

Single-RF Index Shift Keying Aided Differential Space-Time Block Coding

Chao Xu, *Member, IEEE*, Rakshith Rajashekar, *Member, IEEE*, Naoki Ishikawa, *Student Member, IEEE*, Shinya Sugiura, *Senior Member, IEEE*, and Lajos Hanzo* *Fellow, IEEE*

Abstract—We propose a new single-RF Differential Space-Time Block Coding using Index Shift Keying (DSTBC-ISK), which is the first in the family of Differential Space-Time Modulation (DSTM) schemes that can simultaneously achieve the following three imperative objectives of (1) forming a finite-cardinality transmit-signals set under the matrix multiplications of differential encoding; (2) retaining a single-stream ML detection complexity that does not grow with the constellation size; (3) offering a beneficial transmit diversity gain over the recently developed Differential Spatial Modulation (DSM). In order to make a fair comparison, we also conceive a low-complexity single-stream detector for DSM, which does not impose any performance loss in comparison to the existing solutions in open literature. Furthermore, in order to improve the performance of finite-cardinality DSTM schemes at higher throughputs, we propose to generalize both Differential Amplitude Shift Keying (DASK) and Amplitude Shift Keying (ASK), which form the generic multi-level-ring star QAM constellation that subsumes the existing two/four-level-ring DASK solutions as special cases. Although the DASK approach has been popularly used in DSTM schemes, we demonstrate that our generalized ASK technique achieves a higher capacity and a better performance in channel coding assisted systems. Moreover, since the employment of star QAM constellations imposes the ring-amplitude dependent signal power problem for detection, we further develop bespoke Maximum-Likelihood (ML), Minimum Mean Squared Error (MMSE) and Least Square (LS) detectors for DSTM using DASK/ASK, which exhibit different performance versus complexity tradeoffs. Our simulation results demonstrate that the proposed DSTBC-ISK is capable of achieving substantial diversity gains over DSM without eroding its low transceiver complexity.

Index Terms—Differential space-time modulation, differential space-time block code, differential spatial modulation, group code, star QAM, single-RF, single-stream detection.

I. INTRODUCTION

In Single-Input Single-Output (SISO) channels, Differential Phase Shift Keying (DPSK) [1], [2], Differential Amplitude Phase Shift Keying (DAPSK) [3]–[5] and Absolute-amplitude Differential Phase Shift Keying (ADPSK) [4]–[6] constitute low-complexity alternatives to coherent PSK/QAM schemes. In Multiple-Input Multiple-Output (MIMO) channels, Differential Space-Time Modulation (DSTM) that dispenses with high-complexity channel estimation has also attracted substantial research interests. Explicitly, Differential Space-Time Block Code (DSTBC) using Alamouti’s signal structure [7]

was proposed in [8] for two Transmit Antennas (TAs), and then it was extended to multiple TAs in [9], [10]. In order to improve throughput, star QAM and square QAM were also applied to DSTBC in [11] and [12], [13], respectively.

Furthermore, the Differential Group Code (DGC) concept was proposed in [14], [15]. The DGC’s diagonal or anti-diagonal signal matrices form a finite group under multiplication, where the signal elements are given by exponentiation functions of an L -PSK symbol optimized for diversity gain. The entire family of full-diversity finite-group DGCs was summarized in [16], [17]. In order to improve the DGC’s performance at higher throughputs, a variety of nongroups and infinite groups were conceived in [18]–[21]. Moreover, Differential Linear Dispersion Code (DLDC) was proposed in [22] and was further optimized in [23]–[25]. The DLDC combines a total of Q Hermitian matrices multiplied by Q modulated symbols and then the Cayley transform [22] converts the resultant Hermitian matrix to a unitary matrix. Owing to the associated non-linear signal mapping, both DGC and DLDC exhibit detection complexities that increase exponentially with the throughput. As a remedy, the polynomial-complexity lattice decoding and sphere decoding schemes were invoked for DGC and DLDC in [26] and [22], respectively, which however may introduce a performance loss.

Inspired by the recently proposed Spatial Modulation (SM) [27]–[30] and Space-Time Shift Keying (STSK) [31]–[33], Differential SM (DSM) and Differential STSK (DSTSK) have also been developed in [34]–[37] and [31]–[33], respectively. More explicitly, the DSM [34]–[37] activates a single position in each row and column of the signal matrix for transmitting modulated symbols. The DSTSK concept [31]–[33] has evolved from the DLDC arrangement, where only a single matrix out of Q dispersion matrices is activated. In the absence of multiplexing, the DLDC’s Cayley transform was eliminated in DSTSK [32], [33], and single-stream detectors were invoked for both DSM and DSTSK in [32], [38], [39]. As a further advance, the dispersion matrix based DSM was proposed in [40], [41], where a beneficial tradeoff was struck between the multiplexing gain and diversity gain attained.

The coherent MIMO schemes generally have a finite-cardinality set of legitimate signals. However, when coherent MIMO signals are directly used in DSTM, the matrix multiplication invoked by differential encoding typically results in arbitrary signals sets having a cardinality tending to infinity. This imposes extra constraints on the speed, precision and dynamic range of both the signal processing digital circuitry and Digital-to-Analog Converter (DAC). This may be avoided by the DGC [14]–[17] and DSM schemes [34]–[37]. However, DSM does not perform well, when the number of Receive Antennas (RAs) N is small. Moreover, the DGC’s superior

C. Xu, R. Rajashekar and L. Hanzo are with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK (e-mail: {cx1g08,rmr1u14,lh}@soton.ac.uk).

N. Ishikawa and S. Sugiura are with the Department of Computer and Information Sciences, Tokyo University of Agriculture and Technology, Kaganei184-8588, Japan (e-mail: naoki@ishikawa.cc, sugiura@iee.org).

The financial support of the EPSRC projects EP/N004558/1 and EP/L018659/1, as well as of the European Research Council’s Advanced Fellow Grant under the Beam-Me-Up project and of the Royal Society’s Wolfson Research Merit Award is gratefully acknowledged.

performance is achieved at the excessive cost of an exponentially increasing complexity as a function of the throughput. Against this background, we propose a new DSTBC using Index Shift Keying (DSTBC-ISK), which is the first DSTM scheme capable of simultaneously achieving the following three imperative objectives:

- (I) **Finite-cardinality:** the DSTM scheme forms a finite-cardinality transmit-signals set under the matrix multiplications of differential encoding.
- (II) **Single-stream ML detection:** the DSTM scheme retains a single-stream ML detection complexity that does not grow with the constellation size.
- (III) **Diversity gain:** the DSTM scheme achieves a beneficial transmit diversity gain over DSM.

More explicitly, the novel contributions are as follows:

- 1) The proposed DSTBC-ISK activates a single symbol out of a total of Q symbols in the STBC signal structure. In contrast to the dispersion matrix based DSM of [40], [41], the proposed DSTBC-ISK systematically forms a finite-cardinality transmit-signal set under matrix multiplication. Furthermore, thanks to the STBC signal structure, the proposed DSTBC-ISK retains both full transmit diversity as well as a single-stream ML detection complexity.
- 2) In order to make a fair comparison, we also conceive a single-stream ML detector for DSM. The resultant complexity reduction is as high as a factor of **170 times** for the case of using ($M = 2$) TAs and ($N = 1$) RAs at the throughput of ($R = 4.0$). Unlike the solutions in [38], [39], the proposed DSM detector does not impose any performance loss.
- 3) Furthermore, in order to improve the performance of the finite-cardinality DSTM schemes at high throughputs, we propose to generalize Differential Amplitude Shift Keying (DASK) and Amplitude Shift Keying (ASK), which form the classic multi-level-ring star QAM constellation that subsumes the two/four-level-ring DASK solutions of [11], [36], [37] as special cases. Although the DASK approach has been popularly used in DSTM schemes [11], [36], [37], our generalized ASK technique is capable of achieving a higher capacity and a better performance in channel coding assisted systems.
- 4) Moreover, since the employment of star QAM constellations imposes the ring-amplitude dependent signal power problem for detection, we further develop bespoke Maximum-Likelihood (ML), Minimum Mean Squared Error (MMSE) and Least Square (LS) detectors for DSTM using star QAM, which exhibit different performance versus complexity tradeoffs.
- 5) Our simulation results demonstrate that the proposed DSTBC-ISK is capable of achieving a beneficial diversity gain over DSM without eroding its low transceiver complexity. The DSTBC-ISK's diversity gains over DSM are as substantial as **14.2 dB** and **24.5 dB** at $\text{BER}=10^{-4}$ for the cases of ($M = 2, N = 1, R = 2.0$) and ($M = 4, N = 1, R = 1.0$), respectively. Moreover, compared to DGC, which also form finite-cardinality

transmit-signal set, DSTBC-ISK is capable of achieving a competitive performance at the low-cost of a single-stream ML detection complexity.

The rest of this paper is organized as follows. The design objectives are presented in Sec. II. The new DSTBC-ISK is proposed in Sec. III. The single-stream ML detector is conceived for DSM in Sec. IV. The generalized DASK/ASK is modelled in Sec. V. Our simulation results are provided in Sec. VI, while our conclusions are offered in Sec. VII.

The following notations are used throughout the paper. $\Re(\cdot)$ and $\Im(\cdot)$ take the real and imaginary parts of complex numbers, while $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ denote the conjugate of a complex number, the transpose of a matrix and the Hermitian transpose of a complex matrix, respectively. The operation \otimes represents the Kronecker product, while $\text{tr}(\cdot)$, $\text{rank}(\cdot)$ and $\det(\cdot)$ take trace, rank and determinant of a matrix, respectively. Moreover, for a generic DSTM, the notations M , N , T and Q represent the numbers of TAs, RAs, transmission time slots and modulated symbols, respectively. The superscript of a data-carrying symbol (e.g. ' l ' in $\{x^l\}_{l=0}^{L-1}$) or matrix (e.g. ' i ' in $\{\mathbf{X}^i\}_{i=0}^{I-1}$) denotes the data-carrying index, which is directly determined by the source bits, such as $[l = \text{bin2dec}(b_1 \cdots b_{\log_2 L})]$ and $[i = \text{bin2dec}(b_1 \cdots b_{\log_2 I})]$, where the operation $\text{bin2dec}(\cdot)$ converts bits to a decimal integer.

II. PRELIMINARIES ON DESIGN OBJECTIVES

In this section, we use the classic DSTBC as an example in order to introduce the objectives of our new DSTM scheme. First of all, the DSTM transmitter's finite-cardinality constellation property is introduced in Sec. II-A. Secondly, the DSTM receiver's single-stream ML detection property is presented in Sec. II-B. Thirdly, the beneficial diversity gain is introduced in the performance analysis of Sec. II-C.

A. Transmitter Finite-Cardinality Constellation Property

The DSTBC introduced in this section is the Amicable Orthogonal (AO) design [10]. The transmitter firstly constructs a Q -element LPSK vector $\mathbf{x}_{n-1} = [x^{l_1}, \dots, x^{l_Q}]$, where each LPSK symbol conveys $\log_2 L$ source information bits as $x^{l_q} = \exp(j\frac{2\pi}{L}l_q) = w_L^{l_q}$ and we have $w_L = \exp(j\frac{2\pi}{L})$. The phase index l_q is Gray coded \check{l}_q . Following this, a $(T \times T)$ -element STBC matrix \mathbf{X}_{n-1} is modelled as:

$$\mathbf{X}_{n-1} = \frac{1}{\sqrt{Q}} G_T^{\text{AO}}(\mathbf{x}_{n-1}), \quad (1)$$

where the AO-STBC signal structure is given by [10], [42]:

$$G_{2^\iota}^{\text{AO}}([x_1, \dots, x_{\iota+1}]) = \begin{bmatrix} G_{2^{\iota-1}}^{\text{AO}}([x_1, \dots, x_\iota]) & x_{\iota+1} \mathbf{I}_{2^{\iota-1}} \\ -x_{\iota+1}^* \mathbf{I}_{2^{\iota-1}} & G_{2^{\iota-1}}^{\text{AO}}([x_1, \dots, x_\iota])^H \end{bmatrix}. \quad (2)$$

and $\mathbf{I}_{2^{\iota-1}}$ represents a $(2^{\iota-1} \times 2^{\iota-1})$ -element identity matrix. If T is a power of 2 as $T = 2^\iota$ for a positive integer ι , (2) can be constructed from $\iota = 1$ and $G_1^{\text{AO}}([x_1]) = x_1$. The number of modulated symbols in the STBC signal structure of (2) is given by $Q = \iota + 1 = \log_2 T + 1$. The throughput is given by $R = \frac{Q \log_2 L}{T}$. Following this, the $(T \times M)$ -element transmitted signal matrix \mathbf{S}_n may be obtained by differential encoding:

$$\mathbf{S}_n = \mathbf{X}_{n-1} \mathbf{S}_{n-1}, \quad (3)$$

SISO DPSK: $s_n = x_{n-1}s_{n-1}$ DSTBC (M=2): $\mathbf{S}_n = \mathbf{X}_{n-1}\mathbf{S}_{n-1}$ $\begin{bmatrix} s_{n,1} & s_{n,2} \\ -s_{n,2}^* & s_{n,1}^* \end{bmatrix} = \begin{bmatrix} x_{n-1,1} & x_{n-1,2} \\ -x_{n-1,2}^* & x_{n-1,1}^* \end{bmatrix} \begin{bmatrix} s_{n-1,1} & s_{n-1,2} \\ -s_{n-1,2}^* & s_{n-1,1}^* \end{bmatrix}$ $s_{n,1} = x_{n-1,1}s_{n-1,1} - x_{n-1,2}s_{n-1,2}^*$ $s_{n,2} = x_{n-1,1}s_{n-1,2} + x_{n-1,2}s_{n-1,1}^*$
DGC-cyclic (M=2): $\begin{bmatrix} s_{n,1} & 0 \\ 0 & s_{n,2} \end{bmatrix} = \begin{bmatrix} x_{n-1,1} & 0 \\ 0 & x_{n-1,2} \end{bmatrix} \begin{bmatrix} s_{n-1,1} & 0 \\ 0 & s_{n-1,2} \end{bmatrix}$ $s_{n,1} = x_{n-1,1}s_{n-1,1}$ $s_{n,2} = x_{n-1,2}s_{n-1,2}$

Fig. 1. Differential encoding of the SISO DPSK, DSTBC ($M = 2$) using PSK and DGC-cyclic schemes ($M = 2$). The DSTBC and DGC schemes are further explained by Example 1 and Example 4, respectively.

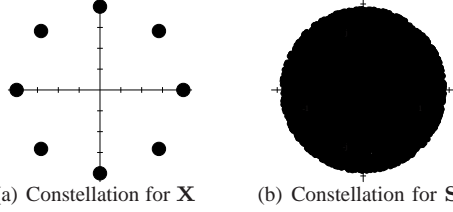


Fig. 2. Constellation diagrams for DSTBC signals in \mathbf{X} and \mathbf{S} of (3), where ($M = 2$) TAs and 8PSK signals are used.

where we have $\text{tr}(\mathbf{S}_n^H \mathbf{S}_n) = T$ and ($M \leq T$). \mathbf{S}_{n-1} in (3) may start from $\mathbf{S}_1 = \sqrt{\frac{T}{M}} [\mathbf{I}_M, \mathbf{0}]^T$, where the all-zero matrix $\mathbf{0}$ has $[M \times (T - M)]$ elements. Moreover, \mathbf{S}_n in (3) becomes \mathbf{S}_{n-1} in the next transmission block, and \mathbf{X}_{n-1} carries source information in each block, as exemplified by (1).

Example 1: The matrix multiplication of (3) is extended for DSTBC ($M = 2$) in Fig. 1. In contrast to SISO DPSK, although the DSTBC's data-signals $x_{n-1,1}$ and $x_{n-1,2}$ in \mathbf{X}_{n-1} of (1) are drawn from an LPSK constellation, the transmit-signals $s_{n,1}$ and $s_{n,2}$ in \mathbf{S}_n of (3) do not remain limited to LPSK due to the additions/subtractions seen in Fig. 1. Instead, the matrix multiplication of (3) results in an infinite-cardinality set of arbitrary signals, which is exemplified in Fig. 2 for the case of DSTBC ($M = 2$) using 8PSK.

More explicitly, it is demonstrated by Fig. 3(a) that the cardinality of the transmit-signals in \mathbf{S}_n of (3) grows with the transmission time index n without limit. This implies that the cardinality of the signals that the TAs should be prepared to transmit tends to infinity, despite the fact that the TAs can only radiate a limited number of patterns [43]–[45]. We note that if the time index n is limited to a shorter transmission frame, the cardinality seen in Fig. 3(a) is indeed finite. However, the constellation of arbitrary transmit-signals exemplified by Fig. 2(b) is still different from frame to frame, which leads to the infinite-cardinality problem for the TAs over time.

This infinite-cardinality problem is a direct consequence of the differential encoding of (3), which does not exist for coherent MIMO. In particular, although Linear Dispersion Codes (LDCs) [46]–[48] and STSK [31], [33] use signals that are not drawn from the classic PSK/QAM constellations, they still always have a finite-cardinality set of legitimate signals.

As a further detrimental effect, it is demonstrated in Fig. 3(b) that the transmit-signals of DSTBC suffer from a substantial Peak-to-Average Power Ratio (PAPR) surge compared to their SISO DPSK/DAPSK counterparts using the classic PSK and star QAM constellations.

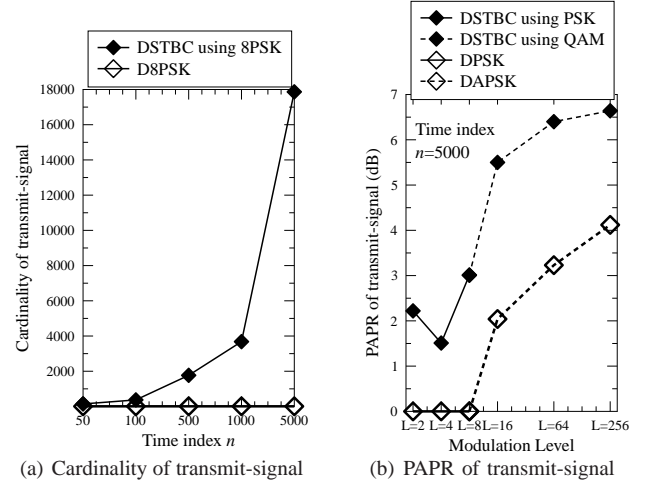


Fig. 3. Cardinality and PAPRs of transmit-signals of DSTBC using PSK [10] and DSTBC using QAM [12], [13], where ($M = 2$) TAs are used.

Consequently, the DSTBC's arbitrary transmit-signals may impose significant constraints on the speed, precision and power consumption of both the digital circuitry and the DAC. In reality, the DSTBC's transmit-signals may have to be quantized to a finite-cardinality set of uniquely distinguishable signals for the TA's analog RF chain to process and transmit. The quantization complexity imposed and the errors are subject to the precision of realistic hardware design. This problem was not considered when the DSTBCs of [8]–[10], the DLDCs of [21], [22] and the DSTSKs of [31]–[33] were conceived.

