Reduced-Complexity Noncoherently Detected Differential Space-Time Shift Keying

Chao Xu, Shinya Sugiura*, Soon Xin Ng and Lajos Hanzo
School of ECS, University of Southampton, SO17 1BJ, United Kingdom.
Email: {cx1g08,sxn,lh}@ecs.soton.ac.uk, http://www-mobile.ecs.soton.ac.uk

***Abstract*—Motivated by the recent development of Spatial Modulation (SM) and Differential Space-Time Shift Keying (DSTSK), we propose a reduced-complexity Conventional Differential Detector (CDD) as well as a reduced-complexity Multiple-Symbol Differential Sphere Detector (MSDSD) for DSTSK. Both schemes operate on a symbol-by-symbol basis in order to reduce the complexity of the classic block-by-block-based CDD and MSDSD, whilst still approaching the optimum performance of the full-search-based Maximum Likelihood (ML) detector.**

I. Introduction

The concept of Spatial Modulation (SM) was proposed in [1], [2], where only one of the M transmit antenna elements was activated for trasmitting a single modulated symbol. Hence a low-complexity single-antenna-based detector can be employed at the receiver. Furthermore, the effect of Inter-Channel Interference (ICI) faced by all MIMO schemes simultaneously activating several antennas, such as for example the Vertical Bell Laboratories Layered Space-Time (V-BLAST) scheme [3] is eliminated. Motivated by this philosophy, a new Space-Time Modulation (STM) scheme termed as Space-Time Shift Keying (STSK) was proposed in [4] based on the conceptual amalgamation of SM and Linear Dispersion Codes (LDC) [5]. More specifically, one out of Q dispersion matrices is activated for each transmitted symbol in order to disperse a single modulated symbol to M transmit antennas as well as to T time-slots, so that a diversity gain can be achieved. As a further benefit, high-rate transmission can be attained by increasing the number of dispersion matrices, instead of increasing the number of transmit antennas. However, the family of coherently detected STSK schemes require accurate Channel State Information (CSI) and its estimation is a challenging task, especially for high-speed vehicles. In order to mitigate the coherent MIMO detection complexity, Differential STSK (DSTSK) was also introduced in [4] based on the concept of the Cayley transform aided Differential LDCs (DLDC) [6].

We will demonstrate that the detection of DSTSK can be carried out on a symbol-by-symbol basis using a reduced-complexity SM decoder. However, it was demonstrated in [7] that unless we appropriately normalize the transmit power, the optimal ML performance of SM detection cannot be achieved, because the erroneous decisions concerning the specific index of the activated antenna would mislead the single-antenna-based detector. As a potential remedy, the technique of Space Shift Keying (SSK) was proposed in [8], where simply the antenna activation index conveys the source information. Hence the only way to improve the SSK MIMO system's throughput is to increase the number of transmit antennas, which is not always practical owing to space-limitations. *Against this background, the novel contributions of this paper are:*

1) *We further develop the DSTSK scheme of [4] for avoiding the employment of the non-linear Cayley transform in order to reduce the encoding complexity imposed, while eliminates the potential performance degradation.*
2) *The low-complexity SM detector of [2] is carefully improved for facilitating the employment of Conventional Differential Detection (CDD) on a symbol-by-symbol basis, while approaching the optimum ML performance. More explicitly, our improved detection scheme opts for a carefully chosen objective function used for deciding upon the activated antenna index by explicitly considering the modem constellation employed.*
3) *We mitigate the potential performance degradation of noncoherent DSTSK in rapidly fading channels with the aid of Multiple-Symbol Differential Sphere Detector (MSDSD). The detection complexity imposed is significantly reduced by operating the block-based MSDSD of [9] on a symbol-by-symbol basis.*

The rest of this paper is organized as follows. The introduction of DSTSK is presented in Sec. II, while our CDD and MSDSD schemes are portrayed in Sections III and IV, respectively. Finally, our conclusions are offered in Sec. IV.

The following notations are used throughout the paper. A DSTSK scheme employing L-level PSK signalling is denoted by the nomenclature of DSTSK(MNTQ)-LPSK, where M and N indicate the number of transmit and receive antennas, while T and Q represent the number of time-slots per transmission block as well as the number of dispersion matrices employed, respectively. Furthermore, N_{wind} refers to the detection window width of the MSDSD.

The financial support of the RC-UK under the auspices of the India-UK Advanced Technology Centre as well as of the China-UK 4th generation wireless systems project and that of the European Union under the auspices of the Optimix project is gratefully acknowledged.

