Jiwook Choi, Yunseo Nam, and Namyoon Lee

#### Abstract

This paper proposes spatial lattice modulation (SLM), a spatial modulation method for multipleinput-multiple-output (MIMO) systems. The key idea of SLM is to jointly exploit spatial, in-phase, and quadrature dimensions to modulate information bits into a multi-dimensional signal set that consists of lattice points. One major finding is that SLM achieves a higher spectral efficiency than the existing spatial modulation and spatial multiplexing methods for the MIMO channel under the constraint of *M*-ary pulseamplitude-modulation (PAM) input signaling per dimension. In particular, it is shown that when the SLM signal set is constructed by using dense lattices, a significant signal-to-noise-ratio (SNR) gain, i.e., a nominal coding gain, is attainable compared to the existing methods. In addition, closed-form expressions for both the average mutual information and average symbol-vector-error-probability (ASVEP) of generic SLM are derived under Rayleigh-fading environments. To reduce detection complexity, a low-complexity detection method for SLM, which is referred to as *lattice sphere decoding*, is developed by exploiting lattice theory. Simulation results verify the accuracy of the conducted analysis and demonstrate that the proposed SLM techniques achieve higher average mutual information and lower ASVEP than do existing methods.

#### **Index Terms**

Multiple-input-multiple-output (MIMO), spatial modulation (SM), lattice modulation.

### I. INTRODUCTION

Spatial modulation (SM) [1] is a transmission method that sends information bits using the index of an active antenna and conventional quadrature-amplitude-modulation (QAM) symbols. SM has been proposed to improve both the spectral and energy efficiency of MIMO systems [2]–[5]. For example, when a transmitter is equipped with  $N_t$  antennas that use a radio frequency

J. Choi, Y. Nam, and N. Lee are with the department of electrical engineering, POSTECH, South Korea, emails:{jiwook,edwin624,nylee}@postech.ac.kr

(RF) chain,  $\log_2 (N_t) + \log_2(|\mathcal{C}_Q|)$  bits can be modulated into a spatial symbol vector, where  $\mathcal{C}_Q$  denotes the QAM constellation set and  $|\mathcal{C}_Q|$  represents its cardinality. A simplified version of SM, referred to as space shift keying (SSK) [6] was presented to improve energy efficiency. SSK only maps information bits into the antenna index, so it is able to achieve the spectral efficiency of  $\log_2 (N_t)$  bits/sec/Hz when signal-to-noise ratio (SNR) is high enough.

The concepts of SM and SSK have been generalized in numerous ways by mapping information bits into multiple indices of the transmit antennas. Generalized spatial modulation (GSM) [7] and generalized space shift keying (GSSK) [8] are representative generalizations of SM. The idea of both GSM and GSSK is to map information bits onto an antenna subset that consists of  $N_a$  elements among  $N_t$ . Therefore, a transmitter is able to modulate  $\log_2 {N_t \choose N_a} + \log_2(|C_Q|)$ information bits when using GSM with constellation set  $C_Q$ . This modulation strategy allows sending of  $\log_2 {N_t \choose N_a} - \log_2 {N_t \choose 1}$  more information bits than both SM and SSK. Multiple active spatial modulation (MA-SM) [9] is another variation of GSM, MA-SM was introduced by harnessing multiplexing gains of the MIMO system. MA-SM sends distinct QAM symbols by choosing  $N_a$  active elements among  $N_t$  transmit antennas; thereby,  $\log_2 {N_t \choose N_a} + N_a \log_2(|C_Q|)$ information bits are modulated to a symbol vector with QAM constellation set  $C_Q$ .

