Multiple–Symbol Differential Sphere Decoding for Unitary Space–Time Modulation

Volker Pauli
Lehrstuhl für Informationsübertragung
Universität Erlangen–Nürnberg, Germany
Email: pauli@lnt.de

Lutz Lampe
Department of Electrical & Computer Engineering
University of British Columbia, Canada
Email: lampe@ece.ubc.ca

***Abstract*— We consider multiple–symbol differential detection (MSDD) for multiple–input multiple–output (MIMO) Rayleigh–fading channels. MSDD, which jointly processes blocks of N received symbols to detect $N-1$ data symbols, allows for power–efficient transmission over rapid–fading channels. However, the complexity of the straightforward approach to find the maximum–likelihood (ML) MSDD solution is exponential in N, the number of transmit antennas N_{T} and the rate R. In this paper, we introduce an MSDD algorithm based on sphere decoding whose average complexity is not exponential in N for interesting ranges of signal–to–noise ratio (SNR) and aribtrary unitary signal constellations. For the interesting special cases of diagonal and orthogonal constellations we achieve a similar complexity reduction in N_{T} and R. Based on an error–rate analysis for MSDD we also propose a variant of MSDD that considerably improves power efficiency in relatively fast fading at a very moderate increase in complexity.**

I. Introduction

Transmission over multiple–input multiple–output (MIMO) fading channels allows for much higher data rates and/or power efficiency than traditional single–input single–output (SISO) communication. Most often, coherent detection based on the assumption of perfect channel state information (CSI) at the receiver is considered for both SISO and MIMO channels. However, the acquisition of CSI for time–varying MIMO channels is difficult and usually requires a considerable overhead of pilot symbols to be transmitted. Thus, the approach of differential space–time modulation (DSTM) in conjunction with noncoherent detection, which does not require CSI, appears attractive, cf. e.g. [1], [2], [3].

Detection of DSTM is usually based on two consecutively received symbols, which is referred to as conventional differential detection (CDD) in the following. DSTM with CDD achieves a power efficiency within $3\,\mathrm{dB}$ of coherent detection in relatively slow–fading environments, cf. e.g. [1], [2], [3]. However, CDD causes an error floor in fast–fading channels, where more sophisticated detection techniques are required for reliable communication. Such detectors perform multiple–symbol differential detection (MSDD) of $(N-1)$ data symbols based on N consecutively received symbols, cf. e.g. [4], [5]. Optimum maximum–likelihood MSDD (ML MSDD) approaches the performance of coherent detection when increasing the so–called observation window size N, but its complexity grows exponentially with N. To benefit from larger N without exponential complexity increase, suboptimum detectors based on decision–feedback differential detection (DFDD) (e.g. [5], [6]) and noncoherent sequence detection (NSD) (e.g. [7]) have been proposed for DSTM.

Recently, the authors have devised a new optimum ML MSDD algorithm for single–antenna differential phase–shift keying (DPSK) in [8]. This algorithm, which is based on sphere decoding (cf. e.g. [9]) and hence referred to as multiple–symbol differential sphere decoding (MSDSD)[1], accomplishes true ML MSDD while its complexity is comparable to that of DFDD in most cases. In this paper, we exploit the close relationship between DPSK modulation and DSTM with unitary constellations and extend MSDSD to unitary DSTM. In this context, we develop different variants of MSDSD, which are tailored to specific DSTM constellations in order to avoid detection complexity being exponential in the number of transmit antennas and data rate. Based on an error–rate analysis for MSDD we also propose "subset MSDSD", which further improves the performance of MSDSD while incurring only a marginal increase in complexity. Simulation results for DSTM transmission show that (subset) MSDSD significantly outperforms DFDD in power efficiency for comparable decoding complexity.

The remainder of this paper is organized as follows. Section II briefly introduces the system model. In Section III, MSDSD is derived and optimized for different DSTM constellations. The error–rate analysis of MSDSD and the formulation of subset MSDSD are given in Section IV. Performance and complexity results are presented and discussed in Section V. Section VI concludes this paper.

Notation: Vectors and matrices are printed in bold lower– and uppercase letters, respectively. $(\cdot)^{\mathsf{H}}$, $(\cdot)^{\mathsf{T}}$, $\mathrm{tr}\{\cdot\}$, $||\cdot||$ and $\det\{\cdot\}$ denote Hermitian transpose, transpose, trace, Frobenius norm and determinant of a matrix (or vector, where applicable), respectively. $\otimes$ denotes the Kronecker product of two matrices or vectors. $\mathrm{j} \triangleq \sqrt{-1}$ and $\mathcal{E}\{\cdot\}$ denote the imaginary unit and the expectation operator.

II. System Model

We consider a transmission scheme using N_{T} transmit and N_{R} receive antennas. At the transmitter $N_{\mathrm{T}}R$ bits are mapped

[1] In the context of this paper, the terms "decoding" and "detection" refer to the same procedure and are used interchangeably.

to $N_\mathrm{T} \times N_\mathrm{T}$ dimensional unitary matrices $\boldsymbol{V}[k]$ which are taken from a set $\boldsymbol{\mathcal{V}} \triangleq \{\boldsymbol{V}^{(l)} \mid l \in \{1,\ldots,L\}, L \triangleq 2^{N_\mathrm{T}R}\}$. In order to facilitate noncoherent detection the data symbols $\boldsymbol{V}[k]$ are differentially encoded yielding transmit symbols

$$\boldsymbol{S}[k] = \boldsymbol{V}[k]\boldsymbol{S}[k-1], \quad \boldsymbol{S}[0] = \boldsymbol{I}_{N_\mathrm{T}}. \tag{1}$$

At time $\kappa = kN_\mathrm{T} + i$ the transmitter radiates from antenna j the element $s_{i,j}[k]$ in the ith row and jth column of $\boldsymbol{S}[k]$. The transmit power is independent of N_T at all times, i.e., $\sum_{j=1}^{N_\mathrm{T}} |s_{i,j}[k]|^2 = 1$, $1 \le i \le N_\mathrm{T}$ holds.