S. Sugiura is currently with the Toyota Central R&D Labs, Inc., Aichi, 480-1192, Japan (sugiura@ieee.org).

II. Differential Space-Time Shift Keying

The differential encoding process of Differential STM (DSTM) is expressed as [10]:

$$\mathbf{S}_n = \begin{cases} \mathbf{S}_1 & n = 1 \\ \mathbf{X}_{n-1}\mathbf{S}_{n-1} & n > 1 \end{cases}, \quad (1)$$

where the $(T \times M)$-element transmission matrix $\mathbf{S}_n$ invokes M transmit antennas and T symbol periods for transmitting the DSTSK symbols, while the $(T \times T)$-element matrix $\mathbf{X}_n$ stores the source information. Eq. (1) requires $\mathbf{X}_n$ to be unitary, but satisfying this constraint constitutes the most challenging task of DSTM design. The DSTSK in [4] is proposed based on DLDCs [6], where the Cayley transform is utilized to generate unitary matrices. In DLDCs, Q modulated symbols are dispersed with the aid of Q dispersion matrices, which is expressed as:

$$\widetilde{\mathbf{X}}_n = \sum_{q=1}^{Q} \widetilde{\mathbf{A}}_q x_n^q, \quad (2)$$

where the dispersion matrices $\{\widetilde{\mathbf{A}}_q\}_{q=1}^{Q}$ are designed to be Hermitian matrices. Provided that a real-valued BPSK/L-PAM scheme is employed for $\{x_n^q\}_{q=1}^{Q}$, the dispersed matrix $\widetilde{\mathbf{X}}_n$ in Eq. (2) is also a Hermitian matrix. By contrast, the sum or superposition of unitary matrices cannot be guaranteed to be unitary [6]. Then the Cayley transform may be invoked for converting the Hermitian matrix $\widetilde{\mathbf{X}}_n$ of Eq. (2) to the required unitary matrix $\mathbf{X}_n$ of Eq. (1), which is formulated as:

$$\mathbf{X}_n = (\mathbf{I_T} + j\widetilde{\mathbf{X}}_n)^{-1}(\mathbf{I_T} - j\widetilde{\mathbf{X}}_n). \quad (3)$$

As an associated predicament, it was demonstrated in [4] that the non-linear Cayley transform of Eq. (3) not only increases the encoding complexity, but owing to its non-linearity it also results in a performance degradation for the linearized symbol-based DSTSK detector. Fortunately, the employment of the Cayley transform may be circumvented, as a benefit of the simple dispersion process of DSTSK, which is formulated as:

$$\mathbf{X}_n = \widetilde{\mathbf{A}}_q x_n, \quad (4)$$

where $\log_2 Q$ bits are assigned to decide which one of the Q dispersion matrices of $\{\widetilde{\mathbf{A}}_q\}_{q=1}^{Q}$ is activated, while $\log_2 L$ bits are mapped to the L-PSK/PAM symbol x_n. Hence a total of $\log_2(QL)$ bits are required to encode a DSTSK information matrix $\mathbf{X}_n$ in Eq. (4). Table I presents an example of the dispersion process obeying Eq. (4), when a fixed number of $log_2(QL) = 3$ bits per block are conveyed by the DSTSK schemes employing L-PSK.

Observe in Eq. (4) that as an explicit benefit of dispersing without the dispersion matrix summation seen in Eq. (2), each of the Q dispersion matrices $\{\widetilde{\mathbf{A}}_q\}_{q=1}^{Q}$ can be designed as a unitary matrix, and the modulated symbol x_n can be drawn from the complex-valued L-PSK constellation[1], while the dispersed matrix $\mathbf{X}_n$ in Eq. (4) remains a unitary matrix.

[1]If QAM schemes are employed, then similar to the Differential STBCs (DSTBCs) employing QAM in [11], the transmit power normalization is required in Eq. (1). The normalization factor has to be known at the receiver, which in practice requires the periodic transmission of side-information.

The generation of a near-optimum DSTSK scheme requires a carefully conducted random search. More specifically, the Q dispersion matrices $\{\widetilde{\mathbf{A}}_q\}_{q=1}^{Q}$ are randomly generated unitary matrices, while the information matrices $\mathbf{X}_n$ in Eq. (4) may be optimized by minimizing the Pairwise Symbol Error Probability (PSEP) given by [12]:

$$p(\mathbf{X} \to \hat{\mathbf{X}}) \leq \frac{1}{\det(\mathbf{I_{TN}} + \frac{E_s}{4N_0}\mathbf{R_x} \otimes \mathbf{I_N})}, \quad (5)$$

where we have $\mathbf{R_x} = (\mathbf{X} - \hat{\mathbf{X}})^H(\mathbf{X} - \hat{\mathbf{X}})$, while $\mathbf{X}$ as well as $\hat{\mathbf{X}}$ denote any pair out of QL legitimate information matrices. The PSEP based optimization requires $\mathbf{R_x}$ to have full rank, and the minimum determinant of $\mathbf{R_x}$ for any two legitimate codewords should be maximized, which is formulated as [6]:

$$\max \quad \left\{\det\left[(\mathbf{X}_f - \mathbf{X}_g)^H \cdot (\mathbf{X}_f - \mathbf{X}_g)\right]\right\}_{min}. \quad (6)$$

III. Conventional Differential Detection

A. Block-based CDD

The received signal at the receiver may be modelled as:

$$\mathbf{Y}_n = \mathbf{S}_n\mathbf{H}_n + \mathbf{V}_n, \quad (7)$$

where the $(T \times N)$-element matrix $\mathbf{Y}_n$ models the signal received at the N receive antennas during T symbol periods. The Additive White Gaussian Noise (AWGN) term $\mathbf{V}_n$ has a size of $(T \times N)$ with a zero mean and a variance of N_0 in each dimension. The Rayleigh fading channel matrix $\mathbf{H}_n$ has $(M \times N)$ elements, which are independent and identically distributed (i.i.d), and their temporal correlation is given by $\varepsilon\{h_n h_{n+k}^*\} = J_0(2\pi k f_d)$ according to Clarke's fading model [13], where J_0 denotes the zero-order Bessel function of the first kind and f_d is the normalized Doppler frequency.