Variable set of active antenna GSM (VA-GSM) [10] is another variation of GSM. VA-GSM allows the number of active antennas to vary from 1 to  $N_t$ , while sending the same transmit symbol for each active antenna. In addition, quadrature spatial modulation (QSM) [11] separately exploits in-phase and quadrature signal dimensions. For instance, when a transmitter has  $N_t$ antennas, QSM is able to modulate  $2 \log_2 (N_t) + \log_2 (|C_Q|)$  information bits. Recently, another generalized method of GSM, called GSM with multiplexing (GSMM) [12], was introduced. GSMM sends S data symbols using a set of precoding matrices  $\mathcal{F}$  in which  $1 \le S \le N_t$  and  $|\mathcal{F}| = 2^{N_t} - 1$ . As a result, GSMM is able to modulate  $\log_2 (2^{N_t} - 1) + S \log_2(|C_Q|)$  information bits. The common limitation of the methods in [7]–[12] is that information bits are separately modulated to the index of active antenna-subsets (or the index of precoding matrices [12]) and to the transmission of QAM symbols. This separated modulation approach generally cannot achieve a higher spectral efficiency than that attained by a joint modulation strategy for a fixed  $N_t$  and  $C_Q$ .

Adaptive-joint-mapping GSM (AJM-GSM) [13] and jointly-mapped SM (JM-SM) [14] both jointly modulate information bits into an active antenna index and transmit symbols using the proposed joint mapping rule. This joint mapping can generate more signal points than the

mapping methods that separately modulate information bits into the indices of antenna subsets and transmit symbols. The limitation of the methods in [11] and [14] is that they do not jointly take into account all possible signaling dimensions, (i.e., spatial, in-phase, quadrature) when constructing signal sets for SM.

Since the minimum Euclidean distance between adjacent symbols in a constellation set affects the detection accuracy, numerous methods for the signal set design of SM have been proposed to maximize the minimum distance [15]–[17]. For instance, the QAM symbols of MA-SM [9] are rotated to a different phase for each active antenna-subset; as a result, the minimum distance increases. Enhanced spatial modulation (ESM) [15], [16] uses a similar approach. The idea of ESM is to construct a primary and a secondary signal constellation set. Then the secondary set is interpolated into the primary set to increase the minimum distance between symbol vectors. Another constellation design method for SM exploits Eisenstein integer [17]; in this method, a transmitter uses a two-dimensional hexagonal lattice to send symbols. This lattice is the densest packing lattice in a two-dimensional complex domain. All these methods [15]–[17] demonstrate that carefully-designed constellation sets is able to achieve higher SNR gains than SM methods that use the conventional QAM constellation set.

In this paper, we consider a MIMO system in which a transmitter is equipped with  $N_t$  transmit antennas and a receiver is equipped with  $N_r$  receive antennas. We assume that the transmitter can send one symbol per in-phase (or quadrature) component of each transmit antenna from the *M*-ary pulse-amplitude-modulation (PAM) signal set  $C = \left\{\frac{-M}{2}, \frac{-M}{2} + 1, \dots, \frac{M}{2} - 1, \frac{M}{2}\right\}$ . The contributions of this paper are summarized as follows:

Our major contribution is to propose a novel multidimensional spatial modulation method called *spatial lattice modulation (SLM)*. Unlike the existing SM techniques in which information bits are separately mapped to a set of antenna indices and modulation symbols [7]–[10], [12], the key idea of the proposed SLM is to modulate information bits into a set of lattice points in ℝ<sup>2Nt</sup> by jointly exploiting spatial, in-phase, and quadrature signal dimensions. An element of each lattice point is chosen from the set C
= {0} ∪ C with |C
= M + 1, where the null element {0} indicates that the input signal for a chosen signal dimension is deactivated for SM. In particular, we present two SLM methods: SLM using a simple cubic lattice and SLM that uses a dense packing lattice that has a large nominal coding gain in a low-dimensional vector space [18], [19]. We show that the proposed SLM methods achieve the spectral efficiency of 2Nt log<sub>2</sub> (M + 1) bits/sec/Hz

when SNR is high enough. This result is interesting because it enables the transmission of  $2N_t (\log_2 (M + 1) - \log_2 (M))$  additional information bits per channel use compared to the conventional spatial multiplexing; this gain is unbounded as  $N_t$  increases. In particular, to attain a nominal coding gain for a given target spectral efficiency, we propose a signal set design algorithm that uses Barnes-Wall lattices for SLM; these are the densest lattices below 16 dimensions (or eight transmit antennas).