We assume a frequency–nonselective MIMO Rayleigh fading channel. Consequently, in the equivalent complex baseband domain the signal $r_{i,j}[k]$ received by antenna j at time $\kappa = kN_\mathrm{T} + i$ is given by

$$r_{i,j}[k] = \sum_{\nu=1}^{N_\mathrm{T}} s_{i,\nu}[k] h_{\nu,j}[\kappa] + n_j[\kappa]. \tag{2}$$

$h_{\nu,j}[\kappa]$ and $n_j[\kappa]$ denote the complex fading gain between transmit antenna ν and receive antenna j and the zero–mean complex additive spatially and temporally white Gaussian noise (AWGN) with variance σ_n^2 effective at the jth receive antenna at time κ, respectively. The fading channels are assumed spatially uncorrelated with identical temporal statistical properties, according to the widely–used model by Clarke, i.e., $\psi_{hh}[\kappa] \triangleq \mathcal{E}\{h_{i,j}[k+\kappa]h_{i,j}^*[k]\} = \mathrm{J}_0(2\pi B_\mathrm{f} T\kappa)$, where $\mathrm{J}_0(\cdot)$ and $B_\mathrm{f}T$ denote the zeroth–order Bessel function of the first kind and the maximum normalized fading bandwidth, respectively.

When deriving the MSDSD receiver structure, we make the additional assumption that the fading is quasistatic (QS) over N_T consecutive modulation intervals. Under this assumption the received signal corresponding to the transmission of a symbol $\boldsymbol{S}[k]$ can be expressed as an $N_\mathrm{T} \times N_\mathrm{R}$ dimensional matrix

$$\boldsymbol{R}[k] = \boldsymbol{S}[k]\boldsymbol{H}[k] + \boldsymbol{N}[k], \tag{3}$$

where $\boldsymbol{R}[k]$, $\boldsymbol{H}[k]$ and $\boldsymbol{N}[k]$ each contain $r_{i,j}[k]$, $h_{i,j}[kN_\mathrm{T}]$ and $n_j[kN_\mathrm{T} + i]$ in the ith row and jth column.

III. Multiple–Symbol Differential Sphere Decoding

To derive the MSDSD algorithm, we first provide a suitable representation of the ML MSDD decision rule in Section III-A. The general MSDSD algorithm is presented in Section III-B and optimized for different DSTM constellations in Section III-C.

A. ML Multiple–Symbol Differential Detection

ML MSDD processes blocks $\bar{\boldsymbol{R}}[k] \triangleq [\boldsymbol{R}^\mathsf{T}[k-N+1],\ldots,\boldsymbol{R}^\mathsf{T}[k]]^\mathsf{T}$ of N consecutively received matrix symbols to find ML estimates $\hat{\bar{S}}[k]$ for corresponding blocks $\bar{\boldsymbol{S}}[k] \triangleq [\boldsymbol{S}^\mathsf{T}[k-N+1],\ldots,\boldsymbol{S}^\mathsf{T}[k]]^\mathsf{T}$ of N transmit symbols or equivalently estimates $\hat{\bar{\boldsymbol{V}}}[k]$ for the $N-1$ data symbols $\bar{\boldsymbol{V}}[k] \triangleq [\boldsymbol{V}^\mathsf{T}[k-N+2],\ldots,\boldsymbol{V}^\mathsf{T}[k]]^\mathsf{T}$. In analogy to the single–antenna case (cf. e.g. [4]) consecutive blocks $\bar{\boldsymbol{R}}[k]$ have to overlap by (at least) one matrix symbol, i.e., the observation window of length N moves forward by (at most) $N-1$ symbols at a time. With an $NN_\mathrm{T} \times NN_\mathrm{T}$ block–diagonal matrix $\bar{\boldsymbol{S}}_\mathrm{D}[k] \triangleq \mathrm{diag}\{\boldsymbol{S}[k-N+1],\ldots,\boldsymbol{S}[k]\}$ and $NN_\mathrm{T} \times N_\mathrm{R}$ matrices $\bar{\boldsymbol{H}}[k]$ and $\bar{\boldsymbol{N}}[k]$ defined in the same way as $\bar{\boldsymbol{R}}[k]$ we can write

$$\bar{\boldsymbol{R}}[k] = \bar{\boldsymbol{S}}_\mathrm{D}[k]\bar{\boldsymbol{H}}[k] + \bar{\boldsymbol{N}}[k]. \tag{4}$$

For sake of readability we will in the sequel omit the reference $[k]$ to time and address submatrices of the above blockmatrices via subscripts with $\boldsymbol{X}_i$ being the ith submatrix of $\bar{\boldsymbol{X}}$ and $1 \le i \le (N$ or $N-1)$.