At the time of writing, only DGC [14]–[17] and DSM [34]–[37] can systematically avoid the infinite-cardinality problem by their sparse matrix design. An example of DGC-cyclic ($M = 2$) is also portrayed in Fig. 1, where the transmit-signals $s_{n,1}$ and $s_{n,2}$ may retain the same LPSK constellation as the data-signals $x_{n-1,1}$ and $x_{n-1,2}$. The schematics of cyclic/decyclic DGCs and DSM will be further explained in Secs. II-C and IV, respectively. In summary, we formulate the following transmitter constellation property, which constitutes the *first objective* of our proposed DSTBC-ISK design:

Property 1: In order to arrive at a finite-cardinality transmit-signal set under matrix multiplication, it is sufficient to ensure that the signal matrices belong to a specific type of sparse matrices that have only a single non-zero element in each row and column, where the non-zero elements assume equi-spaced PSK phases and/or equi-spaced DASK/ASK ring-amplitudes.

We emphasize that condition of Property 1 is sufficient but not necessary. The special cases of differentially encoded G2 using BPSK and QPSK in [8] as well as the real-valued DSTBC using BPSK in [9] have extended but finite cardinality for their transmit-signal sets. Moreover, the Field Extension based DSM (FE-DSM) of [41] using the specific parameter ($u_1 = \frac{2\pi}{L}$) may also form a finite group. Nonetheless, only DGC [14]–[17] and DSM [34]–[37] are capable of always satisfying Property 1, regardless of their parameters.

It is also worth noting that Property 1 does not necessarily lead to the same group property as the DGC of [14]–[17]. For example, DSM [34]–[37] does not always have the same set for \mathbf{X}_{n-1} and \mathbf{S}_n . As a further benefit, Property 1 subsumes Condition 1 of [41], which facilitates single-RF transmission. Without diverting the focus of this paper, interested readers may refer to [49]–[51] for more details in this area.

Moreover, similarly to the coherent MIMO scheme of SM that sacrifices V-BLAST's throughput in exchange for lower transceiver complexity, Property 1 may also result in a throughput loss for the prospective DSTM scheme. Nonetheless, the aim of our DSTM design is to offer a differential alternative to the classic SM scheme that may achieve an improved performance compared to DSM at the same requirement of having a low transceiver complexity.

B. Receiver Single-Stream ML Detection Property

The received signals may be modeled as:

$$\mathbf{Y}_n = \mathbf{S}_n \mathbf{H}_n + \mathbf{V}_n, \quad (4)$$

where the $(T \times N)$ -element \mathbf{Y}_n and \mathbf{V}_n are the received signal and the AWGN matrices, respectively. The $(M \times N)$ -element matrix \mathbf{H}_n models spatially uncorrelated Rayleigh fading channels. We note that transmitter precoding may be required in the presence of transmit and receive antenna correlation [52]–[54], which is not considered in this work.

In quasi-static fading channels, we have $\mathbf{H}_n = \mathbf{H}_{n-1}$, hence (4) becomes $\mathbf{Y}_n = \mathbf{X}_{n-1}(\mathbf{Y}_{n-1} - \mathbf{V}_{n-1}) + \mathbf{V}_n$, which leads to the classic Differential Detection (DD):

$$\hat{\mathbf{X}}_{n-1} = \arg \min_{\mathbf{X}_{n-1}} \|\mathbf{Y}_n - \mathbf{X}_{n-1} \mathbf{Y}_{n-1}\|^2. \quad (5)$$

The family of differential schemes is generally considered as a low-complexity design alternative to their coherent counterparts. If the complexity of channel estimation is to be avoided at the receiver, the exponentially increasing DSTM detection complexity of (5) with respect to throughput should also be avoided without any further performance loss. Therefore, in order to invoke the single-stream ML demodulator for (5), we offer the following property based on the STBC's orthogonal design [10], [42], [55], which constitutes *the second objective* of our work:

Property 2: In order to decouple (5) without imposing any performance loss, it is sufficient to require \mathbf{X}_{n-1} in (3) to be linear and orthogonal:

- ① \mathbf{X}_{n-1} is linear combination of Q modulated symbols $\{x_{n-1,q}\}_{q=1}^Q$ and dispersion matrices $\{\mathbf{A}_q, \mathbf{B}_q\}_{q=1}^Q$ as $\mathbf{X}_{n-1} = \sqrt{P_t} \sum_{q=1}^Q [\mathbf{A}_q \Re(x_{n-1,q}) + j \mathbf{B}_q \Im(x_{n-1,q})]$.
- ② \mathbf{X}_{n-1} in (3) has orthogonal columns as $\mathbf{X}_{n-1}^H \mathbf{X}_{n-1} = \frac{\sum_{q=1}^Q |x_{n-1,q}|^2}{Q} \mathbf{I}_T$.

We note that P_t is introduced in order to guarantee $E[\text{tr}(\mathbf{X}_{n-1}^H \mathbf{X}_{n-1})] = T$. Consequently, as demonstrated in [56], the decision metric in (5) may be decoupled as:

$$\|\mathbf{Y}_n - \mathbf{X}_{n-1} \mathbf{Y}_{n-1}\|^2 = \|\mathbf{Y}_n\|^2 + 2\Re[\text{tr}(\mathbf{Y}_n^H \mathbf{X}_{n-1} \mathbf{Y}_{n-1})] + \text{tr}(\mathbf{Y}_{n-1}^H \mathbf{X}_{n-1}^H \mathbf{X}_{n-1} \mathbf{Y}_{n-1}) \quad (6a)$$

$$= \sum_{q=1}^Q \left(\frac{\|\mathbf{Y}_{n-1}\|^2}{Q} |\bar{z}_{n-1,q} - x_{n-1,q}|^2 \right) + \varrho. \quad (6b)$$

More explicitly, the first term $\|\mathbf{Y}_n\|^2$ in (6a) is a constant that does not depend on \mathbf{X}_{n-1} . The second term and third term of (6a) may be further decoupled into (6b) because of the two conditions $\mathbf{X}_{n-1} = \sqrt{P_t} \sum_{q=1}^Q (\mathbf{D}_q^+ x_{n-1,q} + \mathbf{D}_q^- x_{n-1,q}^*)$ and $\mathbf{X}_{n-1}^H \mathbf{X}_{n-1} = \frac{\sum_{q=1}^Q |x_{n-1,q}|^2}{Q} \mathbf{I}_T$ of Property 2, respectively. The equivalent pairs of dispersion matrices are given by

$\{\mathbf{D}_q^+ = \frac{1}{2}(\mathbf{A}_q + \mathbf{B}_q)\}_{q=1}^Q$ and $\{\mathbf{D}_q^- = \frac{1}{2}(\mathbf{A}_q - \mathbf{B}_q)\}_{q=1}^Q$. As a result, the decision variable of (6b) is given by $\bar{z}_{n-1,q} = \frac{Q\sqrt{P_t}}{\|\mathbf{Y}_{n-1}\|^2} \text{tr}[\mathbf{Y}_n^H \mathbf{D}_q^- \mathbf{Y}_{n-1} + \mathbf{Y}_{n-1}^H (\mathbf{D}_q^+)^H \mathbf{Y}_n]$, and the constant is given by $\varrho = \|\mathbf{Y}_n\|^2 - \sum_{q=1}^Q \frac{\|\mathbf{Y}_{n-1}\|^2}{Q} |\bar{z}_{n-1,q}|^2$. Therefore, the single-stream demodulation $\mathbb{M}^{-1}(\cdot)$ may be directly invoked for each data stream in (5) as:

$$\hat{x}_{n-1,q} = \mathbb{M}^{-1}(\bar{z}_{n-1,q}), \quad \forall q \in \{1, \dots, Q\}. \quad (7)$$

Taking LPSK as an example, (7) is $\hat{x}_{n-1,q} = \exp(j \frac{2\pi}{L} \tilde{l})$, where the phase of $\bar{z}_{n-1,q}$ is directly rounded as $\tilde{l} = \lfloor \frac{L}{2\pi} \angle \bar{z}_{n-1,q} \rfloor$.

According to (1), the DSTBC satisfies Property 2, where we have $P_t = \frac{1}{Q}$. The corresponding $\{\mathbf{A}_q, \mathbf{B}_q\}_{q=1}^Q$ and $\{\mathbf{D}_q^+, \mathbf{D}_q^-\}_{q=1}^Q$ are all real-valued sparse matrices, which only contain 0 and ± 1 . As a result, instead of matrix multiplications, the evaluation of the DSTBC's $\bar{z}_{n-1,q}$ in (6) only involves a total of T complex-valued vector multiplications for demodulating each $x_{n-1,q}$.

Example 2: For the case of DSTBC ($M = 2$) using PSK [10], the data-carrying matrix $\mathbf{X}_{n-1} = \begin{bmatrix} x_{n-1,1} & x_{n-1,2} \\ -x_{n-1,2}^* & x_{n-1,1}^* \end{bmatrix}$ satisfies Property 2. The dispersion matrices are given by $\mathbf{A}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\mathbf{B}_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $\mathbf{A}_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and $\mathbf{B}_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. The equivalent dispersion matrices are given by $\mathbf{D}_1^+ = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $\mathbf{D}_1^- = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, $\mathbf{D}_2^+ = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ and $\mathbf{D}_2^- = \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix}$. Due to the sparsity of $\{\mathbf{D}_q^+, \mathbf{D}_q^-\}_{q=1}^2$, the pair of decision variables in (6) are given by $\bar{z}_{n-1,1} = \frac{\sqrt{Q}}{\|\mathbf{Y}_{n-1}\|^2} [(\mathbf{Y}_n^H)^H \mathbf{Y}_{n-1}^2 + (\mathbf{Y}_{n-1}^1)^H \mathbf{Y}_n^1]$ and $\bar{z}_{n-1,2} = \frac{\sqrt{Q}}{\|\mathbf{Y}_{n-1}\|^2} [-(\mathbf{Y}_n^2)^H \mathbf{Y}_{n-1}^1 + (\mathbf{Y}_{n-1}^2)^H \mathbf{Y}_n^1]$, where $\{\mathbf{Y}_n^t\}_{t=1}^2$ and $\{\mathbf{Y}_{n-1}^t\}_{t=1}^2$ refer to the t -th row in \mathbf{Y}_n and \mathbf{Y}_{n-1} , respectively.

We note that the DGCs [14]–[16] and their variants [17]–[21] as well as the DLDCs [22]–[25] do not satisfy the first condition of Property 2. By contrast, DSTSK [31]–[33] and DSM [34], [35], [40], [41] do satisfy Property 2. The reduced-complexity DSTSK detection algorithms that do not impose any performance loss may be found in [28], [29], [32], [41]. Based on Property 2, we will also conceive the single-stream DSM detection in Sec. IV.

C. Error Probability and Capacity

The DD of (5) aims for detecting \mathbf{X}_{n-1} based on two observations \mathbf{Y}_{n-1} and \mathbf{Y}_n , hence its performance may be characterized by the following equivalent model:

$$\bar{\mathbf{Y}} = \bar{\mathbf{S}} \bar{\mathbf{H}} + \bar{\mathbf{V}}, \quad (8)$$

where the received signal matrix $\bar{\mathbf{Y}} = [\mathbf{Y}_{n-1}^T, \mathbf{Y}_n^T]^T$ and the AWGN matrix $\bar{\mathbf{V}} = [\mathbf{V}_{n-1}^T, \mathbf{V}_n^T]^T$ are both of size $(2T \times N)$, while the $(M \times N)$ -element fading matrix is given by $\bar{\mathbf{H}} = \mathbf{H}_{n-1} = \mathbf{H}_n$. The $(2T \times M)$ -element transmitted signal matrix $\bar{\mathbf{S}} = [\mathbf{S}_{n-1}^T, \mathbf{S}_n^T]^T$ in (8) forms the Unitary Space-Time Modulation (USTM) of [57], [58], where $\bar{\mathbf{S}}$ has orthogonal columns $\bar{\mathbf{S}}^H \bar{\mathbf{S}} = \mathbf{S}_{n-1}^H \mathbf{S}_{n-1} + \mathbf{S}_n^H \mathbf{S}_n = \frac{2T}{M} \mathbf{I}_M$

for the sake of maximizing the attainable capacity¹. This may be satisfied by \mathbf{X}_{n-1} in (3) using LPSK signalling and obeying the second condition of Property 2, which results in $\mathbf{S}_{n-1}^H \mathbf{S}_{n-1} = \mathbf{S}_n^H \mathbf{S}_n = \frac{T}{M} \mathbf{I}_M$. As a result, the non-coherent detection probability based on (8) is given by [57], [58]:

$$p(\bar{\mathbf{Y}}|\bar{\mathbf{S}}) = \frac{\exp\left\{-\text{tr}\left[\bar{\mathbf{Y}}^H (\bar{\mathbf{S}} \bar{\mathbf{S}}^H + N_0 \mathbf{I}_{2T})^{-1} \bar{\mathbf{Y}}\right]\right\}}{\pi^{2TN} \det(\bar{\mathbf{S}} \bar{\mathbf{S}}^H + N_0 \mathbf{I}_{2T})^N} \quad (9a)$$

$$= \frac{(\eta N_0)^{MN}}{(\pi N_0)^{2TN}} \exp\left(-\frac{\kappa_{n-1}^2 + \kappa_n^2}{N_0}\right) \exp\left(\frac{\eta}{N_0} \|(\mathbf{Y}_{n-1} + \mathbf{Y}_n^H \mathbf{X}_{n-1}) \mathbf{S}_{n-1}\|^2\right), \quad (9b)$$

where the constants are $\eta = \frac{M}{MN_0 + 2T}$, $\kappa_{n-1}^2 = \|\mathbf{Y}_{n-1}\|^2$ and $\kappa_n^2 = \|\mathbf{Y}_n\|^2$. Explicitly, according to $\det(\mathbf{I} + \mathbf{A}\mathbf{B}) = \det(\mathbf{I} + \mathbf{B}\mathbf{A})$, we have $\det(N_0 \mathbf{I}_{2T} + \bar{\mathbf{S}} \bar{\mathbf{S}}^H) = \frac{N_0^{2T}}{(\eta N_0)^M}$ in (9). According to $(\mathbf{A} + \mathbf{B}\mathbf{C}\mathbf{D})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}^{-1} + \mathbf{D}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{D}\mathbf{A}^{-1}$, we have $(N_0 \mathbf{I}_{2T} + \bar{\mathbf{S}} \bar{\mathbf{S}}^H)^{-1} = \frac{1}{N_0} \mathbf{I}_{2T} - \frac{\eta}{N_0} \bar{\mathbf{S}} \bar{\mathbf{S}}^H$. Based on (9), the Pairwise Error Probability (PEP) may be expressed as:

$$p(\mathbf{X}^i \rightarrow \mathbf{X}^{i'}) = \frac{1}{2\pi} \int_0^\infty \frac{4}{4\omega^2 + 1} \prod_{m=1}^M \left[1 + \frac{T^2 \lambda_m (4\omega^2 + 1)}{4MN_0(MN_0 + 2T)}\right]^{-N} d\omega, \quad (10)$$

where $\{\lambda_m\}_{m=1}^M$ are eigenvalues of the difference matrix $\Delta = 4(\mathbf{I}_M - \mathbf{D}\mathbf{D}^H)$, while we define $\mathbf{D} = \frac{M}{2T} (\bar{\mathbf{S}}^i)^H \bar{\mathbf{S}}^{i'} = \frac{1}{2} [\mathbf{I}_M + \frac{M}{T} \mathbf{S}_{n-1}^H (\mathbf{X}^i)^H \mathbf{X}^{i'} \mathbf{S}_{n-1}]$. Moreover, \mathbf{S}_{n-1} may be replaced by $\mathbf{S}_1 = \sqrt{\frac{T}{M}} [\mathbf{I}_M, \mathbf{0}]^T$ for evaluating (10), when $\mathbf{X}_{n-2} \cdots \mathbf{X}_1$ between \mathbf{S}_{n-1} and \mathbf{S}_1 are all unitary. For the case of $(M=T)$, the difference matrix becomes $\Delta = (\mathbf{X}^i - \mathbf{X}^{i'})^H (\mathbf{X}^i - \mathbf{X}^{i'})$, and (10) tends to $p(\mathbf{X}^i \rightarrow \mathbf{X}^{i'}) \leq (\frac{1}{8N_0})^{-\text{rank}(\Delta)N} \det(\Delta)^{-N}$ at high SNRs. Therefore, the trend of (10) may be characterized by the concept of *diversity product* [17], [55]:

$$\Lambda = \frac{1}{2} \min_{i \neq i'} \det(\Delta)^{\frac{1}{2M}}. \quad (11)$$

Following this, the average BER may be evaluated by [1]:

$$\bar{P}_{e, \text{bit}} \leq \mathbb{E} \left\{ \sum_{i=0}^{I-1} \sum_{i'=0}^{I-1} \frac{d_H(i, i')}{T \log_2 I} p(\mathbf{X}^i \rightarrow \mathbf{X}^{i'}) \right\}, \quad (12)$$

where $d_H(i, i')$ refers to the Hamming distance between the bit-mappings of \mathbf{X}^i and $\mathbf{X}^{i'}$.

In the absence of transmit diversity, the recently developed DSM scheme of [34], [35] does not perform well, when the number of RAs N is small. In particular, DSM may perform even worse than DPSK for the case of using $N = 1$, as seen in [35]. Against this background, the *third objective* of our design is to achieve a beneficial diversity gain for DSM without compromising its low transceiver complexity. We offer the following diversity gains of DSTBC and DGC as examples:

Example 3: According to the STBC structure of (1), the difference matrix is $\Delta = \frac{1}{Q} G_T^{\text{AO}} (\mathbf{x} - \mathbf{x}')^H G_T^{\text{AO}} (\mathbf{x} - \mathbf{x}') = \frac{\sum_{q=1}^Q |x_q - x'_q|^2}{Q} \mathbf{I}_T$. The minimum determinant in (11) is obtained by the closest codewords pair associated with $x_1 = 1$, $x'_1 = w_L$ and $\{x_q = x'_q\}_{q=2}^Q$, which results in the diversity product of $\Lambda = \frac{1}{2} \det\left(\frac{|1-w_L|^2}{Q} \mathbf{I}_T\right)^{\frac{1}{2T}} = \frac{1}{\sqrt{Q}} \sin \frac{\pi}{L}$ for the case of $(M=T)$.

¹The Continuous-input Continuous-output Memoryless Channel (CCMC) capacity given in [57], [58] corresponds to the maximized mutual information $I(\bar{\mathbf{S}}; \bar{\mathbf{Y}})$ for the idealistic continuous-input signals. In this treatise, we specifically investigate the finite-cardinality DSTBC schemes that have discrete-input signals $\{\bar{\mathbf{S}}^i\}_{i=0}^{I-1}$. The resultant Discrete-input Continuous-output Memoryless Channel (DCMC) capacity will be derived later in this section.

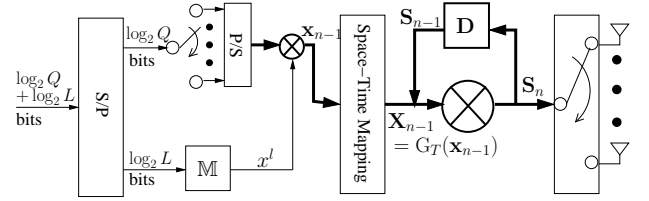


Fig. 4. Schematic of DSTBC-ISK, where the constellation of 8PSK is employed as an example.