- We also analyze the average mutual information and average symbol-vector-error-probability (ASVEP) of the proposed SLM methods under a Rayleigh MIMO channel. Although the mutual information expressions have been characterized for SM and GSM in [21]–[23] for a fixed MIMO channel, the average mutual information expression is unknown for general SM methods. We derive a tight approximation of the average mutual information in a closed-form for the proposed SLM methods. We also derive a closed-form upper bound of ASVEP for SLM to complete our analysis. Simulation results verify the effectiveness of our analysis. One major observation is that SLM using dense lattices provides SNR gains in both the average mutual information and the ASVEP.
- Lastly, we present a low-complexity SLM detection method, which is called *lattice sphere decoding (LSD)*. The key idea of LSD is to exploit the property that a lattice is closed under addition. Using this lattice property, the proposed LSD algorithm reduces the effective search space of SLM by calculating ML metrics only in the closest lattice vectors from an initially estimated lattice vector. We show that the complexity order of the proposed LSD algorithm is  $\mathcal{O}(N_t^3)$ , which is the same detection complexity order with linear-type detection methods such as zero-forcing MIMO detection. Simulation results show that the error performance of LSD closely matches that of ML detection in practical MIMO systems, while significantly diminishing the detection complexity.

## **II. SYSTEM MODEL AND PRELIMINARYIES**

This section presents the system model considered in this paper and provides some useful mathematical definitions that will be used subsequently.

# A. System Model

We consider a MIMO channel in which a transmitter equipped with  $N_t$  transmit antennas sends information symbols to a receiver equipped with  $N_r$  receive antennas. We denote the complex



Fig. 1. System architecture for the proposed SLM in a MIMO system.

baseband transmit vector by  $\bar{\mathbf{x}} \in \mathbb{C}^{N_t \times 1}$ , and the MIMO channel matrix by  $\bar{\mathbf{H}} \in \mathbb{C}^{N_r \times N_t}$ . Then the complex baseband received signal at the receiver is

$$\bar{\mathbf{y}} = \bar{\mathbf{H}}\bar{\mathbf{x}} + \bar{\mathbf{v}},\tag{1}$$

where  $\bar{\mathbf{v}}$  is a complex Gaussian noise vector with zero mean and covariance matrix  $\sigma^2 \mathbf{I}$ . We assume a rich-scattering and frequency-flat channel model, in which all elements of  $\bar{\mathbf{H}}$  are chosen from complex Gaussian random variables with zero mean and unit variance. The transmit vector  $\bar{\mathbf{x}}$  is normalized to satisfy the power constraint  $E_s$ , i.e.,

$$\operatorname{Tr}\left(\mathbb{E}\left[\bar{\mathbf{x}}\bar{\mathbf{x}}^{H}\right]\right) = E_{\mathrm{s}}.$$
(2)

Without loss of generality, the complex-baseband input-output relationship in (1) can be rewritten in real-vector representation as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v},\tag{3}$$

where

$$\mathbf{x} = \begin{bmatrix} \operatorname{Re}\{\bar{\mathbf{x}}\}\\\operatorname{Im}\{\bar{\mathbf{x}}\} \end{bmatrix} \in \mathbb{R}^{2N_{t}}, \mathbf{y} = \begin{bmatrix} \operatorname{Re}\{\bar{\mathbf{y}}\}\\\operatorname{Im}\{\bar{\mathbf{y}}\} \end{bmatrix} \in \mathbb{R}^{2N_{r}}, \mathbf{v} = \begin{bmatrix} \operatorname{Re}\{\bar{\mathbf{v}}\}\\\operatorname{Im}\{\bar{\mathbf{v}}\} \end{bmatrix} \in \mathbb{R}^{2N_{r}}$$

and

$$\mathbf{H} = \begin{bmatrix} \operatorname{Re}\{\bar{\mathbf{H}}\} & -\operatorname{Im}\{\bar{\mathbf{H}}\} \\ \operatorname{Im}\{\bar{\mathbf{H}}\} & \operatorname{Re}\{\bar{\mathbf{H}}\} \end{bmatrix} \in \mathbb{R}^{2N_{\mathrm{r}} \times 2N_{\mathrm{t}}}$$

For notational convenience, we will use this real-value representation of the MIMO system in this paper, except in Section IV. We also assume that *M*-ary pulse-amplitude-modulation (PAM),  $\left\{\frac{-M}{2}, \frac{-M}{2} + 1, \dots, \frac{M}{2} - 1, \frac{M}{2}\right\}$ , is used for the input constellation points per dimension; therefore the input signal set is finite. We also assume that perfect channel state information is known to the receiver, i.e., perfect CSIR, which can be reliably estimated using conventional pilot-transmission.