With $\bar{\boldsymbol{H}}$ and $\bar{\boldsymbol{N}}$ being matrices of zero–mean complex Gaussian random variables and $\bar{\boldsymbol{S}}_\mathrm{D}$ being an unitary matrix the ML MSDD decision rule can be derived in complete analogy to the case of single–antenna DPSK transmission (e.g. [10]). After straightforward manipulations we obtain [11]

$$\hat{\bar{\boldsymbol{S}}} = \underset{\bar{\boldsymbol{S}}}{\mathrm{argmin}} \left\{ \sum_{i=1}^{N} \left\| \sum_{j=i}^{N} \boldsymbol{S}_j \tilde{\boldsymbol{R}}_{i,j}^\mathsf{H} \right\|^2 \right\}, \tag{5}$$

where $\tilde{\boldsymbol{R}}_{i,j} \triangleq \frac{p_{j-i}^{(N-i)}}{\sigma_e^{(N-i)}} \boldsymbol{R}_j$. $p_j^{(i)}$ and $(\sigma_e^{(i)})^2$ denote the jth coefficient of the ith order linear backward minimum mean–squared error (MMSE) predictor for the discrete time random process $h_{\mu,\nu}[kN_\mathrm{T}] + n_\nu[kN_\mathrm{T}]$ and the corresponding error variance, respectively. Solving (5) and reversing (1) we obtain an estimate $\hat{\bar{\boldsymbol{V}}}$ for the vector $\bar{\boldsymbol{V}}$. The brute–force approach would be to evaluate (5) for all $2^{(N-1)N_\mathrm{T}R}$ $\bar{\boldsymbol{S}}$ corresponding to all possible $\bar{\boldsymbol{V}}$. However, this approach would quickly become computationally infeasible.

B. Sphere Decoding Algorithm

It can be observed that the ML metric in (5) is a sum of N nonnegative scalar terms

$$\delta_n^2 \triangleq \left\| \boldsymbol{S}_n \tilde{\boldsymbol{R}}_{n,n}^\mathsf{H} + \boldsymbol{X}_n \right\|^2, \quad 1 \le n \le N, \tag{6}$$

which through $\boldsymbol{X}_n \triangleq \sum_{j=n+1}^{N} \boldsymbol{S}_j \tilde{\boldsymbol{R}}_{n,j}^\mathsf{H}$ depend on symbols $\boldsymbol{S}_j$, $n \le j \le N$. Thus, the ML detection problem (5) lends itself to a SD approach. In particular, a SD algorithm very similar to SD for single–antenna MSDSD devised in [8] can be employed for unitary DSTM considered here.

To formulate the SD algorithm, let us define

$$d_n^2 \triangleq \sum_{i=n}^{N} \left\| \boldsymbol{S}_i \tilde{\boldsymbol{R}}_{i,i}^\mathsf{H} + \boldsymbol{X}_i \right\|^2 = d_{n+1}^2 + \delta_n^2, \quad 1 \le n \le N, \tag{7}$$

where $d_{N+1}^2 \equiv 0$ and d_1^2 equals the ML metric in (5). Starting at $n = N-1$,[2] the SD algorithm selects candidates for symbols $\boldsymbol{S}_n$ based on (tentative) decisions for $\boldsymbol{S}_j$, $n+1 \le j \le N$ and continues to decrement n as long as the current metric d_n^2 does not exceed a given maximum metric ρ^2, i.e.,

$$d_n^2 \le \rho^2. \tag{8}$$

[2]Without loss of generality and due to the differential encoding $\boldsymbol{S}_N \equiv \boldsymbol{I}_{N_\mathrm{T}}$ can be chosen.

If the decoder reaches $n = 1$, the metric of the currently best candidate $\hat{\hat{\boldsymbol{S}}}$ is used to further reduce the size of the search space by updating $\rho = d_1$. If d_n exceeds ρ for any value of n, n is incremented and a new candidate for $\boldsymbol{S}_n$ is examined. If the decoder returns to $n = N$ it means that there are no further candidates inside the current sphere and that the ML solution $\hat{\hat{\boldsymbol{S}}}$ has been found. For the ordering of candidates for a $\boldsymbol{S}_n$ we employ the Schnorr–Euchner (SE) enumeration strategy [12], i.e., we check candidates in order of increasing δ_n^2, as this allows for an initialization with $\rho \to \infty$ and an (usually) fast convergence to the ML solution.

C. Application to Different Signal Constellations

While the above is valid for arbitrary unitary DSTM, MSDSD can be tailored to specific constellations in order to simplify the enumeration of candidates for symbols $\boldsymbol{S}_n$. In the following, we consider two important classes of unitary ST constellations.

1) Orthogonal Designs: First, we consider DSTM based on Alamouti's code [13], where data symbols $\boldsymbol{V}[k]$ are taken from the set

$$\boldsymbol{\mathcal{V}}_{\rm OD} \triangleq \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} a & -b^* \\ b & a^* \end{bmatrix} \middle| a, b \in \sqrt{L}\text{–PSK} \right\}. \quad (9)$$

For this particular constellation δ_n^2 can be expressed as

$$\delta_n^2 = \gamma_n^2 + \mathrm{Re}\{a_n \alpha_n\} + \mathrm{Re}\{b_n \beta_n\}, \quad 1 \le n \le N, \quad (10)$$

with variables $\alpha_n,\ \beta_n$ and γ_n being functions of $\tilde{\boldsymbol{R}}_{n,n}$ and $\boldsymbol{X}_n$. This has two immediate consequences: (i) The $(N-1)$–dimensional ML decoding problem with respect to L–ary ST symbols is transformed into a $(2N-2)$–dimensional one with respect to $\sqrt{L}$–PSK symbols and (ii) the SE strategy can be implemented without explicit sorting of candidate symbols in the same efficient way as for single–antenna MSDSD (cf. subfunctions "findBest(·)" and "findNext(·)" in the pseudo–code description in [8]).