Example 4: The DGC-cyclic of [14]–[17] constructs its data-carrying matrix as $\mathbf{X}^l = \mathbf{G}_c^l$, where the index l is Gray coded from $\log_2 L$ information bits, while the generator matrix is given by $\mathbf{G}_c = \text{diag}([w_L^{u_1}, w_L^{u_2}, \dots, w_L^{u_T}])$. The integer parameters $\{u_t\}_{t=1}^T$ are chosen for the sake of maximizing the diversity product of $\Lambda = \frac{1}{2} \min_{l \neq l'} \det[(\mathbf{I}_T - \mathbf{G}_c^l)^H (\mathbf{I}_T - \mathbf{G}_c^{l'})] = \min_{l \neq l'} \frac{1}{2} (\prod_{t=1}^T |1 - w_L^{u_t l}|)^{\frac{1}{T}} = \min_{l \neq l'} [\prod_{t=1}^T |\sin(\frac{\pi u_t l}{L})|]^{\frac{1}{T}}$. Similarly, DGC-dicyclic of [16], [17] is constructed by $\mathbf{X}^i = \mathbf{G}_c^{l_1} \mathbf{G}_d^{l_2}$, where Gray coded indices l_1 and l_2 carry $\log_2 L$ and 1 information bits, respectively, while the generator matrices are $\mathbf{G}_c = \text{diag}[\bar{\mathbf{G}}_c, \bar{\mathbf{G}}_c^*]$, $\bar{\mathbf{G}}_c = \text{diag}[w_L^{u_1}, w_L^{u_2}, \dots, w_L^{u_{T/2}}]$ and $\mathbf{G}_d = \begin{bmatrix} \mathbf{0} & -\mathbf{I}_{T/2} \\ \mathbf{I}_{T/2} & \mathbf{0} \end{bmatrix}$. The appropriate integer parameters $\{u_t\}_{t=1}^{T/2}$ are chosen to maximize its diversity product of $\Lambda = \min\{\frac{1}{\sqrt{2}}, \min_{l_1 \neq l_2} [\prod_{t=1}^{T/2} |\sin(\frac{\pi u_t l_1}{L})|]^2\}$. It is shown in [16], [17] that DGC-cyclic generally outperforms DGC-dicyclic, except for some exceptions at low rates.

Moreover, according to (9), the DCMC capacity [2], [59] may be expressed by (13) at the top of the next page. More explicitly, (13a) evaluates the mutual information between the discrete-input signal $\bar{\mathbf{S}}^i$ and the continuous-output signal $\bar{\mathbf{Y}}$, which is maximized for equiprobable sources $\{p(\bar{\mathbf{S}}^i) = \frac{1}{I}\}_{i=0}^{I-1}$. Then we further invoke the expression of $p(\bar{\mathbf{Y}}|\bar{\mathbf{S}}^i)$ in (9), which results in the DCMC capacity of (13c). The discrete-input signal $\bar{\mathbf{S}}^i$ in (13b) is unambiguously determined by the data-carrying matrix \mathbf{X}^i in (13c). Once again, \mathbf{S}_{n-1} may be replaced by $\mathbf{S}_1 = \sqrt{\frac{T}{M}} [\mathbf{I}_M, \mathbf{0}]^T$ for evaluating (13), when $\mathbf{X}_{n-2} \cdots \mathbf{X}_1$ between \mathbf{S}_{n-1} and \mathbf{S}_1 are all unitary.

III. DIFFERENTIAL SPACE-TIME BLOCK CODE - INDEX SHIFT KEYING

The schematic of DSTBC-ISK is portrayed in Fig. 4. More explicitly, DSTBC-ISK conveys a total of $(\log_2 Q + \log_2 L)$ bits, where $\log_2 Q$ bits are assigned to activate an ISK position index q in the DSTBC's signal vector \mathbf{x}_{n-1} seen in (1) as:

$$\mathbf{x}_{n-1} = [\underbrace{0 \cdots 0}_q, x^l, \underbrace{0 \cdots 0}_{Q-q-1}], \quad (14)$$

while $\log_2 L$ bits are mapped to the L -PSK symbol x^l . As a result, the DSTBC's space-time mapping of (3) becomes:

$$\mathbf{X}_{n-1} = G_T(\mathbf{x}_{n-1}) = \mathbf{A}_q \Re(x^l) + j \mathbf{B}_q \Im(x^l). \quad (15)$$

More explicitly, as demonstrated in Sec. II-B, the DSTBC's data matrix of (3) may be expressed as $\mathbf{X}_{n-1} = \sqrt{P_t} \sum_{q=1}^Q [\mathbf{A}_q \Re(x_{n-1,q}) + j \mathbf{B}_q \Im(x_{n-1,q})]$. The ISK index q activates a single position in the signal vector \mathbf{x}_{n-1} of (14), which is equivalent to activating the q -th pair of DSTBC dispersion matrices in (15), where we have $P_t = 1$ for the sake of $\text{tr}(\mathbf{X}_{n-1}^H \mathbf{X}_{n-1}) = T$. Following this, the differential

$$C_{\text{DDMC}} = \max_{\{\bar{\mathbf{S}}^i\}_{i=0}^{I-1}} \frac{1}{T} \sum_{i=0}^{I-1} \int p(\bar{\mathbf{Y}}|\bar{\mathbf{S}}^i) p(\bar{\mathbf{S}}^i) \log_2 \frac{p(\bar{\mathbf{Y}}|\bar{\mathbf{S}}^i) p(\bar{\mathbf{S}}^i)}{\sum_{i'=0}^{I-1} p(\bar{\mathbf{Y}}|\bar{\mathbf{S}}^{i'}) p(\bar{\mathbf{S}}^{i'})} d\bar{\mathbf{Y}} \quad (13a)$$

$$= \frac{\log_2 I}{T} - \frac{1}{I \cdot T} \sum_{i=0}^{I-1} \int p(\bar{\mathbf{Y}}|\bar{\mathbf{S}}^i) \log_2 \frac{\sum_{i'=0}^{I-1} p(\bar{\mathbf{Y}}|\bar{\mathbf{S}}^{i'})}{p(\bar{\mathbf{Y}}|\bar{\mathbf{S}}^i)} d\bar{\mathbf{Y}} = R - \frac{1}{I \cdot T} \sum_{i=0}^{I-1} \mathbb{E} \left[\log_2 \frac{\sum_{i'=0}^{I-1} p(\bar{\mathbf{Y}}|\bar{\mathbf{S}}^{i'})}{p(\bar{\mathbf{Y}}|\bar{\mathbf{S}}^i)} \right] \quad (13b)$$

$$= R - \frac{1}{I \cdot T} \sum_{i=0}^{I-1} \mathbb{E} \left\{ \log_2 \left[\sum_{i'=0}^{I-1} \exp \left(\frac{\eta}{N_0} \|(\mathbf{Y}_{n-1}^H + \mathbf{Y}_n^H \mathbf{x}^{i'}) \mathbf{S}_{n-1}\|^2 - \frac{\eta}{N_0} \|(\mathbf{Y}_{n-1}^H + \mathbf{Y}_n^H \mathbf{x}^i) \mathbf{S}_{n-1}\|^2 \right) \right] \right\}. \quad (13c)$$

encoding of (3) is invoked in Fig. 4. Owing to the fact that the DSTBC's dispersion matrices $\{\mathbf{A}_q, \mathbf{B}_q\}_{q=1}^Q$ are all real-valued sparse matrices that only have a single non-zero element of $\{\pm 1\}$ in each row and column, DSTBC-ISK complies with Property 1 of Sec. II-A, which facilitates both finite-cardinality and single-RF transmission, as portrayed in Fig. 4.

Furthermore, the DSTBC-ISK's \mathbf{X}_{n-1} in (15) also satisfies Property 2 of Sec. II-B. Therefore, the DD of (5) may be decoupled for DSTBC-ISK as:

$$\hat{\mathbf{X}}_{n-1} = \arg \min_{l,q} |z_{n-1,q} - x^l|^2, \quad (16)$$

where $z_{n-1,q} = \text{tr}[\mathbf{Y}_n^H \mathbf{D}_q^- \mathbf{Y}_{n-1} + \mathbf{Y}_{n-1}^H (\mathbf{D}_q^+)^H \mathbf{Y}_n]$. As a result, the single-stream L -PSK demodulator may be invoked for each index q as:

$$\hat{l}_q = \mathbb{M}^{-1}(z_{n-1,q}), \quad \forall q \in \{1, \dots, Q\}. \quad (17)$$

Then the optimum activation index q may be obtained by:

$$\hat{q} = \arg \min_{\forall q} |z_{n-1,q} - x^{\hat{l}_q}|^2 = \arg \max_{\forall q} \Re[(z_{n-1,q})^* z_{n-1,q}]. \quad (18)$$

Finally, the optimum PSK index is given by $\hat{l} = \hat{l}_{\hat{q}}$, which is the index obtained from (17) associated with \hat{q} . It can be seen that the DSTBC-ISK may invoke the decoupled DSTBC receiver, where the evaluation of $z_{n-1,q}$ only involves a total of T vector multiplications.

In contrast to the diversity schemes of DGC [14]–[17], DLDC [22]–[25], DSTSK [31]–[33] and dispersion matrix based DSM [40], [41], the proposed DSTBC-ISK using the classic STBC signal structure does not require optimization of signal parameters for attaining a high diversity gain. Although the parameter optimization can be carried out offline, maximizing the diversity product of (11) is generally a non-convex/concave problem, which implies that the global optimum may not necessarily be found. Nonetheless, in the absence of this parameter optimization, the proposed DSTBC-ISK may suffer from a performance loss compared to the diversity-gain-optimized DGC scheme at high throughputs, but DSTBC-ISK exhibits a substantially lower detection complexity. The performance versus complexity comparisons will be thoroughly examined in Sec. VI.

In the rest of this section, we present the details of the novel DSTBC-ISK design that aims for achieving the best attainable performance at the low-cost of using the classic low-complexity DSTBC transceiver, where the cases of $(T = 2)$ and $(T = 4)$ are exemplified in Sec. III-A and Sec. III-B, respectively, then the generalization is offered in Sec. III-C.

A. Spreading Over Two Time Slots ($T = 2$)

For the case of $(T = 2)$, the AO-STBC signal structure of (2) becomes Alamouti's G2 having $(Q = 2)$ [7]. As a result, DSTBC-ISK associated with $(T = 2)$ has $(Q = 2)$ ISK

Input bits	ISK index	QPSK symbol	Data-carrying matrix
000	$q = 1$	$x^0 = 1$	$\mathbf{X}_{n-1} = G_2^{\text{AO}}([1, 0]) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
001	$q = 1$	$x^1 = j$	$\mathbf{X}_{n-1} = G_2^{\text{AO}}([j, 0]) = \begin{bmatrix} j & 0 \\ 0 & -j \end{bmatrix}$
010	$q = 1$	$x^2 = -j$	$\mathbf{X}_{n-1} = G_2^{\text{AO}}([-j, 0]) = \begin{bmatrix} -j & 0 \\ 0 & j \end{bmatrix}$
011	$q = 1$	$x^3 = -1$	$\mathbf{X}_{n-1} = G_2^{\text{AO}}([-1, 0]) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$
100	$q = 2$	$x^0 = 1$	$\mathbf{X}_{n-1} = G_2^{\text{AO}}([0, 1]) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$
101	$q = 2$	$x^1 = j$	$\mathbf{X}_{n-1} = G_2^{\text{AO}}([0, j]) = \begin{bmatrix} 0 & j \\ j & 0 \end{bmatrix}$
110	$q = 2$	$x^2 = -j$	$\mathbf{X}_{n-1} = G_2^{\text{AO}}([0, -j]) = \begin{bmatrix} 0 & -j \\ -j & 0 \end{bmatrix}$
111	$q = 2$	$x^3 = -1$	$\mathbf{X}_{n-1} = G_2^{\text{AO}}([0, -1]) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

TABLE I
EXAMPLE OF DSTBC-ISK ($T = 2$) USING QPSK.

candidates $\mathbf{x}_{n-1} = [x^l, 0]$ and $\mathbf{x}_{n-1} = [0, x^l]$ for (14). The corresponding candidate matrices \mathbf{X}_{n-1} of (15) are given by:

$$\begin{aligned} q = 1: \quad G_2^{\text{AO}}([x^l, 0]) &= \begin{bmatrix} x^l & 0 \\ 0 & (x^l)^* \end{bmatrix}, \\ q = 2: \quad G_2^{\text{AO}}([0, x^l]) &= \begin{bmatrix} 0 & x^l \\ -(x^l)^* & 0 \end{bmatrix}, \end{aligned} \quad (19)$$

where the ISK index q is determined by the first input bit, while the LPSK index l is modulated by the following $\log_2 L$ bits. An example of DSTBC-ISK ($T = 2$) using QPSK is presented in Table I. Although the DSTBC-ISK of (19) appears to have the same constellation as the dicyclic DGC of [14], [16] associated with $(T = 2)$, we will demonstrate in Sec. VI that the DSTBC-ISK has a single-stream detection complexity that does not increase with constellation size.

Proposition 1: The diversity product of DSTBC-ISK ($M = T = 2$) using LPSK is given by $\Lambda = \min\{\sin(\pi/L), \frac{1}{\sqrt{2}}\}$.

Proof. According to (19), there are two candidates for Δ in (11), which are $|x - x'|^2 \mathbf{I}_2$ and $(|x|^2 + |x'|^2) \mathbf{I}_2 = 2\mathbf{I}_2$. For the first case, an instance of the closest constellation pair is given by $x = 1$ and $x' = w_L$, which results in $\frac{1}{2} \min_{\forall i \neq i'} \det(|1 - w_L|^2 \mathbf{I}_2)^{\frac{1}{4}} = \sin(\pi/L)$ in (11). The second case is a constant of $\frac{1}{2} \min_{\forall i \neq i'} \det(2\mathbf{I}_2)^{\frac{1}{4}} = \frac{1}{\sqrt{2}}$. \square

B. Spreading Over Four Time Slots ($T = 4$)

For the case of $(T = 4)$, the AO-STBC of (2) has $(Q = 3)$. As a result, DSTBC-ISK of (15) using (2) for $(T = 4)$ can only convey $\lceil \log_2 Q \rceil = 1$ bit on the ISK index q . Against this background, we propose to use the Quasi-Orthogonal (QO) STBC of [60], [61], so that a higher number of $(\log_2 Q = 2)$ bits may be conveyed by ISK, while the full diversity may be retained by invoking phase rotations [62]–[64]. More explicitly, DSTBC-ISK associated with $(T = 4)$ may invoke the QO-STBC structure of the so-called "ABBA" form [62], [63]:

$$G_4^{\text{QO}}([x_1, x_2, x_3, x_4]) = \begin{bmatrix} G_2^{\text{AO}}([x_1, x_2]) & G_2^{\text{AO}}([x_3, x_4]) \\ G_2^{\text{AO}}([x_3, x_4]) & G_2^{\text{AO}}([x_1, x_2]) \end{bmatrix}, \quad (20)$$

$G_4^{QO}([x^l, 0, 0, 0]) =$	$\begin{bmatrix} x^l & 0 & 0 & 0 \\ 0 & (x^l)^* & 0 & 0 \\ 0 & 0 & x^l & 0 \\ 0 & 0 & 0 & (x^l)^* \end{bmatrix}$
$G_4^{QO}([0, x^l, 0, 0]) =$	$\begin{bmatrix} 0 & x^l & 0 & 0 \\ -(x^l)^* & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -(x^l)^* & 0 \end{bmatrix}$
$G_4^{QO}([0, 0, x^l e^{j\theta}, 0]) =$	$\begin{bmatrix} 0 & 0 & x^l e^{j\theta} & 0 \\ 0 & 0 & 0 & (x^l e^{j\theta})^* \\ x^l e^{j\theta} & 0 & 0 & 0 \\ 0 & (x^l e^{j\theta})^* & 0 & 0 \end{bmatrix}$
$G_4^{QO}([0, 0, 0, x^l e^{j\theta}]) =$	$\begin{bmatrix} 0 & 0 & 0 & x^l e^{j\theta} \\ 0 & 0 & -(x^l e^{j\theta})^* & 0 \\ 0 & x^l e^{j\theta} & 0 & 0 \\ -(x^l e^{j\theta})^* & 0 & 0 & 0 \end{bmatrix}$

TABLE II

THE SET OF \mathbf{X}_{n-1} OF (15) CORRESPONDING TO INDICES $q \in \{0, 1, 2, 3\}$.

As a result, the set of \mathbf{X}_{n-1} of (15) corresponding to ISK indices of $q \in \{0, 1, 2, 3\}$ may be expressed as:

$$\begin{aligned} q = 1 : & G_4^{QO}([x^l, 0, 0, 0]) = \mathbf{I}_2 \otimes G_2^{AO}([x^l, 0]), \\ q = 2 : & G_4^{QO}([0, x^l, 0, 0]) = \mathbf{I}_2 \otimes G_2^{AO}([0, x^l]), \\ q = 3 : & G_4^{QO}([0, 0, x^l e^{j\theta}, 0]) = \mathbf{J}_2 \otimes G_2^{AO}([x^l e^{j\theta}, 0]), \\ q = 4 : & G_4^{QO}([0, 0, 0, x^l e^{j\theta}]) = \mathbf{J}_2 \otimes G_2^{AO}([0, x^l e^{j\theta}]), \end{aligned} \quad (21)$$

where \mathbf{J}_2 is a (2×2) -element anti-diagonal matrix with all anti-diagonal entries equal to 1, while a phase rotation of θ is introduced in (21) for achieving full diversity [63]. The detailed matrix expressions of (21) are listed in Table II.

Proposition 2: The diversity product of DSTBC-ISK ($M = T = 4$) using LPSK is $\Lambda = \min\{\sin(\pi/L), \frac{1}{\sqrt{2}}\sqrt{|\sin\theta|}\}$.

Proof. The matrices of (21) have three candidates for Δ in (11), which are $|x - x'|^2 \mathbf{I}_4$, $(|x|^2 + |x'|^2) \mathbf{I}_4$ and $2\mathbf{I}_4 - 2\Re\{\mathbf{J}_2 \otimes G_2^{AO}([x^* x' e^{j\theta}, 0])\}$. As same as the case of ($M = T = 2$) in Proposition 1, the first two candidates of $\Delta = |x - x'|^2 \mathbf{I}_4$ and $\Delta = (|x|^2 + |x'|^2) \mathbf{I}_4$ results in the diversity product candidates of $\sin(\pi/L)$ and $\frac{1}{\sqrt{2}}$, respectively. Moreover, the third candidate may be further expressed as $\Delta = 2\mathbf{I}_4 - 2\Re\{\mathbf{J}_2 \otimes G_2^{AO}([x^* x' e^{j\theta}, 0])\} = \begin{bmatrix} 2\mathbf{I}_2 & -G_2([2\Re(x^* x' e^{j\theta}), 0]) \\ -G_2([2\Re(x^* x' e^{j\theta}), 0]) & 2\mathbf{I}_2 \end{bmatrix}$. An instance of the closest constellation pair is given by $q = 1$, $q' = 3$ and $x = x'$, which results in the diversity product candidate of $\frac{1}{\sqrt{2}}\sqrt{|\sin\theta|}$. Since we have $\frac{1}{\sqrt{2}}\sqrt{|\sin\theta|} \leq \frac{1}{\sqrt{2}}$, the diversity product is given by $\Lambda = \min\{\sin(\pi/L), \frac{1}{\sqrt{2}}\sqrt{|\sin\theta|}\}$. \square

Since the phase rotation θ is effective in the interval $[0, \frac{\pi}{L}]$, it is sensible to choose $\theta = \frac{\pi}{L}$, which maximizes Λ , while retaining the equi-spaced PSK phases. The resultant transmitted signals in (3) are drawn from a $2L$ -PSK constellation. We note that although QO-STBC [60]–[64] is not unitary, DSTBC-ISK using the QO-STBC signal structure exemplified by Table II is unitary, as it satisfies Property 1 of Sec. II-A.

