# Rectangular Differential Spatial Modulation for Open-Loop Noncoherent Massive-MIMO Downlink

Naoki Ishikawa, *Student Member, IEEE*, and Shinya Sugiura, *Senior Member, IEEE*

***Abstract*— In this paper, a novel differential space-time coding scheme is conceived for open-loop noncoherent multiple-input multiple-output (MIMO) downlink scenarios, where the transmission rate increases logarithmically in a scalable manner upon increasing the number of transmit antennas. More specifically, the proposed scheme relies on the projection of a differentially encoded square matrix to its rectangular counterpart and so is capable of reducing the number of symbol intervals needed for block transmission. This is especially beneficial for massive MIMO scenarios, in which the number of transmit antennas is very high. Another advantage exclusive to the presented scheme is that no channel state information (CSI) is required at either the transmitter or the receiver, which eliminates pilot overhead, CSI estimation, CSI feedback, and time-division duplex reciprocity. Furthermore, the rectangular transmission matrix of the proposed scheme contains only a single non-zero element per column, and hence, the transmitter may rely on only a single RF chain, similar to the conventional coherent spatial modulation scheme.**

***Index Terms*— differential modulation, differential space-time codes, massive MIMO, noncoherent detection, spatial modulation.**

## I. Introduction

Massive multiple-input multiple-output (MIMO) [1]–[4] is a promising technique that can significantly increase both spectral and energy efficiencies in multiuser scenarios and is achieved with the aid of a large-scale antenna array equipped at a base station (BS). For example, in the conjugate beamforming (BF) massive MIMO downlink [1], each user receives near-interference-free symbol streams without the need for any complicated detection. As a result, the sum rate increases in a scalable manner with the number of transmit antennas. This advantage is achieved under the assumption that the channel state information (CSI) associated with all of the links is available at the base station. The use of time-division duplex (TDD) reciprocity is typically assumed for closed-loop massive MIMO systems. Furthermore, the simultaneous use of a high number of pilot symbols may cause severe pilot-contamination.

An alternative highly-scalable MIMO technique is spatial modulation (SM) [5]–[9], where additional information is modulated by activating one of multiple transmit antenna elements, and the transmission rate increases upon increasing the number of transmit antennas. Another merit of the SM scheme is that the transmitter activates only a single antenna element per time slot, and so needs only a single radio frequency (RF) chain. Unlike the conjugate-BF massive MIMO counterpart [1]–[4], the SM scheme operates in an open-loop manner, where CSI at the transmitter (CSIT) is not needed. However, the SM receiver has to estimate the CSI associated with all of the links, which may also require high pilot overhead in the single-user massive-MIMO scenario. Irrespective of the above-mentioned closed-loop conjugate-BF massive MIMO system [1]–[4] or the open-loop SM [5]–[7] counterpart, the overhead required for attaining CSI becomes significantly high, especially in a fast fading scenario, since the CSI must be updated more frequently than the channel coherence time.

A number of differential space-time coding (DSTC) schemes [10]–[19] have been developed in order to eliminate the above-mentioned limitations imposed on CSI estimation at the receiver while maintaining the gain of the MIMO system. The differential detection available at the receiver eliminates any pilot overhead as well as the CSI-estimation-error induced performance degradation, which is typically imposed on the coherent counterpart. However, conventional DSTC schemes have focused primarily on achieving the transmit diversity gain, rather than on increasing the transmission rate. More specifically, in conventional DSTC schemes, $(M \times M)$–sized square unitary space-time matrices are generated according to the differential space-time encoding process. The number of symbol intervals in each transmission block is the same as the number of transmit antennas $M$. For this reason, conventional square DSTC schemes are unable to scale up MIMO in terms of the transmission rate. In the present study, we focus on the massive-MIMO downlink scenario.

Against this background, the novel contribution of the present study is the conception of a novel high-rate DSTC scheme, referred to as a rectangular differential SM (RDSM) scheme, for open-loop massive MIMO systems, which has rate scalability with respect to the number of transmit antennas.

Manuscript received July 20, 2016; revised November 21, 2016; accepted January 18, 2017. Date of publication January 24, 2017; date of current version March 8, 2017. The work of N. Ishikawa was supported by the Japan Society for the Promotion of Science (JSPS) Research Fellowship under Grant 16J05344. The work of S. Sugiura was supported by the JSPS KAKENHI under Grant 26709028. The associate editor coordinating the review of this paper and approving it for publication was P. A. Martin.

The authors are with the Department of Computer and Information Sciences, Tokyo University of Agriculture and Technology, Tokyo 184-8588, Japan (e-mail: naoki@ishikawa.cc; sugiura@ieee.org).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2017.2657497

More specifically, in the proposed RDSM scheme, $(M \times M)$–sized square matrices are generated according to the differential encoding process, which is similar to the conventional DSTC scheme. Then, the differentially-encoded square matrix is projected to an $(M \times 1)$–sized rectangular matrix, which is transmitted from the $M$ transmit antennas over a single symbol interval, thereby increasing the transmission rate by a factor of $M$ as compared to the classic DSTC scheme. Thus, the proposed RDSM scheme simultaneously realizes the following beneficial characteristics for massive MIMO downlink:

- *High Transmission Rate Scalable for Massive MIMO:* Unlike conventional DSTC schemes, the proposed scheme is capable of maintaining a low number of symbol intervals per transmission block, which can be reduced to as low as $T = 1$, in contrast to the classic square DSTC, which requires $T \geq M$ symbol intervals. This rectangular transmission matrix structure of the RDSM scheme allows us to logarithmically increase the transmission rate by increasing the number of transmit antennas. In this sense, the proposed RDSM scheme is the first antenna-scalable differential MIMO.
- *Open-Loop Massive MIMO:* The proposed RDSM scheme does not rely on the use of CSIT, whereas its conjugate-BF massive MIMO transmitter does. Hence, the constraint of the TDD reciprocity, imposed on the conjugate BF scheme, is eliminated in the proposed scheme. Furthermore, since the transmitter of the proposed scheme does not have to wait to gather CSIT before transmission, the packet delay becomes lower than for the conjugate-BF counterpart. Moreover, owing to the absence of the effects of CSIT estimation errors, the base station of the proposed scheme can maintain robust connections with fast-moving vehicles.
- *Noncoherent Detection:* In the proposed RDSM scheme, noncoherent detection is carried out at the receiver, which is similar to classic DSTC schemes, implying that the receiver does not estimate CSI. Naturally, the transmitter eliminates any pilot overhead. Considering that the complexity and overhead required for CSI estimation linearly increases with the number of transmit antennas, the advantages of this noncoherent detection as well as the above-mentioned open-loop operation are explicit in the massive MIMO scenario.
- *Flexible Rate-Diversity Tradeoff:* By generalizing the RDSM scheme to that supporting an arbitrary symbol intervals per block, the proposed scheme is capable of striking a further flexible design tradeoff between the transmission rate and the transmit diversity gain. As a result, the generalized RDSM scheme subsumes several conventional DSTC schemes.
- *Single-RF Transmitter:* The rectangular differentially encoded transmission matrices are designed to contain only a single non-zero element in each column, and hence the transmitter may only require a single RF chain in a narrowband scenario, similar to the coherent [5]–[7] and noncoherent SM schemes [15]–[19]. This allows for a low-cost, high-energy-efficiency transmitter structure, which is especially beneficial for the massive MIMO transmitter.[1]

In contrast to the above-mentioned advantages, a potential disadvantage of the proposed non-square DSTC scheme over conventional square DSTC schemes is the error propagation effects induced in the differential decoding process at the receiver. However, the comprehensive performance results obtained herein demonstrate that the above-mentioned advantages of the proposed scheme remain unchanged, even when taking into account the detrimental effects induced by the error propagation.

The remainder of the present paper is organized as follows. A review of related research is presented in Section II. Section III introduces the system model of the proposed scheme, and Section IV presents the design guidelines of the system model. In Sections V and VI, we compare the achievable performances of the proposed scheme and other conventional schemes, in terms of complexity and BER performance, respectively. Finally, the present paper is concluded in Section VII.

## II. Related Research

In this section, we review four conventional MIMO schemes, i.e., the classic square DSTC scheme [10], the conventional differential SM (DSM) scheme [17], [19], the semi-blind joint channel estimation (CE) and data detection (DD) scheme [22], and the closed-loop conjugate BF scheme [1], each of which has its own advantages, when operating in a massive MIMO system. These four schemes are used as benchmarks in the performance comparisons of Section VI.

### A. Classic DSTC

In the classic square DSTC scheme [10], $B$ information bits are modulated onto an $(M \times M)$–sized unitary symbol matrix $\mathbf{X}(i) \in \mathbb{C}^{M \times M}$, which is then differentially encoded as follows:

$$\mathbf{S}(i) = \mathbf{S}(i-1)\mathbf{X}(i) \in \mathbb{C}^{M \times M}, \tag{1}$$

where $i$ is the block index, and $\mathbf{S}(i)$ is the differentially encoded transmission matrix, which is transmitted from $M$ transmit antennas over $M$ symbol intervals. Note that the initial transmission matrix $\mathbf{S}(1)$ is typically set to the identity matrix $\mathbf{I}_M \in \mathbb{R}^{M \times M}$. In the DSTC scheme, the $2^B$ unitary symbol matrices $\mathbf{X}_j$ $(j = 1, \cdots, 2^B)$ are assigned at the transmitter in advance of the transmissions. Note that the $2^B$ legitimate unitary matrices $\mathbf{X}(i)$ are constructed based on several concepts, such as differentially encoded orthogonal space-time block coding [11], column-orthogonal coding [10], Cayley unitary coding [14], [15], and the DSM scheme [17], [18], [23].

[1] As mentioned in [20], the advantage of the single-RF SM transmitter structure is achieved only when using single-carrier transmissions rather than multicarrier transmissions. The same limitation is imposed on the RDSM scheme. Furthermore, similar to other differential MIMO schemes, the proposed RDSM scheme may not operate in a channel having a delay tap. This implies that, in the proposed RDSM scheme, the advantage of the single-RF transmitter is attained in a narrowband scenario. Naturally, all of these limitations may be eliminated by using the full-RF transmitter.

The normalized transmission rate of the conventional DSTC scheme is given by

$$R_{\mathrm{DSTC}} = B/M. \tag{2}$$

In order to maintain a specific transmission rate $R_{\mathrm{DSTC}}$ while linearly increasing the number of transmission antennas $M$, the number of legitimate unitary symbol matrices must be increased exponentially. This makes configuring a DSTC-based massive-MIMO system a challenging task.

Furthermore, denoting the $i$th received signal block as $\mathbf{Y}(i)$, we obtain the optimal maximum likelihood (ML) detection for the square DSTC scheme as follows: [24]

$$\hat{\mathbf{X}}(i) = \arg\min_{\mathbf{X}_j} \left\| \mathbf{Y}(i) - \mathbf{Y}(i-1) \cdot \mathbf{X}_j \right\|_{\mathrm{F}}^2. \tag{3}$$

Note that the ML detection of (3) suffers from a well-known noise-doubling effect, in comparison to its coherently detected counterpart.