# B. Preliminaries

We provide some useful definitions which also can be found in [18], [19].

**Definition 1 (Lattice):** Let  $\{g_1, \ldots, g_n\}$  be a set of linearly independent vectors in  $\mathbb{R}^n$ , in which each vector constitutes a basis for the lattice. A real lattice is the countably infinite set defined by integer combinations of basis vectors, i.e.,

$$\mathbf{\Lambda} = \left\{ \mathbf{x} \in \mathbb{R}^n : \mathbf{x} = c_1 \mathbf{g}_1 + c_2 \mathbf{g}_2 + \dots + c_n \mathbf{g}_n \right\},\tag{4}$$

where  $c_i \in \mathbb{Z}$ . Thus, a lattice,  $\Lambda$ , is a discrete additive subgroup of  $\mathbb{R}^n$ , and is closed under addition and reflection. Matrix  $\mathbf{G} = [\mathbf{g}_1, \dots, \mathbf{g}_n]^{\top}$  is referred to as a generating matrix of the lattice.

**Definition 2** (Nearest neighbor quantizer): The nearest-neighbor quantizer associated with  $\Lambda$  is defined as

$$Q(\mathbf{v}) = \mathbf{x}_i \in \mathbf{\Lambda} \text{ if } \mathbf{v} \in \{\mathbf{v} \in \mathbb{R}^n : \|\mathbf{v} - \mathbf{x}_i\|_2 \le \|\mathbf{v} - \mathbf{x}_j\|_2\},$$
(5)

for any other point  $\mathbf{x}_j \in \mathbf{\Lambda}$ . The Voronoi cell associated with  $\mathbf{x}_i \in \mathbf{\Lambda}$  is the set of points in  $\mathbb{R}^n$ closest to  $\mathbf{x}_i$ , i.e.,  $\mathcal{V}_i(\mathbf{\Lambda}) = {\mathbf{v} : Q(\mathbf{v}) = \mathbf{x}_i}.$ 

**Definition 3 (Nested Lattice):** A pair of *n*-dimensional lattices  $(\Lambda_1, \Lambda_2)$  is called nested if  $\Lambda_2 \subset \Lambda_1$ , i.e., corresponding generator matrices  $\mathbf{G}_1$  and  $\mathbf{G}_2$  exist such that  $\mathbf{G}_2 = \mathbf{G}_1 \mathbf{A}$ , where  $\mathbf{A}$  is an  $n \times n$  integer matrix that has a determinant > 1. The volumes of the Voronoi cells of  $\Lambda_1$  and  $\Lambda_2$  satisfy  $\mathcal{V}_i(\Lambda_1) = \det(\mathbf{A})\mathcal{V}_i(\Lambda_2)$ .

**Definition 4 (Normalized Coding Gain):** Let  $d_{\min}^2(\Lambda)$  be the minimum squared distance of the lattice points in  $\Lambda$ . The Hermite parameter of  $\Lambda$  is the normalized density parameter or the normalized coding gain, which is defined as

$$\gamma_c(\mathbf{\Lambda}) = \frac{d_{\min}^2(\mathbf{\Lambda})}{\mathcal{V}_0(\mathbf{\Lambda})^{\frac{2}{n}}}.$$
(6)

Asymptotically, for very high SNR, determining the maximum possible nominal coding gain of an *n*-dimensional lattice code is equivalent to finding the densest lattice in a sphere-packing sense.

## **III. SPATIAL LATTICE MODULATION**

In this section, we present the idea of SLM. Unlike the existing SM techniques in which information bits are separately mapped to a set of antenna indices and modulation symbols, the key idea of SLM is to jointly map K information bits to one of  $2^{K}$  lattice vectors in  $\mathbb{R}^{2N_{t}}$ . This joint mapping strategy using lattices makes it possible to obtain the maximum entropy of input symbol vectors for an given M-ary PAM condition per dimension. Also, by using a dense lattice, we are able to achieve the largest nominal coding gain for a given  $N_{t}$ . Depending on different lattice structures, we propose two SLM methods: SLM using a cubic lattice and SLM using a dense lattice.