DSTBC-ISK ($T = 4$) may also invoke the low-complexity detector of (17) and (18), where the decision variables $\{z_{n-1,q}\}_{q=3}^4$ are replaced by their rotations $\{\tilde{z}_{n-1,q} = z_{n-1,q} e^{-j\theta_q}\}_{q=3}^4$ in order to take into account the phase rotation on input signal in (21).

C. Generalization

The generalization is based on the QO-STBC structure constructed recursively from Alamouti's $G_2^{AO}([x_1, x_2])$,

	$\bar{q}=0$	$\bar{q}=1$	$\bar{q}=2$	$\bar{q}=3$	$\bar{q}=4$	$\bar{q}=5$	$\bar{q}=6$	$\bar{q}=7$
$Q=4$	0	$\frac{\pi}{L}$						
$Q=8$	0	$\frac{\pi}{L}$	$\frac{\pi}{2L}$	$\frac{3\pi}{2L}$				
$Q=16$	0	$\frac{\pi}{L}$	$\frac{\pi}{2L}$	$\frac{3\pi}{2L}$	$\frac{\pi}{4L}$	$\frac{5\pi}{4L}$	$\frac{3\pi}{4L}$	$\frac{7\pi}{4L}$

TABLE III

THE ARRANGEMENT OF PHASE ROTATIONS $\{\theta_{\bar{q}}\}_{\bar{q}=0}^{Q/2-1}$ FOR DSTBC-ISK.

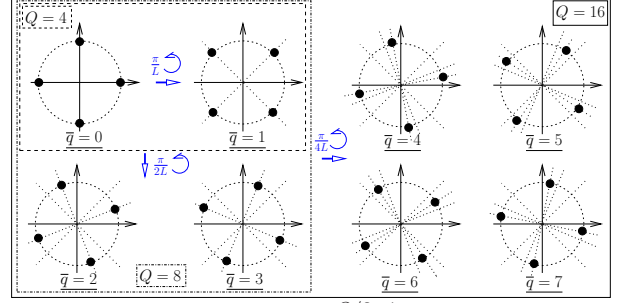


Fig. 5. Example of phase rotations $\{\theta_{\bar{q}}\}_{\bar{q}=0}^{Q/2-1}$ for DSTBC-ISK using ($Q = 8$) and QPSK.

so that the cases of ($T = 2$) and ($T = 4$) in Sec. III-A and Sec. III-B may be subsumed as special cases: [60]–[64]

$$G_T^{QO}([x_1, \dots, x_Q]) = \begin{bmatrix} G_{T/2}^{QO}([x_1, \dots, x_{Q/2}]) & G_{T/2}^{QO}([x_{Q/2+1}, \dots, x_Q]) \\ G_{T/2}^{QO}([x_{Q/2+1}, \dots, x_Q]) & G_{T/2}^{QO}([x_1, \dots, x_{Q/2}]) \end{bmatrix}, \quad (22)$$

which may start from ($T = 4$) using $G_2^{QO}([x_1, x_2]) = G_2^{AO}([x_1, x_2])$ as seen in (20). The QO-STBC of (22) offers ($Q = T$) for T being a power of two, and the DSTBC-ISK's data matrix of (15) is now given by $\mathbf{X}_{n-1} = G_T^{QO}(\mathbf{x}_{n-1})$. Furthermore, in order to retain full diversity, a $Q/2$ -element set of phase rotations $\{\theta_{\bar{q}}\}_{\bar{q}=0}^{Q/2-1}$ may be introduced into (14):

$$\mathbf{x}_{n-1} = [\underbrace{0 \dots 0}_q, x^l e^{j\theta_{\bar{q}}}, \underbrace{0 \dots 0}_{Q-q-1}]. \quad (23)$$

where the phase rotation index is given by $\bar{q} = \lceil \frac{q}{2} \rceil - 1$. In order to comply with Property 1 of Sec. II-A, instead of exploring all possible phase rotations [62], [63], we propose to recursively generate the equi-spaced set of $\{\theta_{\bar{q}}\}_{\bar{q}=0}^{Q/2-1}$. More explicitly, when ($Q = T$) is doubled from $Q/2$, the new $Q/2$ -element phase rotation set is given by all the $Q/4$ elements in the former set of $\{\theta_{\bar{q}}\}_{\bar{q}=0}^{Q/4-1}$, together with all the $Q/4$ of them adding an increment $\frac{4\pi}{QL}$, which starts from $\theta_0 = 0$ and $Q = 4$. For example, when G_4^{QO} is constructed based on G_2^{QO} according to (22), the new two-element phase rotations set is given by the former set of $\{\theta_0 = 0\}$ together with $\{\theta_0 = 0\} + \frac{\pi}{L}$, which results in $\{0, \frac{\pi}{L}\}$ as seen in Sec. III-B. Moreover, for the case of ($Q = 8$), the increment is $\frac{\pi}{2L}$, hence the new set is given by $\{0, \frac{\pi}{L}\}$ together with $\{0, \frac{\pi}{L}\} + \frac{\pi}{2L}$, which results in $\{0, \frac{\pi}{L}, \frac{\pi}{2L}, \frac{3\pi}{2L}\}$. The arrangement of phase rotations $\{\theta_{\bar{q}}\}_{\bar{q}=0}^{Q/2-1}$ is detailed in Table III, and an example for DSTBC-ISK using ($Q = 8$) and QPSK is portrayed by Fig. 5. As a result, the signals for \mathbf{X}_{n-1} and \mathbf{S}_n in (3) are now both drawn from a $QL/2$ -PSK constellation. As a result, when DSTBC-ISK ($T \geq 4$) invokes the low-complexity detector of (17) and (18), the decision variables $\{z_{n-1,q}\}_{q=1}^Q$ are replaced by their rotations of $\{\tilde{z}_{n-1,q} = z_{n-1,q} e^{-j\theta_{\bar{q}}}\}_{q=1}^Q$.

Proposition 3: The diversity product of generic DSTBC-ISK associated with ($M = T \geq 4$) using LPSK is given by $\Lambda = \min\{\sin(\pi/L), \frac{1}{\sqrt{2}}\sqrt{\sin(\frac{4\pi}{QL})}\}$.

DGC	Cyclic	$\Lambda = \min_{l \neq 0} \prod_{t=1}^T \sin(\frac{\pi}{L} u_t l) ^{\frac{1}{T}}$
	Dicyclic	$\Lambda = \min\{\frac{1}{\sqrt{2}}, \min_{l_1 \neq 0} \prod_{t=1}^{T/2} \sin(\frac{\pi}{L} u_t l_1) ^{\frac{2}{T}}\}$
DSTBC	$\Lambda = \frac{\sin(\frac{\pi}{L})}{\sqrt{Q}}$	
DSTBC-ISK	$T = 2$	$\Lambda = \min\{\sin(\pi/L), \frac{1}{\sqrt{2}}\}$
	$T \geq 4$	$\Lambda = \min\{\sin(\pi/L), \frac{1}{\sqrt{2}} \sqrt{\sin(\frac{4\pi}{QL})}\}$

TABLE IV

SUMMARY OF DIVERSITY PRODUCTS OF (11) FOR DGC, DSTBC AND DSTBC-ISK ASSOCIATED WITH $(M = T)$.

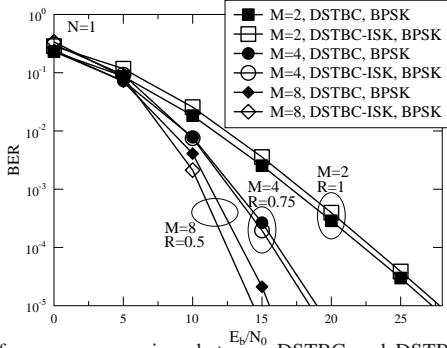


Fig. 6. Performance comparison between DSTBC and DSTBC-ISK using BPSK associated with $(M = T = \{2, 4, 8\})$ and $(N = 1)$.

Proof. In line with the QO-STBC structure of (22), the difference matrix Δ in (11) may also be expressed in the same "ABBA" form [62], [63] as $\Delta = \begin{bmatrix} \mathbf{E}_1 & \mathbf{E}_2 \\ \mathbf{E}_2 & \mathbf{E}_1 \end{bmatrix}$ for $G_{T/2}^{\text{QO}}(\mathbf{x}_{n-1})$, where \mathbf{E}_1 and \mathbf{E}_2 are both $(T/4 \times T/4)$ -element submatrices. When G_T^{QO} is constructed based on $G_{T/2}^{\text{QO}}$ according to (22), the updated difference matrix is given by $\Delta = \begin{bmatrix} \mathbf{F}_1 \otimes \mathbf{E}_1 & \mathbf{F}_2 \otimes \mathbf{E}_2 \\ \mathbf{F}_2 \otimes \mathbf{E}_2 & \mathbf{F}_1 \otimes \mathbf{E}_1 \end{bmatrix}$, where $\{\mathbf{F}_1, \mathbf{F}_2\} \in \{\mathbf{I}_2, \mathbf{J}_2\}$. This results in a squared value for $\det(\Delta)$, which is normalized to the same diversity product according to (11). As a result, the diversity product of DSTBC-ISK associated with $(M = T > 4)$ remains the same as Proposition 2. Taking into account all phase rotations, the updated diversity product is now given by $\Lambda = \min\{\sin(\frac{\pi}{L}), \min_{\forall q} \frac{1}{\sqrt{2}} \sqrt{|\sin \theta_q|}\} = \min\{\sin(\frac{\pi}{L}), \frac{1}{\sqrt{2}} \sqrt{\sin(\frac{4\pi}{QL})}\}$. \square

The diversity products of (11) for DGC, DSTBC and DSTBC-ISK are summarized in Table IV. It can be seen in Table IV that the diversity product of DSTBC-ISK may be even higher than that of DSTBC, when they employ the same L -PSK scheme. This feature is exemplified by Fig. 6, where DSTBC-ISK using BPSK may even outperform DSTBC using BPSK for $(M = 4)$ and $(M = 8)$. However, the DSTBC-ISK's throughput of $R = \frac{\log_2 LQ}{T}$ is generally lower than the DSTBC's throughput of $R = \frac{Q \log_2 L}{T}$, except for the BPSK case of Fig. 6. This implies that DSTBC-ISK has to employ higher-order L -PSK, when it is compared to its DSTBC counterpart at a higher throughput. Our detailed performance comparisons will be further discussed in Sec. VI.

It was demonstrated in [24] that the diversity product of (11) that features the worst PEP of (10) does not reflect the averaged BER accurately. Considering the BER expression of (12), we propose the concept of *average diversity product* as:

$$\bar{\Lambda} = \frac{1}{2} \left(\frac{2}{\log_2 T + 1} \sum_{\forall i} \sum_{\forall i' \neq i} \frac{d_H(i, i')}{T \log_2 T \det(\Delta)} \right)^{-\frac{1}{2M}}. \quad (24)$$

The averaged diversity products of DGC, DSTBC and DSTBC-ISK are compared in Fig. 7, where DSTBC-ISK

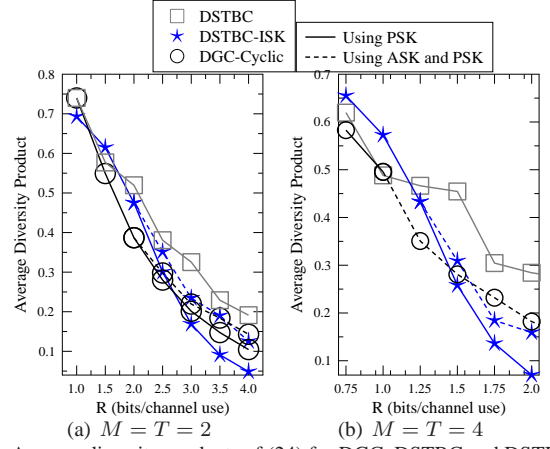


Fig. 7. Average diversity products of (24) for DGC, DSTBC and DSTBC-ISK associated with $(M = T = \{2, 4\})$.

$$\bar{\mathbf{A}}_{\bar{m},q} = \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{bmatrix}$$

$a_{\bar{m},q}$ th column

Fig. 8. Schematic of DSM dispersion matrix $\{\{\bar{\mathbf{A}}_{\bar{m},q}\}_{\bar{m}=1}^{\bar{M}}\}_{q=1}^Q$.

exhibits an even better diversity advantage than DGC-cyclic at low throughputs, but its average diversity product decays rapidly as the throughput increases. In order to circumvent this problem without compromising Property 1 of Sec. II-A, we will further introduce DASK/ASK modulation in Sec. V.

IV. SINGLE-STREAM DIFFERENTIAL SPATIAL MODULATION DETECTION

In order to provide a fair comparison to DSTBC-ISK, we devise a single-stream ML detector for the DSM schemes of [34], [35] in this section based on the fact that DSM is also linear and unitary as defined by Property 2 of Sec. II-B. The DSM scheme has $(M = T = Q)$. The transmitter firstly modulates a vector of $\{L_q\}_{q=1}^Q$ -PSK symbols $\{x^{l_q} = \exp(j\frac{2\pi}{L_q} \tilde{l}_q)\}_{q=1}^Q$ from a total of $\sum_{q=1}^Q \log_2 L_q$ bits. Moreover, a total of $\lceil \log_2 M! \rceil = \log_2 \bar{M}$ bits are assigned for determining the symbol activation order, which is represented by the activation sequence $\mathbf{a}_{\bar{m}} = [a_{\bar{m},1}, a_{\bar{m},2}, \dots, a_{\bar{m},Q}]$. The range for the permutation index \bar{m} is $1 \leq \bar{m} \leq \bar{M}$. According to Property 1 in Sec. II-A, the activation sequence obeys $1 \leq \{a_{\bar{m},q}\}_{q=1}^Q \leq M$ and $a_{\bar{m},1} \neq a_{\bar{m},2} \neq \dots \neq a_{\bar{m},Q}$. This presentation indicates that the $a_{\bar{m},q}$ -th element on the q -th row of the $(T \times T)$ -element data-carrying matrix \mathbf{X}_{n-1} in (3) is activated to transmit x^{l_q} . In summary, DSM modulates \mathbf{X}_{n-1} in (3) as:

$$\mathbf{X}_{n-1}(r, c) = \begin{cases} x^{l_q}, & \text{if } r = q \text{ and } c = a_{\bar{m},q} \\ 0, & \text{all the other elements} \end{cases}, \quad (25)$$

where $\mathbf{X}_{n-1}(r, c)$ denotes the element on the r -th row and c -th column in \mathbf{X}_{n-1} .

The data-carrying matrix \mathbf{X}_{n-1} of (25) may also be represented in the form of linear combinations of Q modulated symbols $\{x_{n-1,q}\}_{q=1}^Q$ and Q dispersion matrices $\{\bar{\mathbf{A}}_{\bar{m},q}\}_{q=1}^Q$:

$$\mathbf{X}_{n-1} = \sum_{q=1}^Q \bar{\mathbf{A}}_{\bar{m},q} x^{l_q}, \quad (26)$$

Permutation Index \bar{m}	Activation Sequence \mathbf{a}	Data-carrying matrix \mathbf{X}_{n-1}	Dispersion Matrices $\{\{\bar{\mathbf{A}}_{\bar{m},q}\}_{\bar{m}=1}^{\bar{M}}\}_{q=1}^Q$	
			$q=1$	$q=2$
$\bar{m}=1$	$\mathbf{a}_1=[1, 2]$	$\mathbf{X}_{n-1}=\begin{bmatrix} x^{l_1} & 0 \\ 0 & x^{l_2} \end{bmatrix}$	$\bar{\mathbf{A}}_{1,1}=\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$	$\bar{\mathbf{A}}_{1,2}=\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$
$\bar{m}=2$	$\mathbf{a}_2=[2, 1]$	$\mathbf{X}_{n-1}=\begin{bmatrix} 0 & x^{l_1} \\ x^{l_2} & 0 \end{bmatrix}$	$\bar{\mathbf{A}}_{2,1}=\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$	$\bar{\mathbf{A}}_{2,2}=\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$

TABLE V
EXAMPLE OF DSM ASSOCIATED WITH $M = 2$

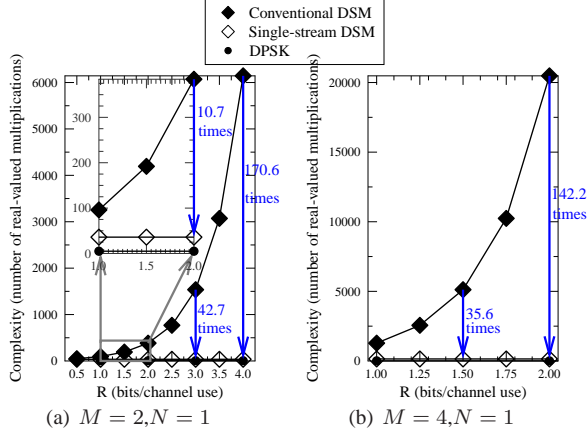


Fig. 9. Comparison of complexity (real-valued multiplications) of conventional DSM detector of (5) and single-stream DSM detector of (29) and (30). The complexity of the SISO's DPSK is also portrayed as benchmark.

where the PSK indices $\{l_q\}_{q=1}^Q$ and permutation index \bar{m} carries source information. The dispersion matrix is given by:

$$\bar{\mathbf{A}}_{\bar{m},q}(r, c) = \begin{cases} 1, & \text{if } r = q \text{ and } c = a_{\bar{m},q} \\ 0, & \text{all the other elements} \end{cases}, \quad (27)$$

which is virtualized in Fig. 8. The example of DSM associated with $(M = 2)$ is further characterized by Table V. According to (26), Property 2 in Sec. II-B is validated for DSM. Therefore, the analysis in (6) may also be applied to DSM, which simplifies the DD of (5) as:

$$\hat{\mathbf{X}}_{n-1} = \arg \max_{\forall \{l_q\}_{q=1}^Q, \forall \bar{m}} \sum_{q=1}^Q \Re \left[(x^{l_q})^* \text{tr} \left(\bar{\mathbf{A}}_{\bar{m},q}^H \mathbf{Z} \right) \right], \quad (28)$$

where $\mathbf{Z} = \mathbf{Y}_n \mathbf{Y}_{n-1}^H$, while the constants of $\kappa_{n-1}^2 = \|\mathbf{Y}_{n-1}\|^2$ and $\kappa_n^2 = \|\mathbf{Y}_n\|^2$ in (5) are omitted in (28). According to (27), we also have $\text{tr}(\bar{\mathbf{A}}_{\bar{m},q}^H \mathbf{Z}) = \mathbf{Z}(q, a_{\bar{m},q})$, where $\mathbf{Z}(q, a_{\bar{m},q})$ denotes the element on the q -th row and $a_{\bar{m},q}$ -th column in \mathbf{Z} . As a result, the single-stream L PSK demodulator may be invoked $\bar{M}Q$ times as:

$$\hat{l}_{\bar{m},q} = \mathbb{M}^{-1}(\mathbf{Z}(q, a_{\bar{m},q})), \quad \forall \bar{m} \in \{1, \bar{M}\}, \forall q \in \{1, \dots, Q\}. \quad (29)$$

Upon obtaining the PSK indices $\{\hat{l}_{\bar{m},q}\}_{\bar{m}=0}^{\bar{M}-1}\}_{q=1}^Q$, the optimum \hat{m} may be obtained by:

$$\hat{m} = \arg \max_{\forall \bar{m}} \sum_{q=1}^Q \Re \left[(x^{\hat{l}_{\bar{m},q}})^* \mathbf{Z}(q, a_{\bar{m},q}) \right]. \quad (30)$$

Finally, the optimum PSK indices $\{\hat{l}_q\}_{q=1}^Q$ are the ones in (29) associated with \hat{m} of (30) as $\hat{l}_q = \hat{l}_{\hat{m},q}\}_{q=1}^Q$.

