \\ &\lesssim \gamma^{LN_TN_R}\det(\mathbf{\Phi}). \end{aligned} \quad (20)$$

Obviously, GDD as proposed in [1] achieves the upper bound for $d(\mathbf{g})$ for high SNRs. Therefore, in the considered special case, GDD is indeed optimum with respect to the optimization criterion adopted in this paper. This result holds for arbitrary (nonsingular) correlation matrices $\mathbf{\Phi}$, i.e., for arbitrary spatial and temporal correlation of the CIR coefficients. Note that the asymptotic optimality of GDD was not proved in [1].

On the other hand, as long as not all eigenvalues of $\mathbf{\Phi}$ are equal, for small-to-moderate SNRs $Q < LN_TN_R$ is valid and GDD is no longer optimum.

### C. Example for $N < (N_T-1)L+1$

This section compares GDD and ODD for $N_R = 1$ receive antenna, $N_T = 2$ transmit antennas with a spatial correlation of $\rho_{12} = 0.5$, and the EQ profile with $L = 7$. The comparison is based on $P_d \triangleq 1/d(\mathbf{g})$ instead of $d(\mathbf{g})$ itself since $P_d$ is related to the PEP for MLSE [cf. (9)–(12)]. The upper bound on $d(\mathbf{g})$ in (18) corresponds to a lower bound on $P_d$.

Fig. 2 shows $P_d$ versus $10\log_{10}(\gamma)$ for GDD and ODD with $N = 3$, respectively. For ODD, the FIR transmit filters were optimized for each $\gamma$ value that is marked by a "$\circ$" using the SD algorithm described in Section III-C. Also shown in Fig. 2 is the lower bound on $P_d$. Obviously, ODD significantly outperforms GDD, e.g., at $P_d = 10^{-3}$ the performance gain of ODD over GDD is 2.2 dB, whereas the remaining gap to the lower bound is 1.5 dB. In order to better understand why ODD achieves large gains for low-to-moderate SNRs, the optimized FIR filter coefficients for different values of $\gamma$ are shown in Table I. Although the ODD FIR filters change with increasing $\gamma$, for $-3$ dB $\leq 10\log_{10}(\gamma) \leq 7$ dB, the filters for both transmit antennas are identical. It is easy to show that this means that unlike GDD the optimized TD scheme has only a

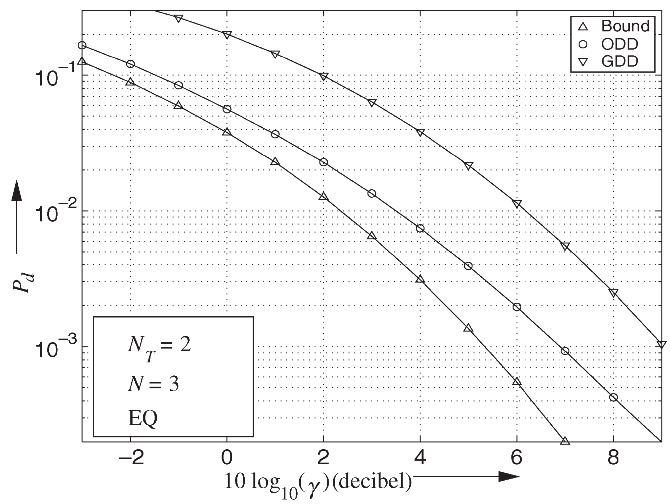


Fig. 2. $P_d$ versus $10\log_{10}(\gamma)$ for EQ profile with $L = 7$ and $\rho_{12} = 0.5$.

TABLE I
ODD FIR TRANSMIT FILTER COEFFICIENTS FOR EQ PROFILE WITH $L = 7$ AND $N_T = 2$ TRANSMIT ANTENNAS WITH $\rho_{12} = 0.5$. $N = 3$ IS VALID