Assuming that the channel's envelope is nearly time-invariant over two consecutive DSTSK transmission blocks' duration, i.e. we have $\mathbf{H}_{n+1} \approx \mathbf{H}_n$, the next received signal block may be expressed as:

$$\begin{aligned} \mathbf{Y}_{n+1} &\approx \mathbf{S}_{n+1}\mathbf{H}_n + \mathbf{V}_{n+1} \\ &\approx \mathbf{X}_n\mathbf{S}_n\mathbf{H}_n + \mathbf{V}_{n+1} \\ &\approx \mathbf{X}_n\mathbf{Y}_n + \widetilde{\mathbf{V}}_{n+1}, \end{aligned} \quad (8)$$

where the equivalent noise term $\widetilde{\mathbf{V}}_{n+1} = -\mathbf{X}_n\mathbf{V}_n + \mathbf{V}_{n+1}$ has a zero mean and a variance of $2N_0$, which results in a 3 dB performance degradation for employing CDD compared to its coherent counterpart. The Maximum Likelihood (ML) CDD based on Eq. (8) may be formulated as:

$$\mathbf{X}_n = \arg\min_{\mathbf{X}_n} \|\mathbf{Y}_{n+1} - \mathbf{X}_n\mathbf{Y}_n\|^2. \quad (9)$$

The classic CDD of Eq. (9) is block-based, which means that all QL legitimate codewords have to be checked for finding the ML estimate of a DSTSK information block. If we quantify the decoding complexity in terms of the total number of complex additions, multiplications and absolute value calculations, then the complexity of the block-based CDD may be represented by $[(2T + 1)NT] \cdot (LQ)$.

TABLE I
EXAMPLES OF THE INFORMATION MATRIX GENERATION FOR DSTSK SCHEMES EMPLOYING L-PSK OF EQ. (4) [4].

	Q=1, L=8		Q=2, L=4		Q=4, L=2		Q=8, L=1	
Input bits	$\widetilde{\mathbf{A}}_q$	x_n	$\widetilde{\mathbf{A}}_q$	x_n	$\widetilde{\mathbf{A}}_q$	x_n	$\widetilde{\mathbf{A}}_q$	x_n
000	$\widetilde{\mathbf{A}}_1$	1	$\widetilde{\mathbf{A}}_1$	1	$\widetilde{\mathbf{A}}_1$	1	$\widetilde{\mathbf{A}}_1$	1
001	$\widetilde{\mathbf{A}}_1$	$\frac{(1+j)}{\sqrt{2}}$	$\widetilde{\mathbf{A}}_1$	j	$\widetilde{\mathbf{A}}_1$	-1	$\widetilde{\mathbf{A}}_2$	1
010	$\widetilde{\mathbf{A}}_1$	$\frac{(-1+j)}{\sqrt{2}}$	$\widetilde{\mathbf{A}}_1$	$-j$	$\widetilde{\mathbf{A}}_2$	1	$\widetilde{\mathbf{A}}_3$	1
011	$\widetilde{\mathbf{A}}_1$	j	$\widetilde{\mathbf{A}}_1$	-1	$\widetilde{\mathbf{A}}_2$	-1	$\widetilde{\mathbf{A}}_4$	1
100	$\widetilde{\mathbf{A}}_1$	$\frac{(1-j)}{\sqrt{2}}$	$\widetilde{\mathbf{A}}_2$	1	$\widetilde{\mathbf{A}}_3$	1	$\widetilde{\mathbf{A}}_5$	1
101	$\widetilde{\mathbf{A}}_1$	$-j$	$\widetilde{\mathbf{A}}_2$	j	$\widetilde{\mathbf{A}}_3$	-1	$\widetilde{\mathbf{A}}_6$	1
110	$\widetilde{\mathbf{A}}_1$	-1	$\widetilde{\mathbf{A}}_2$	$-j$	$\widetilde{\mathbf{A}}_4$	1	$\widetilde{\mathbf{A}}_7$	1
111	$\widetilde{\mathbf{A}}_1$	$\frac{(-1-j)}{\sqrt{2}}$	$\widetilde{\mathbf{A}}_2$	-1	$\widetilde{\mathbf{A}}_4$	-1	$\widetilde{\mathbf{A}}_8$	1

B. Modified CDD

Since the index q of the activated dispersion matrix and of the L-PSK symbol x_n are separately encoded using $\log_2 Q$ and $\log_2 L$ bits, respectively, the motivation of developing a low-complexity detector is to reduce the complexity of the block-based CDD of Eq. (9) from the order of LQ to the order of $(L+Q)$ by decoding q and x_n separately.