Fig. 9 demonstrates complexity comparison between the conventional DSM detector of (5) and the single-stream DSM detector of (29) and (30) in terms of the total number of real-valued multiplications. More explicitly, the complexity of (5)

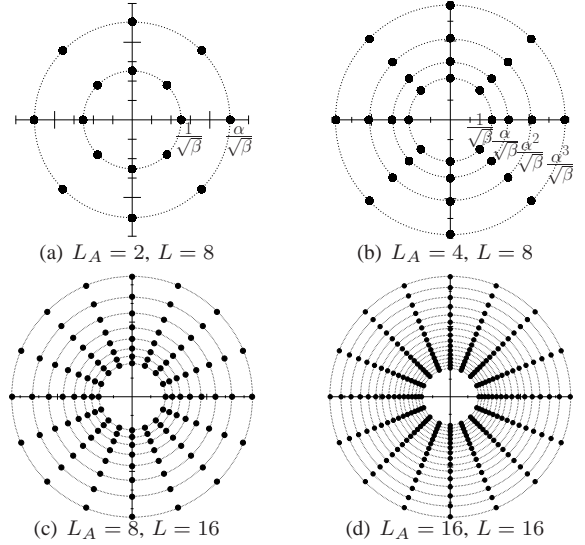


Fig. 10. Constellation diagrams for \tilde{L} -level star QAM associated with $L_A = \{2, 4, 8, 16\}$ -level ring-amplitudes and $L = \{8, 16\}$ -level PSK phases.

is given by² $[(4T^2N + 2TN)I]$, while the complexity of (29) and (30) is given by $(4M^2N + 5\bar{M}Q)$. It is confirmed by Fig. 9 that the complexity of the conventional DSM detector grows exponentially with throughput, while the complexity of the single-stream DSM detector has a linearized complexity that is in line with the SISO's DPSK complexity. The resultant complexity reductions achieved are as high as a factor of **170** and **142** for the cases of $(M = 2, N = 1, R = 4.0)$ and $(M = 4, N = 1, R = 2.0)$ in Figs. 9(a) and 9(b), respectively. In contrast to the solutions in [38], [39], the single-stream DSM detection of (29)-(30) is directly streamlined from the ML detector of (5), which does not impose any performance loss for any DSM scheme associated with all parameter values.

This optimal detection capability is especially important in channel coded scenarios [2], [28], [65], where the binary decisions are replaced by the soft-decision Log-Likelihood Ratios (LLRs) that reflect the Max-A-Posteriori (MAP) probabilities. More explicitly, sub-optimal detectors may produce large LLR values that deviate from the true probabilities, which are hard to correct by the channel decoder, especially after a few decoding iterations between the sub-optimal detector and the channel decoder.

V. DIFFERENTIAL AMPLITUDE SHIFT KEYING AND AMPLITUDE SHIFT KEYING

In order to improve the performance of the finite-cardinality DSTM schemes of DGC, DSM and the proposed DSTBC-ISK at high throughputs, in this section, we propose to extend the L PSK constellation used in the previous sections to the classic \tilde{L} -level star QAM constellation associated with L_A -level ring-amplitudes and L -level PSK phases, which is exemplified by Fig. 10. The encoding on star QAM's ring-amplitude is introduced in Sec. V-A. As a result of using star QAM signalling, the power of the DSTM's data-carrying matrix becomes variable depending on the data-carrying ring-amplitude. Against this background, we further develop be-

²We note that the DSM detector in [34], [35] is extended from (5) as $\hat{\mathbf{X}}_{n-1} = \arg \min_{\forall \mathbf{X}_{n-1}} \Re[\text{tr}(\mathbf{X}_{n-1} \mathbf{Y}_{n-1} \mathbf{Y}_n^H)]$, but its complexity of $(8T^2NI)$ is higher than that of (5).

Transmit signal matrix	$\tilde{\mathbf{S}}_n = \Gamma_n \mathbf{S}_n$
Transmit ring-amplitude	$\Gamma_n = \frac{\alpha^{\mu_n}}{\sqrt{\beta}}$
Transmit unitary signal matrix	\mathbf{S}_n
Data-carrying signal matrix	$\tilde{\mathbf{X}}_{n-1} = \gamma_{n-1} \mathbf{X}_{n-1} = \{\tilde{\mathbf{X}}^i\}_{i=0}^{I-1}$
Data-carrying ring-amplitude	$\gamma_{n-1} = \{\gamma^a\}_{a=0}^{L_A-1}$
Data-carrying unitary signal matrix	$\mathbf{X}_{n-1} = \{\mathbf{X}^i\}_{i=0}^{I-1}$

TABLE VI
NOTATIONS OF DSTM USING STAR QAM.

	Data-carrying ring-amplitude	Transmit ring-amplitude
DASK	$\gamma^a = \alpha^{[(\mu_{n-1} + \tilde{a}) \bmod L_A] - \mu_{n-1}}$	$\Gamma_n = \gamma_{n-1} \Gamma_{n-1}$
ASK	$\gamma^a = \frac{\alpha^{\tilde{a}}}{\sqrt{\beta}}$	$\Gamma_n = \gamma_{n-1}$

TABLE VII
SUMMARY ON DASK AND ASK.

spoke ML, MMSE and LS detectors for DSTM using star QAM in Secs. V-B, V-C and V-D, respectively.

A. Differential and Absolute Amplitude Encoding

For the sake of clarification, the notations of DSTM using star QAM are summarized in Table VI, where the superscript of a symbol/matrix (e.g. \mathbf{X}^i) denotes the data-carrying index mapped from the source information bits, while the subscript of a symbol/matrix (e.g. \mathbf{X}_{n-1}) denotes the temporal index. More explicitly, the transmit-signal matrix is now given by $\tilde{\mathbf{S}}_n = \Gamma_n \mathbf{S}_n$, where \mathbf{S}_n is the PSK-aided unitary transmit-signal matrix of (3) using PSK signalling. The extra term is the ring-amplitude of $\Gamma_n = \frac{\alpha^{\mu_n}}{\sqrt{\beta}}$, where α and $\beta = \frac{\sum_{\mu=0}^{L_A-1} \alpha^{2\mu}}{L_A}$ respectively represent the ring ratio and the associated normalization factor. In this paper, the ring-amplitude ratios of $\alpha = \{2.0, 1.4, 1.2, 1.1\}$ are used for $L_A = \{2, 4, 8, 16\}$, respectively. Moreover, the data-carrying matrix is now given by $\tilde{\mathbf{X}}_{n-1} = \gamma_{n-1} \mathbf{X}_{n-1}$, where $\log_2 I$ source bits are assigned to encode the PSK-aided unitary data-carrying matrix $\mathbf{X}_{n-1} = \{\mathbf{X}^i\}_{i=0}^{I-1}$ in (3), while $\log_2 L_A$ bits are assigned to modulate a ring-amplitude $\gamma_{n-1} = \{\gamma^a\}_{a=0}^{L_A-1}$. As a result, there are a total of $\tilde{I} = L_A I$ combinations for the data-carrying matrix $\tilde{\mathbf{X}}_{n-1} = \{\tilde{\mathbf{X}}^i\}_{i=0}^{\tilde{I}-1}$, and the throughput is increased to $R = \frac{\log_2 \tilde{I}}{T} = \frac{\log_2 I + \log_2 L_A}{T}$.

In order to guarantee that all the transmitted signals in $\tilde{\mathbf{S}}_n$ are drawn from the equi-spaced star QAM constellation, we propose for the DSTM scheme to employ the ring-amplitude encoding methods of DASK and ASK in [11], [36], [37], which are summarized in Table VII. More explicitly, for DSTM using DASK, differential encoding is applied both to the unitary matrices and to the ring-amplitudes by:

$$\tilde{\mathbf{S}}_n = \tilde{\mathbf{X}}_{n-1} \tilde{\mathbf{S}}_{n-1}, \quad (31)$$

which may start from $\tilde{\mathbf{S}}_1 = \sqrt{\frac{T}{\beta M}} [\mathbf{I}_M, \mathbf{0}]^T$. The differential encoding applied to the unitary matrices in (31) is still given by (3) as $\mathbf{S}_n = \mathbf{X}_{n-1} \mathbf{S}_{n-1}$. Moreover, the differential encoding applied to the ring-amplitudes in (31) is the same as in classic DAPSK [3]–[5], which is given by $\Gamma_n = \gamma_{n-1} \Gamma_{n-1}$. The L_A -level transmitted ring-amplitude and data-carrying ring-amplitude are given by $\Gamma_n = \frac{\alpha^{\mu_n}}{\sqrt{\beta}}$ and $\gamma_{n-1} = \alpha^{\tilde{a}}$, respectively. According to $\Gamma_n = \gamma_{n-1} \Gamma_{n-1}$, the ring-amplitude indices have the relationship of $[\mu_n = (\mu_{n-1} + \tilde{a}) \bmod L_A]$, where the data-carrying index a is Gray coded \tilde{a} . As a result, the modulation of $\gamma_{n-1} = \{\gamma^a\}_{a=0}^{L_A-1}$ is determined both by the

data-carrying index a as well as by the previous transmitted ring-amplitude $\Gamma_{n-1} = \frac{\alpha^{\mu_{n-1}}}{\sqrt{\beta}}$, which may be expressed as $\gamma^a = \alpha^{[(\mu_{n-1} + \tilde{a}) \bmod L_A] - \mu_{n-1}}$.

Alternatively, for DSTM using ASK, the differential encoding of (31) may be revised as:

$$\tilde{\mathbf{S}}_n = \frac{1}{\Gamma_{n-1}} \tilde{\mathbf{X}}_{n-1} \tilde{\mathbf{S}}_{n-1}, \quad (32)$$

where the differential encoding of the unitary matrices is also given by (3), while the absolute-amplitude is invoked in the same way as in ADPSK [4]–[6]. More explicitly, γ_{n-1} in $\tilde{\mathbf{X}}_{n-1} = \gamma_{n-1} \mathbf{X}_{n-1}$ of (32) is an L_A -level data-carrying ring-amplitude formulated as $\gamma^a = \frac{\alpha^{\tilde{a}}}{\sqrt{\beta}}$, and then thanks to the normalization of $\frac{1}{\Gamma_{n-1}}$ in (32), the transmitted ring-amplitudes always assume the absolute-amplitude of $\Gamma_n = \gamma_{n-1}$.

B. Maximum-Likelihood Differential Detection (ML-DD)

The received signal model of (4) may be revised for DSTM using star QAM as:

$$\mathbf{Y}_n = \tilde{\mathbf{S}}_n \mathbf{H}_n + \mathbf{V}_n, \quad (33)$$

As a result, for the equivalent received signal model of (8), we have $\bar{\mathbf{S}} = [\tilde{\mathbf{S}}_{n-1}^T, \tilde{\mathbf{S}}_n^T]^T$ and hence $\bar{\mathbf{S}}^H \bar{\mathbf{S}} = \tilde{\mathbf{S}}_{n-1}^H \tilde{\mathbf{S}}_{n-1} + \tilde{\mathbf{S}}_n^H \tilde{\mathbf{S}}_n = (\Gamma_{n-1}^2 + \Gamma_n^2) \frac{T}{M} \mathbf{I}_M$, which results in a variable power of $\text{tr}(\bar{\mathbf{S}}^H \bar{\mathbf{S}}) = (\Gamma_{n-1}^2 + \Gamma_n^2) T$ depending on ring-amplitudes Γ_{n-1} and Γ_n . Furthermore, the probability $p(\bar{\mathbf{Y}}|\bar{\mathbf{S}})$ of (9a) may be revised for DSTM using DASK and ASK as:

$$p(\bar{\mathbf{Y}}|\bar{\mathbf{S}}) = \frac{(\eta N_0)^{MN}}{(\pi N_0)^{2TN}} \exp\left(-\frac{\kappa_{n-1}^2 + \kappa_n^2}{N_0}\right) \exp\left(-\frac{\eta}{N_0} \|\tilde{\mathbf{S}}_{n-1}^H (\mathbf{Y}_{n-1} + \tilde{\mathbf{X}}_{n-1} \mathbf{Y}_n)\|^2\right), \quad (34a)$$

$$p(\bar{\mathbf{Y}}|\bar{\mathbf{S}}) = \frac{(\eta N_0)^{MN}}{(\pi N_0)^{2TN}} \exp\left(-\frac{\kappa_{n-1}^2 + \kappa_n^2}{N_0}\right) \exp\left(-\frac{\eta}{N_0} \|\tilde{\mathbf{S}}_{n-1}^H (\mathbf{Y}_{n-1} + \frac{1}{\Gamma_{n-1}} \tilde{\mathbf{X}}_{n-1} \mathbf{Y}_n)\|^2\right), \quad (34b)$$

respectively, where $\eta = \frac{M}{MN_0 + (\Gamma_{n-1}^2 + \Gamma_n^2) T}$ is now a variable. Then the average BER of (12) may be rewritten for DSTM using DASK and ASK as:

$$\bar{P}_{e, \text{bit}} \leq \mathbb{E} \left\{ \sum_{i=0}^{\tilde{I}-1} \sum_{i'=0}^{\tilde{I}-1} \sum_{\mu_{n-1}=0}^{L_A-1} \frac{d_H(\tilde{i}, \tilde{i}')}{L_A \tilde{I} \log_2 \tilde{I}} p(\bar{\mathbf{S}}^{\mu_{n-1}, \tilde{i}} \rightarrow \bar{\mathbf{S}}^{\mu_{n-1}, \tilde{i}'}), \right\} \quad (35)$$

where the equivalent transmit-signal matrix $\bar{\mathbf{S}}$ is now determined by both the previous transmitted ring-amplitude $\Gamma_{n-1} = \frac{\alpha^{\mu_{n-1}}}{\sqrt{\beta}}$ and the data-carrying index \tilde{i} . The PEP $p(\bar{\mathbf{S}}^{\mu_{n-1}, \tilde{i}} \rightarrow \bar{\mathbf{S}}^{\mu_{n-1}, \tilde{i}'})$ in (35) is still expressed by the right hand side of (10), where we have $\mathbf{D} = \frac{M}{2T} (\bar{\mathbf{S}}^{\mu_{n-1}, \tilde{i}})^H \bar{\mathbf{S}}^{\mu_{n-1}, \tilde{i}'}$. Based on this PEP, we offer the following two propositions concerning the performance of DSTM using star QAM:

Proposition 4: The PEP $p(\bar{\mathbf{S}}^{\mu_{n-1}, \tilde{i}} \rightarrow \bar{\mathbf{S}}^{\mu_{n-1}, \tilde{i}'})$ of DSTM using DASK is as same as that of DSTM using ASK.

Proof. The difference matrix that determines the PEP of (10) is given by $\mathbf{\Delta} = 4[\mathbf{I}_M - \mathbf{D}\mathbf{D}^H]$. For DSTM using DASK, we have $\mathbf{D} = \frac{1}{2}(\Gamma_{n-1}^2) \mathbf{I}_M + \frac{M}{2T}(\Gamma_{n-1} \gamma^a)(\Gamma_{n-1} \gamma^{a'}) \mathbf{S}_1^H (\mathbf{X}^i)^H \mathbf{X}^{i'} \mathbf{S}_1$. For DSTM using ASK, we have $\mathbf{D} = \frac{1}{2}(\Gamma_{n-1}^2) \mathbf{I}_M + \frac{M}{2T} \gamma^a \gamma^{a'} \mathbf{S}_1^H (\mathbf{X}^i)^H \mathbf{X}^{i'} \mathbf{S}_1$. Owing to the fact that DASK's $\{\Gamma_{n-1} \gamma^a =$

$\frac{\alpha^{(\mu_{n-1}+\hat{a}) \bmod L_A}}{\sqrt{\beta}}\}_{\forall a}$ and ASK's $\{\gamma^a = \frac{\alpha^a}{\sqrt{\beta}}\}_{\forall a}$ results in the same set of ring-amplitudes, DSTM using DASK and DSTM using ASK have exactly the same PEP values. \square

However, the different bit-to-ring-amplitude mapping of DASK and ASK may result in different $\bar{P}_{e,bit}$ of (35). Nonetheless, for the mapping of the data-carrying matrix $\tilde{\mathbf{X}}^i = \gamma^a \mathbf{X}^i$, the $\log_2 L_A$ number of bits assigned to ring-amplitude γ^a is generally smaller than the $\log_2 I$ bits mapped to the PSK-aided unitary matrix \mathbf{X}^i , hence the BER performance results for DSTM using DASK and ASK are expected to be similar without the assistance of channel coding.

Proposition 5: An optimized diversity product for a DSTM scheme using PSK also results in an optimized diversity product³ for its star QAM counterpart of DSTM using DASK/ASK.