| $10\log_{10}(\gamma)$(decibel) | $n_T$ | $g_{n_T}[0]$ | $g_{n_T}[1]$ | $g_{n_T}[2]$ |
|---|---|---|---|---|
| -3 | 1 | 0.4501 | 0.7712 | 0.4501 |
| | 2 | 0.4501 | 0.7712 | 0.4501 |
| -1 | 1 | 0.4256 | 0.7986 | 0.4256 |
| | 2 | 0.4256 | 0.7986 | 0.4256 |
| 1 | 1 | 0.1713 | 0.7770 | 0.6057 |
| | 2 | 0.1713 | 0.7770 | 0.6057 |
| 3 | 1 | -0.0036 | 0.7053 | 0.7089 |
| | 2 | -0.0036 | 0.7053 | 0.7089 |
| 5 | 1 | -0.1031 | 0.6499 | 0.7530 |
| | 2 | -0.1031 | 0.6499 | 0.7530 |
| 7 | 1 | -0.1704 | 0.6063 | 0.7767 |
| | 2 | -0.1704 | 0.6063 | 0.7767 |
| 9 | 1 | 0.0085 | 0.5525 | 0.8335 |
| | 2 | -0.4999 | 0.3504 | 0.7921 |
| 17 | 1 | 0.8236 | -0.1176 | -0.5548 |
| | 2 | 0.8972 | 0.0409 | 0.4396 |
| 27 | 1 | 0.9503 | -0.1245 | -0.2854 |
| | 2 | 0.6238 | -0.0513 | 0.7799 |
| 37 | 1 | 0.9470 | -0.1566 | -0.2041 |
| | 2 | 0.5960 | -0.0535 | 0.8012 |
| 47 | 1 | 0.9904 | -0.0914 | -0.1037 |
| | 2 | -0.1037 | -0.0914 | 0.9904 |

diversity order of 7 and does not exploit the maximum possible diversity order of $L_{\text{eq}} = 9$. Interestingly, for calculation of the lower bound, we also obtained $Q = 7$ in (18) for that range of $\gamma$. For $10\log_{10}(\gamma) > 7$ dB, the two ODD transmit filters are different and ODD makes use of the full available diversity of the channel. From Table I, it is also observed that for $10\log_{10}(\gamma) = 47$ dB the ODD FIR filter coefficients are practically identical to the GDD coefficients. This example clearly shows that, in contrast to the frequency-nonselective case [3], for frequency-selective channels and SNRs of practical interest, diversity order should not be the primary TD design criterion. The reason for this difference is the inherent frequency diversity of frequency-selective channels.

In Fig. 3(a) and (b), $P_d$ versus $N$ is depicted for $10\log_{10}(\gamma) = 7$ dB and $10\log_{10}(\gamma) = 47$ dB, respectively. For $10\log_{10}(\gamma) = 47$ dB, the performance of GDD is identical to that of ODD for $N \geq 3$ and for $N \geq 5$ GDD achieves the lower bound. On the other hand, for the practically relevant case of $10\log_{10}(\gamma) = 7$ dB, ODD achieves considerable gains over GDD. Note that increasing the FIR filter length $N$ does not necessarily improve the performance of GDD. In fact, for GDD and $10\log_{10}(\gamma) = 7$ dB, $N = 1$ is optimum and outperforms GDD with $N = L + 1 = 8$.[1] Fig. 3(a) also shows that even if transmit filters of maximum length $N = L + 1 = 8$ are used, the GDD scheme in [1] is not optimum for finite SNRs.

## V. SIMULATION RESULTS

This section presents some simulation results for GDD and ODD. The practically most important case of $N_R = 1$ receive antenna was only considered. For equalization, the same schemes as commonly used in the single transmit antenna case were adopted, i.e., we employ MLSE with Viterbi decoding [24], DFSE [16], and DFE [17]. Perfect knowledge of the CIR at the receiver was assumed, and for ODD, perfect knowledge of the CIR ACM at the transmitter was also assumed. The ODD FIR filters are designed for a certain target $E_b/N_0$ ratio, i.e., they are not optimized for every SNR value. The channel model described in Section II-B was used, and for each bit error rate (BER) curve, at least 10 000 CIRs have been randomly generated in accordance with the respective power delay profile. In order to closely approximate the GSM and EDGE system parameters, respectively, we adopt binary PSK (BPSK) and 8-ary PSK (8PSK) modulation in the following.

### A. BPSK Transmission

In Fig. 4, the EQ channel $(L = 7)$ and $N_T = 2$ transmit antennas with $\rho_{12} = 0.5$ were considered. We show the performance of ODD and GDD with $N = 3$, respectively. The ODD FIR filters were optimized for $10\log_{10}(E_b/N_0) = 10$ dB (corresponding to $10\log_{10}(\gamma) = 7$ dB). The same filters were used for all simulation points and their coefficients can be found in Table I. Obviously, ODD outperforms GDD for all considered equalization schemes. For $10\log_{10}(E_b/N_0) \approx 10$ dB, ODD gains approximately 1.4, 1.4, and 0.9 dB for MLSE, DFSE with $S = 4$ states, and DFE, respectively. It is observed that the gain expected for MLSE is also retained for suboptimum low-complexity equalization schemes. In addition, it is interesting to note that ODD also outperforms GDD for $E_b/N_0$ ratios different from the target value of 10 dB. From Fig. 3(b), we know however that GDD is optimum for very high SNRs. This shows again that a TD design based on the assumption $\gamma \to \infty$ is not advisable for frequency-selective channels. For comparison, in Fig. 4, the performance of GDD with $N = L + 1 = 8$

[1]Note that $N = L + 1 = 8$ is the value proposed in [1].