The first step of simplifying the block-based CDD is to express the decision metric of Eq. (9), in a vectorial form, which may be formulated as [4]:

$$\{q, x_n\} = \arg \min_{\{q,x_n\}} \left\| \bar{\mathbf{Y}}_n - \bar{\mathbf{H}}_n \chi \bar{\mathbf{K}}_n \right\|^2, \tag{10}$$

where the notations are given by:

$$\begin{aligned} \bar{\mathbf{Y}}_n &= \mathit{rvec}(\mathbf{Y}_{n+1}), \\ \bar{\mathbf{H}}_n &= \mathbf{I_T} \otimes \mathbf{Y}_n^T, \\ \chi &= \left[\mathit{rvec}(\widetilde{\mathbf{A}}_1) \ \cdots \ \mathit{rvec}(\widetilde{\mathbf{A}}_Q) \right], \\ \bar{\mathbf{K}}_n &= [\underbrace{0 \ \cdots \ 0}_{q-1}, x_n, \underbrace{0 \ \cdots \ 0}_{Q-q}], \end{aligned} \tag{11}$$

where the operation $\mathit{rvec}(\cdot)$ takes the successive rows of the matrix considered, in order to form a vector, while $\otimes$ denotes the Kronecker product. In the absence of the Cayley transform in the codeword generation of Eq. (4), the linearization operation seen in Eq. (10) does not impose any performance penalty compared to the classic ML CDD characterized in Eq. (9) [4].

It may be readily shown that the Euclidean norm calculation of Eq. (10) leads to the following decision variable:

$$\mathbf{Z}_n = \left(\bar{\mathbf{H}}_n \chi\right)^H \bar{\mathbf{Y}}_n, \tag{12}$$

where $\mathbf{Z}_n$ is a vector having Q components, which is used for decoding the Q-component vector $\bar{\mathbf{K}}_n$. It can be readily seen that the detection vector $\mathbf{Z}_n$ is equivalent to that of a SM scheme [2] having Q transmit antennas. The benefit of our scheme is, however, that DSTSK does not have to increase the number of transmit antennas for attaining a higher throughput.

Therefore, the low-complexity SM detector proposed in [2] can be employed for detecting DSTSK using the detection vector of Eq. (12). The simplified CDD first makes a decision concerning the index q of the activated dispersion matrix by testing which particular antenna element has the highest power. This may be expressed as:

$$q = \arg \max |\mathbf{Z}_n|. \tag{13}$$

Then a L-PSK demodulator is invoked to demodulate the q-th symbol in $\mathbf{Z}_n$, which may be written as:

$$x_n = \mathit{round}(\angle Z_n^q). \tag{14}$$

However, it was recognize in [7] that the simplified detector of Eqs. (13) and (14) cannot approach the ML decoding capability of Eq. (9). For example, given $\mathbf{Z}_n = [2.2 + 3.0j, \ -3.4 + 1.1j]^T$ and assuming that QPSK was employed, Eqs. (13) and (14) would make a decision of $\{q, x_n\} = \{1, j\}$, because the symbol at the $q = 1$-st position has the highest power. Hence $Z_n^1 = 2.2 + 3.0j$ is demodulated to j. By contrast, Eq. (9) gives $\{q, x_n\} = \{2, -1\}$. The problem of the simplified CDD was that the erroneous decision concerning the index q made by Eq. (13) misled the the L-PSK demodulator of Eq. (14) to detect the wrong symbol.

Since the QPSK constellation set is $\{\pm 1, \pm j\}$, the estimation of the activation index q would become more reliable, if the decoder could aim for finding the specific symbol in $\mathbf{Z}_n$, which has the highest absolute real or imaginary value. Revisiting the previous example, the index $q = 2$ would be selected with an absolute real value of 3.4. Then Eq. (14) is invoked to demodulate $Z_n^2 = -3.4 + 1.1j$ yielding -1, which is the same as the ML CDD result. Therefore, the ML CDD's decoding capability can be retained by using the appropriate objective function for estimating the activation index q.

In summary, we propose to modify Eq. (13), explicitly depending on which particular L-PSK signalling scheme was employed, as summarized below:

$$\begin{aligned} &L = 1, \quad q = \arg\max \mathit{Re}(\mathbf{Z}_n), \\ &L = 2, \quad q = \arg\max |\mathit{Re}(\mathbf{Z}_n)|, \\ &L = 4, \quad q = \arg\max \{|\mathit{Re}(\mathbf{Z}_n)|, |\mathit{Im}(\mathbf{Z}_n)|\}, \\ &L = 8, \quad q = \arg\max \{|\mathit{Re}(\mathbf{Z}_n)|, |\mathit{Im}(\mathbf{Z}_n)|, \tfrac{|\mathit{Re}(\mathbf{Z}_n)| + |\mathit{Im}(\mathbf{Z}_n)|}{\sqrt{2}}\}. \end{aligned} \tag{15}$$

The decoding complexity of our modified CDD is constituted by two distinct contributions. Quantitatively, the first part is represented by $(2T^2 - 1)NTQ + (2NT - 1)Q$ imposed by calculating $\mathbf{Z}_n$ in Eq. (12), while the second part arises from the simplified decoding operation of Eqs. (14) and (15), which is given by Q, $(Q+2)$, $(2Q+4)$ and $(3Q+8)$ for employing 1-PSK[2], BPSK, QPSK and 8PSK, respectively.