# A. SLM using Cubic Lattices

This proposed SLM method uses a joint mapping strategy to map K information bits into a set of information symbol vectors, each in a cubic lattice. Let  $S^{CB} = \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_{2^K}\}$  be a set of transmit symbol vectors where  $\mathbf{s}_k \in \mathbb{R}^{2N_t}$  is the real-representation of  $\bar{\mathbf{s}}_k \in \mathbb{C}^{N_t}$ . We denote the number of activated dimensions in an  $2N_t$ -dimensional space by  $N_a$ , where  $0 \le N_a \le 2N_t$ . Under the premise that  $N_a$  dimensions are active, a set of all possible transmit vectors using generalized spatial modulation method under the constraint of the *M*-ary PAM input signal per dimension is

$$\mathcal{S}^{N_{\mathrm{a}}} = \left\{ \mathbf{s}_{1}^{N_{\mathrm{a}}}, \dots, \mathbf{s}_{L_{N_{\mathrm{a}}}}^{N_{\mathrm{a}}} \right\},\tag{7}$$

where the cardinality of  $\mathcal{S}^{N_{\mathrm{a}}}$  is

$$L_{N_{\rm a}} = \binom{2N_{\rm t}}{N_{\rm a}} M^{N_{\rm a}}.$$
(8)

Because  $0 \le N_a \le 2N_t$ , we construct an entire signal set for the joint mapping by the union of  $S^{N_a}$  as

$$\mathcal{S}^{\rm CB}(N_{\rm t}, M) = \bigcup_{N_{\rm a}=0}^{2N_{\rm t}} \mathcal{S}^{N_{\rm a}}.$$
(9)

 $S^i$  and  $S^j$  are disjoint for all  $i \neq j$ , so the cardinality of  $S^{CB}$  is

$$|\mathcal{S}^{\rm CB}(N_{\rm t},M)| = \sum_{i=0}^{2N_{\rm t}} \binom{2N_{\rm t}}{i} M^{i} = (M+1)^{2N_{\rm t}},\tag{10}$$

where the last equality follows from the binomial expansion.

8	

Index	$\mathbf{x}^{\mathrm{T}}$	$\mathcal{S}^{N_{\mathrm{a}}}$	Index	$\mathbf{x}^{\mathrm{T}}$	$\mathcal{S}^{N_{\mathrm{a}}}$
1	[0, 0, 0, 0]	$\mathcal{S}^0$	22-25	$[0, \pm 1, \pm 1, 0]$	
2-3	$[\pm 1, 0, 0, 0]$		26-29	$[0,\pm 1,0,\pm 1]$	$\mathcal{S}^2$
4-5	$[0,\pm 1,0,0]$	<b>c</b> <sup>1</sup>	30-33	$[0,0,\pm 1,\pm 1]$	
6-7	$[0, 0, \pm 1, 0]$	0	34-41	$[\pm 1,\pm 1,\pm 1,0]$	
8-9	$[0, 0, 0, \pm 1]$		42-49	$[\pm1,\pm1,0,\pm1]$	<b>c</b> <sup>3</sup>
10-13	$[\pm 1, \pm 1, 0, 0]$		50-57	$[\pm 1, 0, \pm 1, \pm 1]$	0
14-17	$[\pm1,0,\pm1,0]$	$\mathcal{S}^2$	58-65	$[0,\pm1,\pm1,\pm1]$	
18-21	$[\pm 1, 0, 0, \pm 1]$		66-81	$[\pm 1, \pm 1, \pm 1, \pm 1]$	$\mathcal{S}^4$

TABLE I Look-up Table for SLM with Cubic lattices ( $N_{\rm t}=2,~M=2$ ).

**Example 1:** Suppose  $N_t = 2$  and M = 2. In this case, we can create 81 lattice vectors by using the joint mapping strategy. The corresponding 81 vectors are listed in Table I<sup>1</sup>. These cubic lattice vectors can be also generated by using the following generating matrix of  $\Lambda_4^{CB}$ , i.e.,

$$\mathbf{G}_{4}^{\mathrm{CB}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$
 (11)

Using this set of vectors, one can send information bits at  $\log_2(81) = 6.3$  bits/sec/Hz if SNR is high enough.