### B. Conventional Square DSM [17], [19]

The DSM [17], [19], [25]–[28] is obtained by the recently proposed square unitary DSTC scheme, where each column of the unitary symbol matrices $\mathbf{X}_j$ $(j = 1, \cdots, 2^B)$ includes a single non-zero coefficient that has a unit absolute value. Hence, the differentially encoded transmission matrices $\mathbf{S}(i)$ also have the same matrix structure. This implies that only a single transmit antenna element is activated in each symbol duration at the transmitter. More specifically, $Q$ unitary dispersion matrices (DMs) $\mathbf{A}_q$ $(q = 1, \cdots, Q) \in \mathbb{C}^{M \times M}$, having a single non-zero element that has a unit absolute value in each row and in each column, are assigned to the DSM transmitter in advance of transmissions. In each block index $i$, a square unitary matrix is generated based on $B$ information bits as follows: $\mathbf{X}(i) = \mathrm{diag}\left[\mathbf{s}(i)\right] \mathbf{A}(i)$, where $\mathbf{s}(i) \in \mathbb{C}^M$ are the $M$ symbols, which are modulated onto the $\mathcal{L}$-point APSK constellations, and $\mathbf{A}(i)$ is an activated DM out of the $Q$ DMs. Furthermore, diag[·] represents the diagonalization operation. Here, we have the relationship $B = \log_2 Q + M \log_2 \mathcal{L}$.[2]

Since the DSM scheme is a unitary square DSTC scheme, the normalized transmission rate of the DSM scheme is given by (2). This implies that the square DSM scheme also suffers from the above-mentioned limitation of DSTC schemes.

### C. Semi-Blind Joint CE and DD [22], [30]

In [30], the semi-blind CE and DD schemes were developed for MIMO systems, where a short training block is transmitted before the transmissions of data symbols. The training block is used to provide a rough initial CE, and the data symbols are detected with the aid of the estimated channels. The detected data symbols are used to refine the CE. Here, since the number of data symbols is typically greater than the number of pilot symbols, more accurate channel estimates tend to be attainable. By iteratively carrying out CE and DD, near-perfect channel state information (near-PCSI) performance is achievable. Hence, the semi-blind CE and DD scheme is used to achieve PCSI performance while minimizing the pilot overhead. This is beneficial for reducing the overhead of the massive MIMO scenario at the cost of increased complexity due to the iterative CE and DD process. Both the DSTC scheme and the semi-blind CE and DD scheme minimize the reference (pilot) overhead per frame.

Recently, the joint CE and DD concept was used for the SM scheme in [16], [22] to achieve a high transmission rate and a low-cost transmitter. Note that, since the channels are assumed to remain constant over the data block in [16], [22], modifications are needed for tracking the time-varying channels in the rapidly fading environment.

### D. Conjugate BF [1]

The conjugate BF [1] is a pioneering massive MIMO scheme that was designed for the closed-loop multiuser downlink, where the BS, equipped with a large number of antenna elements, transmits the multiplexed symbol streams to multiple receivers. The symbol streams are simply precoded by the conjugate of the channel matrix, which results in the focusing of each symbol stream on each user. More specifically, upon increasing the number of transmit antennas, multiuser interference diminishes, and hence near-capacity performance is achievable. This allows relaying to be performed with simple low-complexity noncoherent detection at each user. Furthermore, since PCSI of all the links is assumed to be available at the BS, the TDD constraint is typically imposed on the conjugate BF scheme.

The fundamental characteristics of the four above-mentioned MIMO schemes, the coherent SM scheme [5], [6], and the single-antenna differential scheme modulated by star quadrature amplitude modulation (QAM) [21] are summarized in Table I.[3]

## III. Proposed Rectangular DSM

In this section, we introduce the system model of the proposed RDSM scheme, in which the number of symbol intervals per transmission block is $T = 1$. Then, we extend the RDSM scheme to support the generalized scenario with an arbitrary number of symbol intervals per block in the range of $1 \leq T \leq M$. Finally, we introduce a suboptimal noncoherent detector for the proposed scheme.

### A. Proposed RDSM ($T = 1$)

Fig. 1(a) shows a schematic diagram of the proposed RDSM transmitter, where the number of symbol intervals per block is $T = 1$. Let us consider that each RDSM frame has $W$ symbols, consisting of $M$-length reference blocks and $(W - M)$-length data blocks. At the data transmission block index $i$ $(> M)$, $B$-length information bits $\mathbf{b}(i) \in \mathbb{Z}^B$ are serial-to-parallel (S/P) converted into $B_1$ and $B_2$ bits. The first $B_1$ bits $\mathbf{b}_1(i) \in \mathbb{Z}^{B_1}$ are modulated by selecting a single dispersion vector (DV) $\mathbf{a}(i) \in \mathbb{C}^M$ out of $Q$ DVs $\mathbf{a}_q$ $(q = 1, \cdots, Q) \in \mathbb{C}^M$, where each

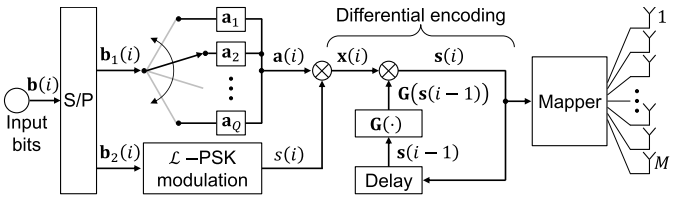
[2]Optionally, the number of symbols in $\mathbf{s}(i)$ may be selected in the range of $1 \leq P \leq M$, in order to strike the rate-diversity tradeoff [17], [29]. The relationship $P = M$ maximizes the transmission rate of the DSM scheme. For details, refer to [17]

[3]The detail of the computational complexity listed in Table I is provided in Section V, where $N$ is the number or receive antennas, and $J$ is the number of iterations used in the semi-blind CE and DD scheme.

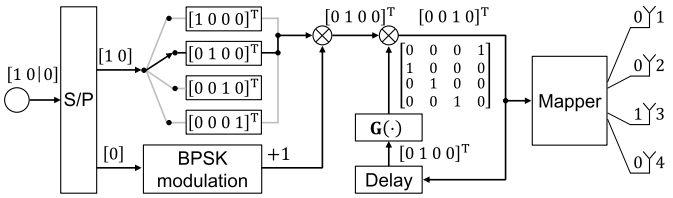
TABLE I

FUNDAMENTAL CHARACTERISTICS OF THE PROPOSED SCHEME, CONVENTIONAL DIFFERENTIAL SCHEMES, AND THE CONJUGATE BF SCHEME

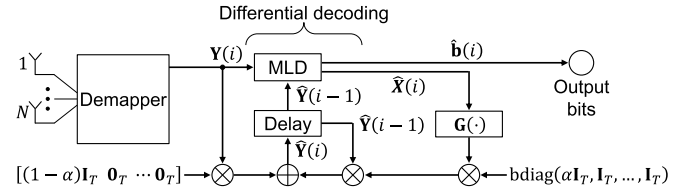
| Type | CSIT | Pilot overhead (Reference block) | Error propagation | Computational complexity |
|---|---|---|---|---|
| Coherent SM [5, 6] | No | Yes | No | $6N2^{RT} \approx \mathcal{O}(N2^{RT})$ |
| Single-antenna differential star-QAM [21] | No | Short block | No | $4(N+1)2^R + 8 \approx \mathcal{O}(N2^R)$ |
| Square DSTC [10] | No | Short block | No | $2NM(2M+1)2^{RM} \approx \mathcal{O}(NM^2 2^{RM})$ |
| Square DSM [17, 19] | No | Short block | No | $6NM2^{RM} \approx \mathcal{O}(NM2^{RM})$ |
| Semi-blind CE and DD scheme of [22] | No | Short block | Yes | $2J(3N2^{RT} + 2N + 2) \approx \mathcal{O}(JN2^{RT})$ |
| Conjugate BF of [1] | Yes | Partial | No | $2N(2N+1)2^{RT} + 4N^2M \approx \mathcal{O}(N^2 2^{RT})$ |
| Proposed RDSM scheme | No | Short block | Yes | $6N2^{RT} + 4N(M/T+1) \approx \mathcal{O}(N2^{RT})$ |




(a) Transmitter




(b) Transmitter example ($M = 4, T = 1, \mathcal{L} = 2$)




(c) Receiver

Fig. 1. System model of (a) the proposed RDSM transmitter and (c) the RDSM receiver, and (b) an example of the RDSM encoding process with $M = 4$ antennas and $T = 1$ symbol per block.

DV is designed to have a single non-zero element that has a unit absolute value. Based on the $B_2$ information $\mathbf{b}_2(i) \in \mathbb{Z}^{B_2}$, bits are modulated onto an $\mathcal{L}$–PSK symbol $s(i) \in \mathbb{C}$. Then, a symbol vector $\mathbf{x}(i) \in \mathbb{C}^M$ is calculated by

$$\mathbf{x}(i) = s(i)\mathbf{a}(i) \quad (i > M). \tag{4}$$

Finally, the transmission vector $\mathbf{s}(i) \in \mathbb{C}^M$ is differentially encoded as follows:

$$\mathbf{s}(i) = \mathbf{G}\left[\mathbf{s}\left(i-1\right)\right]\mathbf{x}(i) \quad (i > M), \tag{5}$$

where $\mathbf{G}[\boldsymbol{\alpha}]$ is the function that transforms a vector $\boldsymbol{\alpha} \in \mathbb{C}^M$ to a unitary matrix, which is defined by

$$\mathbf{G}[\boldsymbol{\alpha}] = \left[\boldsymbol{\alpha}\ \mathbf{E}\boldsymbol{\alpha}\ \mathbf{E}^2\boldsymbol{\alpha}\ \cdots\ \mathbf{E}^{M-1}\boldsymbol{\alpha}\right] \in \mathbb{C}^{M\times M}, \tag{6}$$

assuming that in $\boldsymbol{\alpha}$ there is only one non-zero element that has a unit absolute value. Furthermore, $\mathbf{E}$ is the lower shift matrix defined by

$$\mathbf{E} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \in \mathbb{Z}^{M\times M}. \tag{7}$$

Note that the inverse operation of $\mathbf{G}$ is given by

$$\mathbf{G}^{-1}\left[\mathbf{G}[\boldsymbol{\alpha}]\right] = \mathbf{G}[\boldsymbol{\alpha}]\left[1\ 0\ \cdots\ 0\right]^{\mathrm{T}} = \boldsymbol{\alpha}. \tag{8}$$

In the differential encoding of (5), the vector $\mathbf{s}(i-1)$ is transformed to an $(M \times M)$–sized unitary matrix $\mathbf{G}[\mathbf{s}(i-1)]$. Hence, at the transmitter, the differentially encoded transmission vector $\mathbf{s}(i)$ is transmitted from $M$ transmit antennas over a single symbol duration.

As mentioned above, in each RDSM frame, $M$ reference symbols $\mathbf{s}(i)$ $(i = 1, \cdots, M)$ must be transmitted before the transmission of differentially encoded symbols $\mathbf{s}(i)$ $(i > M)$ of (5). More specifically, the $i$th reference symbols $\mathbf{s}(i)$ given by the $i$th column of $\mathbf{E}$, while having the corresponding symbol vectors of $\mathbf{x}(i) = \mathbf{s}(1)$ for $i = 1, \cdots, M$.