[2]The 1-PSK modulation scheme simply represents the absence of source information assigned to the L-PSK modulation.

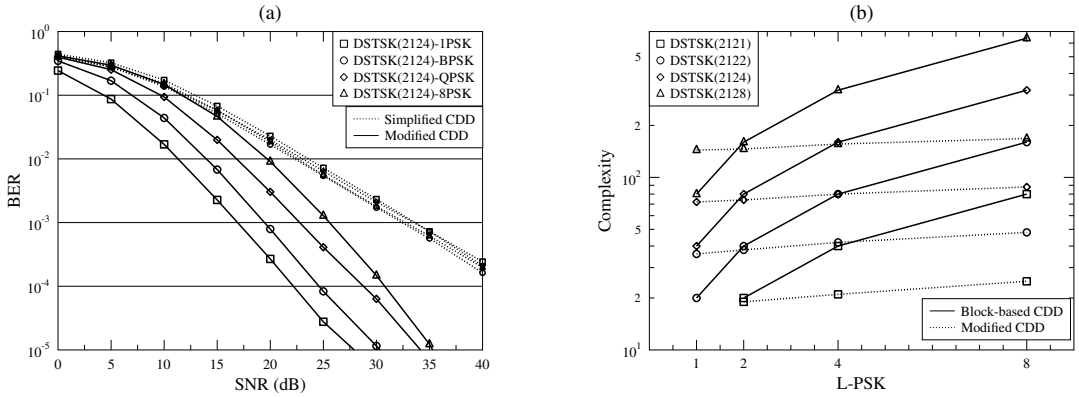

Fig. 1. BER performance and Complexity of DSTSK(212Q)-LPSK employing the proposed CDD of Eqs. (14) and (15), for $f_d = 0.001$.

C. Simulation Results

The BER performance of CDD aided DSTSK is characterized in Fig. 1(a). Observe that employing the simplified CDD adopted from [2] of Eqs. (13) and (14) results in a degraded performance, regardless of which particular L-PSK modulation scheme is employed. This is because any potential erroneous decisions on q mislead the L-PSK demodulator. By contrast, the proposed CDD of Eqs. (14) and (15) exhibits optimum ML detection capability. We have investigated both the proposed CDD and the ML CDD by arranging for them to decode the same channel output at different SNRs, while employing different DSTSK configurations. We found that they always make exactly the same hard decisions.

The complexity comparison between the ML CDD of Eq. (9) and the proposed CDD of Eqs. (14) and (15) is portrayed in Fig. 1(b). Note that the proposed CDD reduces the complexity of ML CDD from the order of QL to $(Q+L)$. However, since the correlation calculation yielding $\mathbf{Z}_n$ in Eq. (12) also imposes a substantial complexity, it can be seen in Fig. 1(b) that the overall complexity is only modestly reduced, when 1-PSK or BPSK is employed. However, observe furthermore in Fig. 1(b) that the complexity curves recorded for the proposed CDD remain near-horizontal, as the levels L of the L-PSK modulation increases. This suggests that the proposed CDD arrangement is capable of decoding DSTSK schemes employing QPSK/8PSK at a similar complexity as decoding BPSK. Indeed, this is expected, because the associated additional complexity is only on the order of $(Q+L)$, which remains modest compared to the complexity of block-based decoding that was shown to be proportional to QL.

IV. Multiple-Symbol Differential Detection

A. Block-based MSDSD

The CDD introduced in Sec. III performs well under the assumption of slow fading channels, but upon increasing the Doppler frequency an irreducible error floor is formed. Therefore, the Multiple-Symbol Differential Detection (MSDD) scheme detailed in [13], [14] observes multiple received signal blocks in order to make a joint decision on the source information for the sake of mitigating the potential performance degradation of noncoherent receivers encountered in rapidly fading channels. Furthermore, the Multiple-Symbol Differential Sphere Detection (MSDSD) was proposed in [15] in order to mitigate the excessive complexity of MSDD.

In order to observe N_{wind} received signal blocks, the noise-contaminated complex channel envelope stretching over N_{wind} blocks may be characterized by the correlation matrix of:

$$\mathbf{C} = \mathbf{R_{HH}} + \mathbf{R_{VV}}, \tag{16}$$

where the correlation of the fading channel is given by:

$$\mathbf{R_{HH}} = \begin{bmatrix} \rho_0 & \rho_1 & \cdots & \rho_{N_{wind}-1} \\ \rho_1 & \rho_0 & \cdots & \rho_{N_{wind}-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{N_{wind}-1} & \rho_{N_{wind}-2} & \cdots & \rho_0 \end{bmatrix},$$

with $\rho_k = J_0(2\pi k f_d)$. Furthermore, the correlation matrix of the AWGN term $\mathbf{V}_n$ in Eq. (7) is given by $\mathbf{R_{VV}} = N_0 \mathbf{I_{N_{wind}}}$. It was shown in [9] that the inversion of the channel's correlation matrix $\mathbf{C}^{-1}$ may be used for predicting the channel. Upon applying Cholesky factorization, we arrive at the channel predictor, which is formulated as:

$$\mathbf{C}^{-1} = \mathbf{L}\mathbf{L}^H, \tag{17}$$

where $\mathbf{L}$ is a lower triangular matrix. Then the block-based MSDSD aided DSTM may be expressed as [9]:

$$\sum_{i=1}^{N_{wind}-1} \left\| \sum_{j=i}^{N_{wind}} l_{ji} \mathbf{S}_j^H \mathbf{Y}_j \right\|^2 \leq R^2, \tag{18}$$

where l_{ji} represents the predictor coefficients hosted by the corresponding elements in the lower triangular matrix $\mathbf{L}$, while R denotes the MSDSD's decoding sphere radius, which is minimized by the sphere decoder.