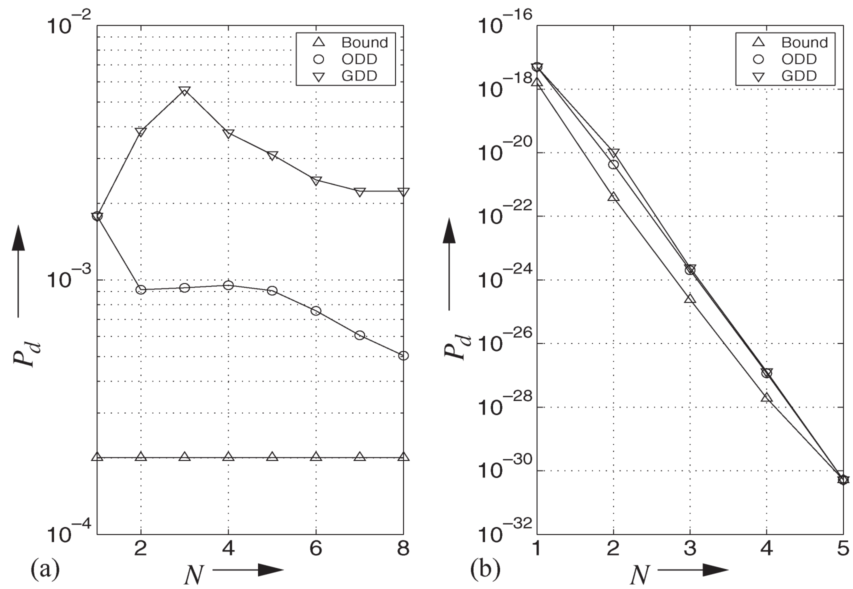


Fig. 3. $P_d$ versus $N$ for (a) $10\log_{10}(\gamma) = 7$ dB and (b) $10\log_{10}(\gamma) = 47$ dB. EQ profile with $L = 7$, $N_T = 2$, and $\rho_{12} = 0.5$.

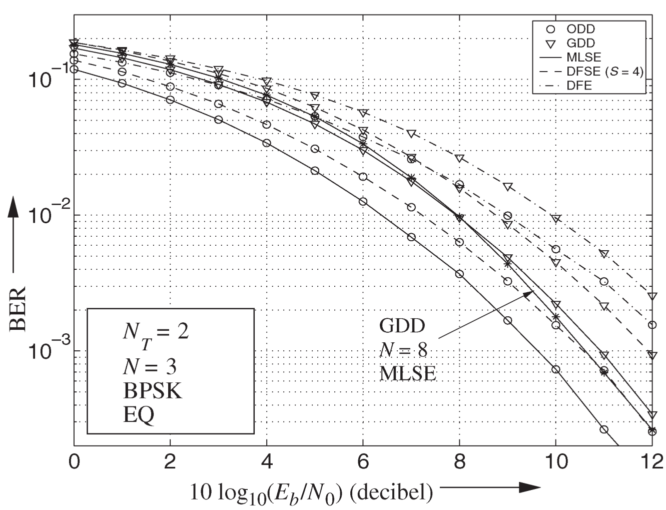


Fig. 4. BER of BPSK versus $10\log_{10}(E_b/N_0)$ for ODD (optimized for $10\log_{10}(E_b/N_0) = 10$ dB) and GDD (both with $N = 3$). $\rho_{12} = 0.5$.