Proof. Since DSTM using DASK and ASK have the same PEP according to Proposition 4, let's now consider DSTM using ASK associated with ($M = T$). The difference matrix that determines the diversity product Λ of (11) may be extended as $\Delta = [4 - (\Gamma_{n-1}^2 + \gamma^a \gamma^{a'})^2] \mathbf{I}_T + \Gamma_{n-1}^2 \gamma^a \gamma^{a'} \Delta_0$, where $\Delta_0 = (\mathbf{X}^i - \mathbf{X}^{i'})^H (\mathbf{X}^i - \mathbf{X}^{i'})$ is the difference matrix used for its PSK aided DSTM counterpart in Sec. II-C. Therefore, the difference matrices Δ and Δ_0 are in direct proportion, which results in the same relationship for their diversity products. This implies that the parameters used for optimizing DGC using PSK in [14]–[17] may be directly used for star QAM. \square

In summary, the ML-DD aims for maximizing the probability of (34a) for DSTM using DASK, which is equivalent to maximizing the following decision metric:

$$\tilde{d}_{ML}^i \leq \frac{\eta^a}{N_0} \{ \Gamma_{n-1}^2 \kappa_{n-1}^2 + (\Gamma_{n-1} \gamma^a)^2 \kappa_n^2 + 2\Gamma_{n-1} \gamma^a \Re[\text{tr}(\mathbf{X}^i \mathbf{Y}_{n-1} \mathbf{Y}_n^H)] \} + MN \ln \eta^a, \quad (36)$$

where the equality holds for ($M = T$), while we have $\eta^a = \frac{M}{MN_0 + [\Gamma_{n-1}^2 + (\Gamma_{n-1} \gamma^a)^2]T}$. We discovered in [66] that the detection of the DAPSK's phase is independent of the ring-amplitude detection. Similarly, the detection of the unitary matrix \mathbf{X}^i in (36) may be carried out independently as:

$$\hat{i} = \arg \max_{\forall i} \tilde{d}^i, \quad \text{where } \tilde{d}^i = \Re[\text{tr}(\mathbf{X}^i \mathbf{Y}_{n-1} \mathbf{Y}_n^H)]. \quad (37)$$

Upon obtaining the maximum \tilde{d}^i from (37), the ring-amplitude index detection may proceed as:

$$\hat{a} = \arg \max_{\forall a} \frac{\eta^a}{N_0} [\Gamma_{n-1}^2 \kappa_{n-1}^2 + (\Gamma_{n-1} \gamma^a)^2 \kappa_n^2 + 2\Gamma_{n-1} \gamma^a \tilde{d}^i] + MN \ln \eta^a. \quad (38)$$

The updated $\Gamma_n = \gamma^{\hat{a}} \Gamma_{n-1}$ may be passed on to the next ML-DD block as the decision-feedback on Γ_{n-1} .

Similarly, based on (34b), the ML-DD aided DSTM using ASK aims for maximizing the following decision metric:

$$\tilde{d}_{ML}^i \leq \frac{\eta^a}{N_0} \{ \Gamma_{n-1}^2 \kappa_{n-1}^2 + (\gamma^a)^2 \kappa_n^2 + 2\Gamma_{n-1} \gamma^a \Re[\text{tr}(\mathbf{X}^i \mathbf{Y}_{n-1} \mathbf{Y}_n^H)] \} + MN \ln \eta^a, \quad (39)$$

³For DSTM using star QAM, the expression of diversity product Λ of (11) and that of the average diversity product $\bar{\Lambda}$ of (24) may be revised by replacing i, i' and I by \tilde{i}, \tilde{i}' and \tilde{I} , respectively.

where we have $\eta^a = \frac{M}{MN_0 + [\Gamma_{n-1}^2 + (\gamma^a)^2]T}$. As a result, the detection of the unitary matrix \mathbf{X}^i is also given by (37), while the detection of the ring-amplitude γ^a may be rewritten as:

$$\hat{a} = \arg \max_{\forall a} \frac{\eta^a}{N_0} [\Gamma_{n-1}^2 \kappa_{n-1}^2 + (\gamma^a)^2 \kappa_n^2 + 2\Gamma_{n-1} \gamma^a \tilde{d}^i] + MN \ln \eta^a. \quad (40)$$

The updated decision concerning $\Gamma_n = \gamma^{\hat{a}}$ may be fed back to the next ML-DD block as Γ_{n-1} .

Since the detection of the unitary matrix \mathbf{X}^i and the detection of the ring-amplitude γ^a are separated, the single-stream DSTBC-ISK detection of Sec. III and the single-stream DSM detection of Sec. IV may both be directly applied for detecting \mathbf{X}^i in (37).

Following the same steps as (13), the DCMC capacity of DSTM using DASK/ASK derived based on the conditional probabilities of (34) is expressed in (41) at the top of next page, where we have $\bar{\mathbf{S}}^{\mu_{n-1}, \tilde{i}} = [\Gamma_{n-1} \mathbf{S}_1^T, \Gamma_{n-1} (\tilde{\mathbf{X}}^i \mathbf{S}_1)^T]^T$ for DSTM using DASK and $\bar{\mathbf{S}}^{\mu_{n-1}, i} = [\Gamma_{n-1} \mathbf{S}_1^T, (\tilde{\mathbf{X}}^i \mathbf{S}_1)^T]^T$ for DSTM using ASK. However, in fact, (41) can only be applied to DSTM using DASK. For the case of DSTM using ASK, the dominant metric in (41) at high SNRs becomes $\frac{\eta^{\mu_{n-1}, \tilde{i}}}{N_0} \|(\bar{\mathbf{S}}^{\mu_{n-1}, \tilde{i}})^H \bar{\mathbf{Y}}\|^2 \approx \frac{M}{[1 + (\frac{\gamma^a}{\Gamma_{n-1}})^2] \cdot T \cdot N_0} \|\mathbf{S}_{n-1}^H [\mathbf{Y}_{n-1} + \frac{\gamma^a}{\Gamma_{n-1}} (\mathbf{X}^i)^H \mathbf{Y}_n]\|^2$, where $\Gamma_{n-1} = \gamma^a = \frac{1}{\sqrt{\beta}}$ and $\Gamma_{n-1} = \gamma^a = \frac{\alpha}{\sqrt{\beta}}$ would result in the same metric, but they correspond to different bit-mappings. Therefore, the maximized DCMC capacity for DSTM using ASK is given by the situation of perfect decision-feedback concerning $\Gamma_{n-1} = \frac{\alpha^{\mu_{n-1}}}{\sqrt{\beta}}$ in (42).

C. Minimum Mean Squared Error Differential Detection (MMSE-DD)

Alternative to (9), the MMSE model of [67] may also be applied to (8), where the coherent probability based on (8) may be expressed as:

$$p(\bar{\mathbf{Y}}|\bar{\mathbf{S}}, \bar{\mathbf{H}}) = \frac{\exp(-\|\bar{\mathbf{Y}} - \bar{\mathbf{S}} \bar{\mathbf{H}}\|^2 / N_0)}{(\pi N_0)^{2TN}}. \quad (43)$$

As a result, maximizing (43) over $\{\tilde{\mathbf{X}}^i\}_{i=0}^{\tilde{I}-1}$ is equivalent to minimizing $d_{\text{MMSE}} = \|\mathbf{Y}_{n-1} - \tilde{\mathbf{S}}_{n-1} \bar{\mathbf{H}}\|^2 + \|\mathbf{Y}_n - \tilde{\mathbf{S}}_n \bar{\mathbf{H}}\|^2$.

For DSTM using DASK, the decision metric of (43) may be extended as $d_{\text{MMSE}}^i = \kappa_{n-1}^2 - 2\Re[\text{tr}(\bar{\mathbf{H}}^H \tilde{\mathbf{S}}_{n-1}^H \mathbf{Y}_{n-1})] + \text{tr}(\bar{\mathbf{H}}^H \tilde{\mathbf{S}}_{n-1}^H \tilde{\mathbf{S}}_{n-1} \bar{\mathbf{H}}) + \kappa_n^2 - 2\Re[\text{tr}(\bar{\mathbf{H}}^H \tilde{\mathbf{S}}_{n-1}^H (\tilde{\mathbf{X}}^i)^H \mathbf{Y}_n)] + (\gamma^a)^2 \text{tr}(\bar{\mathbf{H}}^H \tilde{\mathbf{S}}_{n-1}^H \tilde{\mathbf{S}}_{n-1} \bar{\mathbf{H}})$. Since $\tilde{\mathbf{S}}_{n-1} \bar{\mathbf{H}}$ is unknown, an MMSE estimate may be obtained by taking $\frac{\partial d_{\text{MMSE}}^i}{\partial \tilde{\mathbf{S}}_{n-1} \bar{\mathbf{H}}} = 0$, which leads to $\tilde{\mathbf{S}}_{n-1} \bar{\mathbf{H}} = \frac{1}{\tau_{\text{DASK}}^a} [\mathbf{Y}_{n-1} + (\tilde{\mathbf{X}}^i)^H \mathbf{Y}_n]$, where we define $\tau_{\text{DASK}}^a = 1 + (\gamma^a)^2$. Using this MMSE estimate of $\tilde{\mathbf{S}}_{n-1} \bar{\mathbf{H}}$, the decision metric of (43) becomes:

$$\begin{aligned} \tilde{d}_{\text{MMSE}}^i &= \frac{1}{\tau_{\text{DASK}}^a} \|\mathbf{Y}_n - \tilde{\mathbf{X}}^i \mathbf{Y}_{n-1}\|^2 \\ &= \frac{\kappa_n^2}{\tau_{\text{DASK}}^a} + \frac{(\gamma^a)^2 \kappa_{n-1}^2}{\tau_{\text{DASK}}^a} - \frac{2\gamma^a}{\tau_{\text{DASK}}^a} \Re[\text{tr}(\mathbf{X}^i \mathbf{Y}_{n-1} \mathbf{Y}_n^H)]. \end{aligned} \quad (44)$$

As a result, the detection of the unitary matrix \mathbf{X}^i is also given by (37). Upon obtaining the maximum \tilde{d}^i from (37), the detection of the ring-amplitude γ^a is formulated as:

$$\hat{a} = \arg \min_{\forall a} \frac{\kappa_n^2}{\tau_{\text{DASK}}^a} + \frac{(\gamma^a)^2 \kappa_{n-1}^2}{\tau_{\text{DASK}}^a} - \frac{2\gamma^a}{\tau_{\text{DASK}}^a} \tilde{d}^i. \quad (45)$$

$$C_{\text{DCMC}}^{\text{DASK}} = \frac{1}{L_A \cdot T} \sum_{\tilde{\mathbf{Y}}_i} \sum_{\mu_{n-1}} \mathbb{E} \left\{ \log_2 \left[\frac{\tilde{I} \sum_{\mu'_{n-1}} (\eta^{\mu'_{n-1}, \tilde{\mathbf{Y}}_i})^{MN} \exp(\frac{\eta^{\mu'_{n-1}, \tilde{\mathbf{Y}}_i}}{N_0} \|(\tilde{\mathbf{S}}^{\mu'_{n-1}, \tilde{\mathbf{Y}}_i})^H \tilde{\mathbf{Y}}\|^2)}{\sum_{\mu'_{n-1}} \sum_{\mu'_{n-1}} (\eta^{\mu'_{n-1}, \tilde{\mathbf{Y}}_i})^{MN} \exp(\frac{\eta^{\mu'_{n-1}, \tilde{\mathbf{Y}}_i}}{N_0} \|(\tilde{\mathbf{S}}^{\mu'_{n-1}, \tilde{\mathbf{Y}}_i})^H \tilde{\mathbf{Y}}\|^2)} \right] \right\}, \quad (41)$$

$$C_{\text{DCMC}}^{\text{ASK}} = \frac{1}{L_A \cdot T} \sum_{\tilde{\mathbf{Y}}_i} \sum_{\mu_{n-1}} \mathbb{E} \left\{ \log_2 \left[\frac{\tilde{I} (\eta^{\mu_{n-1}, \tilde{\mathbf{Y}}_i})^{MN} \exp(\frac{\eta^{\mu_{n-1}, \tilde{\mathbf{Y}}_i}}{N_0} \|(\tilde{\mathbf{S}}^{\mu_{n-1}, \tilde{\mathbf{Y}}_i})^H \tilde{\mathbf{Y}}\|^2)}{\sum_{\mu'_{n-1}} (\eta^{\mu'_{n-1}, \tilde{\mathbf{Y}}_i})^{MN} \exp(\frac{\eta^{\mu'_{n-1}, \tilde{\mathbf{Y}}_i}}{N_0} \|(\tilde{\mathbf{S}}^{\mu'_{n-1}, \tilde{\mathbf{Y}}_i})^H \tilde{\mathbf{Y}}\|^2)} \right] \right\} \quad (42)$$

$$= R - \frac{1}{L_A \cdot T} \sum_{\tilde{\mathbf{Y}}_i} \sum_{\mu_{n-1}} \mathbb{E} \{ \log_2 [\sum_{\mu'_{n-1}} (\eta^{\mu'_{n-1}, \tilde{\mathbf{Y}}_i})^{MN} \exp(\frac{\eta^{\mu'_{n-1}, \tilde{\mathbf{Y}}_i}}{N_0} \|(\tilde{\mathbf{S}}^{\mu'_{n-1}, \tilde{\mathbf{Y}}_i})^H \tilde{\mathbf{Y}}\|^2 - \frac{\eta^{\mu_{n-1}, \tilde{\mathbf{Y}}_i}}{N_0} \|(\tilde{\mathbf{S}}^{\mu_{n-1}, \tilde{\mathbf{Y}}_i})^H \tilde{\mathbf{Y}}\|^2)] \}.$$

Similarly, for DSTM using ASK, the decision metric in (43) may be extended as $\tilde{d}_{\text{MMSE}}^i = \kappa_{n-1}^2 - 2\Re[\text{tr}(\tilde{\mathbf{H}}^H \tilde{\mathbf{S}}_{n-1}^H \mathbf{Y}_{n-1})] + \text{tr}(\tilde{\mathbf{H}}^H \tilde{\mathbf{S}}_{n-1}^H \tilde{\mathbf{S}}_{n-1} \tilde{\mathbf{H}}) + \kappa_n^2 - \frac{2}{\Gamma_{n-1}} \Re[\text{tr}(\tilde{\mathbf{H}}^H \tilde{\mathbf{S}}_{n-1}^H (\tilde{\mathbf{X}}^i)^H \mathbf{Y}_n)] + \frac{(\gamma^a)^2}{\Gamma_{n-1}^2} \text{tr}(\tilde{\mathbf{H}}^H \tilde{\mathbf{S}}_{n-1}^H \tilde{\mathbf{S}}_{n-1} \tilde{\mathbf{H}})$. Evaluating $\frac{\partial \tilde{d}_{\text{MMSE}}^i}{\partial \tilde{\mathbf{S}}_{n-1}^H} = 0$ leads to $\tilde{\mathbf{S}}_{n-1}^H \tilde{\mathbf{H}} = \frac{\Gamma_{n-1}^2}{\tau_{\text{ASK}}^a} [\mathbf{Y}_{n-1} + \frac{1}{\Gamma_{n-1}} (\tilde{\mathbf{X}}^i)^H \mathbf{Y}_n]$, where we define $\tau_{\text{ASK}}^a = \Gamma_{n-1}^2 + (\gamma^a)^2$. Then the decision metric of (43) becomes:

$$\tilde{d}_{\text{MMSE}}^i = \frac{\Gamma_{n-1}^2}{\tau_{\text{ASK}}^a} \|\mathbf{Y}_n - \frac{1}{\Gamma_{n-1}} \tilde{\mathbf{X}}^i \mathbf{Y}_{n-1}\|^2. \quad (46)$$

Therefore, the detection of \mathbf{X}^i based on (46) is also given by (37), while the ring-amplitude detection is given by:

$$\hat{a} = \arg \min_{\forall a} \frac{\Gamma_{n-1}^2 \kappa_n^2}{\tau_{\text{ASK}}^a} + \frac{(\gamma^a)^2 \kappa_{n-1}^2}{\tau_{\text{ASK}}^a} - \frac{2\Gamma_{n-1} \gamma^a}{\tau_{\text{ASK}}^a} \tilde{d}^i. \quad (47)$$

D. Least Square Differential Detection (LS-DD)

In contrast to the MMSE-DD, the LS solution directly takes $\mathbf{Y}_{n-1} = \tilde{\mathbf{S}}_{n-1} \tilde{\mathbf{H}}$ and hence $\|\mathbf{Y}_{n-1} - \tilde{\mathbf{S}}_{n-1} \tilde{\mathbf{H}}\|^2 = 0$ for (43). Consequently, for DSTM using DASK, the decision metric in $p(\tilde{\mathbf{Y}}|\tilde{\mathbf{S}}, \tilde{\mathbf{H}})$ of (43) becomes $\tilde{d}_{\text{LS}}^i = \|\mathbf{Y}_n - \tilde{\mathbf{X}}^i \mathbf{Y}_{n-1}\|^2 = \kappa_n^2 + (\gamma^a)^2 \kappa_{n-1}^2 - 2\gamma^a \Re[\text{tr}(\mathbf{X}^i \mathbf{Y}_{n-1} \mathbf{Y}_n^H)]$. As a result, the detection of the unitary matrix \mathbf{X}^i is given by (37), while the ring-amplitude detection may be expressed as:

$$\hat{a} = \arg \min_{\forall a} \kappa_{n-1}^2 (\gamma^a - \frac{\tilde{d}^i}{\kappa_{n-1}^2})^2 = \mathbb{M}_{\text{DASK}}^{-1}(\tilde{z}_{\text{DASK}}), \quad (48)$$

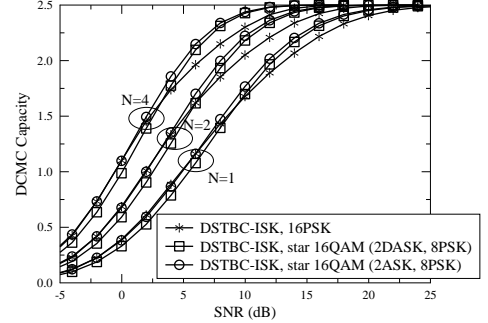
where the decision variable is $\tilde{z}_{\text{DASK}} = \frac{\tilde{d}^i}{\kappa_{n-1}^2}$. Moreover, the single-stream DASK demodulation $\hat{a} = \mathbb{M}_{\text{DASK}}^{-1}(\tilde{z}_{\text{DASK}})$ in (48) may directly deliver \hat{a} as the Gray coded index $\tilde{a} = \{\max[\min(\lfloor \log_{\alpha} \tilde{z}_{\text{DASK}} \rfloor + \mu_{n-1}, 0), L_A - 1] - \mu_{n-1}\} \bmod L_A$.

Similarly, the LS-DD of DSTM using ASK also relies on $\mathbf{Y}_{n-1} = \tilde{\mathbf{S}}_{n-1} \tilde{\mathbf{H}}$, which results in $\tilde{d}_{\text{LS}}^i = \|\mathbf{Y}_n - \frac{1}{\Gamma_{n-1}} \tilde{\mathbf{X}}^i \mathbf{Y}_{n-1}\|^2$. The detection of \mathbf{X}^i is also given by (37), while the ring-amplitude detection is given by:

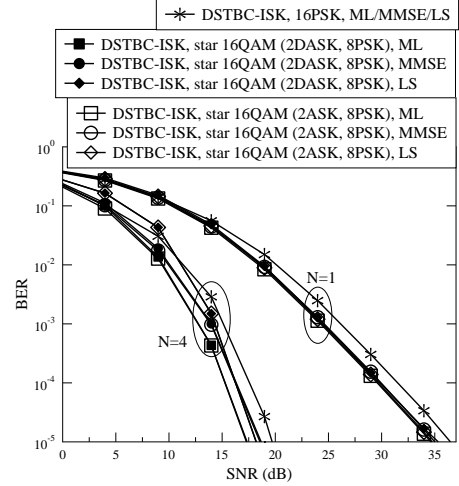
$$\hat{a} = \arg \min_{\forall a} \kappa_{n-1}^2 (\frac{\gamma^a}{\Gamma_{n-1}} - \frac{\tilde{d}^i}{\kappa_{n-1}^2})^2 = \mathbb{M}_{\text{ASK}}^{-1}(\Gamma_{n-1} \tilde{z}_{\text{ASK}}), \quad (49)$$

where we define $\tilde{z}_{\text{ASK}} = \frac{\tilde{d}^i}{\kappa_{n-1}^2}$, while the ASK demodulation $\hat{a} = \mathbb{M}_{\text{ASK}}^{-1}(\Gamma_{n-1} \tilde{z}_{\text{ASK}})$ in (49) directly obtains \hat{a} as the Gray coded $\tilde{a} = \max[\min(\lfloor \log_{\alpha} \sqrt{\beta} \Gamma_{n-1} \tilde{z}_{\text{ASK}} \rfloor, 0), L_A - 1]$.