2) Diagonal Constellations: Second, we consider diagonal constellations, where data symbols are taken from the set [3]

$$\boldsymbol{\mathcal{V}}_{\rm D} \triangleq \left\{ \mathrm{diag}\left\{ \mathrm{e}^{\mathrm{j}\frac{2\pi}{L}c_1}, \ldots, \mathrm{e}^{\mathrm{j}\frac{2\pi}{L}c_{N_{\rm T}}} \right\}^l \middle| l \in \{0, \ldots, L-1\} \right\}. \quad (11)$$

a) Special Case $[c_1, \ldots, c_{N_{\rm T}}] = [1, \ldots, 1]$: In this particular case, which is optimal for $N_{\rm T} = 2$ and $R = 1$ [3], δ_n^2 can be written in the form of

$$\delta_n^2 = \gamma_n^2 + \mathrm{Re}\{a_n \alpha_n\}, \quad (12)$$

where again α_n and γ_n are functions of $\tilde{\boldsymbol{R}}_{n,n}$ and $\boldsymbol{X}_n$ and a_n denotes the L–PSK symbol in $\boldsymbol{S}_n = a_n \boldsymbol{I}_{N_{\rm T}}$. This means, that the decoding problem is reduced to an $(N-1)$–dimensional one with respect to L–PSK symbols and that the SE strategy can again be implemented as in the single antenna–case.

b) MSDSD and Lattice Decoding: For arbitrary coefficients $[c_1, \ldots, c_{N_{\rm T}}]$, a simple sorting is not feasible, but we have to compute δ_n^2 for all elements of $\boldsymbol{\mathcal{V}}_{\rm D}$ and sort them according to increasing δ_n^2 in order to accomplish ML decoding. In this case, MSDSD benefits from the efficiency of the SE–SD approach, but its complexity is still exponential in R and $N_{\rm T}$.

A closer examination of expression (6) for δ_n^2 reveals that minimizing δ_n^2 is equivalent to the problem of finding the ML solution for CDD with $N = 2$. It is shown in [14] that using the cosine approximation for small arguments this problem can be turned into an $N_{\rm T}$–dimensional lattice–decoding problem and can thus be solved by means of e.g. SD. Applying this approximation to the MSDD problem at hand, we build a MSDSD decoder consisting of two hierarchical SDs, where the inner SD (for finding the minimizer of δ_n^2) is initiated by the outer SD (for solving (5)) with matrices $\tilde{\boldsymbol{R}}_{n,n}$ and $\boldsymbol{X}_n$ whenever a new candidate for $\boldsymbol{S}_n$ is required. In addition the outer SD provides the inner SD with a start radius computed from ρ and d_{n+1} and a list of candidates that have been examined previously given the same tentative decisions $\hat{\boldsymbol{S}}_j,\ n+1 \le j \le N$ and thus are to be excluded from the search. We note that due to the cosine–approximation this decoder, which we will refer to as lattice–decoding (LD) MSDSD, is suboptimum in principle. In Section V, however, we will see that the corresponding performance degradation is very small.

IV. Symbol–Error Rate Analysis and Subset MSDSD

We next derive an approximation for the symbol–error rate (SER) for ML MSDD and thus for MSDSD in Section IV-A. Thereby, we consider the individual SERs associated with the $(N-1)$ data symbols contained in one observation window. The results of this analysis suggest a subset MSDSD algorithm presented in Section IV-B.

A. SER Analysis

Due to different correlations between the received samples in $\bar{\boldsymbol{R}}$, symbol decisions on the $(N-1)$ data symbols in $\bar{\boldsymbol{V}}$ are not equally reliable. Especially in relatively fast fading environments it can be expected that symbols located in the center of the observation window can be detected more reliably than those at the edges. In order to substantiate this intuitive reasoning, we will derive analytical expressions for the individual symbol–error rates SER_n for $\boldsymbol{V}_n$ at position n in the observation window, $1 \le n \le N-1$.

To this end, let us briefly introduce a different description of our system. We define vectors $\bar{\boldsymbol{r}}$, $\bar{\boldsymbol{h}}$ and $\bar{\boldsymbol{n}}$ each filled in the same way with elements from the respective variables of (4) as $\bar{r}_i = r_{\mu,\nu}[k-\kappa],\ 1 \le i \le N N_{\rm T} N_{\rm R}$, with $\mu = \mathrm{mod}\,(i, N_{\rm T}) + 1,\ \nu = \lfloor \frac{i}{N_{\rm T} N} \rfloor + 1,\ \kappa = \mathrm{mod}(\lfloor \frac{i}{N_{\rm T}} \rfloor, N)$. This way the transmission of N symbols is described by $\bar{\boldsymbol{r}} = \left(\boldsymbol{I}_{N_{\rm R}} \otimes \bar{\boldsymbol{S}}_{\rm D}\right) \bar{\boldsymbol{h}} + \bar{\boldsymbol{n}}$ and the corresponding ML metric can be written as