The first transmission matrix $\mathbf{S}_1$ of each observation window is a common multiplier for all the transmission matrices, hence fixing $\mathbf{S}_1$ or fixing the most recent transmission matrix $\mathbf{S}_{N_{wind}}$ does not affect the search result. Therefore we can define the accumulated information matrix as:

$$\mathbf{A}_n = \begin{cases} \mathbf{S}_n \mathbf{S}_{N_{wind}}^H = \prod_{i=n}^{N_{wind}-1} \mathbf{X}_i^H & 1 \leqslant n < N_{wind} \\ \mathbf{I_T} & n = N_{wind}. \end{cases} \tag{19}$$

Let us now define the Partial Euclidean Distance (PED) component seen in Eq. (18) as:

$$d_i^2 = \sum_{t=i}^{N_{wind}-1} \left\| \sum_{j=t}^{N_{wind}} l_{jt} \mathbf{A}_j^H \mathbf{Y}_j \right\|^2 = d_{i+1}^2 + \Delta_i, \quad (20)$$

with $i = 1, 2, ..., (N_{wind} - 1)$. The PED increment Δ_i in Eq. (20) is expressed as:

$$\Delta_i = \left\| l_{ii} \mathbf{X}_i \mathbf{Y}_i + \mathbf{A}_{i+1} \left(\sum_{j=i+1}^{N_{wind}} l_{ji} \mathbf{A}_j^H \mathbf{Y}_j \right) \right\|^2. \quad (21)$$

When sphere decoder visits a specific index i for the first time, the optimum codeword $\mathbf{X}_i^{opt}$ corresponding to the smallest PED increment Δ_i of Eq. (21) is found. Furthermore, all the QL codewords $\{\mathbf{X}_i^l\}_{l=1}^{QL}$ have to be ordered according to the increasing values of Δ_i [9], so that when the sphere decoder returns to the same index i, only the second-best sub-optimum codeword has to be checked. Therefore, the complexity of ordering the QL candidate codewords on the basis of the block-based PED increment Δ_i of Eq. (21) is represented by $[(2T + 2)NT] \cdot QL$.

B. Modified MSDSD

Observe in Eq. (21) that the PED increment Δ_i has a similar structure to that of the block-based CDD of Eq. (9), hence the 'vectorization' of Eq. (10) may also be applied to Δ_i as:

$$\Delta_i = \left\| \hat{\mathbf{Y}}_i - \hat{\mathbf{H}}_i \chi \hat{\mathbf{K}}_i \right\|^2, \quad (22)$$

where the corresponding new notations are given by:

$$\begin{aligned} \hat{\mathbf{Y}}_i &= rvec\left(\mathbf{A}_{i+1} \left(\textstyle\sum_{j=i+1}^{N_{wind}} l_{ji} \mathbf{A}_j^H \mathbf{Y}_j \right) \right), \\ \hat{\mathbf{H}}_i &= -l_{ii} \cdot \mathbf{I_T} \otimes \mathbf{Y}_i^T, \\ \hat{\mathbf{K}}_i &= [\underbrace{0 \ \cdots \ 0}_{\hat{q}-1}, \hat{x}_i, \underbrace{0 \ \cdots \ 0}_{Q-\hat{q}}]. \end{aligned} \quad (23)$$

Similar to Eq. (12), the correlation operation generates a detection vector, which may be expressed as:

$$\hat{\mathbf{Z}}_i = (\hat{\mathbf{H}}_i \chi)^H \hat{\mathbf{Y}}_i. \quad (24)$$

The Q-component vector $\hat{\mathbf{Z}}_i$ may be used to rank all the QL candidate codewords, so that the complexity of the MSDSD ranking is reduced from the order of QL to $(Q + L)$.

More explicitly, we propose the symbol-by-symbol based MSDSD ranking as follows:

1) In order to rank the dispersion matrix activation indices $\hat{q}$ first, the constellation points of the L-PSK modulation scheme employed have to be separately considered in groups, which is represented by:

$$\begin{aligned} &L = 1, \ dec\{Re(\mathbf{Z}_i)\}, \\ &L = 2, \ dec\{|Re(\mathbf{Z}_i)|\}, \\ &L = 4, \ dec\{|Re(\mathbf{Z}_i)|, |Im(\mathbf{Z}_i)|\}, \\ &L = 8, \ dec\{|Re(\mathbf{Z}_i)|, |Im(\mathbf{Z}_i)|, \tfrac{|Re(\mathbf{Z}_i)+Im(\mathbf{Z}_i)|}{\sqrt{2}}, \\ &\qquad\qquad \tfrac{|Re(\mathbf{Z}_i)-Im(\mathbf{Z}_i)|}{\sqrt{2}}\}, \end{aligned} \quad (25)$$

where *dec* denotes decreasingly ordering the elements of the vector. For example, if QPSK is employed, then the constellation points are grouped into $\{\pm 1\}$ and $\{\pm j\}$, which requires to order the absolute value of the real and the imaginary part of the detection vector, respectively.