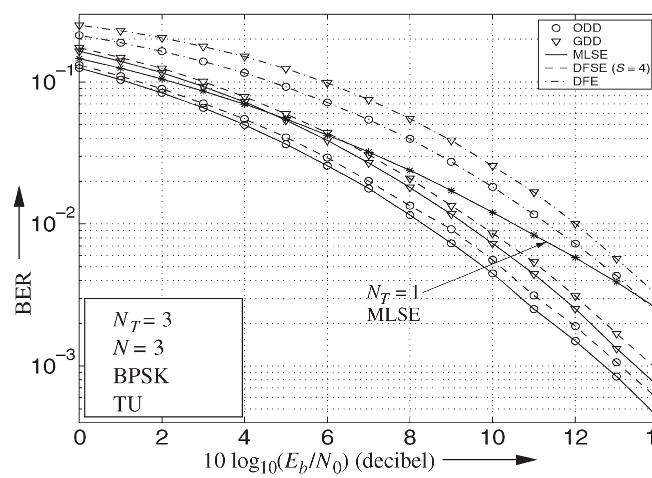


Fig. 5. BER of BPSK versus $10\log_{10}(E_b/N_0)$ for ODD (optimized for $10\log_{10}(E_b/N_0) = 10$ dB) and GDD (both with $N = 3$). $\rho_{12} = \rho_{23} = 0.5$ and $\rho_{13} = 0.2$.

as proposed in [1] is also depicted. For $10\log_{10}(E_b/N_0) > 8$ dB, GDD with $N = 8$ outperforms GDD with $N = 3$ but has a worse performance than ODD with $N = 3$. Note that the Viterbi algorithm for $N = 8$ requires $S = 2^{13} = 8192$ states, whereas for $N = 3$ only $S = 2^8 = 256$ states are necessary.

In Fig. 5, the TU channel $(L = 5)$ and $N_T = 3$ transmit antennas with spatial correlations $\rho_{12} = \rho_{23} = 0.5$ and $\rho_{13} = 0.2$ are assumed. Again, the ODD FIR filters optimized for $10\log_{10}(E_b/N_0) = 10$ dB are used for all considered $E_b/N_0$ ratios. Similar observations as in Fig. 4 can be made. In particular, for $E_b/N_0$ ratios around 10 dB, the performance gain of ODD over GDD is approximately 0.95, 0.8, and 0.8 dB for MLSE, DFSE with $S = 4$ states, and DFE, respectively. For comparison, the performance of single-antenna transmission $(N_T = 1)$ with MLSE is also shown. Both GDD and ODD yield large performance gains over the single-antenna scheme, especially at high SNRs.

### B. 8PSK Transmission

Finally, in Fig. 6, we consider the performance of 8PSK for the EQ channel $(L = 7)$ and $N_T = 3$ transmit antennas with $\rho_{12} = \rho_{13} = \rho_{23} = 0.5$. Again, $N = 3$ is adopted and the ODD FIR filters were optimized for $10\log_{10}(E_b/N_0) = 16$ dB. Since MLSE would require $S = 8^8 = 16\,777\,216$ states, we only show the results for suboptimum equalization schemes. For $E_b/N_0$ ratios around 16 dB, the performance gain of ODD over GDD is approximately 2.5 and 1.9 dB for DFSE with $S = 64$ states and DFE, respectively. For comparison, the performance for a single transmit antenna and DFSE with $S = 64$ states is also depicted in Fig. 6. Interestingly, in the considered SNR range, the scheme with $N_T = 1$ has a better performance than GDD with $N_T = 3$ transmit antennas. In addition, for 10 dB $\leq 10\log_{10}(E_b/N_0) \leq 17$ dB, the BER curves for $N_T = 1$ and GDD and ODD with $N_T = 3$,

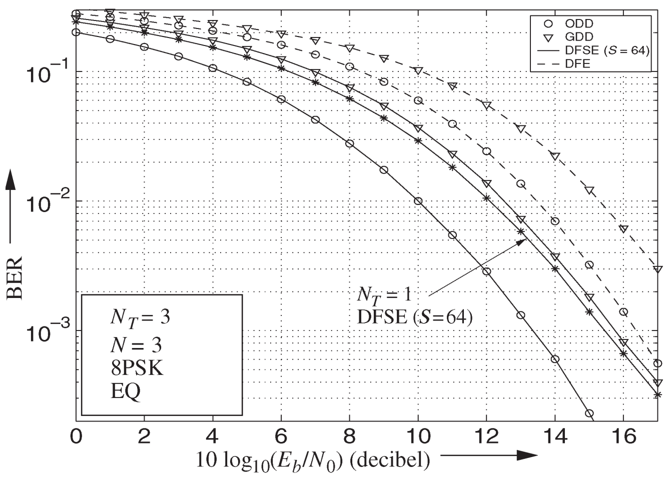


Fig. 6. BER of 8PSK versus $10\log_{10}(E_b/N_0)$ for ODD (optimized for $10\log_{10}(E_b/N_0) = 16$ dB) and GDD (both with $N = 3$). $\rho_{12} = \rho_{13} = \rho_{23} = 0.5$.

respectively, are practically parallel. This clearly shows that for this example and SNRs of practical interest, diversity order is not important.