In order to provide further insight, the differential encoding process of (5) is rewritten as follows:

$$\begin{aligned} \mathbf{s}(i) &= \mathbf{G}[\mathbf{s}(i-1)] \cdot \mathbf{x}(i) \\ &= \mathbf{G}[\mathbf{G}[\mathbf{s}(i-2)] \cdot \mathbf{x}(i-1)] \cdot \mathbf{x}(i) \\ &= \mathbf{G}[\mathbf{s}(i-2)] \cdot \mathbf{G}[\mathbf{x}(i-1)] \cdot \mathbf{x}(i) \\ &= \mathbf{G}[\mathbf{s}(1)] \cdot \mathbf{G}[\mathbf{x}(2)] \cdots \mathbf{G}[\mathbf{x}(i-1)] \cdot \mathbf{x}(i) \\ &= \mathbf{G}[\mathbf{x}(1)] \cdot \mathbf{G}[\mathbf{x}(2)] \cdots \mathbf{G}[\mathbf{x}(i-1)] \cdot \mathbf{x}(i) \end{aligned} \tag{9}$$

Furthermore, by considering $\mathbf{x}(i) = \mathbf{G}^{-1}[\mathbf{G}[\mathbf{x}(i)]]$ of (8), we obtain

$$\begin{aligned} \mathbf{s}(i) &= \mathbf{G}[\mathbf{x}(1)] \cdots \mathbf{G}[\mathbf{x}(i-1)] \cdot \mathbf{G}^{-1}[\mathbf{G}[\mathbf{x}(i)]] \\ &= \mathbf{G}[\mathbf{x}(1)] \cdots \mathbf{G}[\mathbf{x}(i-1)] \cdot \mathbf{G}[\mathbf{x}(i)]\left[1\ 0\ \cdots\ 0\right]^{\mathrm{T}} \\ &= \mathbf{G}^{-1}[\ \underbrace{\mathbf{G}[\mathbf{x}(1)] \cdots \mathbf{G}[\mathbf{x}(i)]}_{\text{unitary differential encoding}}\ ]. \end{aligned} \tag{10}$$

Interestingly, (10) clarifies the relationship between the proposed RDSM scheme's differential encoding process and that of the conventional unitary DSTC scheme. More specifically, as shown in (10), similar to the conventional unitary DSTC scheme, square unitary differential encoding is also performed in the proposed RDSM scheme. The only difference is that,

in the RDSM scheme, the differentially encoded unitary matrix, which is $\mathbf{G}[\mathbf{x}(1)] \cdot \mathbf{G}[\mathbf{x}(2)] \cdots \mathbf{G}[\mathbf{x}(i)]$ in (10), is projected to an $(M \times 1)$–sized rectangular matrix using $\mathbf{G}^{-1}$, before its transmission. This projection operation allows us to increase the transmission rate by a factor of $T$.

By ignoring the rate loss caused by the reference-block insertion, the normalized transmission rate of the proposed scheme is formulated as follows:

$$R = \log_2 Q + \log_2 \mathcal{L}. \tag{11}$$

Furthermore, the effective rate is given by $R^{\text{eff}} = (1-\eta) \cdot R$, where we have an overhead ratio of $\eta = M/W$. The proposed RDSM scheme's transmission rate of (11) does not include the parameter $M$, and hence is free from the above-mentioned limitation imposed on the conventional DSTC scheme's transmission rate.

As an illustrative example, Fig. 1(b) shows the encoding process of the proposed scheme, employing BPSK modulation, $M = 4$ transmit antennas, and $Q = 4$ DVs of $\mathbf{a}_1 = [1\ 0\ 0\ 0]^{\text{T}}$, $\mathbf{a}_2 = [0\ 1\ 0\ 0]^{\text{T}}$, $\mathbf{a}_3 = [0\ 0\ 1\ 0]^{\text{T}}$, and $\mathbf{a}_4 = [0\ 0\ 0\ 1]^{\text{T}}$. Next, we consider the $\mathbf{b}(i) = [1\ 0\ 0]$ bit sequence to be input at block index $i$. Then, since we have the $B_1 = \log_2 M = 2$ bits and the $B_2 = \log_2 \mathcal{L} = 1$ bit, the input information bits [1 0 0] are S/P converted to [1 0] and [0] bit sequences. Based on the first [1 0] bits, the DV $\mathbf{a}_2 = [0\ 1\ 0\ 0]^{\text{T}}$ is selected, and, based on the second [0] bit, the BPSK symbol $s(i) = +1$ is modulated. Then, the symbol vector is calculated as $\mathbf{x}(i) = +1 \cdot \mathbf{a}_2$. Finally, the differentially encoded transmission vector $\mathbf{s}(i)$ is generated as $\mathbf{G}[\mathbf{s}(i-1)] \cdot \mathbf{x}(i)$. Suppose that the previous transmission vector is $\mathbf{s}(i-1) = [0\ 1\ 0\ 0]^{\text{T}}$, then the transmission vector at the $i$th block is calculated as $\mathbf{s}(i) = \mathbf{G}[\mathbf{s}(i-1)] \cdot [0\ 1\ 0\ 0]^{\text{T}} = [0\ 0\ 1\ 0]^{\text{T}}$, which is transmitted from $M = 4$ antennas over $T = 1$ symbol intervals per block.

### *B. Generalized RDSM* $(1 \le T \le M)$

In this section, we generalize the RDSM scheme described in Section III-A to support an arbitrary number of symbol intervals per block T. More specifically, the G-RDSM scheme is capable of striking a flexible balance between the transmission rate and the diversity gain, while subsuming several conventional DSTC schemes, such as the classic square DSTC scheme $(T = M)$, the conventional square DSM scheme $(T = M)$, and the RDSM scheme $(T = 1)$ proposed in Section III-A.

In the G-RDSM scheme, the DVs $\mathbf{a}_q$ $(q = 1, \cdots, Q)$, which are used in the RDSM scheme, are replaced by the DMs $\mathbf{A}_q$ $(q = 1, \cdots, Q) \in \mathbb{C}^{M \times T}$, and so the number of symbols per transmission block $T$ is in the range of $1 \le T \le M$. By increasing the number of symbols per transmission block, the transmit diversity gain increases, which is achieved at the cost of the reduction in a transmission rate. The DMs of the proposed G-RDSM scheme must satisfy the following constraints:

- Only a single non-zero element that has a unit absolute value is contained in each column of the DMs $\mathbf{A}_q$ $(q = 1, \cdots, Q)$, and $M/T$ is an integer value.
- Next, let us define the operation $\mathbf{G}'[\bullet]$, which transforms an $(M \times T)$–sized rectangular matrix $\Psi \in \mathbb{C}^{M \times T}$ to a unitary matrix $\mathbf{G}'[\Psi] \in \mathbb{C}^{T \times T}$, by $\mathbf{G}'[\Psi] = \left[\Psi\ \mathbf{E}^T\Psi\ \mathbf{E}^{2T}\Psi\ \cdots\ \mathbf{E}^{M-T}\Psi\right] \in \mathbb{C}^{M \times M}$. Note that the $\mathbf{G}'[\bullet]$ function becomes equivalent to the $\mathbf{G}[\bullet]$ function of (6) for the $T = 1$ scenario. The DMs are designed so that the matrices $\mathbf{G}'[\mathbf{A}_q]$ $(q = 1, \cdots, Q)$ are unitary.
- Finally, the number of DMs $Q$ is less than $M!^{(M/T)}$, and $\log_2 Q$ is an integer value.[4]

In order to attain the best achievable performance in the proposed G-RDSM scheme, the DMs have to be designed in an appropriate manner. The design guideline of the DMs is presented later in Section IV.

The differential encoding process of the G-RDSM scheme is similar to that of the RDSM scheme, except that the DMs, rather than the DVs, are used. Each G-RDSM frame contains $W/T$ transmission blocks $\mathbf{S}(i)$ $(i = 1, \cdots, W/T) \in \mathbb{C}^{M \times T}$, which span over $W$ symbol intervals. Furthermore, the G-RDSM frame consists of the $M/T$ reference blocks $\mathbf{S}(i)$ $(i = 1, \cdots, M/T)$ and the $(W/T - M/T)$-length symbol blocks $\mathbf{S}(i)$ $(i = M/T+1, \cdots, W/T)$, where $W/T$ and $M/T$ are assumed to be integer values.

At the block index $i$ $(> M/T)$, a rectangular symbol matrix $\mathbf{X}(i) \in \mathbb{C}^{M \times T}$ is generated based on $B = B_1 + B_2 = \log_2 Q + \log_2 \mathcal{L}$ information bits as follows: $\mathbf{X}(i) = s(i)\mathbf{A}(i)$ $(i > M/T)$, where $\mathbf{A}(i)$ is the activated DM among $\mathbf{A}_q$ $(q = 1, \cdots, Q)$ DMs based on the $B_1$ bits, whereas $s(i) \in s_l$ $(l = 1, \cdots, \mathcal{L})$ is the $\mathcal{L}$-point APSK symbol, modulated according to the $B_2$ bits. Finally, a transmission matrix $\mathbf{S}(i)$ is differentially encoded by $\mathbf{S}(i) = \mathbf{G}'[\mathbf{S}(i-1)]\mathbf{X}(i)$ $(i > M/T)$, which is transmitted from the $M$ transmit antennas over $T$ symbol intervals. Furthermore, the inverse operation $\mathbf{G}'^{-1}[\cdot]$ is represented by

$$\mathbf{G}'^{-1}\left[\mathbf{G}'[\mathbf{S}(i)]\right] = \mathbf{G}'[\mathbf{S}(i)]\underbrace{[\mathbf{I}_T\ \mathbf{0}_T\ \cdots\ \mathbf{0}_T]^{\text{T}}}_{M \text{ rows}} = \mathbf{S}(i),$$

where $\mathbf{0}_T$ denotes a $(T \times T)$–sized matrix having $T^2$ zeros.

The reference blocks $\mathbf{S}(i)$ $(i = 1, \cdots, M/T)$ are given by

$$\mathbf{S}(1) = \begin{bmatrix} \mathbf{0}_T \\ \mathbf{I}_T \\ \mathbf{0}_T \\ \vdots \\ \mathbf{0}_T \end{bmatrix}, \quad \mathbf{S}(2) = \begin{bmatrix} \mathbf{0}_T \\ \mathbf{0}_T \\ \mathbf{I}_T \\ \vdots \\ \mathbf{0}_T \end{bmatrix}, \cdots, \quad \mathbf{S}(M/T) = \begin{bmatrix} \mathbf{I}_T \\ \mathbf{0}_T \\ \mathbf{0}_T \\ \vdots \\ \mathbf{0}_T \end{bmatrix}. \tag{12}$$

which correspond to the symbol matrices of $\mathbf{X}(i) = \mathbf{S}(1)$ for $i = 1, \cdots, M/T$.

Similar to the RDSM scheme, the effective transmission rate of the G-RDSM scheme is represented as $R^{\text{eff}} = (1-\eta) \cdot (\log_2 Q + \log_2 \mathcal{L})/T$.

Note that the proposed G-RDSM scheme with $T = M$ is equivalent to the classic square DSM proposed in [17]. Furthermore, the G-RDSM scheme with $T = 1$ is equivalent to the RDSM scheme described in Section III-A, and hence we only consider the G-RDSM scheme in the remainder of

[4] $n!^{(k)}$ represents the multifactorial, which is defined as $n!^{(k)} = n \cdot (n-k) \cdot (n-2k) \cdots k$.

the present paper, instead of separately treating the RDSM and G-RDSM schemes.