2) The ordered dispersion matrix activation indices $\hat{q}$ should be labelled to indicate, which particular group of constellation points they refer to.
3) For all the QL legitimate DSTSK codewords, the first half of the candidate codewords represents the ordered activation indices $\hat{q}$ along with their corresponding optimum constellation points. The second half is given by the indices $\hat{q}$ stored in reversed order, along with their corresponding non-optimum constellation points.

For example, assuming that QPSK was employed and that the detection vector in Eq. (24) is given by $\mathbf{Z}_i = [2.2 + 3.0j, \ -3.4 + 1.1j]^T$, Eq. (25) orders the activation indices $\hat{q}$ as $[2, 1, 1, 2]$ and labels them as $[Re, Im, Re, Im]$, because the real part of the 2nd element in $\mathbf{Z}_i$ has the highest absolute value of 3.4, while the second highest one is given by the imaginary part of the 1st element with a value of 3.0, and so on. Then the sphere decoder is able to order all the legitimate codewords according to the ordered indices and to their labelling. The first four candidate codewords are ordered as $\{(\hat{q}, \hat{x}_i)\} = \{(2, -1), (1, j), (1, 1), (2, j)\}$, because the real part of the 2nd element in $\mathbf{Z}_i$ is demodulated to -1, while the imaginary part of the 1st element is demodulated to be j, and so on. The second half of the candidate codewords are ordered as $\{(\hat{q}, \hat{x}_i)\} = \{(2, -j), (1, -1), (1, -j), (2, 1)\}$. Note that the top ranking activation index $\hat{q} = 2$ along with its optimum demodulated symbol $\hat{x}_i = -1$ is most likely to be the ML solution, while this index along with its non-optimum demodulated symbol $\hat{x}_i = 1$ is in the opposite situation, constituting the bottom ranking candidate of $(\hat{q}, \hat{x}_i) = (2, 1)$.

The complexity of the proposed MSDSD ordering contains two contributions. The first one is represented by $(2T^2 - 1)NTQ + (2NT - 1)Q$ owing to the calculation of $\mathbf{Z}_i$ in Eq. (24). The second contribution is imposed by the index ordering of Eq. (25) and by the L-PSK demodulation, which is given by Q, $(Q + 2)$, $(2Q + 4)$ and $(4Q + 8)$ for 1-PSK, BPSK, QPSK and 8PSK, respectively.

C. Simulation Results

The BER performance of the proposed MSDSD is characterized in Figure (2(a)). It can be seen that the CDD aided DSTSK(2124)-8PSK scheme suffers from the usual 3 dB performance degradation compared to its coherent counterpart using idealized perfect channel estimation in slow fading channels. By contrast, when the normalized Doppler frequency increases to $f_d = 0.03$, an irreducible error floor is formed. Note that the MSDSD associated with $N_{wind} = 2$ is equivalent to the CDD. However, upon increasing the detection window length N_{wind}, the proposed MSDSD successfully mitigates the error floor, and it even partially mitigates the 3 dB performance loss of differential detection, when a sufficiently long window of $N_{wind} = 11$ is employed. Furthermore, our MSDSD

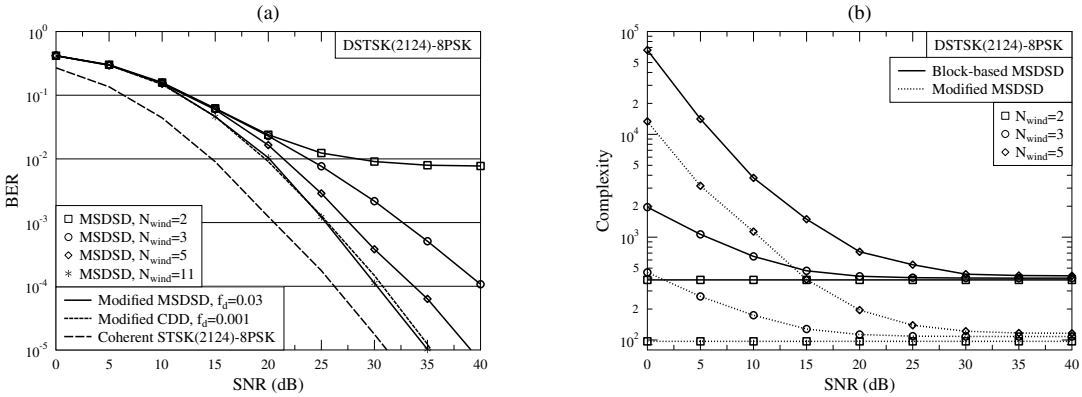

Fig. 2. BER performance and Complexity of DSTSK(2124)-8PSK employing the proposed modified MSDSD in Sec. IV-B, for $f_d = 0.03$.