$$\tilde{d}_1^2(\bar{\boldsymbol{r}}, \bar{\boldsymbol{S}}) \triangleq \bar{\boldsymbol{r}}^{\mathsf{H}} \left(\boldsymbol{I}_{N_{\rm R}} \otimes \left(\bar{\boldsymbol{S}}_{\rm D} \left(\boldsymbol{C}^{-1} \otimes \boldsymbol{I}_{N_{\rm T}} \right) \bar{\boldsymbol{S}}_{\rm D}^{\mathsf{H}} \right) \right) \bar{\boldsymbol{r}}, \quad (13)$$

where $\boldsymbol{C}$ is defined as the $N \times N$ correlation matrix of $h_{\mu,\nu}[kN_\mathrm{T}] + n_\nu[kN_\mathrm{T}]$. If we further define

$$\Delta \triangleq \tilde{d}_1^2(\bar{\boldsymbol{r}}, \hat{\bar{\boldsymbol{S}}}) - \tilde{d}_1^2(\bar{\boldsymbol{r}}, \bar{\boldsymbol{S}}), \tag{14}$$

the pairwise error probability $\mathrm{PEP}(\bar{\boldsymbol{S}} \to \hat{\bar{\boldsymbol{S}}})$ that a ML MSDD detects $\hat{\bar{\boldsymbol{S}}}$ while $\bar{\boldsymbol{S}} \neq \hat{\bar{\boldsymbol{S}}}$ was transmitted is given by [15]

$$\mathrm{PEP}(\bar{\boldsymbol{S}} \to \hat{\bar{\boldsymbol{S}}}) = \Pr(\Delta \le 0) = -\sum_{\mathrm{RH\,poles}} \mathrm{Res}\left(\frac{\Phi_\Delta(s)}{s}\right), \tag{15}$$

where

$$\Phi_\Delta(s) \triangleq \mathcal{E}\{\mathrm{e}^{-s\Delta}\} = \det\{\boldsymbol{I}_{N_\mathrm{T}N} + s\boldsymbol{F}\}^{-N_\mathrm{R}} \tag{16}$$

with

$$\boldsymbol{F} \triangleq \bar{\boldsymbol{S}}_\mathrm{D}\left(\boldsymbol{C} \otimes \boldsymbol{I}_{N_\mathrm{T}}\right)\bar{\boldsymbol{S}}_\mathrm{D}^\mathsf{H}\hat{\bar{\boldsymbol{S}}}_\mathrm{D}\left(\boldsymbol{C}^{-1} \otimes \boldsymbol{I}_{N_\mathrm{T}}\right)\hat{\bar{\boldsymbol{S}}}_\mathrm{D}^\mathsf{H} - \boldsymbol{I}_{NN_\mathrm{T}} \tag{17}$$

denotes the characteristic function of the random variable Δ, and the summation is taken over all residues corresponding to poles located in the right–hand (RH) side of the complex s–plane. From (13) we see that Δ is a Hermitian quadratic form of complex Gaussian distributed random variables and hence, we have a closed–form solution for $\Phi_\Delta(s)$ and can solve (15) as discussed in e.g. [16], [15].

Having found an expression for the PEP, the SER can be upper bounded using the union bound over all possible error events. However, since there are $L^{N-1} - 1$ relevant error events, we need to restrict ourselves to the dominating error events to render an SER analysis for large values of N feasible. Following a similar argumentation as in [16] for single–antenna DPSK and utilizing the results of [17] for the design of unitary space–time constellations, these dominant error events are found to correspond to matrix pairs $(\bar{\boldsymbol{S}}, \hat{\bar{\boldsymbol{S}}})$ with highest correlation

$$\zeta \triangleq \left|\left|\hat{\bar{\boldsymbol{S}}}^\mathsf{H}\bar{\boldsymbol{S}}\right|\right|. \tag{18}$$

A brief examination of (18) reveals that error events with only a single non–identical transmit matrix $\hat{\boldsymbol{S}}_n \neq \boldsymbol{S}_n$ such that

$$\zeta_n \triangleq \mathrm{Re}\left\{\mathrm{tr}\left\{\hat{\boldsymbol{S}}_n^\mathsf{H}\boldsymbol{S}_n\right\}\right\}, \; 1 \le n \le N \tag{19}$$

is maximized also maximize correlation ζ. We collect all symbols $\hat{\boldsymbol{S}}_n$ that maximize ζ_n in sets $\hat{\boldsymbol{\mathcal{S}}}_n$ and denote the corresponding set of matrices $\hat{\bar{\boldsymbol{S}}}$ maximizing ζ by $\hat{\bar{\boldsymbol{\mathcal{S}}}}_n$, $1 \le n \le N$.

If we further assume that the error probability is independent of the transmitted matrix, which is the case for DSTM group constellations including diagonal constellations [3] and also for orthogonal designs, we can approximate SER_n with respect to data symbols via

$$\mathrm{SER}_n \approx \sum_{\hat{\bar{\boldsymbol{S}}} \in \hat{\bar{\boldsymbol{\mathcal{S}}}}_n} \mathrm{PEP}(\bar{\boldsymbol{S}} \to \hat{\bar{\boldsymbol{S}}}) + \sum_{\hat{\bar{\boldsymbol{S}}} \in \hat{\bar{\boldsymbol{\mathcal{S}}}}_{n+1}} \mathrm{PEP}(\bar{\boldsymbol{S}} \to \hat{\bar{\boldsymbol{S}}}), \; 1 \le n < N. \tag{20}$$