In order to provide further insight, the proposed G-RDSM scheme may be viewed as the noncoherent version of the previous space-time shift keying (STSK) scheme [15], where the size of the DMs is determined in a flexible manner.[5] The DM optimization algorithms in the literature [6], [15], [34], [35] are readily applicable to the proposed G-RDSM scheme.

### C. Channel Model

For the proposed G-RDSM receiver, the signals received at the $i$th block are represented by $\mathbf{Y}(i) = \mathbf{H}(i)\mathbf{S}(i) + \mathbf{V}(i) \in \mathbb{C}^{N\times T}$, where $\mathbf{V}(i)$ are the additive white Gaussian noise (AWGN) components, each of which follows a zero-mean complex-valued Gaussian distribution of $\mathcal{CN}(0, \sigma_v^2)$, with a variance of $\sigma_v^2$. Furthermore, $\mathbf{H}(i) \in \mathbb{C}^{N\times M}$ denotes the complex-valued channel coefficients.

For the time-invariant quasi-static Rayleigh fading scenario, the channel coefficients of $\mathbf{H}(i)$ are given by random variables, following the zero-mean unit-variance complex-valued Gaussian distribution of $\mathcal{CN}(0, 1)$, which remain constant over each G-RDSM frame, i.e., $1 \le i \le W/T$.

For the scenario of time-varying Rayleigh fading channels, the Jakes channel model [36], [37] is used to generate the channel coefficients, where the $n$th row and the $m$th column of the channel matrix $\mathbf{H}(i)$ is defined by $\mathbf{H}(i)^{(n,m)} = \frac{1}{\sqrt{N_s}} \sum_{n=1}^{N_s} \cos\left[2\pi \cos\left(\xi_n^{(n,m)}\right) F_d T_s i + \zeta_n^{(n,m)}\right] + \frac{j}{\sqrt{N_s}} \sum_{n=1}^{N_s} \sin\left[2\pi \cos\left(\xi_n^{(n,m)}\right) F_d T_s i + \bar{\zeta}_n^{(n,m)}\right]$. Here, $\xi_n^{(n,m)}$, $\zeta_n^{(n,m)}$, and $\bar{\zeta}_n^{(n,m)}$ are random variables that are uniformly distributed in the range of $[0, 2\pi)$. Furthermore, $N_s$ represents the number of scatters, while $F_d$ denotes the maximum Doppler frequency, and $T_s$ is the symbol duration.

### D. Suboptimal Noncoherent Detection

In the proposed G-RDSM receiver, $B$ information bits $\mathbf{b}(i)$ are detected by identifying the symbol matrix $\mathbf{X}(i) \in \mathbf{X}_j$ $(j = 1, \cdots, 2^B)$. More specifically, Fig. 1(c) shows the proposed suboptimal detection algorithm, which is formulated as follows:

$$\hat{\mathbf{X}}(i) = \arg\min_{\mathbf{X}_j} \left\| \mathbf{Y}(i) - \hat{\mathbf{Y}}(i-1)\cdot \mathbf{X}_j \right\|_{\mathrm{F}}^2 \ (i > M/T), \quad (13)$$

where we have

$$\hat{\mathbf{Y}}(i) = \hat{\mathbf{Y}}(i-1)\mathbf{G}'\left[\hat{\mathbf{X}}(i)\right] \mathrm{bdiag}\underbrace{[\alpha\mathbf{I}_T, \mathbf{I}_T, \cdots, \mathbf{I}_T]}_{M/T} + \mathbf{Y}(i)\left[(1-\alpha)\mathbf{I}_T \ \mathbf{0}_T \ \cdots \ \mathbf{0}_T\right] (i \ge 1), \quad (14)$$

Moreover, the initial value of (14) is given by $\hat{\mathbf{Y}}(0) = \mathbf{0}$. Here, bdiag$(\cdot)$ represents the block diagonalization, and $\alpha \in \mathbb{R}$ $(0 \le \alpha \le 1)$ is a forgetting factor, which reduces the effects of error propagation. For the block index of $i = 1, \cdots, M/T$, which corresponds to the reference symbol transmissions, $\alpha$ in (14) is set to zero.

The forgetting factor contributes to the mitigation of the effects of error propagation, which is caused by the rapidly changing channels as well as the detection error in the previous blocks. More specifically, as shown in (14), $\hat{\mathbf{Y}}(i)$ of (13) contains the $(i-1)$ previously estimated blocks $\hat{\mathbf{X}}(i-1), \cdots, \hat{\mathbf{X}}(1)$ and the $(i-1)$ previously received signal blocks $\mathbf{Y}(i-1), \cdots, \mathbf{Y}(1)$. Hence, the proposed suboptimal detection may suffer from the effects of error propagation. The design guideline of $\alpha$ is also provided later in Section IV. Note that in the classic DSTC scheme, relying on the differential encoding of square matrices, a symbol matrix $\mathbf{X}(i)$ is estimated from only two successive received signal blocks, $\mathbf{Y}(i-1)$ and $\mathbf{Y}(i)$, as shown in (3). Hence, unlike the proposed G-RDSM scheme, the detection error in the previous detection block does not affect the detection at the $i$th block.

## IV. Design Guidelines of the G-RDSM System

In this section, we provide the design guidelines for the system parameters $(M, N, T, Q, \mathcal{L}, \alpha)$ and $\mathbf{A}_q$ $(q = 1, \cdots, Q)$ in the proposed G-RDSM scheme. More specifically, the proposed scheme is designed according to the following four steps:

1) Determine the number of transmit antennas $M$ and the number of receive antennas $N$, based on the specifications of the BS and mobile terminals.
2) Based on the specific target rate $R$, determine the constellation size $\mathcal{L}$, the number of DMs $Q \le M!^{(M/T)}$, and the number of symbol intervals $T$ per transmission block, while considering the transmit diversity gain required.
3) Optimize the DMs $\mathbf{A}_q$ $(q = 1, \cdots, Q)$ based on the criterion employed, such as the rank- and determinant-criterion (RDC) [24] and constrained capacity [35].
4) Based on the frame length $W$ and the channel statistics, determine $\alpha$ at the receiver.

In the remainder of this section, we describe in detail Steps 3) and 4) presented in Sections IV-A and IV-B, respectively. Furthermore, in Section IV-C, we also derive the closed-form near-optimal value of $\alpha$ for the time-invariant channels.

### A. DM Optimization

First, $T$ non-zero positions in each DM $\mathbf{A}_q$ are determined, while satisfying the constraint that $\mathbf{G}[\mathbf{A}_q]$ $(i = 1, \cdots, Q)$ must be unitary matrices. Under these constraints, the number of DMs $Q$ cannot exceed $M!^{(M/T)}$, as mentioned in Section III-B. In order to determine the non-zero positions of the $Q$ DMs while satisfying the above-mentioned constraints, we maximize the minimum Hamming distance (MHD) between $\mathbf{A}_q$ and $\mathbf{A}_{q'}$ for $1 \le q < q' \le Q$.

Next, the coefficients of the non-zero elements in the DMs are designed based on the RDC criterion [24], which maximizes the coding gain while maintaining a high diversity order. The pairwise error probability is upper-bounded by [38]

$$\mathrm{P}\left(\mathbf{X}(i) \to \mathbf{X}'(i)\right) \le \frac{1}{\prod_{i=1}^{\min(M,T)} \mu_i^N} \left(\frac{1}{4\sigma_v^2}\right)^{-m'N}, \quad (15)$$

[5]The SM schemes employing $T \ge 2$ have been proposed in the literature [31]–[33]. Combining the proposed encoding and decoding process with the conventional schemes [31]–[33] is left for future work.

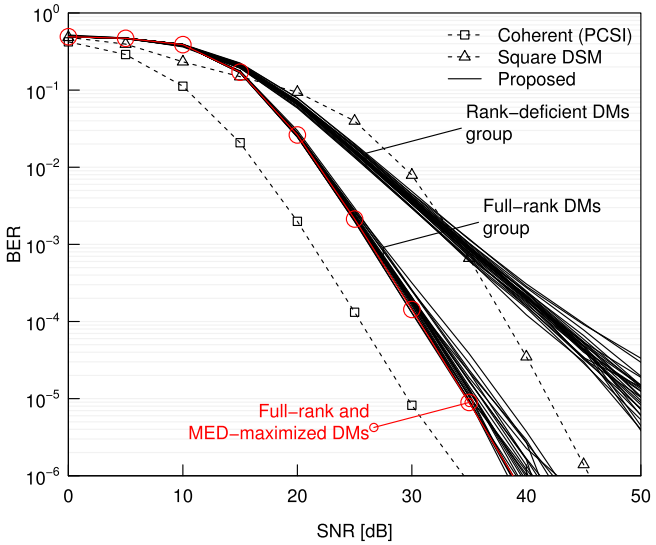



Fig. 2. BER examples of the proposed scheme employing 40 rank-deficient and full-rank DMs sets, where $M = 6$, $N = 1$, $T = 3$, and $R = 3$. The channel coefficients follow quasi-static Rayleigh fading, and the pilot insertion ratio $\eta$ was set to 5%.

which corresponds to the scenario in which the receiver misdetects codeword $\mathbf{X}(i)$ as $\mathbf{X}'(i)$. Here, $\mu_i$ represents the $i$th eigenvalue of $\mathbf{D} = (\mathbf{X}(i) - \mathbf{X}'(i))(\mathbf{X}(i) - \mathbf{X}'(i))^{\mathrm{H}}$, and $m'$ is the rank of matrix $\mathbf{D}$. In (15), the rank $m'$ and the determinant $\prod_{m=1}^{\min(M,T)} \mu_i^N$ correspond to the diversity order and the coding gain, respectively. An exhaustive search is carried out to determine the non-zero coefficients of the DMs by maximizing the minimum determinant $\prod_{m=1}^{\min(M,T)} \mu_i^N$ for the legitimate DM pairs while attaining the full rank of $m' = T$.[6]

Fig. 2 shows the BER curves of the proposed G-RDSM systems having system parameters $(M, N, T, Q, \mathcal{L}) = (6, 1, 3, 32, 16)$, where non-zero coefficients corresponding to the 40 DM sets were randomly generated. The normalized transmission rate for all the schemes was $R = 3$ bps/Hz. The BER curve of the classic DSM scheme having $Q = 64$ DMs is also shown as a benchmark. The red curve corresponds to the G-RDSM scheme that exhibited the best performance. As shown in Fig. 2, the full-rank DMs achieved a higher diversity order than the rank-deficient DMs.