Further simulations not presented here have shown that the performance gain achieved by ODD is usually large for strong spatial correlations and large inherent frequency diversity of the channel.

## VI. Conclusion

This paper has shown that DD is an attractive TD technique since the same receivers as for single-antenna transmission can be adopted, and therefore, existing systems such as GSM and Enhanced Data rates for EDGE can be easily upgraded. In particular, based on the PEP of MLSE, a cost function for optimization of the DD FIR transmit filters has been derived. Since a closed-form maximization of this cost function seems to be not tractable, an SD algorithm for filter optimization has been provided. Based on an upper bound on the cost function, it has been proved that the GDD scheme in [1] is optimum for very high SNRs and FIR filters of maximum length. However, for SNRs of practical interest and reasonable FIR filter lengths, the proposed ODD scheme yields significant performance gains over GDD. Fortunately, these performance gains cannot only be observed for optimum MLSE but are also preserved for good suboptimum equalization schemes such as DFSE and DFE.

Finally, the investigations have shown that for TD design for frequency-selective channels and relevant SNR values, the diversity order is not the most important parameter. Therefore, unlike the frequency-nonselective case [3], for frequency-selective channels, TD schemes optimized for high SNRs may perform poorly for low-to-moderate SNRs.

## Appendix
## Upper Bound on $d(\mathbf{g})$

This appendix maximizes the cost function

$$d_B(\bar{\mathbf{G}}) = \det\left(\mathbf{I}_{L_{\mathrm{eq}}N_R} + \gamma\bar{\mathbf{G}}\mathbf{\Phi}\bar{\mathbf{G}}^{\mathrm{H}}\right) \tag{21}$$

for general (unstructured) $L_{\mathrm{eq}}N_R \times LN_TN_R$ matrices $\bar{\mathbf{G}}$ subject to the energy constraint $\mathrm{tr}\{\bar{\mathbf{G}}\bar{\mathbf{G}}^{\mathrm{H}}\} = LN_TN_R$. The maximum value of $d_B(\bar{\mathbf{G}})$ will be an upper bound for the maximum of the cost function $d(\mathbf{g}_{\mathrm{opt}})$.

For the following, we introduce the eigenvalue decomposition of matrix $\mathbf{\Phi}$ [20]:

$$\mathbf{\Phi} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{\mathrm{H}} \tag{22}$$

where $\mathbf{Q}$ is an $LN_TN_R \times LN_TN_R$ unitary matrix and $\mathbf{\Lambda}$ is a diagonal matrix, whose $q$th main diagonal element is the eigenvalue $\lambda_q(\mathbf{\Phi})$, $\lambda_1(\mathbf{\Phi}) \geq \cdots \geq \lambda_{LN_TN_R}(\mathbf{\Phi}) \geq 0$, of matrix $\mathbf{\Phi}$. Furthermore, we introduce the singular value decomposition of matrix $\bar{\mathbf{G}}$ [20]

$$\bar{\mathbf{G}} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathrm{H}} \tag{23}$$

where $\mathbf{U}$ and $\mathbf{V}$ denote $L_{\mathrm{eq}}N_R \times L_{\mathrm{eq}}N_R$ and $LN_TN_R \times LN_TN_R$ unitary matrices, respectively. $\mathbf{\Sigma}$ is an $L_{\mathrm{eq}}N_R \times LN_TN_R$ matrix whose only nonzero elements are $[\mathbf{\Sigma}]_{qq} = \sigma_q(\bar{\mathbf{G}})$, $1 \leq q \leq L_{\mathrm{eq}}N_R$, where $\sigma_q(\bar{\mathbf{G}})$, $\sigma_1(\bar{\mathbf{G}}) \geq \sigma_2(\bar{\mathbf{G}}) \geq \cdots \geq \sigma_{L_{\mathrm{eq}}N_R}(\bar{\mathbf{G}}) \geq 0$, is the $q$th singular value of matrix $\bar{\mathbf{G}}$ ($[\mathbf{X}]_{\nu\mu}$ denotes the element of matrix $\mathbf{X}$ in row $\nu$ and column $\mu$). Using (22) and (23) in (21), we obtain