We note that the LS-DD of DSTM using DASK may also be obtained based on the extension of the received signal model of (33) using the differential encoding of (31) as $\mathbf{Y}_n = \tilde{\mathbf{X}}_{n-1}(\mathbf{Y}_{n-1} - \mathbf{V}_{n-1}) + \mathbf{V}_n$, which was used in [11], [36], [37]. Similarly, the LS-DD of DSTM using ASK may be obtained according to the extension of (33) using



(a) DCMC Capacity



(b) BER performance

Fig. 11. DCMC capacity and BER performance of DSTBC-ISK using PSK and star QAM associated with $(M = 2)$ and $(R = 2.5)$.

(32) as $\mathbf{Y}_n = \frac{1}{\Gamma_{n-1}} \tilde{\mathbf{X}}_{n-1}(\mathbf{Y}_{n-1} - \mathbf{V}_{n-1}) + \mathbf{V}_n$, which was employed in [13], [33], [68]. Therefore, the detectors relying on $\mathbf{Y}_{n-1} = \tilde{\mathbf{S}}_{n-1} \tilde{\mathbf{H}}$ that have been popularly used in the literature [11], [13], [33], [36], [37], [68] in fact correspond to the sub-optimal LS solution, which suffers from a performance loss compared to the ML-DD and MMSE-DD.

In summary, the ML-DD of Sec. V-B, MMSE-DD of Sec. V-C and LS-DD of Sec. V-D share the same unitary matrix detection concerning \mathbf{X}^i , which may be simplified for DSTBC-ISK and DSM as devised in Sec. III and Sec. IV, respectively. However, the three DDs use different metrics for detecting the ring-amplitude γ^a . As a result, the ML-DD of (38) and (40) that involves calculations in the logarithmic domain exhibits the highest complexity, followed by the MMSE-DD of (45) and (47) and then by the LS-DD of (48) and (49). Their performance results will be further compared in Sec. VI.

VI. PERFORMANCE RESULTS

First of all, it is demonstrated by Fig. 11 that DSTBC-ISK using star QAM outperforms its PSK counterpart in terms of both the DCMC capacity and the BER performance at high

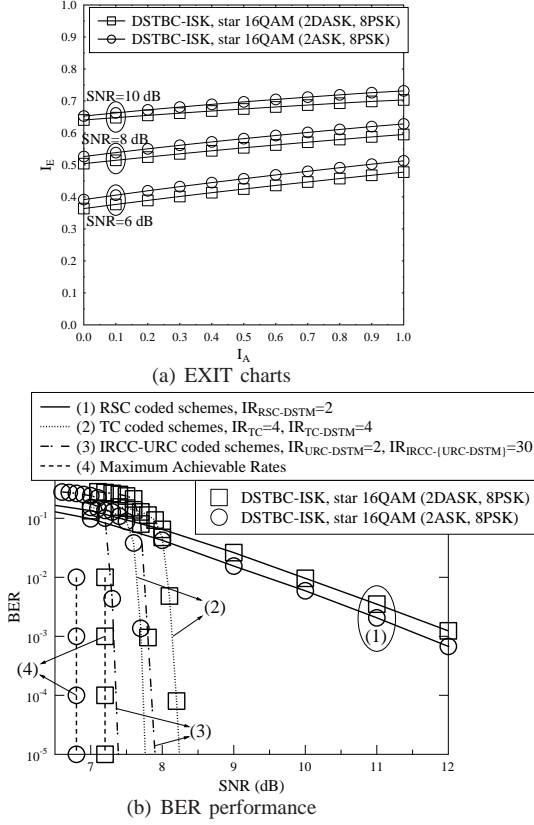


Fig. 12. EXIT charts and BER performance of DSTBC-ISK using star QAM associated with $(M=2)$ and $(R=2.5)$ in RSC/TC/IRCC-URC coded systems.

SNRs. Furthermore, although it was demonstrated in [12] that ML-DD and LS-DD invoked for DSTBC using QAM exhibit a similar performance, we can see from Fig. 11(b) that as N increases, the performance differences between ML-DD, MMSE-DD and LS-DD for DSTBC-ISK using star QAM become more significant, where the ML-DD achieves the best performance, followed by MMSE-DD and then by LS-DD. Therefore, we opt for employing ML-DD for all DSTM schemes in the rest of this paper.

Moreover, it is demonstrated in Fig. 11(a) that DSTBC-ISK using ASK exhibits a higher DCMC capacity than DSTBC-ISK using DASK, but their BER performance results are almost the same in Fig. 11(b), which is indeed expected according to Proposition 4. Nonetheless, the advantage of ASK becomes more significant in Fig. 12, where its performance potential is further improved with the aid of channel coding. More explicitly, the powerful tool of EXtrinsic Information Transfer (EXIT) charts [2], [69], [70] is used in Fig. 12(a), which quantifies the extrinsic information (I_E) produced by the soft-decision DSTM detectors, given the *a priori* information I_A provided by a channel decoder. To avoid losing focus, the soft-decision DSTM detectors are derived in Appendix A. Moreover, Fig. 12(b) examines the performance of DSTBC-ISK using DASK/ASK in conjunction with Recursive Convolutional Codes (RSCs), Turbo Codes (TCs) as well as IRregular Convolutional Codes (IRCCs) and Unity Rate Codes (URCs), where the simulation parameters are the same as those summarized in Table V of [71]. The number of iterations between a channel decoder and the soft-decision DSTM detector is indicated by the acronym “IR” in Fig. 12(b). In summary, although the DASK approach has been popularly

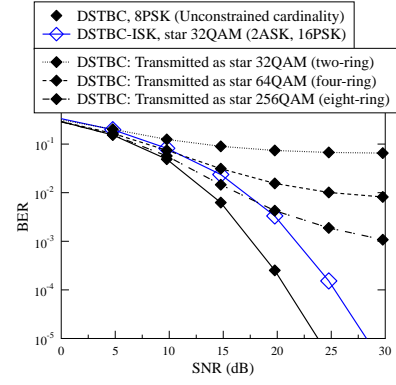


Fig. 13. Performance comparison between DSTBC and DSTBC-ISK associated with $(M=2)$, $(N=2)$ and $(R=3.0)$.

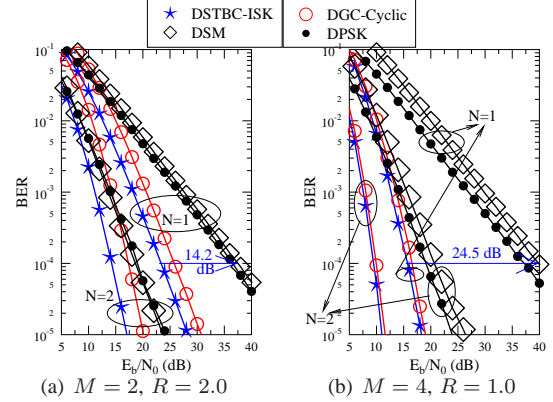


Fig. 14. BER performance results of DSTBC-ISK's diversity gains over DSM associated with $(M=2, R=2.0)$ and $(M=4, R=1.0)$. The performance results of DGC and DPSK are also depicted as benchmarks.

used in DSTM schemes [11], [36], [37], it is demonstrated by Fig. 12 that the ASK technique is capable of offering a better performance in coded scenarios, which confirms the DCMC capacity comparison results of Fig. 11(a). Therefore, in the rest of the paper, we opt for employing the ASK approach for our DSTM schemes using star QAM signalling.

Fig. 13 demonstrates that DSTBC associated with unconstrained transmit-signal cardinality outperforms DSTBC-ISK. However, a severe performance loss is recorded in Fig. 13, when the DSTBC's transmitted signals are quantized and transmitted as a finite-cardinality set of star QAM signals. Therefore, we only consider the finite-cardinality DSTM schemes of DSTBC-ISK, DGC and DSM in the rest of this paper, and their parameters are summarized in Table VIII.

The BER performance comparison between DSTBC-ISK and DSM is exemplified in Fig. 14 for the cases of $(M=2, R=2.0)$ and $(M=4, R=1.0)$. First of all, Fig. 14 shows that DSM performs worse than DPSK in these two cases using $(N=\{1, 2\})$. Secondly, Fig. 14 evidences that DSTBC-ISK achieves substantial diversity gains over DSM, which are as high as **14.2 dB** and **24.5 dB** at $BER=10^{-4}$ for two these cases associated with $(N=1)$. Furthermore, Fig. 14 also shows that DSTBC-ISK is even capable of outperforming the classic DGC for these two cases, despite the fact that DSTBC-ISK simply use the classic DSTBC signal structure without invoking DGC's parameters optimization.

We note that the proposed DSTBC-ISK does not share the same transmitter property as the existing DSTM schemes of DSTBCs [8]–[13], DLDCs [22]–[25] and DSTSKs [31]–[33], [40], [41], which suffer from the infinite-cardinality problem.

(a) DSTBC-ISK, $T = 2$, $R = \{1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0\}$									
$L=2$	$L=4$	$L=8$	$L_A=2, L=8$	$L_A=2, L=16$	$L_A=4, L=16$	$L_A=4, L=32$	$L_A=8, L=32$	$L_A=8, L=64$	
(b) DSTBC-ISK, $T = 4$, $R = \{0.75, 1.0, 1.25, 1.5, 1.75, 2.0, 2.25, 2.5\}$									
$L=2$	$L=4$	$L=8$	$L_A=2, L=8$	$L_A=2, L=16$	$L_A=4, L=16$	$L_A=4, L=32$	$L_A=8, L=32$		
(c) DGC-cyclic, $T = 2$, $R = \{1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0\}$									
$L=4, \mathbf{u}=[1, 1]$	$L=8, \mathbf{u}=[1, 3]$	$L_A=2, L=8, \mathbf{u}=[1, 3]$	$L_A=2, L=16, \mathbf{u}=[1, 7]$	$L_A=2, L=32, \mathbf{u}=[1, 7]$					
$L_A=4, L=32, \mathbf{u}=[1, 7]$	$L_A=4, L=64, \mathbf{u}=[1, 19]$	$L_A=4, L=128, \mathbf{u}=[1, 47]$	$L_A=8, L=128, \mathbf{u}=[1, 47]$						
(d) DGC-cyclic, $T = 4$, $R = \{0.75, 1.0, 1.25, 1.5, 1.75, 2.0, 2.25, 2.5\}$									
$L=8, \mathbf{u}=[1, 1, 3, 3]$	$L=16, \mathbf{u}=[1, 3, 5, 7]$	$L=32, \mathbf{u}=[1, 7, 9, 15]$	$L=64, \mathbf{u}=[1, 15, 27, 29]$						
$L_A=2, L=8, \mathbf{u}=[1, 1, 3, 3]$	$L_A=2, L=16, \mathbf{u}=[1, 3, 5, 7]$	$L_A=2, L=32, \mathbf{u}=[1, 7, 9, 15]$	$L_A=2, L=64, \mathbf{u}=[1, 15, 27, 29]$						
(e) DSM, $T = 2$, $R = \{1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0\}$									
$L_1=1, L_2=2$	$L_1=2, L_2=2$	$L_1=2, L_2=4$	$L_1=4, L_2=4$	$L_1=4, L_2=8$	$L_1=8, L_2=8$	$L_A=2, L_1=8, L_2=8$	$L_A=4, L_1=8, L_2=8$	$L_A=4, L_1=8, L_2=16$	
(f) DSM, $T = 4$, $R = \{1.0, 1.25, 1.5, 1.75, 2.0, 2.25, 2.5\}$									
$L_1=1, L_2=1, L_3=1, L_4=1$	$L_1=1, L_2=1, L_3=1, L_4=2$	$L_1=1, L_2=1, L_3=2, L_4=2$	$L_1=1, L_2=2, L_3=2, L_4=2$						
$L_1=2, L_2=2, L_3=2, L_4=2$	$L_1=2, L_2=2, L_3=2, L_4=4$	$L_1=2, L_2=2, L_3=4, L_4=4$							

TABLE VIII

PARAMETERS FOR FINITE-CARDINALITY DSTM SCHEMES USED IN THIS SECTION.

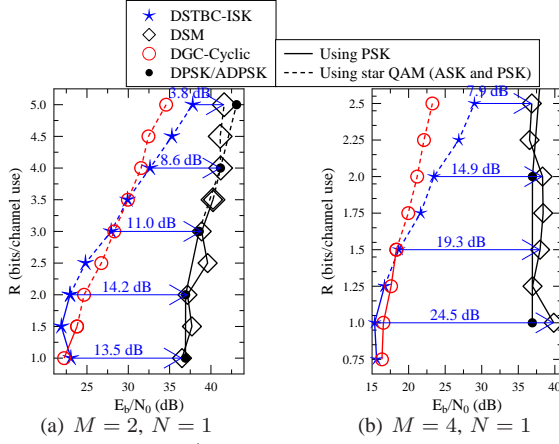
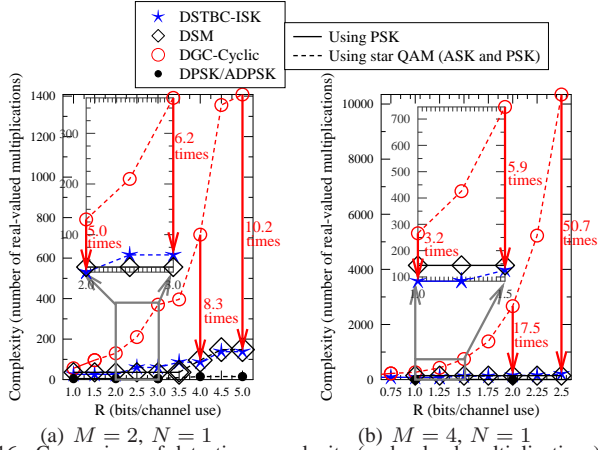
Fig. 15. Comparison of E_b/N_0 required for DSTBC-ISK and DSM to achieve $\text{BER}=10^{-4}$. DGC and DPSK are also depicted as benchmarks.

Fig. 16. Comparison of detection complexity (real-valued multiplications) of DSTBC-ISK and DSM. DGC and DPSK are also depicted as benchmarks.

Moreover, the proposed DSTBC-ISK also does not share the same receiver property as DGCs [14]–[16], which exhibit an exponentially increasing complexity with the throughput. The design objective of the new DSTBC-ISK scheme is to achieve a beneficial diversity gain over DSM [34]–[37] at the same requirement of having a low transceiver complexity.

The simulation results of DSTBC-ISK and DSM are summarized in terms of the E_b/N_0 required for achieving $\text{BER}=10^{-4}$ in Fig. 15, and the corresponding complexity results are presented in Fig. 16. It is evidenced by Fig. 15 that the proposed DSTBC-ISK achieves substantial diversity gains over DSM across the wide throughput range of $(1.0 \leq R \leq 5.0)$

and $(1.0 \leq R \leq 2.5)$ for the cases of $(M=2, N=1)$ and $(M=4, N=1)$, respectively. Furthermore, Fig. 16 confirms that DSTBC-ISK and DSM exhibit the same linearized single-stream ML detection complexity, which is comparable to the SISO DPSK/ADPSK scheme's detection complexity.

Moreover, considering that the finite-cardinality property of Sec. II-A is the first and foremost objective of our design, we have also presented the classic finite-cardinality scheme of DGC in Figs. 15 and 16. First of all, Figs. 15 and 16 evidence that DSTBC-ISK achieves an even better performance than DGC in the lower throughput ranges of $(1.0 < R \leq 3.0)$ and $(0.75 \leq R < 1.5)$ for the cases of $(M=2, N=1)$ and $(M=4, N=1)$, respectively, despite the DSTBC-ISK's lower complexity. Secondly, Fig. 15 confirms that DGC achieves a superior performance in the higher throughput ranges of $(R \geq 4)$ and $(R \geq 1.75)$ for these two cases. However, it is demonstrated by Fig. 16 that the DGC's detection complexity grows exponentially with the throughput, where the detection complexity of DGC becomes rather excessive, which is **10.2 times** and **50.7 times** higher than that of DSTBC-ISK for the cases of $(M=2, N=1, R=5.0)$ and $(M=4, N=1, R=2.5)$, respectively. By contrast, DSTBC-ISK is capable of achieving a beneficial diversity gain over DSM without eroding the DSM's low transceiver complexity.

VII. CONCLUSIONS AND FUTURE WORK

In this paper, we propose a new DSTBC-ISK scheme, which achieves a beneficial diversity gain over DSM at the same requirement of having a low transceiver complexity. The proposed DSTBC-ISK is the first DSTM scheme capable of simultaneously achieving the three imperative objectives of a finite-cardinality transmit-signals set, a single-stream ML detection complexity and the full diversity gain. In order to make a fair comparison, we have also conceived a single-stream ML detector for DSM. Furthermore, we propose to generalize DASK and ASK for DSTM using star QAM. The three bespoke detectors of ML-DD, MMSE-DD and LS-DD exhibit a performance versus complexity tradeoff. We have also demonstrated that our generalized ASK technique is capable of achieving a higher capacity and a better performance than its DASK counterpart in channel coding assisted systems.

In recent years, the technique of hybrid beamforming conceived for massive MIMO systems has attracted a significant

research interest [72]–[74]. Specifically, the recent developments in the millimeter-wave band facilitate the employment of a large number of antennas having fewer RF chains, which results in a significantly reduced power consumption compared to the conventional massive MIMO. In this context, our proposed single-RF finite-cardinality DSTBC-ISK may constitute a suitable low-complexity alternative to the existing coherent schemes due to the following reasons: (1) Analog precoding can be employed based on channel statistics such as the channel's covariance [74] instead of full channel knowledge at the transmitter. This would enable us to employ DSTBC-ISK, which eliminates the need for baseband channel tracking, especially when the channel coefficients fluctuate rapidly [75]. (2) In addition to the low-resolution ADC at the receiver, thanks to beamforming, the finite-cardinality design of DSTBC-ISK also helps us to reduce the power consumption of the DAC, especially at the transmitting base stations. (3) The single-RF feature and transmit diversity design of DSTBC-ISK may improve the performance of its single-antenna-based counterparts without increasing the number of RF chains.