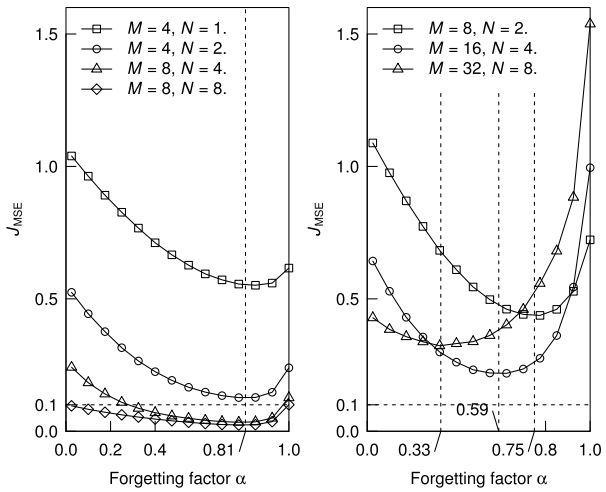
To elaborate a little further, let us exemplify the DM set of the G-RDSM scheme having $(M, T, Q, \mathcal{L}) = (4, 2, 4, 4)$, where we have the $Q = 4$ DMs of

$$\mathbf{A}_1 = \begin{bmatrix} e^{j1.041\pi} & 0.000 \\ 0.000 & e^{j1.571\pi} \\ 0.000 & 0.000 \\ 0.000 & 0.000 \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} 0.000 & e^{j1.503\pi} \\ e^{j1.609\pi} & 0.000 \\ 0.000 & 0.000 \\ 0.000 & 0.000 \end{bmatrix},$$

$$\mathbf{A}_3 = \begin{bmatrix} 0.000 & 0.000 \\ 0.000 & 0.000 \\ e^{j0.633\pi} & 0.000 \\ 0.000 & e^{j0.182\pi} \end{bmatrix}, \quad \mathbf{A}_4 = \begin{bmatrix} 0.000 & 0.000 \\ 0.000 & 0.000 \\ 0.000 & e^{j1.304\pi} \\ e^{j0.011\pi} & 0.000 \end{bmatrix},$$

which are searched based on the RDC criterion [24].

[6]Let us note that, practically, the non-zero coefficients in each DM may be set to ones rather than complex values. This does not impart any significant performance loss, especially for the high-$M$ massive-MIMO scenario.




(a) Quasi-static channel (b) Time-varying channel

Fig. 3. Relationship between the MSE of $\hat{\mathbf{Y}}(W)$ and $\alpha$ in the quasi-static and time-varying Rayleigh fading channels, where we considered $(T, \mathcal{L}) = (1, 2)$, and the pilot insertion ratio $\eta$ was set to 5%.

### B. Forgetting Factor Optimization

The forgetting factor is designed to minimize the mean-square error (MSE) of $\hat{\mathbf{Y}}(i)$, which is given as

$$J_{\mathrm{MSE}} = \frac{T}{W - M} \frac{1}{MN} \sum_{i=M/T+1}^{W/T} \mathrm{E}\left[ \left\| \tilde{\mathbf{Y}}(i) - \hat{\mathbf{Y}}(i) \right\|_{\mathrm{F}}^2 \right], \quad (16)$$

where the channel matrix $\mathbf{H}(i)$, the AWGNs $\mathbf{V}(i)$, and the information bits $\mathbf{b}(i)$ are randomly generated over all data blocks. Note that $\hat{\mathbf{Y}}(i)$ includes symbol blocks $\hat{\mathbf{X}}(i)$, which are estimated by carrying out the proposed detection algorithm of (13) for blocks in the range of $M/T + 1 \leq i \leq W/T$. Furthermore, $\tilde{\mathbf{Y}}(i)$ of (16) is equivalent to $\hat{\mathbf{Y}}(i)$, except that $\tilde{\mathbf{Y}}(i)$ does not contain AWGN components and that the symbol block estimates $\hat{\mathbf{X}}(i)$ are replaced by their error-free versions $\mathbf{X}(i)$. Note that the conventional square DSTC scheme exhibits the MSE of $J_{\mathrm{MSE}} = \sigma_v^2$, which causes the well-known 3 dB BER degradation. Also, the coherent scheme exhibits the MSE of $J_{\mathrm{MSE}} = 0$ because the perfect CSI acquisition is assumed at the receiver.

Figs. 3(a) and 3(b) show the relationship between the MSE of $\hat{\mathbf{Y}}(W)$ and $\alpha$ in the quasi-static and time-varying Rayleigh fading channels, respectively, where we considered $(T, \mathcal{L}) = (1, 2)$ and the pilot insertion ratio $\eta$ was set to 5%. For the quasi-static Rayleigh fading channels, the number of transmit antennas was given as $M = 4$ and 8, whereas the number of receive antennas was varied from $N = 1$ to 8. For the time-varying Rayleigh fading channels, the number of transmit and receive antennas was given as $(M, N) = (8, 2)$, $(16, 4)$, and $(32, 8)$, while maintaining the normalized Doppler frequency to be $F_d T_s = 10^{-3}$. The received SNR was set to 10 dB, which corresponded to $\sigma_v^2 = 0.1$. As shown in Fig. 3(a), for the quasi-static Rayleigh fading channels, the best $\alpha$ was 0.81, regardless of the parameters $(M, N)$ used. In contrast, Fig. 3(b) shows that for the time-varying Rayleigh fading channels, the optimum $\alpha$ changed depending on parameters $(M, N)$.

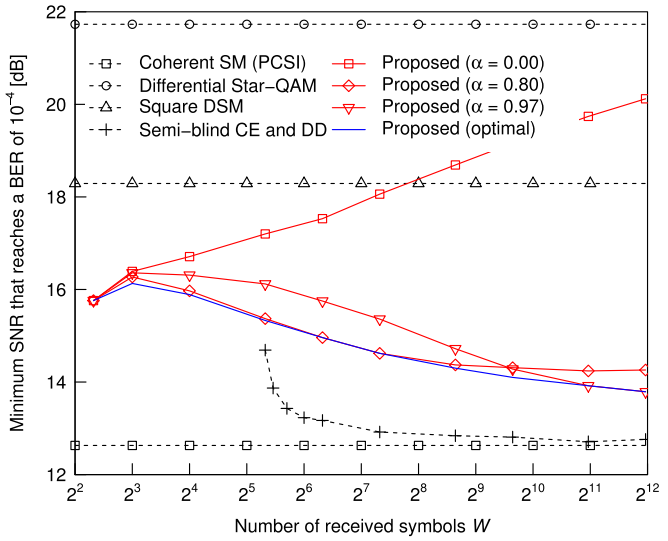



Fig. 4. Effective SNRs required for achieving a BER of $10^{-4}$ by the proposed G-RDSM scheme with parameters $(M, N, T, Q, \mathcal{L}) = (4, 4, 1, 4, 4)$, along with those of its coherent SM counterpart, the square DSM scheme, and the semi-blind CE and DD scheme. The frame length was varied from $W = 5$ to 4000. For example, the frame lengths of $W = 8$, 40, 80 and 400 corresponded to the pilot insertion ratios of $\eta = 50\%$, 10%, 5%, and 1%. Furthermore, the quasi-static Rayleigh fading channel was assumed. The forgetting factor was given as $\alpha = 0$, 0.80, and 0.97.

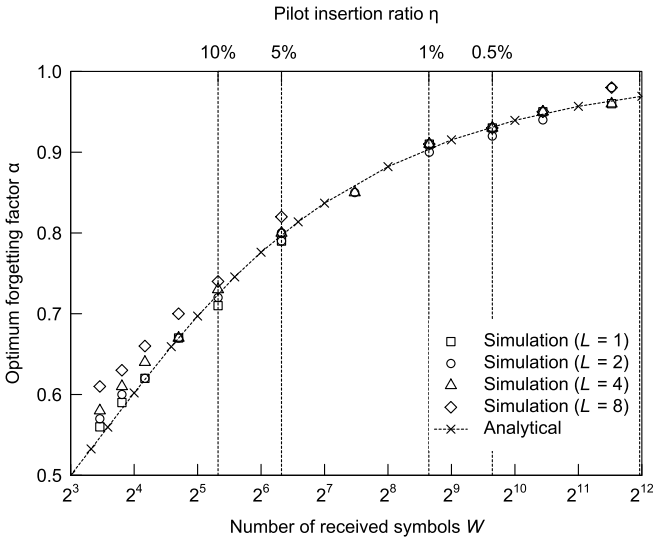



Fig. 5. Relationship between $W$ and $\alpha$ for system parameters $(M, N, T, Q) = (4, 1, 1, 4)$, while varying the constellation size from $\mathcal{L} = 1$ to 8. We also assumed a quasi-static Rayleigh fading channel. The approximation curve of (18) was plotted.

Fig. 4 shows the effective SNRs required in order to achieve a BER of $10^{-4}$ by the proposed G-RDSM scheme with parameters of $(M, N, T, Q, \mathcal{L}) = (4, 4, 1, 4, 4)$, along with those of its coherent SM counterpart, the square DSM scheme having $Q = 16$ DMs, and the semi-blind CE and DD scheme. The frame length was varied from $W = 5$ to 4000. Furthermore, the quasi-static Rayleigh fading channel was assumed. The forgetting factor was given as $\alpha = 0$, 0.80, and 0.97. For reference, we also plotted the optimal bound of the proposed scheme, where the best $\alpha$ was used, depending on the $W$ value. As shown in Fig. 4, the results obtained using the proposed scheme with $\alpha = 0.80$ were better than those obtained for $\alpha = 0$, and 0.97 for $W \leq 400$, which corresponds to $\eta \geq 1\%$. Furthermore, when deactivating the forgetting factor ($\alpha = 0$), the effective SNR increased as $W$ increased. Note that the achievable performance of the classic square DSM scheme remains constant, regardless of $W$, because it is capable of carrying out error-propagation-free block-by-block detection of (3). Moreover, upon increasing $W$, the performance of the semi-blind CE and DD scheme improved, outperforming the proposed G-RDSM counterpart over the entire range. This is because the longer data blocks resulted in better feedback channel estimates in the semi-blind CE and DD scheme.

### C. Analytical Derivation of the Best Forgetting Factor

The $\alpha$ optimization introduced in the previous section is carried out offline. However, calculating the MSE $J_{\mathrm{MSE}}$ of (16) for the high-$M$ massive MIMO scenario may be a challenging task. Therefore, we next approximate the calculation required for the MSE for the quasi-static Rayleigh fading channel.

First, let us simplify the symbol matrices $\mathbf{X}(i)$ $(1 \leq i \leq W/T)$ to be $[\mathbf{0}_T\ \mathbf{I}_T\ \mathbf{0}_T \cdots \mathbf{0}_T]^{\mathrm{T}} \in \mathbb{R}^{M \times T}$, similar to the reference blocks, while assuming $\hat{\mathbf{X}}(i) = \mathbf{X}(i)$. Then, the $J_{\mathrm{MSE}}$ of (16) may be simplified significantly to $J'_{\mathrm{MSE}} = \sum_{i=2}^{K} \left[\alpha^{2(i-1)} + (i-1)(1-\alpha)^2\alpha^{2(K-i)}\right]$, where we have $K = W/M = 1/\eta$. The derivative of $J'_{\mathrm{MSE}}$ with respect to $\alpha$ is given by

$$\frac{d}{d\alpha}J'_{\mathrm{MSE}} = \sum_{i=2}^{K}\Big[2(i-1)\Big(\alpha^{2i-3} - (1-\alpha)\alpha^{2(K-i)} + (1-\alpha)^2(K-i)\alpha^{2(K-i)-1}\Big)\Big]. \quad (17)$$

By considering $\frac{d}{d\alpha}J'_{\mathrm{MSE}} = 0$ in (17), we obtain

$$2(K-1)\alpha^{2K+2} + \alpha^{2K+1} - (2K+1)\alpha^{2K} - K\alpha^3 + (K+2)\alpha^2 + (K-1)\alpha - K + 1 = 0. \quad (18)$$

Fig. 5 shows the relationship between $W$ and $\alpha$ in the quasi-static Rayleigh fading scenario, comparing the numerical search and the analysis calculated by the approximated equation of (18). As shown in Fig. 5, the analytical curve closely matches that of the numerical search. More specifically, for $W = 80$ (i.e., $\eta = 5\%$), the forgetting factor calculated by (18) was $\alpha = 0.80$, whereas that of the numerical search was $\alpha = 0.81$, as shown in Fig. 3.