$$d_B(\bar{\mathbf{G}}) = \det\left(\mathbf{I}_{L_{\mathrm{eq}}N_R} + \gamma\mathbf{\Sigma}\mathbf{M}^{\mathrm{H}}\mathbf{\Lambda}\mathbf{M}\mathbf{\Sigma}^{\mathrm{H}}\right) \tag{24}$$

where $\mathbf{M} \stackrel{\triangle}{=} \mathbf{Q}^{\mathrm{H}}\mathbf{V}$ is also unitary. For the constraint, we get $\mathrm{tr}\{\bar{\mathbf{G}}\bar{\mathbf{G}}^{\mathrm{H}}\} = \sum_{q=1}^{L_{\mathrm{eq}}N_R}\sigma_q^2(\bar{\mathbf{G}}) = LN_TN_R$. Obviously, both the cost function $d_B(\bar{\mathbf{G}})$ and the constraint are independent of matrix $\mathbf{U}$. Therefore, for maximization of $d_B(\bar{\mathbf{G}})$, we can restrict their attention to the optimization of $\mathbf{\Sigma}$ and $\mathbf{V}$.

Using Hadamard's inequality [25, Th. 7.8.1], we get from (24)

$$d_B(\bar{\mathbf{G}}) \leq \prod_{q=1}^{L_{\mathrm{eq}}N_R} f_q \tag{25}$$

with

$$f_q \stackrel{\triangle}{=} \left[\mathbf{I}_{L_{\mathrm{eq}}N_R} + \gamma\mathbf{\Sigma}\mathbf{M}^{\mathrm{H}}\mathbf{\Lambda}\mathbf{M}\mathbf{\Sigma}^{\mathrm{H}}\right]_{qq} = 1 + \gamma\sigma_q^2(\bar{\mathbf{G}})\mathbf{m}_q^{\mathrm{H}}\mathbf{\Lambda}\mathbf{m}_q \tag{26}$$

where $\mathbf{m}_q$ denotes the $q$th column of matrix $\mathbf{M}$. Taking into account that $\mathbf{M}$ is unitary, it can be easily shown that for any set of singular values $\sigma_q(\bar{\mathbf{G}})$, $1 \leq q \leq L_{\mathrm{eq}}N_R$, the product $\prod_{q=1}^{L_{\mathrm{eq}}N_R} f_q$ is maximized for $\mathbf{m}_q = \mathbf{e}_q$, $1 \leq q \leq L_{\mathrm{eq}}N_R$, where $\mathbf{e}_q$ is a column vector of length $LN_TN_R$ with a 1 in position $q$ and zeros in all other positions. Therefore, $f_q$ simplifies to $f_q = 1 + \gamma\sigma_q^2(\bar{\mathbf{G}})\lambda_q(\mathbf{\Phi})$, $1 \leq q \leq L_{\mathrm{eq}}N_R$. On the other hand, for this particular choice of $\mathbf{m}_q$, equality holds in (25) and we obtain

$$d_B(\bar{\mathbf{G}}) = \prod_{q=1}^{L_{\mathrm{eq}}N_R}\left(1 + \gamma\sigma_q^2(\bar{\mathbf{G}})\lambda_q(\mathbf{\Phi})\right). \tag{27}$$

Now, we can restate their optimization problem as follows. Optimize the singular values $\sigma_q^2(\bar{\mathbf{G}})$ for maximization of $d_B(\bar{\mathbf{G}})$ subject to the constraints $\sum_{q=1}^{L_{\rm eq}N_R} \sigma_q^2(\bar{\mathbf{G}}) = LN_TN_R$ and $\sigma_q^2(\bar{\mathbf{G}}) \geq 0,\ 1 \leq q \leq L_{\rm eq}N_R$. The additional inequality constraint is necessary to prevent negative $\sigma_q^2(\bar{\mathbf{G}})$, which do not constitute admissible solutions since singular values are always nonnegative [20]. The solution to this optimization problem can be obtained using the Kuhn–Tucker method, cf., e.g., [20, Sec. 18.9], and is given by

$$\sigma_q^2(\bar{\mathbf{G}}_{\rm opt}) = \max\left\{\frac{1}{Q}\left(LN_TN_R + \frac{1}{\gamma}\sum_{i=1}^{Q}\frac{1}{\lambda_i(\mathbf{\Phi})}\right) - \frac{1}{\gamma\lambda_q(\mathbf{\Phi})}, 0\right\}, \qquad 1 \leq q \leq L_{\rm eq}N_R \quad (28)$$