We have evaluated (20) and results for the example of diagonal DSTM with parameters $N_\mathrm{T} = 4$, $R = 1$, $B_\mathrm{f}T = 0.03$ and $N_\mathrm{R} = 1$ are presented in Fig. 1. It shows the

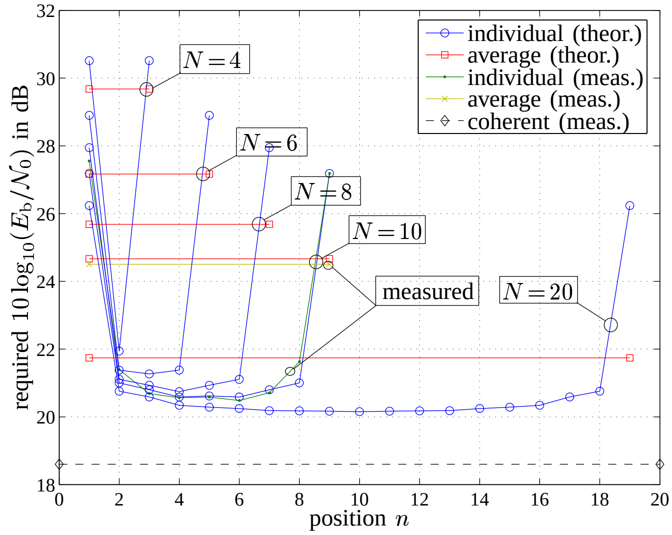

Fig. 1. Required SNR to achieve $\mathrm{SER} = 10^{-5}$ for position n in observation window of MSDSD. Parameters: Diagonal DSTM, $N_\mathrm{T} = 4, N_\mathrm{R} = 1, R = 1, B_\mathrm{f}T = 0.03$.

SNR $10\log_{10}(E_\mathrm{b}/\mathcal{N}_0)^3$ required for MSDSD with different values of N to achieve individual $\mathrm{SER}_n = 10^{-5}$ (blue) and average $\mathrm{SER} = 10^{-5}$ (red) as function of the position n, $1 \le n \le N-1$. Also included are simulated SERs for $N = 10$, and the SER for coherent detection assuming perfect CSI.

First, we note a good agreement between the SER from (20) and simulated SER, which verifies the precision of our SER approximation. Second, it can be seen, that the individual SER_n are almost identical for positions $2 \le n \le N-2$, but deteriorate significantly for positions $n = 1$ and $n = N-1$. In fact, there are gaps of $6-8\,\mathrm{dB}$ in power efficiency when comparing non–edge and edge positions. These observations strongly suggest a variant of MSDSD, which is introduced in the next section.

B. Subset MSDSD

As an immediate consequence of the results from the SER analysis presented in the previous section, we propose subset MSDSD. In subset MSDSD, only $N'-1 \le N-1$ decisions per decoder use corresponding to $\boldsymbol{V}_n$ with $(N-N')/2 < n < (N+N')/2$, $N \pm N'$ even, located in the middle of the observation window are passed to the sink, whereas the remaining $N-N'$ are discarded. Consequently, the observation window must slide forward by $N'-1$ received samples at a time and the decoding complexity is increased by a factor of $(N-1)/(N'-1)$.

For the example considered above, the results from the SER analysis depicted in Fig. 1 suggest that only the decisions corresponding to the edge symbols at positions $n = 1$ and $n = N-1$ should be discarded, i.e., $N' = N-2$ should be used.

[3] E_b and $\mathcal{N}_0$ denote the average received energy per information bit and the two–sided equivalent baseband noise power density, respectively.

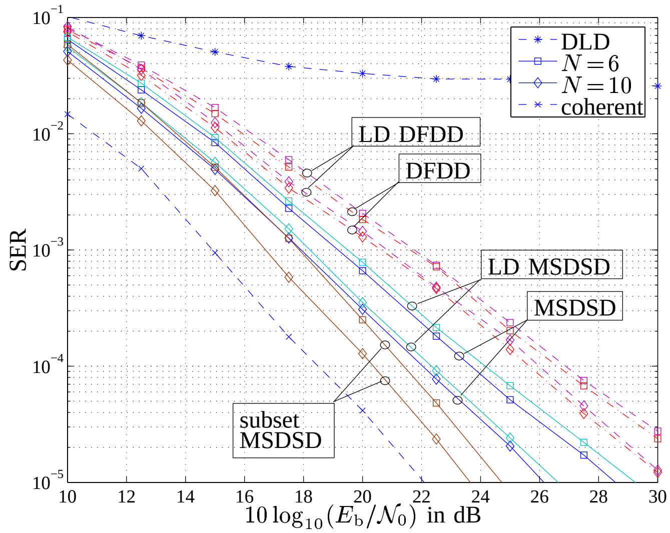

Fig. 2. SER of MSDSD (w/o LD), subset MSDSD ($N' = N - 2$), DFDD (w/o LD), DLD and coherent detection. Parameters: Diagonal DSTM, $N_T = 3$, $N_R = 1$, $R = 1$, $B_f T = 0.03$.