## V. Complexity Analysis

In this section, we compare the decoding complexity of the proposed and conventional schemes, which is imparted by the ML detection per symbol duration. Here, similar to [39], the number of real-valued multiplications needed for the decoding process is determined. For example, the multiplication of two complex values requires four real-valued multiplications.

For a given rate $R$, the computational complexity of the coherent SM scheme is given by $6N2^{RT} \approx O(N2^{RT})$,

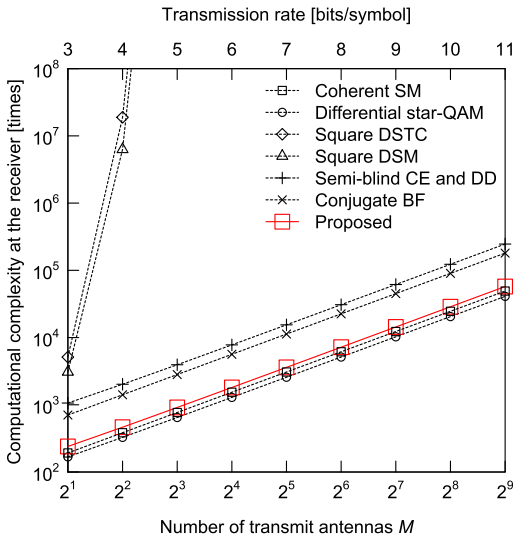



Fig. 6. Decoding complexities of the MIMO schemes listed in Table I. For the proposed G-RDSM scheme, the system parameters are $(N, T, \mathcal{L}) = (4, 1, 4)$. The number of transmit antennas was changed from $M = 2$ to 512 while maintaining $Q = M$.

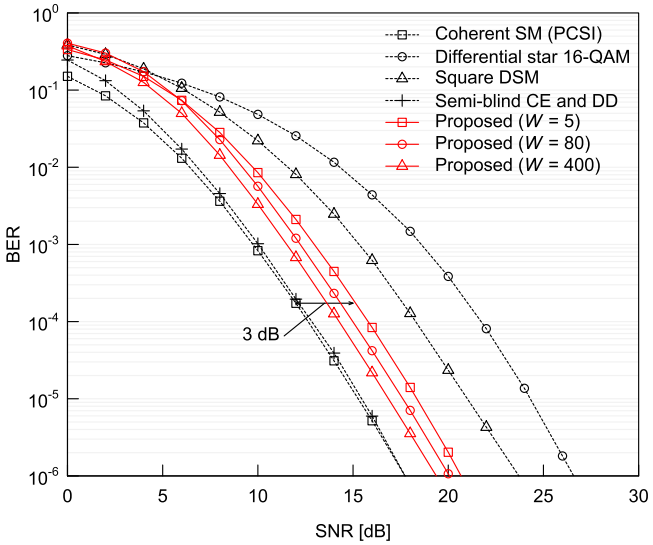



Fig. 7. BER comparisons in the quasi-static Rayleigh fading channel, where $M = N = 4$ transmit and receive antennas were used, while considering a transmission rate of $R = 4$ bps/Hz.

since the ML detection of $\hat{\mathbf{S}} = \arg\min_{\mathbf{S}_j} \|\mathbf{Y} - \mathbf{H} \cdot \mathbf{S}_j\|_{\mathrm{F}}^2$ requires $2^B$ $N \cdot T$ complex-valued multiplications as well as norm calculations, where $\mathbf{S}_j$ $(1 \leq j \leq 2^B)$ represents the legitimate symbol matrices. Then, the total number of real-valued multiplications per symbol interval is calculated as $6NT2^B/T = 6N2^B = 6N2^{RT}$.

In the proposed G-RDSM scheme, the decoding complexity of the suboptimal detection is formulated by $6N2^{RT} + 4N\,(M/T + 1)$, which is $O(N2^{RT})$. For the purpose of comparison, the computational complexities of the coherent SM scheme, the single-antenna differential star-QAM scheme, the square DSTC scheme, the square DSM scheme, the semi-blind CE and DD scheme, and the conjugate BF scheme are listed in Table I. Note that the complexity needed for channel estimations in the coherent SM and the conjugate BF are not included in the evaluations.

Fig. 6 shows the decoding complexities of the proposed G-RDSM scheme and the benchmark schemes listed in Table I. For the proposed G-RDSM scheme, the system parameters were given as $(N, T, \mathcal{L}) = (4, 1, 4)$, and the number of transmit antennas was varied from $M = 2$ to 512 while maintaining $Q = M$. As shown in Fig. 6, the proposed G-RDSM scheme, the coherent SM scheme, and the single-antenna differential star-QAM scheme exhibited comparable low decoding complexities, specifically, approximately 10 times lower than the complexities of the semi-blind CE and DD scheme and the conjugate BF scheme. As expected, the ML complexities of the square DSM scheme and the square DSTC scheme were prohibitively high, especially for $R > 5$ bps/Hz. For example, the ML decoding complexity of the DSTC was $1.2 \times 10^{15}$ for $M = 8$ and $3.35 \times 10^{32}$ for $M = 16$.

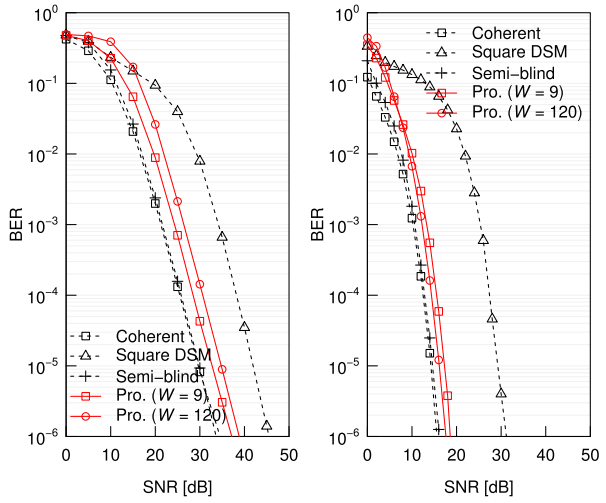
## VI. Performance Results

In this section, we present our performance results in order to characterize the proposed G-RDSM scheme in uncoded scenarios. We considered both the quasi-static Rayleigh fading channel and the time-varying Rayleigh fading channel obeying the classic Jakes model of Section III-C. The single-antenna differential star-QAM scheme [21], the conventional square DSM scheme [17], the semi-blind CE and DD scheme [16], and the conjugate BF scheme [1] were considered as benchmarks for the proposed G-RDSM scheme. The coherent counterparts of the proposed scheme were equivalent to the SM scheme [5] for $T = 1$ and the asynchronous STSK scheme [6] for $T \geq 2$. The DMs of the conventional DSM scheme and those of the proposed G-RDSM scheme were obtained by the RDC criterion [24]. In our simulations, the total transmit power per symbol interval was fixed to unity for all of the schemes.

### A. Scenarios Involving a Small Number of Transmit Antennas ($M \leq 8$)

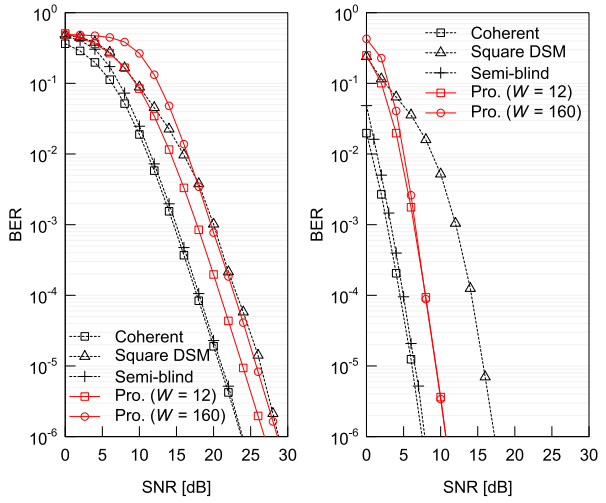
First, in Figs. 7 through 9, the achievable BER performance for a scenario involving a small number of transmit antennas ($M \leq 8$) was investigated assuming a quasi-static Rayleigh fading channel. Note that the effects of the power penalty induced by the reference overhead were not reflected.

Fig. 7 shows the achievable BER performance of the proposed G-RDSM scheme for system parameters $(M, N, T, Q, \mathcal{L}) = (4, 4, 1, 4, 4)$ and the benchmark schemes, i.e., the coherent SM scheme, the single-antenna differential star-QAM scheme, the square DSM scheme having $Q = 16$ DMs, and the semi-blind CE and DD scheme. The transmission rate was $R = 4$ bps/Hz. The BER curves of the proposed scheme were plotted for frame lengths of $W = 5$, 80, and 400. Note that the $W = 5$ case corresponds to the minimum frame length scenario, including $M = 4$ reference blocks and a single data block. The curve of $W = 5$ was included as an idealistic reference because the relative overhead was as high as $\eta = 80\%$. As shown in Fig. 7, the proposed scheme,




(a) $N = 1$ (b) $N = 4$

Fig. 8. BER performance of the proposed G-RDSM scheme with parameters (a) $(M, N, T, Q, \mathcal{L}) = (6, 1, 3, 32, 16)$ and (b) $(6, 4, 3, 32, 16)$, which achieved a transmission rate of $R = 3$ bps/Hz.




(a) $N = 1$ (b) $N = 4$

Fig. 9. BER performance of the proposed G-RDSM scheme with the parameters (a) $(M, N, T, Q, \mathcal{L}) = (8, 1, 4, 128, 2)$ and (b) $(8, 4, 4, 128, 2)$, which yields a transmission rate of $R = 2$ bps/Hz.

having a minimum frame length of $W = 5$, exhibited a performance 3 dB lower than the coherent SM counterpart, while outperforming the conventional square DSM scheme and the single-antenna differential star-QAM scheme. Upon increasing $W$ for the proposed scheme, the achievable BER improved, as expected from Fig. 4. The semi-blind CE and DD scheme achieved a performance similar to that of the coherent SM scheme.

Next, as shown in Figs. 8(a) and 8(b), we plotted the BER performance of the proposed G-RDSM scheme with parameters $(M, N, T, Q, \mathcal{L}) = (6, 1, 3, 32, 16)$ and $(6, 4, 3, 32, 16)$, respectively, which achieved a transmission rate of $R = 3$ bps/Hz. The frame lengths for the proposed G-RDSM scheme were $W/T = 3$ and 40. The BER curves of the coherent counterpart of the proposed scheme and the conventional DSM scheme having $Q = 64$ DMs were plotted for reference. As shown in Fig. 8, the proposed G-RDSM scheme outperformed the conventional square DSM scheme while achieving a performance similar to those of the coherent SM scheme and the semi-blind CE and DD scheme. Note that the diversity order of the proposed scheme was $\mathcal{D} = 2.8$ for $N = 1$ and $\mathcal{D} = 6.9$ for $N = 4$.