where $Q$ is the maximum integer $1 \leq Q \leq L_{\rm eq}N_R$ that yields $\sigma_Q^2(\bar{\mathbf{G}}_{\rm opt}) > 0$. The resulting maximum $d_B(\bar{\mathbf{G}}_{\rm opt})$ is given by

$$d_B(\bar{\mathbf{G}}_{\rm opt}) = \left(\frac{LN_TN_R}{Q}\gamma\right)^Q \times \left(1 + \frac{1}{\gamma LN_TN_R}\sum_{q=1}^{Q}\frac{1}{\lambda_q(\mathbf{\Phi})}\right)^Q \prod_{q=1}^{Q}\lambda_q(\mathbf{\Phi}) \quad (29)$$

that is an upper bound for $d(\mathbf{g}_{\rm opt})$.

## REFERENCES


[1] D. Gore, S. Sandhu, and A. Paulraj, "Delay diversity codes for frequency selective channels," in *Proc. IEEE Int. Conf. Communications (ICC)*, New York, May 2002, pp. 1949–1953.

[2] A. Wittneben, "A new bandwidth efficient transmit antenna modulation diversity scheme for linear digital modulation," in *Proc. Int. Conf. Communications (ICC)*, Geneva, Switzerland, May 1993, pp. 1630–1634.

[3] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space–time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[4] S. M. Alamouti, "A simple transmitter diversity scheme for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[5] Y. Li, J. C. Chuang, and N. R. Sollenberger, "Transmit diversity for OFDM systems and its impact on high-rate data wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 7, pp. 1233–1243, Jul. 1999.

[6] H. Bölcskei and A. J. Paulraj, "Space-frequency coded broadband OFDM systems," in *Proc. IEEE Wireless Communications and Networking Conf. (WCNC)*, Chicago, IL, Sep. 2000, pp. 1–6.

[7] E. Lindskog and A. Paulraj, "A transmit diversity scheme for channels with intersymbol interference," in *Proc. IEEE Int. Conf. Communications (ICC)*, New Orleans, LA, 2000, pp. 307–311.

[8] N. Al-Dhahir, "Single-carrier frequency-domain equalization for space–time block-coded transmission over frequency-selective fading channels," *IEEE Commun. Lett.*, vol. 5, no. 7, pp. 304–306, Jul. 2001.

[9] S. Zhou and G. B. Giannakis, "Single-carrier space–time block-coded transmission over frequency-selective fading channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 1, pp. 164–179, Jan. 2003.

[10] W.-J. Choi and J. M. Cioffi, "Space–time block codes over frequency selective Rayleigh fading channels," in *Proc. IEEE Vehicular Technology Conf.*, Houston, TX, Nov. 1999, pp. 2541–2545.

[11] W. H. Gerstacker, F. Obernosterer, R. Schober, A. Lehmann, A. Lampe, and P. Gunreben, "Symbol-by-symbol and trellis-based equalization with widely linear processing for space–time block-coded transmission over frequency selective channels," in *Proc. IEEE Global Telecommunications Conf. (GLOBECOM)*, Taipei, Taiwan, R.O.C., Dec. 2002, pp. 365–369.

[12] A. F. Naguib, "Equalization of transmit diversity space–time coded signals," in *Proc. IEEE Global Telecommunications Conf. (GLOBECOM)*, San Francisco, CA, Dec. 2000, pp. 1077–1082.

[13] N. Al-Dhahir, "Overview and comparison of equalization schemes for space–time-coded signals with application to EDGE," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2477–2488, Oct. 2002.

[14] Y. Liu, M. P. Fitz, and O. Y. Takeshita, "Space–time codes performance criteria and design for frequency selective fading channels," in *Proc. IEEE Int. Conf. Communications (ICC)*, Helsinki, Finland, Jun. 2001, pp. 2800–2804.

[15] R. Schober, W. H. Gerstacker, and L. H.-J. Lampe, "Performance analysis and design of STBCs for fading ISI channels," in *Proc. Int. Conf. Communications (ICC)*, New York, May 2002, pp. 1451–1455.

[16] A. Duel-Hallen and A. C. Heegard, "Delayed decision-feedback sequence estimation," *IEEE Trans. Commun.*, vol. 37, no. 5, pp. 428–436, May 1989.