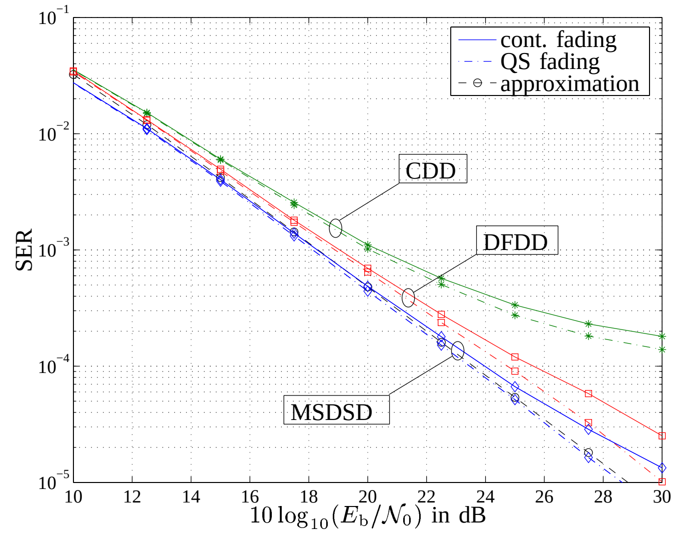

Fig. 3. SER of MSDSD, DFDD and CDD for continuous and QS fading. Parameters: Orthogonal DSTM, $N_T = 2$, $N_R = 1$, $R = 1$, $N = 10$, $B_f T = 0.01$.

When comparing the resulting average SER of subset MSDSD with that of MSDSD, we observe gains in power efficiency of 7.5 dB, 4 dB, and 1.5 dB for $N = 4$, $N = 10$, and $N = 20$, while complexity is increased by a factor of 3, 1.29, and 1.12, respectively. For a reasonable value of $N = 10$, subset MSDSD with $N' = 8$ approaches coherent detection with perfect CSI within 2 dB, which is quite remarkable considering the large normalized fading bandwidth $B_f T = 0.03$.

V. Performance Evaluation

In this section, we present further SER results for and discuss the complexity of the proposed MSDSD and subset MSDSD. In particular, we compare MSDSD with DFDD as proposed in [5] and CDD with $N = 2$ as benchmark algorithms. For results on diagonal constellations we also implemented computationally efficient algorithms for both DFDD and CDD based on lattice decoding (cf. [14], [18]), which we refer to as LD DFDD and DLD, respectively.

A. SER Results

Fig. 2 compares the power efficiency of the various detectors with that of coherent detection with perfect CSI. We assumed diagonal DSTM with $N_T = 3$ and $R = 1$, $N_R = 1$ and $B_f T = 0.03$. We observe a high error floor for DLD in this rapid–fading regime, and also a large gap of 9.5 dB and 8.5 dB in power efficiency at SER $= 10^{-5}$ between DFDD with $N = 6$ and $N = 10$ and coherent detection, respectively. Using MSDSD power efficiency is improved by approximately 3 dB for $N = 6$ and 4 dB for $N = 10$ compared to DFDD. Further improvements are accomplished with subset MSDSD and $N' = N - 2$, whose performance is within 1.6 dB ($N = 10$) to 2.6 dB ($N = 6$) of that for coherent detection with perfect CSI, which means a total improvement of about 7 dB over DFDD. It can also be seen from Fig. 2 that the cosine–approximation required for LD–based symbol search incurrs only a very small degradation compared to the optimum, more complex search.

In Fig. 3 we consider orthogonal DSTM with $N_T = 2$, $R = 1$, $B_f T = 0.01$. Again, we compare MSDSD $(N = 10)$ with DFDD $(N = 10)$ and CDD, this time for the two cases where the fading of the channel is either continuous (solid) or quasistatic (dash–dotted). Whereas channel variations during the transmission of one transmit symbol $\boldsymbol{S}[k]$ are irrelevant for diagonal constellations, they do affect performance for nondiagonal constellations. In particular, we observe a performance degradation for all detectors in continuous fading for error rates below SER $\approx 10^{-3}$, as all are based on the assumption of QS fading. Nevertheless, MSDSD consistently outperforms DFDD, e.g. by approximately 2 dB in power efficiency for SER $\lesssim 10^{-4}$.

B. Computational Complexity

Next, we compare the computational complexity of the proposed MSDSD with those of the DFDD and CDD benchmark decoders. In order to present meaningful results, we consider the average number of real–valued flops per decoded symbol.

Fig. 4 compares the average complexity of MSDSD and DFDD, both w/o lattice–decoder–based symbol search, and DLD for diagonal DSTM assuming the same system and channel parameters as for Fig. 2. As can be seen, the complexity of (LD) MSDSD decreases rapidly with increasing SNR, since the search quickly terminates for small enough noise. For $10\log_{10}(E_b/\mathcal{N}_0) \gtrsim 16$ dB ($N = 6$) and $10\log_{10}(E_b/\mathcal{N}_0) \gtrsim 20$ dB ($N = 10$) the complexity is well in the order of that of (LD) DFDD, which is (almost) independent of the SNR. When considering Fig. 2 we note that MSDSD with larger N is particularly advantageous for moderate to high SNR, where