**Proposition 1:** Let  $\mathbf{x} \in \mathbb{R}^{2N_t}$  be an input symbol vector that is chosen uniformly from the proposed signal set  $\mathcal{S}^{CB}(N_t, M)$ . Then the maximum input entropy of the proposed signal set is

$$\mathcal{H}(\mathbf{x}) = 2N_{\rm t}\log_2(M+1). \tag{12}$$

*Proof.* The proof follows from the definition of the entropy and the uniform input distribution, i.e.,  $P(\mathbf{x} = \mathbf{s}_k) = \frac{1}{(M+1)^{2N_t}}$  for  $k \in \{1, 2..., (M+1)^{2N_t}\}$ .

We offer some remarks and examples on our joint mapping strategy to clarify understanding.

**Remark 1:** The proposed joint mapping method has an effective modulation size of M + 1 per dimension. Because we assume that M-ary PAM is used for the input constellation set per

<sup>&</sup>lt;sup>1</sup>In case of high bit-rates transmission, look-up table method is not suitable to implement practical system. Alternatively, we can avoid this look-up table by using combinatorial method [24].

dimension, the modulation size of M is trivial. However, because we use spatial modulation, we can add one hidden constellation point of '0' as an input vector by physically deactivating the input per dimension; this process provides the effective modulation size of M+1. This increment provides a considerable gain in the input entropy. For example, suppose a spatial multiplexing transmission method with uniform M-ary PAM input signaling. The maximum input entropy of this method is  $2N_t \log_2(M)$ . Therefore, one can achieve the gain of

$$2N_{\rm t}\log_2(M+1) - 2N_{\rm t}\log_2(M) = 2N_{\rm t}\log_2\left(1 + \frac{1}{M}\right).$$

This entropy gain implies that one can send  $2N_t \log_2 \left(1 + \frac{1}{M}\right)$  more information bits per transmission using the proposed method when SNR is high enough. The gain increases linearly with  $N_t$  and is therefore unbounded.

**Remark 2:** The proposed joint mapping technique generalizes the existing SM methods in [1], [6], [7], [9], [11], [14] and spatial multiplexing method. For example, supposing that  $N_t = 2$  and M = 2, the signal sets generated by the conventional SM and QSM method in [1], [11] are contained in  $S^2$ , which is a subset of  $S^{CB}(2, 2)$ . Similarly, the signal set of the conventional spatial multiplexing method is  $\{1 + j, 1 - j, -1 + j, -1 - j\}^2$ , which is same with  $S^4$  of  $S^{CB}(2, 2)$ .

**Remark 3:** When the effective modulation size M + 1 is a prime number, the proposed constellation set  $S^{CB}(N_t, M)$  generated by the joint spatial mapping can also be constructed by a nested lattice modulation method. Letting  $\epsilon = \frac{1}{\sqrt{E_s}}$ , we consider two nested  $2N_t$ -dimensional cubic lattices with the generating matrices  $\mathbf{G}_c = \epsilon \mathbf{I}_{2N_t \times 2N_t}$  and  $\mathbf{G}_s = \epsilon (M + 1) \mathbf{I}_{2N_t \times 2N_t}$ respectively:,

$$\Lambda_{c} = \left\{ \mathbf{x} \in \mathbb{R}^{2N_{t}} : \mathbf{x} = \mathbf{c}^{\top} \mathbf{G}_{c} \right\} \text{ and}$$
$$\Lambda_{s} = \left\{ \mathbf{x} \in \mathbb{R}^{2N_{t}} : \mathbf{x} = \mathbf{c}^{\top} \mathbf{G}_{s} \right\},$$
(13)

where  $\mathbf{c} \in \mathbb{Z}^{2N_{\mathrm{t}}}$  and  $\Lambda_c \subset \Lambda_s$ . Using these nested lattices, the same constellation set is generated by

$$\mathcal{S}^{\rm CB}(N_{\rm t}, M) = \Lambda_c \cap \mathcal{V}_0(\Lambda_s), \tag{14}$$

where  $\mathcal{V}_0(\Lambda_s)$  is the Voronoi region associated with  $\mathbf{0} \in \Lambda_s$ , i.e., the volume  $\left[-\frac{\epsilon(M+1)}{2}, \frac{\epsilon(M+1)}{2}\right]^{2N_t}$ . Therefore, the joint spatial mapping method can be interpreted as the nested lattice modulation using the corresponding cubic generating matrices.