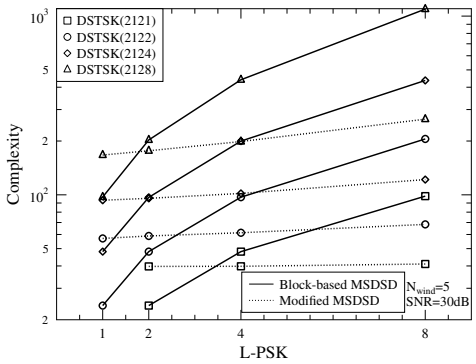

Fig. 3. Complexity comparison between the DSTSK(212Q)-LPSK employing the block-based MSDSD of Eq. (20) and employing our proposed MSDSD in Sec. IV-B with $N_{wind} = 5$, for $SNR = 30$ dB and $f_d = 0.03$.

introduced in Sec. IV-B has the same decoding capability, as the block-based MSDSD of Eq. (20).

However, the proposed MSDSD has beneficially reduced complexity compared to the block-based MSDSD, which is evidenced by Fig. 2(b). It can also be seen that the complexity gradually decreases as SNR increases, because the sphere decoder typically terminates sooner at high SNRs.

Fig. 3 portrays the complexity comparison of the block-based MSDSD of Eq. (20) and of the proposed MSDSD introduced in Sec. IV-B for different DSTSK schemes. Similar to the CDD complexity comparison shown in Fig. 1(b), the modified MSDSD also has a significantly reduced complexity compared to the block-based MSDSD, when QPSK/8PSK is employed, which is evidenced by Fig. 3.

V. Conclusions

In this paper, we proposed a reduced-complexity CDD and MSDSD aided DSTSK scheme. Our solution facilitates the employment of arbitrary complex-valued constellations, while avoiding the non-linear calamities of the Cayley transform. Our simulation results demonstrate that the proposed CDD and MSDSD are capable of achieving the optimum ML decoding performance of the block-based CDD and MSDSD, despite the fact that our proposed CDD and MSDSD are capable of decoding DSTSK schemes employing QPSK/8PSK at a similar complexity as decoding BPSK.

References

[1] Y. A. Chau and S.-H. Yu, "Space modulation on wireless fading channels," in *IEEE 54th Vehicular Technology Conference (VTC2001-Fall)*, vol. 3, pp. 1668 –1671, Aug. 2001.

[2] R. Mesleh, H. Haas, S. Sinanovic, C. W. Ahn, and S. Yun, "Spatial modulation," *IEEE Transactions on Vehicular Technology*, vol. 57, pp. 2228 –2241, July 2008.

[3] C. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs. Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.

[4] S. Sugiura, S. Chen, and L. Hanzo, "Coherent and differential space-time shift keying: A dispersion matrix approach," *IEEE Transactions on Communications*, vol. 58, pp. 3219 –3230, Nov. 2010.

[5] R. Heath and A. Paulraj, "Linear dispersion codes for MIMO systems based on frame theory," *IEEE Transactions on Signal Processing*, vol. 50, pp. 2429 – 2441, Oct. 2002.

[6] B. Hassibi and B. Hochwald, "Cayley differential unitary space-time codes," *IEEE Transactions on Information Theory*, vol. 48, pp. 1485 –1503, June 2002.

[7] J. Jeganathan, A. Ghrayeb, and L. Szczecinski, "Spatial modulation: optimal detection and performance analysis," *IEEE Communications Letters*, vol. 12, pp. 545 –547, Aug. 2008.

[8] J. Jeganathan, A. Ghrayeb, L. Szczecinski, and A. Ceron, "Space shift keying modulation for MIMO channels," *IEEE Transactions on Wireless Communications*, vol. 8, pp. 3692 –3703, July 2009.

[9] V. Pauli and L. Lampe, "On the complexity of sphere decoding for differential detection," *IEEE Transactions on Information Theory*, vol. 53, pp. 1595 –1603, Apr. 2007.

[10] L. Hanzo, O. Alamri, M. El-Hajjar, and N. Wu, *Near-Capacity Multi-Functional MIMO Systems - Sphere-Packing, Iterative Detection and Cooperation*. Wiley-IEEE Press, 2009.

[11] C.-S. Hwang, S. H. Nam, J. Chung, and V. Tarokh, "Differential space time block codes using nonconstant modulus constellations," *IEEE Transactions on Signal Processing*, vol. 51, pp. 2955 – 2964, Nov. 2003.

[12] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, pp. 744 –765, Mar. 1998.

[13] 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*, vol. 40, pp. 1566–1569, Oct. 1992.

[14] L. Hanzo, Y. Akhtman, M. Jiang, and L. Wang, *MIMO-OFDM for LTE, WIFI and WIMAX: Coherent versus Non-Coherent and Cooperative Turbo-Transceivers*. John Wiley & Sons, 2010.

[15] L. Lampe, R. Schober, V. Pauli, and C. Windpassinger, "Multiple-symbol differential sphere decoding," *IEEE Transactions on Communications*, vol. 53, pp. 1981–1985, Dec. 2005.