APPENDIX A SOFT-DECISION DSTM DETECTION

Based on (9), the *a posteriori* LLR produced by the Max-Log-MAP algorithm [2], [65] may be expressed as:

$$L_p(b_k) = \max_{\mathbf{X}^i \in \mathbf{X}_{b_k=1}} d^i - \max_{\mathbf{X}^i \in \mathbf{X}_{b_k=0}} d^i, \quad (50)$$

where $\mathbf{X}_{b_k=1}$ and $\mathbf{X}_{b_k=0}$ refer to the signal set \mathbf{X} , when the specific bit b_k is set to 1 and 0, respectively. The probability metric in (50) is given by:

$$d^i = \frac{\eta}{N_0} \|\mathbf{Y}_{n-1}^H + \mathbf{Y}_n^H \mathbf{X}^i\|^2 + \sum_{k=1}^{\log_2 I} \tilde{b}_k L_a(b_k), \quad (51)$$

where $\{\tilde{b}_k\}_{k=1}^{\log_2 I}$ represents the bit-mapping corresponding to the data-carrying matrix \mathbf{X}^i , while $\{L_a(b_k)\}_{k=1}^{\log_2 I}$ refers to the *a priori* LLRs gleaned from a channel decoder.

For DSTM using DASK, the soft-decision ML-DD using Max-Log-MAP may be expressed based on (34a) as:

$$L_p(b_k) = \max_{\forall \mu_{n-1}} \max_{\tilde{\mathbf{X}}^i \in \tilde{\mathbf{X}}_{b_k=1}} d_{\text{DASK}}^{\mu_{n-1}, \tilde{i}} - \max_{\forall \mu_{n-1}} \max_{\tilde{\mathbf{X}}^i \in \tilde{\mathbf{X}}_{b_k=0}} d_{\text{DASK}}^{\mu_{n-1}, \tilde{i}}, \quad (52)$$

where the probability metric is given by:

$$d_{\text{DASK}}^{\mu_{n-1}, \tilde{i}} = \frac{\eta^{\mu_{n-1}, \tilde{i}}}{N_0} \|\tilde{\mathbf{S}}_{n-1} [\mathbf{Y}_{n-1} + (\tilde{\mathbf{X}}^i)^H \mathbf{Y}_n]\|^2 + MN \ln \eta^{\mu_{n-1}, \tilde{i}} + \sum_{k=1}^{\log_2 \tilde{I}} \tilde{b}_k L_a(b_k). \quad (53)$$

Similar to the DCMC capacity evaluation of (41), $\tilde{\mathbf{S}}_{n-1}$ in (53) may use $\Gamma_{n-1} \mathbf{S}_1 = \frac{\alpha^{\mu_{n-1}}}{\sqrt{\beta}} \sqrt{\frac{T}{M}} [\mathbf{I}_M, \mathbf{0}]^T$.

As explained for the DCMC capacity evaluation of (42), DSTM using ASK relies on decision-feedback concerning $\Gamma_{n-1} = \frac{\alpha^{\mu_{n-1}}}{\sqrt{\beta}}$. Therefore, (52) may be revised for ASK as:

$$L_p(b_k) = \max_{\tilde{\mathbf{X}}^i \in \tilde{\mathbf{X}}_{b_k=1}} \tilde{d}_{\text{ASK}}^i - \max_{\tilde{\mathbf{X}}^i \in \tilde{\mathbf{X}}_{b_k=0}} \tilde{d}_{\text{ASK}}^i, \quad (54)$$

where the probability metric is expressed based on (34b) as:

$$\tilde{d}_{\text{ASK}}^i = \frac{\eta^i}{N_0} \|\tilde{\mathbf{S}}_{n-1} [\mathbf{Y}_{n-1} + \frac{1}{\Gamma_{n-1}} (\tilde{\mathbf{X}}^i)^H \mathbf{Y}_n]\|^2 + MN \ln \eta^i + \sum_{k=1}^{\log_2 \tilde{I}} \tilde{b}_k L_a(b_k). \quad (55)$$

Upon obtaining the optimum \tilde{d}_{ASK}^i of (55) over all $\{\tilde{\mathbf{X}}^i\}_{i=0}^{\tilde{I}-1}$, the updated $\Gamma_n = \gamma_{n-1}$ may be passed on to the next ML-DD block as the decision-feedback on Γ_{n-1} .

REFERENCES

- [1] A. Goldsmith, *Wireless communications*. Cambridge University Press, 2005.
- [2] L. Hanzo, O. Alamri, M. El-Hajjar, and N. Wu, *Near-Capacity Multi-Functional MIMO Systems: Sphere-Packing, Iterative Detection and Cooperation*. John Wiley & Sons, May 2009.
- [3] W. Webb, L. Hanzo, and R. Steele, "Bandwidth efficient QAM schemes for Rayleigh fading channels," *IEE Proc. I, Commun., Speech and Vision*, vol. 138, no. 3, pp. 169–175, June 1991.
- [4] C. Xu, S. X. Ng, and L. Hanzo, "Multiple-symbol differential sphere detection and decision-feedback differential detection conceived for differential qam," *IEEE Transactions on Vehicular Technology*, vol. 65, pp. 8345–8360, Oct 2016.
- [5] C. Xu, L. Wang, S. X. Ng, and L. Hanzo, "Soft-decision multiple-symbol differential sphere detection and decision-feedback differential detection for differential QAM dispensing with channel estimation in the face of rapidly fading channels," *IEEE Trans. Wireless Commun.*, vol. 15, pp. 4408–4425, June 2016.
- [6] L. Lampe and R. Schober, "Low-complexity iterative demodulation for noncoherent coded transmission over Ricean-fading channels," *IEEE Trans. Veh. Technol.*, vol. 50, no. 6, pp. 1481–1496, Nov. 2001.
- [7] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.
- [8] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity," *IEEE J. Sel. Areas Commun.*, vol. 18, pp. 1169–1174, July 2000.
- [9] H. Jafarkhani and V. Tarokh, "Multiple transmit antenna differential detection from generalized orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 47, pp. 2626–2631, Sept. 2001.
- [10] G. Ganesan and P. Stoica, "Differential modulation using space-time block codes," *IEEE Signal Process. Lett.*, vol. 9, pp. 57–60, Feb 2002.
- [11] X. Xia, "Differentially en/decoded orthogonal space-time block codes with APSK signals," *IEEE Commun. Lett.*, vol. 6, pp. 150–152, 2002.
- [12] M. Tao and R. S. Cheng, "Differential space-time block codes," in *IEEE Global Telecommun. Conf.*, vol. 2, pp. 1098–1102 vol.2, 2001.
- [13] C. Hwang, S. Nam, J. Chung, and V. Tarokh, "Differential space time block codes using nonconstant modulus constellations," *IEEE Trans. Signal Process.*, vol. 51, pp. 2955–2964, Nov 2003.
- [14] B. L. Hughes, "Differential space-time modulation," *IEEE Trans. Inf. Theory*, vol. 46, pp. 2567–2578, Nov 2000.
- [15] B. M. Hochwald and W. Sweldens, "Differential unitary space-time modulation," *IEEE Trans. Commun.*, vol. 48, pp. 2041–2052, Dec 2000.
- [16] B. L. Hughes, "Optimal space-time constellations from groups," *IEEE Trans. Inf. Theory*, vol. 49, pp. 401–410, Feb 2003.
- [17] A. Shokrollahi, B. Hassibi, B. M. Hochwald, and W. Sweldens, "Representation theory for high-rate multiple-antenna code design," *IEEE Trans. on Inf. Theory*, vol. 47, pp. 2335–2367, Sep 2001.
- [18] X. Liang and X. Xia, "Unitary signal constellations for differential space-time modulation with two transmit antennas: parametric codes, optimal designs, and bounds," *IEEE Trans. Inf. Theory*, vol. 48, pp. 2291–2322, Aug 2002.
- [19] Y. Jing and B. Hassibi, "Three-transmit-antenna space-time codes based on SU(3)," *IEEE Trans. Signal Process.*, vol. 53, pp. 3688–3702, 2005.
- [20] J. Abarbanel, A. Averbuch, S. Rosset, and J. Zlotnick, "Unitary nongroup STBC from cyclic algebras," *IEEE Trans. on Inf. Theory*, vol. 52, pp. 3903–3912, Sept 2006.
- [21] F. Oggier, "Cyclic algebras for noncoherent differential space-time coding," *IEEE Trans. Inf. Theory*, vol. 53, pp. 3053–3065, Sept 2007.
- [22] B. Hassibi and B. M. Hochwald, "Cayley differential unitary space-time codes," *IEEE Trans. Inf. Theory*, vol. 48, pp. 1485–1503, Jun 2002.
- [23] H. E. Gamal and M. O. Damen, "Universal space-time coding," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1097–1119, May 2003.
- [24] J. Wang, X. Wang, and M. Madhian, "Design of minimum error-rate Cayley differential unitary space-time codes," *IEEE J. Sel. Areas Commun.*, vol. 23, pp. 1779–1787, Sept 2005.
- [25] F. Oggier and B. Hassibi, "Algebraic Cayley differential space-time codes," *IEEE Trans. Inf. Theory*, vol. 53, pp. 1911–1919, May 2007.
- [26] K. L. Clarkson, W. Sweldens, and A. Zheng, "Fast multiple-antenna differential decoding," *IEEE Trans. Commun.*, vol. 49, pp. 253–261, Feb 2001.

- [27] R. Y. Mesleh, H. Haas, S. Sinanovic, C. W. Ahn, and S. Yun, "Spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 57, pp. 2228–2241, 2008.
- [28] C. Xu, S. Sugiura, S. X. Ng, and L. Hanzo, "Spatial modulation and space-time shift keying: Optimal performance at a reduced detection complexity," *IEEE Trans. Commun.*, vol. 61, pp. 206–216, January 2013.
- [29] R. Rajashekar, K. V. S. Hari, and L. Hanzo, "Reduced-complexity ML detection and capacity-optimized training for spatial modulation systems," *IEEE Trans. Commun.*, vol. 62, pp. 112–125, January 2014.
- [30] L. L. Yang, "Signal detection in antenna-hopping space-division multiple-access systems with space-shift keying modulation," *IEEE Trans. Signal Process.*, vol. 60, pp. 351–366, Jan 2012.
- [31] S. Sugiura, S. Chen, and L. Hanzo, "Coherent and differential space-time shift keying: A dispersion matrix approach," *IEEE Trans. Commun.*, vol. 58, pp. 3219–3230, November 2010.
- [32] C. Xu, S. Sugiura, S. X. Ng, and L. Hanzo, "Reduced-complexity non-coherently detected differential space-time shift keying," *IEEE Signal Process. Lett.*, vol. 18, pp. 153–156, March 2011.
- [33] S. Sugiura, C. Xu, S. X. Ng, and L. Hanzo, "Reduced-complexity coherent versus non-coherent QAM-aided space-time shift keying," *IEEE Trans. Commun.*, vol. 59, pp. 3090–3101, November 2011.
- [34] Y. Bian, M. Wen, X. Cheng, H. V. Poor, and B. Jiao, "A differential scheme for spatial modulation," in *IEEE Global Commun. Conf.*, pp. 3925–3930, Dec 2013.
- [35] Y. Bian, X. Cheng, M. Wen, L. Yang, H. V. Poor, and B. Jiao, "Differential spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 64, pp. 3262–3268, July 2015.
- [36] P. A. Martin, "Differential spatial modulation for APSK in time-varying fading channels," *IEEE Commun. Lett.*, vol. 19, pp. 1261–1264, 2015.
- [37] J. Liu, L. Dan, P. Yang, L. Xiao, F. Yu, and Y. Xiao, "High-rate APSK-aided differential spatial modulation: Design method and performance analysis," *IEEE Commun. Lett.*, vol. PP, no. 99, pp. 1–1, 2016.
- [38] M. Wen, X. Cheng, Y. Bian, and H. V. Poor, "A low-complexity near-ML differential spatial modulation detector," *IEEE Signal Process. Lett.*, vol. 22, pp. 1834–1838, Nov 2015.
- [39] L. Xiao, P. Yang, X. Lei, Y. Xiao, S. Fan, S. Li, and W. Xiang, "A low-complexity detection scheme for differential spatial modulation," *IEEE Commun. Lett.*, vol. 19, pp. 1516–1519, Sept 2015.
- [40] N. Ishikawa and S. Sugiura, "Unified differential spatial modulation," *IEEE Wireless Commun. Lett.*, vol. 3, pp. 337–340, Aug 2014.
- [41] R. Rajashekar, N. Ishikawa, S. Sugiura, K. V. S. Hari, and L. Hanzo, "Full-diversity dispersion matrices from algebraic field extensions for differential spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 66, pp. 385–394, Jan 2017.
- [42] G. Ganesan and P. Stoica, "Space-time block codes: a maximum SNR approach," *IEEE Trans. Inf. Theory*, vol. 47, pp. 1650–1656, May 2001.
- [43] A. Kalis, A. G. Kanatas, and C. B. Papadias, "A novel approach to MIMO transmission using a single RF front end," *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 972–980, August 2008.
- [44] O. N. Alrabadi, C. B. Papadias, A. Kalis, and R. Prasad, "A universal encoding scheme for MIMO transmission using a single active element for PSK modulation schemes," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 5133–5142, October 2009.
- [45] S. Sugiura, "Coherent versus non-coherent reconfigurable antenna aided virtual MIMO systems," *IEEE Signal Process. Lett.*, vol. 21, pp. 390–394, April 2014.
- [46] R. W. Heath and A. J. Paulraj, "Linear dispersion codes for MIMO systems based on frame theory," *IEEE Trans. Signal Process.*, vol. 50, pp. 2429–2441, Oct 2002.
- [47] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1804–1824, Jul 2002.
- [48] B. A. Sethuraman, B. S. Rajan, and V. Shashidhar, "Full-diversity, high-rate space-time block codes from division algebras," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2596–2616, Oct 2003.
- [49] P. Yang, Y. Xiao, Y. L. Guan, K. V. S. Hari, A. Chockalingam, S. Sugiura, H. Haas, M. D. Renzo, C. Masouros, Z. Liu, L. Xiao, S. Li, and L. Hanzo, "Single-carrier SM-MIMO: A promising design for broadband large-scale antenna systems," *IEEE Commun. Surveys Tutorials*, vol. 18, pp. 1687–1716, thirdquarter 2016.
- [50] E. Basar, "On multiple-input multiple-output OFDM with index modulation for next generation wireless networks," *IEEE Trans. Signal Process.*, vol. 64, pp. 3868–3878, Aug 2016.
- [51] M. Wen, X. Cheng, M. Ma, B. Jiao, and H. V. Poor, "On the achievable rate of OFDM with index modulation," *IEEE Trans. on Signal Process.*, vol. 64, pp. 1919–1932, April 2016.
- [52] M. R. Bhatnagar, A. Hjørungnes, and L. Song, "Precoded DOSTBC over Rayleigh channels," *J. Electr. and Comput. Eng.*, 2007.
- [53] M. R. Bhatnagar, A. Hjørungnes, and L. Song, "Precoded differential orthogonal space-time modulation over correlated Ricean MIMO channels," *IEEE J. Sel. Topics Signal Process.*, vol. 2, pp. 124–134, April 2008.
- [54] M. R. Bhatnagar, A. Hjørungnes, and L. Song, "Double-differential orthogonal space-time block codes for arbitrarily correlated Rayleigh channels with carrier offsets," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 145–155, January 2010.
- [55] M. Z. A. Khan and B. S. Rajan, "Single-symbol maximum likelihood decodable linear STBCs," *IEEE Trans. Inform. Theory*, vol. 52, pp. 2062–2091, May 2006.
- [56] O. Tirkkonen and A. Hottinen, "Square-matrix embeddable space-time block codes for complex signal constellations," *IEEE Trans. Inf. Theory*, vol. 48, pp. 384–395, Feb. 2002.
- [57] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. on Inf. Theory*, vol. 45, pp. 139–157, Jan 1999.
- [58] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 46, pp. 543–564, Mar 2000.
- [59] S. X. Ng and L. Hanzo, "On the MIMO channel capacity of multidimensional signal sets," *IEEE Trans. Veh. Technol.*, vol. 55, pp. 528–536, March 2006.
- [60] O. Tirkkonen, A. Boariu, and A. Hottinen, "Minimal non-orthogonality rate 1 space-time block code for 3+ Tx antennas," in *2000 IEEE 6th Int. Symp. Spread Spectrum Techniques and Applicat.*, vol. 2, pp. 429–432 vol.2, 2000.
- [61] H. Jafarkhani, "A quasi-orthogonal space-time block code," *IEEE Trans. Commun.*, vol. 49, pp. 1–4, Jan 2001.
- [62] N. Sharma and C. B. Papadias, "Improved quasi-orthogonal codes through constellation rotation," *IEEE Trans. Commun.*, vol. 51, pp. 332–335, March 2003.
- [63] W. Su and X. Xia, "Signal constellations for quasi-orthogonal space-time block codes with full diversity," *IEEE Trans. Inf. Theory*, vol. 50, pp. 2331–2347, Oct 2004.
- [64] L. A. Dalton and C. N. Georgiades, "A full-rate, full-diversity four-antenna quasi-orthogonal space-time block code," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 363–366, March 2005.
- [65] C. Xu, S. Sugiura, S. X. Ng, P. Zhang, L. Wang, and L. Hanzo, "Two decades of MIMO design tradeoffs and reduced-complexity MIMO detection in near-capacity systems," *IEEE Access*, 2017.
- [66] C. Xu, D. Liang, S. X. Ng, and L. Hanzo, "Reduced-complexity noncoherent soft-decision-aided DAPSK dispensing with channel estimation," *IEEE Trans. Veh. Technol.*, vol. 62, no. 6, pp. 2633–2643, Feb. 2013.
- [67] M. R. Bhatnagar, A. Hjørungnes, and L. Song, "Differential coding for non-orthogonal space-time block codes with non-unitary constellations over arbitrarily correlated Rayleigh channels," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 3985–3995, August 2009.
- [68] C. Xu, L. Wang, S. X. Ng, and L. Hanzo, "Multiple-symbol differential sphere detection aided differential space-time block codes using QAM constellations," *IEEE Signal Process. Lett.*, vol. 18, pp. 497–500, Sept 2011.
- [69] S. ten Brink, "Convergence of iterative decoding," *Electron. Lett.*, vol. 35, pp. 806–808, May 1999.
- [70] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, pp. 1727–1737, Oct 2001.
- [71] C. Xu, X. Zuo, S. X. Ng, R. G. Maunder, and L. Hanzo, "Reduced-complexity soft-decision multiple-symbol differential sphere detection," *IEEE Trans. Commun.*, vol. 63, pp. 3275–3289, Sept 2015.
- [72] F. Söhrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE J. Sel. Topics Signal Process.*, vol. 10, pp. 501–513, April 2016.
- [73] R. Rajashekar and L. Hanzo, "Hybrid beamforming in mm-Wave MIMO systems having a finite input alphabet," *IEEE Trans. on Commun.*, vol. 64, pp. 3337–3349, Aug 2016.
- [74] S. Park, J. Park, A. Yazdan, and R. W. Heath, "Exploiting spatial channel covariance for hybrid precoding in massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 65, pp. 3818–3832, July 2017.
- [75] Z. Xiao, X. G. Xia, D. Jin, and N. Ge, "Iterative eigenvalue decomposition and multipath-grouping Tx/Rx joint beamformings for millimeter-wave communications," *IEEE Trans. Wireless Commun.*, vol. 14, pp. 1595–1607, March 2015.