Similarly, Figs. 9(a) and 9(b) show the BER performance of the proposed G-RDSM scheme with parameters $(M, N, T, Q, \mathcal{L}) = (8, 1, 4, 128, 2)$ and $(8, 4, 4, 128, 2)$, respectively, which achieves a transmission rate of $R = 2$ bps/Hz. The number of DMs for the conventional DSM scheme was $Q = 256$. Similar to in Fig. 8, the performance advantage of the proposed G-RDSM scheme over the conventional square DSM scheme was confirmed. The diversity order of the G-RDSM scheme was $\mathcal{D} = 3.3$ for $N = 1$ and $\mathcal{D} = 6.7$ for $N = 4$.

### *B. Massive-MIMO Scenarios*

Next, we evaluated the proposed G-RDSM scheme in massive-MIMO scenarios while considering time-varying Rayleigh fading channels. The single-antenna differential star-QAM scheme, the conventional square DSM scheme, and the conjugate BF scheme were considered as benchmarks. Since the semi-blind CE and DD scheme of [22] was originally proposed for the time-invariant channels, it was not included here.[7] In Figs. 10 through 12, the proposed G-RDSM scheme used the $T = 1$ time slot per block and the constellation size $\mathcal{L} = 4$.

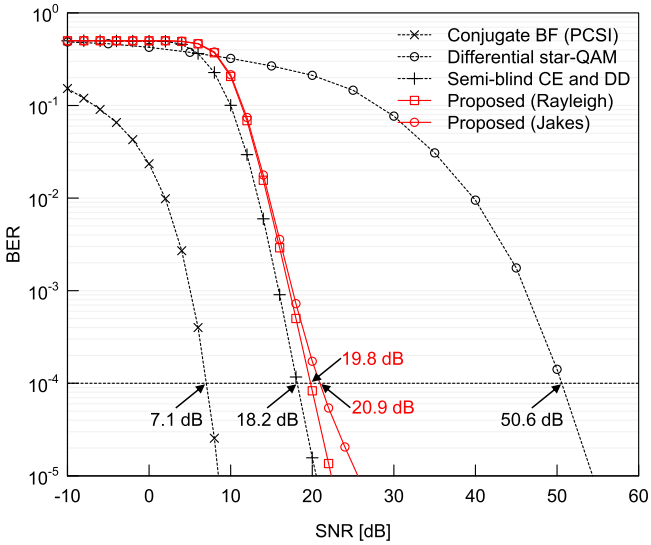
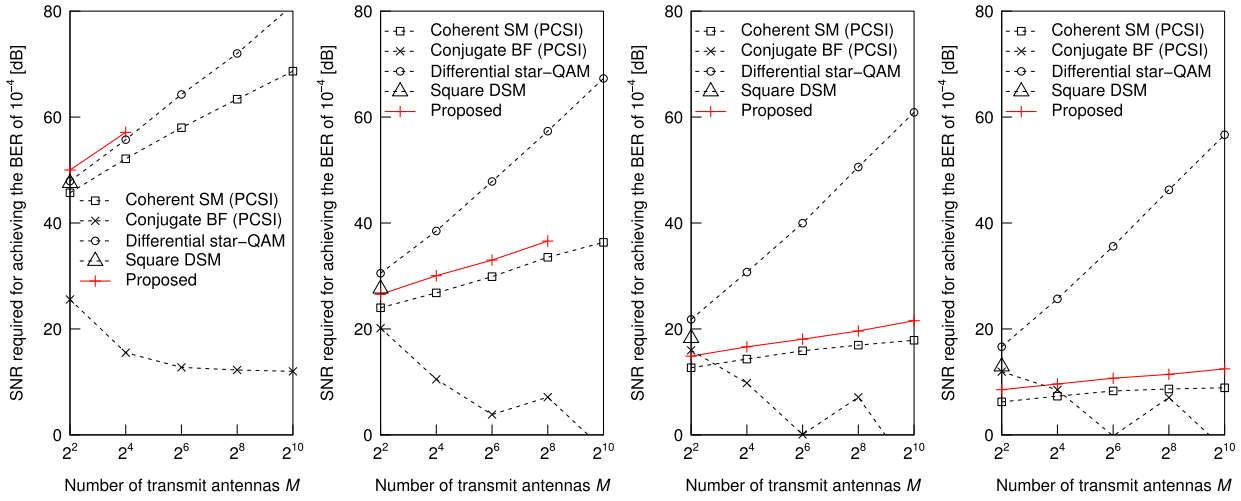



Fig. 10. BER of the proposed scheme and the single-antenna differential star-QAM scheme, where $(M, N) = (256, 8)$ and $R = 10$ bps/Hz. The normalized Doppler frequency was given by $F_d T_s = 10^{-5}$.

Fig. 10 shows the BER performance of the proposed G-RDSM scheme for $M = 256$ transmit antennas, which corresponds to transmission rates of $R = 10$ bps/Hz. Here, the normalized Doppler frequency was given by $F_d T_s = 10^{-5}$. For comparison, the BER curves of the single-antenna

[7]Note that in the quasi-static channels, the semi-blind CE and DD scheme is capable of achieving the near-PCSI performance, as shown in Figs. 7, 8, and 9, which tends to outperform the proposed G-RDSM scheme.




(a) $N = 1$ (b) $N = 2$ (c) $N = 4$ (d) $N = 8$

Fig. 11. Effective SNRs recorded for BER= $10^{-4}$, where the number of transmit antennas was set to $M = 4, 16, 64, 256$, and 1024 corresponding to transmission rates of $R = 4, 6, 8, 10$, and 12 bps/Hz, as well as the frame lengths of $W = 80$, 320, 1280, 5120, and 20480, respectively. The normalized Doppler frequency was set to $F_d T_s = 1 \times 10^{-6}$.

TABLE II

EFFECTIVE SNRS RECORDED FOR BER $= 10^{-4}$, WHERE WE HAVE $(M, N) = (64, 4)$

| $F_d T_s$ | Coherent SM (PCSI) | Conjugate BF (PCSI) | Differential star-QAM | Square DSM | Proposed |
|---|---|---|---|---|---|
| 0 | 15.88 | 0.080 | 39.97 | N/A | 18.36 |
| $1 \times 10^{-6}$ | 15.88 | 0.080 | 39.97 | N/A | 18.36 |
| $1 \times 10^{-5}$ | 15.88 | 0.080 | 39.97 | N/A | 18.36 |
| $7 \times 10^{-5}$ | 15.88 | 0.080 | 39.97 | N/A | 25.56 |
| $1 \times 10^{-4}$ | 15.88 | 0.080 | 39.97 | N/A | $+\infty$ |

differential star-QAM scheme, the conjugate-BF scheme, the semi-blind CE and DD scheme were plotted. As shown in Fig. 10, in the proposed scheme, the effective SNR recorded for BER= $10^{-4}$ was 19.8 dB for the quasi-static channel, whereas that for time-varying channel was 20.9 dB, which clearly outperformed the single-antenna star-QAM scheme. Again, let us note that the conventional DSTC and DSM schemes do not operate in the high-rate scenario, such as $R = 10$ bps/Hz, since they do not have any rate scalability with respect to the number of transmit antennas.

Fig. 11 shows the effective SNRs for various numbers of transmit antennas (BER= $10^{-4}$). The number of transmit antennas was set to $M = 4$, 16, 64, 256, and 1024, corresponding to transmission rates of $R = 4$, 6, 8, 10, and 12 bps/Hz, as well as the frame lengths of $W = 80$, 320, 1280, 5120, and 20480, respectively. The normalized Doppler frequency was $F_d T_s = 1 \times 10^{-6}$. Here, the conventional square DSM scheme was only plotted for $R = 4$ bps/Hz because its decoding complexity was prohibitively high for $R > 4$ bps/Hz. As shown in Fig. 11, the proposed G-RDSM scheme exhibited better performance than two other differentially encoded benchmark schemes, i.e., the conventional square DSM scheme and the single-antenna differential star-QAM scheme. The performance advantage of the G-RDSM scheme was particularly notable for $N \geq 2$ scenarios. Furthermore, the performance gap between the proposed G-RDSM scheme and the coherent SM scheme remained at 3 dB for the $F_d T_s = 1 \times 10^{-6}$ scenario. Note that the conjugate BF scheme achieved the best performance in all of the scenarios owing to the idealistic assumption of PCSI at the transmitter.

Additionally, Table II compares the effective SNR required for achieving the BER of $10^{-4}$, where the normalized Doppler frequency was given as $F_d T_s = 0$, $1 \times 10^{-6}$, $1 \times 10^{-5}$, $7 \times 10^{-5}$, and $1 \times 10^{-4}$. The numbers of transmit and receive antennas were fixed to $M = 64$ and $N = 4$, respectively. Other simulation parameters were the same with those used in Fig. 11. Since we considered $M = 64$ transmit antennas, the effective SNRs of the conventional square DSM were not available (N/A) due to the complexity issue analyzed in Section V. The performance gap between the proposed scheme and its coherent counterpart was approximately 3 dB. Upon increasing the normalized Doppler frequency $F_d T_s$, the effective SNR of the proposed scheme increased, and finally exhibited an error floor for $F_d T_s = 1 \times 10^{-4}$.

Finally, Fig. 12 compares the constrained capacity of the proposed scheme, the conjugate-BF scheme, and the single-antenna differential star-QAM scheme, where the number of transmit antennas was given by $M = 4$, 64, and 1024, which corresponded to the transmission rates of $R = 4$, 8, and 12 bps/Hz, respectively.[8] The number of receive antennas was set to $N = 4$. For the proposed G-RDSM scheme,

[8]To be more specific, the constrained capacity of the proposed G-RDSM scheme was calculated, by assuming that the associated capacity is 3-dB lower than its coherent counterpart according to [40].

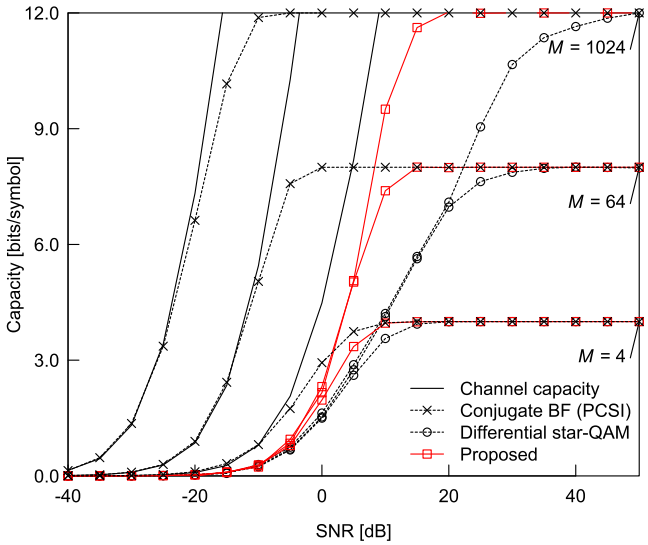



Fig. 12. Constrained capacity comparisons between the proposed scheme, the conjugate-BF scheme, and the single-antenna differential star-QAM scheme, where the number of transmit antennas was given by $M = 4$, 64, and 1024, which corresponded to the transmission rates of $R = 4$, 8, and 12 bps/Hz, respectively. The number of receive antennas was set to $N = 4$. The number of symbols per block was maintained to be $T = 1$ in the proposed scheme. The associated curves of the unconstrained channel capacity were also plotted.

the number of symbols per block was maintained to be $T = 1$ in the proposed scheme, while considering the minimum reference block of $W = M + 1$, for simplicity. The associated unconstrained capacity curves were also plotted. Observe in Fig. 12 that the proposed G-RDSM scheme exhibited a clear rate scalability with respect to the number of transmit antennas. It should be emphasized that this is exclusive benefit for the proposed scheme, which cannot be achieved in other noncoherent DSTC schemes.