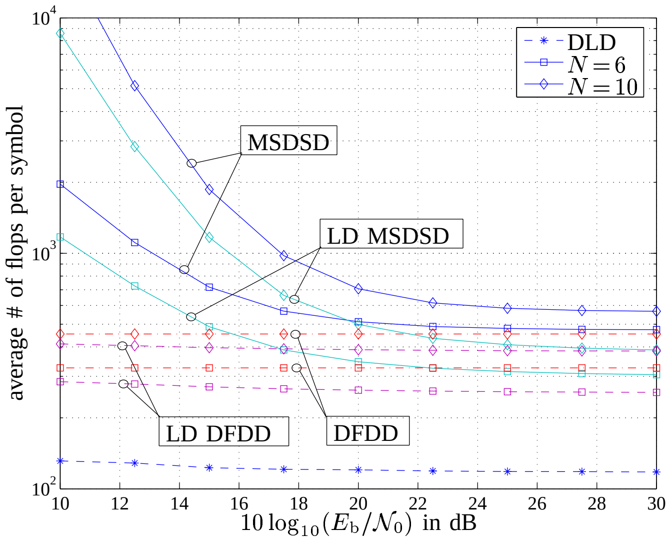

Fig. 4. Complexity of MSDSD (w/o LD), DFDD (w/o LD) and DLD. Parameters: Diagonal DSTM, $N_\mathrm{T}=3$, $N_\mathrm{R}=1$, $R=1$, $B_\mathrm{f}T=0.03$.

the average complexity per symbol is in the same order of magnitude as that of DLD, which assumes $N=2$.

It is also worth pointing out that the complexity of MSDSD strongly depends on the fading rate $B_\mathrm{f}T$ (not shown in the figure) with lower complexities for slower fading.

VI. Conclusions

In this paper, we have extended the concept of MSDSD recently developed for single–antenna DPSK to unitary DSTM. For the interesting special cases of orthogonal and diagonal signal constellations, we have devised optimum and LD–based suboptimum MSDSD algorithms, whose average complexity is not exponential in the number of antennas and the data rate for relevant target error rates. Especially for moderate to high SNRs these decoders achieve significant improvements in power efficiency over DFDD while having comparable complexity. Through an error–rate analysis and simulations we have shown that the performance can be further improved by the proposed subset MSDSD with only a marginal increase in decoder complexity.

References

[1] B.L. Hughes. Differential Space–Time Modulation. *IEEE Transactions on Information Theory*, 46(7):2567–2578, November 2000.

[2] V. Tarokh and H. Jafarkhani. A Differential Detection Scheme for Transmit Diversity. *IEEE Journal on Selected Areas in Communications*, 18(7):1168–1174, July 2000.

[3] B.M. Hochwald and W. Sweldens. Differential Unitary Space–Time Modulation. *IEEE Transactions on Communications*, 48(12):2041–2052, December 2000.

[4] D. Divsalar and M.K. Simon. Multiple–Symbol Differential Detection of MPSK. *IEEE Transactions on Communications*, 38(3):300–308, March 1990.

[5] R. Schober and L. Lampe. Noncoherent Receivers for Differential Space–Time Modulation. *IEEE Transactions on Communications*, 50(5):768–777, May 2002.

[6] C. Ling, K.H. Li, and A.C. Kot. On Decision–Feedback Detection of Differential Space–Time Modulation in Continuous Fading. *IEEE Transactions on Information Theory*, 52(10):1613–1617, October 2004.

[7] E. Chiavaccini and G.M. Vitetta. Further Results on Tarokh's Space–Time Differential Technique. In *Proc. of IEEE International Conference on Communications (ICC)*, New York, April-May 2002.

[8] L. Lampe, R. Schober, V. Pauli, and C. Windpassinger. Multiple–Symbol Differential Sphere Decoding. *Accepted for publication in IEEE Transactions on Communications*, 2005.

[9] O. Damen, H. El Gamal, and G. Caire. On Maximum Likelihood Detection and the Search for the Closest Lattice Point. *IEEE Transactions on Information Theory*, 49(10):2389–2402, October 2003.

[10] G.M. Vitetta and D.P. Taylor. Maximum Likelihood Decoding of Uncoded and Coded PSK Signal Sequences Transmitted over Rayleigh Flat–Fading Channels. *IEEE Transactions on Communications*, 43(11):2750–2758, November 1995.

[11] L. Lampe and R. Schober. Low–Complexity Iterative Decoding for Coded Differential Transmission. In *Proc. of IEEE Wireless Communications and Networking Conference*, New Orleans, March 2003.

[12] C.P. Schnorr and M. Euchner. Lattice Basis Reduction: Improved Practical Algorithms and Solving Subset Sum Problems. *Mathematical Programming*, 66:181–191, 1994.

[13] S.M. Alamouti. A Simple Transmitter Diversity Scheme for Wireless Communications. *IEEE Journal on Selected Areas in Communications*, 16(7):1451–1458, October 1998.

[14] K.L. Clarkson, W. Sweldens, and A. Zheng. Fast Multiple Antenna Differential Decoding. *IEEE Transactions on Communications*, 49(2):253–261, February 2001.

[15] E. Biglieri, G. Caire, G. Taricco, and J. Ventura–Traveset. Computing Error Probabilities over Fading Channels: A Unified Approach. *European Transactions on Telecommunications*, 9:15–25, January/February 1998.

[16] P. Ho and D. Fung. Error Performance of Multiple–Symbol Differential Detection of PSK Signals Transmitted over Correlated Rayleigh Fading Channels. *IEEE Transactions on Communications*, 40:25–29, October 1992.

[17] B.M. Hochwald, T.L. Marzetta, T.J. Richardson, W. Sweldens, and R. Urbanke. Systematic Design of Unitary Space–Time Constellations. *IEEE Transactions on Information Theory*, 46:1962–1973, September 2000.

[18] C. Ling, W.H. Mow, K.H. Li, and A.C. Kot. Multiple–Antenna Differential Lattice Decoding. *Accepted for publication in IEEE J. Select. Areas Commun.*, 2005.