[17] C. A. Belfiore and J. H. Park, "Decision-feedback equalization," *Proc. IEEE*, vol. 67, no. 8, pp. 1143–1156, Aug. 1979.

[18] ETSI, *Tdoc SMG2 WPB 325/98*, Nov. 1998.

[19] GSM Recommendation 05.05: "Propagation Conditions," Vers. 5.3.0, Release 1996.

[20] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*. New York: Prentice-Hall, 2000.

[21] P. Balaban and J. Salz, "Optimum diversity combining and equalization in digital data transmission with applications to cellular mobile radio—Part I and II," *IEEE Trans. Commun.*, vol. 40, no. 5, pp. 885–907, May 1992.

[22] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 1996.

[23] H. L. Van Trees, *Optimum Array Processing*. New York: Wiley, 2002.

[24] G. D. Forney, Jr., "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inf. Theory*, vol. 18, no. 3, pp. 363–378, May 1972.

[25] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1999.


**Thorsten Hehn** (S'04) was born in 1979. He received the Diplom (University) degree in electrical engineering from the University of Erlangen-Nürnberg, Erlangen, Germany, in 2004.

In 2002–2003, he was a Visiting Student at the University of British Columbia, Vancouver, BC, Canada. He is currently a Research Assistant in the Telecommunications Laboratory at the University of Erlangen-Nürnberg. His research interests include space–time processing, multiple-input multiple-output (MIMO) systems, and iterative equalization.

**Robert Schober** (S'98–M'01) was born in Neuendettelsau, Germany, in 1971. He received the Diplom (University) and Ph.D. degrees in electrical engineering from the University of Erlangen-Nürnberg, Erlangen, Germany, in 1997 and 2000, respectively.

From May 2001 to April 2002, he was a Postdoctoral Fellow at the University of Toronto, Canada, sponsored by the German Academic Exchange Service (DAAD). Since May 2002, he has been an Assistant Professor and Canada Research Chair (Tier II) in Wireless Communication at the University of British Columbia (UBC), Vancouver, BC, Canada. His research interests include noncoherent detection, equalization, multiuser detection, multiple-input multiple-output (MIMO) systems, space–time processing and coding, and ultra-wideband communication.

Dr. Schober was a co-recipient of the Best Paper Award of the German Information Technology Society (ITG) in 2001. In 2002, he received the Heinz Maier–Leibnitz Award of the German Science Foundation (DFG). He was also a co-recipient of the 2004 Innovations Award of the Vodafone Foundation for Research in Mobile Communications. He currently serves as an Editor for Detection, Equalization, and MIMO for the IEEE TRANSACTIONS ON COMMUNICATIONS and as a Guest Editor for a Special Issue of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS on "Differential and Noncoherent Wireless Communications."

**Wolfgang H. Gerstacker** (S'93–M'99) was born in Nürnberg, Germany, in 1966. He received the Dipl.-Ing. degree in electrical engineering, the Dr.-Ing. degree for a thesis on equalization concepts for fast digital transmission over twisted pair lines, and the Habilitation degree for work on digital transmission concepts for the EDGE mobile communications system from the University of Erlangen-Nürnberg, Erlangen, Germany, in 1991, 1998, and 2004, respectively.

From 1992 to 1998, he was a Research Assistant at the Telecommunications Institute of the University of Erlangen-Nürnberg. From 1998 to 2002, he was a Consultant for mobile communications and an External Lecturer in digital communications at the University of Erlangen-Nürnberg. From 1999 to 2000, he was a Postdoctoral Research Fellow for a six-month period at the University of Canterbury, Christchurch, New Zealand, sponsored by a fellowship from the German Academic Exchange Service (DAAD). Since 2002, he has been a Senior Researcher and Lecturer at the Institute for Mobile Communications of the University of Erlangen-Nürnberg. His current research interests include wireless communications, detection, equalization and parameter estimation, blind techniques, space–time processing, multiple-input multiple-output (MIMO) systems, enhanced data rates for global evolution (EDGE), and wireless local area networks (WLANs).

Dr. Gerstacker was a co-recipient of the Research Award of the German Society for Information Technology (ITG) in 2001. For work on single-antenna interference cancellation for GSM, he was a co-recipient of the Electrical and Electronic Engineering for Communications (EEEfCOM) Innovation Award 2003 and of the Vodafone Innovation Award 2004. He is a Member of the Editorial Board of the European Association for Signal Processing (EURASIP) *Journal on Wireless Communications and Networking*.