## VII. Conclusions

In the present study, we proposed a novel RDSM principle that allows us to construct a rectangular transmission matrix rather than a square unitary matrix, which enables a high transmission rate that is scalable with respect to the number of transmit antennas. Furthermore, we generalized the RDSM scheme to subsume the conventional square unitary DSTC scheme and the square DSM scheme. The performance results of the present study demonstrated that, as a direct result of the proposed differentially encoded rectangular matrix structure, the proposed RDSM scheme outperformed the conventional DSTC family for high-rate massive-MIMO scenarios.

## References

[1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[2] H. Yang and T. L. Marzetta, "Performance of conjugate and zero-forcing beamforming in large-scale antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 172–179, Feb. 2013.

[3] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[4] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[5] M. Di Renzo, H. Haas, A. Ghrayeb, S. Sugiura, and L. Hanzo, "Spatial modulation for generalized MIMO: Challenges, opportunities and implementation," *Proc. IEEE*, vol. 102, no. 1, pp. 1–47, Jan. 2014.

[6] S. Sugiura, S. Chen, and L. Hanzo, "A universal space-time architecture for multiple-antenna aided systems," *IEEE Commun. Surveys Tuts.*, vol. 14, no. 2, pp. 401–420, 2nd Quart., 2012.

[7] P. Yang *et al.*, "Single-carrier SM-MIMO: A promising design for broadband large-scale antenna systems," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1687–1716, 3rd Quart., 2016.

[8] D. A. Basnayaka, M. Di Renzo, and H. Haas, "Massive but few active MIMO," *IEEE Trans. Veh. Technol.*, vol. 65, no. 9, pp. 6861–6877, Sep. 2016.

[9] P. Yang, M. Di Renzo, Y. Xiao, S. Li, and L. Hanzo, "Design guidelines for spatial modulation," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 6–26, 1st Quart., 2015.

[10] 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, no. 2, pp. 543–564, Mar. 2000.

[11] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 7, pp. 1169–1174, Jul. 2000.

[12] B. M. Hochwald and W. Sweldens, "Differential unitary space-time modulation," *IEEE Trans. Commun.*, vol. 48, no. 12, pp. 2041–2052, Dec. 2000.

[13] B. L. Hughes, "Differential space-time modulation," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2567–2578, Nov. 2000.

[14] B. Hassibi and B. M. Hochwald, "Cayley differential unitary space-time codes," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1485–1503, Jun. 2002.

[15] S. Sugiura, S. Chen, and L. Hanzo, "Coherent and differential space-time shift keying: A dispersion matrix approach," *IEEE Trans. Commun.*, vol. 58, no. 11, pp. 3219–3230, Nov. 2010.

[16] S. Sugiura and L. Hanzo, "Effects of channel estimation on spatial modulation," *IEEE Signal Process. Lett.*, vol. 19, no. 12, pp. 805–808, Dec. 2012.

[17] N. Ishikawa and S. Sugiura, "Unified differential spatial modulation," *IEEE Wireless Commun. Lett.*, vol. 3, no. 4, pp. 337–340, Aug. 2014.

[18] S. Sugiura, "Coherent versus non-coherent reconfigurable antenna aided virtual MIMO systems," *IEEE Signal Process. Lett.*, vol. 21, no. 4, pp. 390–394, Apr. 2014.

[19] Y. Bian, X. Cheng, M. Wen, L. Yang, H. V. Poor, and B. Jiao, "Differential spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 64, no. 7, pp. 3262–3268, Jul. 2015.

[20] S. Sugiura and L. Hanzo, "Single-RF spatial modulation requires single-carrier transmission: Frequency-domain turbo equalization for dispersive channels," *IEEE Trans. Veh. Technol.*, vol. 64, no. 10, pp. 4870–4875, Oct. 2015.

[21] W. T. Webb, L. Hanzo, and R. Steele, "Bandwidth efficient QAM schemes for Rayleigh fading channels," *IEE Proc. I-Commun., Speech Vis.*, vol. 138, no. 3, pp. 169–175, Jun. 1991.

[22] S. Chen, S. Sugiura, and L. Hanzo, "Semi-blind joint channel estimation and data detection for space-time shift keying systems," *IEEE Signal Process. Lett.*, vol. 17, no. 12, pp. 993–996, Dec. 2010.

[23] F. Oggier, "Cyclic algebras for noncoherent differential space–time coding," *IEEE Trans. Inf. Theory*, vol. 53, no. 9, pp. 3053–3065, Sep. 2007.

[24] L. Hanzo, O. Alamri, M. El-Hajjar, and N. Wu, *Near-Capacity Multi-Functional MIMO Systems: Sphere-Packing, Iterative Detection and Cooperation*. Hoboken, NJ, USA: Wiley, 2009.

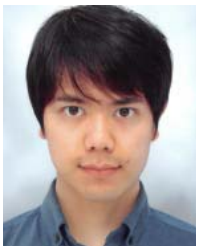
[25] M. Wen, X. Cheng, Y. Bian, and H. V. Poor, "A low-complexity near-ML differential spatial modulation detector," *IEEE Signal Process. Lett.*, vol. 22, no. 11, pp. 1834–1838, Nov. 2015.

[26] J. Li, M. Wen, X. Cheng, Y. Yan, S. Song, and M. H. Lee, "Differential spatial modulation with Gray coded antenna activation order," *IEEE Commun. Lett.*, vol. 20, no. 6, pp. 1100–1103, Jun. 2016.

[27] M. Zhang, M. Wen, X. Cheng, and L. Yang, "A dual-hop virtual MIMO architecture based on hybrid differential spatial modulation," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6356–6370, Sep. 2016.

[28] P. A. Martin, "Differential spatial modulation for APSK in time-varying fading channels," *IEEE Commun. Lett.*, vol. 19, no. 7, pp. 1261–1264, Jul. 2015.

[29] 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, no. 1, pp. 385–394, Jan. 2017.
[30] S. Buzzi, M. Lops, and S. Sardellitti, "Performance of iterative data detection and channel estimation for single-antenna and multiple-antennas wireless communications," *IEEE Trans. Veh. Technol.*, vol. 53, no. 4, pp. 1085–1104, Jul. 2004.
[31] E. Basar, U. Aygolu, E. Panayirci, and H. V. Poor, "Space-time block coded spatial modulation," *IEEE Trans. Commun.*, vol. 59, no. 3, pp. 823–832, Mar. 2011.
[32] M. T. Le, V. D. Ngo, H. A. Mai, X. N. Tran, and M. Di Renzo, "Spatially modulated orthogonal space-time block codes with non-vanishing determinants," *IEEE Trans. Commun.*, vol. 62, no. 1, pp. 85–99, Jan. 2014.
[33] A. G. Helmy, M. Di Renzo, and N. Al-Dhahir, "Enhanced-reliability cyclic generalized spatial-and-temporal modulation," *IEEE Commun. Lett.*, vol. 20, no. 12, pp. 2374–2377, Dec. 2016.
[34] S. Sugiura, "Dispersion matrix optimization for space-time shift keying," *IEEE Commun. Lett.*, vol. 15, no. 11, pp. 1152–1155, Nov. 2011.
[35] S. Sugiura and L. Hanzo, "On the joint optimization of dispersion matrices and constellations for near-capacity irregular precoded space-time shift keying," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 380–387, Jan. 2013.
[36] J. Proakis, *Digital Communications*. New York, NY, USA: McGraw-Hill, 2001.
[37] C.-S. Hwang, S. H. Nam, J. Chung, and V. Tarokh, "Differential space time block codes using nonconstant modulus constellations," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2955–2964, Nov. 2003.
[38] R. W. Heath and A. J. Paulraj, "Linear dispersion codes for MIMO systems based on frame theory," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2429–2441, Oct. 2002.
[39] S. Sugiura, S. Chen, and L. Hanzo, "MIMO-aided near-capacity turbo transceivers: Taxonomy and performance versus complexity," *IEEE Commun. Surveys Tuts.*, vol. 14, no. 2, pp. 421–442, 2nd Quart., 2012.
[40] H. V. Nguyen, C. Xu, S. X. Ng, and L. Hanzo, "Near-capacity wireless system design principles," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 1806–1833, 4th Quart., 2015.



**Naoki Ishikawa** (S'13) was born in Kanagawa, Japan, in 1991. He received the B.E. and M.E. degrees (Hons.) in computer and information sciences from the Tokyo University of Agriculture and Technology, Tokyo, Japan, in 2014 and 2015, respectively, where he is currently pursuing the Ph.D. degree. In 2015, he was an Academic Visitor with the School of Electronics and Computer Science, University of Southampton, U.K., fully funded by the Ministry of Education, Culture, Sports, Science and Technology, Japan.

Dr. Ishikawa was a Research Fellow of the Japan Society for the Promotion of Science. He has received seven domestic student awards, including the Telecom System Technology Student Award (honorable mention) from the Telecommunications Advancement Foundation of Japan in 2014, the Outstanding Paper Award for Young C&C Researchers from NEC Computer and Communications Foundation in 2014, the Young Researcher's Encouragement Award from the IEEE VTS Japan Chapter in 2014, and four honors from the Tokyo University of Agriculture and Technology.

**Shinya Sugiura** (M'06–SM'12) received the B.S. and M.S. degrees in aeronautics and astronautics from Kyoto University, Kyoto, Japan, in 2002 and 2004, respectively, and the Ph.D. degree in electronics and electrical engineering from the University of Southampton, Southampton, U.K., in 2010.

From 2004 to 2012, he was a Research Scientist with Toyota Central Research and Development Laboratories, Inc., Aichi, Japan. Since 2013, he has been an Associate Professor with the Department of Computer and Information Sciences, Tokyo University of Agriculture and Technology, Tokyo, Japan, where he heads the Wireless Communications Research Group. He authored or co-authored over 70 refereed research publications, including 47 IEEE journal and magazine papers. His research has covered a range of areas in wireless communications, networking, signal processing, and antenna technology.

Dr. Sugiura was a recipient of a number of awards, including the Sixth RIEC Award from the Foundation for the Promotion of Electrical Communication in 2016, the Young Scientists' Prize by the Minister of Education, Culture, Sports, Science and Technology of Japan in 2016, the 14th Funai Information Technology Award (First Prize) from the Funai Foundation in 2015, the 28th Telecom System Technology Award from the Telecommunications Advancement Foundation in 2013, the Sixth IEEE Communications Society Asia-Pacific Outstanding Young Researcher Award in 2011, the 13th Ericsson Young Scientist Award in 2011, and the 2008 IEEE Antennas and Propagation Society Japan Chapter Young Engineer Award. He was also certified as an Exemplary Reviewer of the IEEE COMMUNICATIONS LETTERS in 2013 and 2014.