### B. SLM with Low-Dimensional Dense Lattices

The cubic lattice  $\Lambda^{\text{CB}}$  used for the previous SLM is a baseline lattice, because by the definition it offers no nominal coding gain, i.e.,  $\gamma_c \left( \Lambda^{\text{CB}} \right) = 1$ . Therefore, a natural extension is to use a dense lattice  $\Lambda$  to construct a set of multi-dimensional lattice vectors that yield a higher nominal coding gain than  $\Lambda^{\text{CB}}$  in a given number of dimensions, i.e.,  $\gamma_c \left( \Lambda \right) > \gamma_c \left( \Lambda^{\text{CB}} \right)$ .

Finding the densest lattice packing in an arbitrary number of dimensions is a difficult mathematical problem. The lattices that have the largest nominal coding gain are well characterized up to 24 dimensions [18], [19], [26]. The Barnes-Wall lattice is a good one due to its simplicity of lattice construction and its tractability to analyze. It also provides high nominal coding gains in a low-dimensional signal space.

The following lemma provides a method to construct the Barnes-Wall lattice in  $\mathbb{R}^{2^{m+1}}$  [20]; this lattice is used in the proposed SLM.

**Lemma 1** (Barnes-Wall Lattice Construction [20]). Let  $M_1$  denote the generating matrix of a balanced Barnes-Wall lattice in  $\mathbb{R}^2$ :

$$\mathbf{M}_1 = \begin{bmatrix} \sqrt{2} & 0\\ 1 & 1 \end{bmatrix}. \tag{15}$$

The generating matrix of  $\Lambda_{2^{m+1}}^{BW}$  lattice in  $\mathbb{R}^{2^{m+1}}$  is obtained from  $\mathbf{M}_{m+1}$ , which is the m+1 times Kronecker products of  $\mathbf{M}_1$ , i.e.,

$$\mathbf{M}_{m+1} = \underbrace{\mathbf{M}_1 \otimes \mathbf{M}_1 \otimes \cdots \otimes \mathbf{M}_1}_{m+1}.$$
(16)

By rescaling the irrational elements in  $M_{m+1}$ , the generating matrix of  $\Lambda_{2^{m+1}}^{BW}$  is obtained as

$$\mathbf{G}_{2^{m+1}}^{\mathrm{BW}}(i,j) = \begin{cases} \mathbf{M}_{m+1}(i,j), & \text{if } \mathbf{M}_{m+1}(i,j) \text{ is rational,} \\ \frac{1}{\sqrt{2}}\mathbf{M}_{m+1}(i,j), & \text{if } \mathbf{M}_{m+1}(i,j) \text{ is irrational.} \end{cases}$$
(17)

The following lemma shows some coding theoretic properties of Barnes-Wall lattices.

**Lemma 2** (Coding Properties of Barnes-Wall Lattices [18], [19]). For all integer  $m \ge 0$ , a  $2^{m+1}$ dimensional Barnes-Wall lattice  $\Lambda_{2^{m+1}}^{BW}$  exists that has the minimum squared Euclidean distance  $d_{\min}^2 \left( \Lambda_{2^{m+1}}^{BW} \right) = 2^m$  and the normalized volume  $\mathcal{V} \left( \Lambda_{2^{m+1}}^{BW} \right)^{\frac{1}{2^m}} = 2^{\frac{m}{2}}$ ; therefore its nominal coding gain is  $\gamma_c \left( \Lambda_{2^{m+1}}^{BW} \right) = 2^{\frac{m}{2}}$ . In addition, the kissing number of  $\Lambda_{2^{m+1}}^{BW}$  is  $K_{\min} \left( \Lambda_{2^{m+1}}^{BW} \right) =$  $\prod_{i=1}^{m+1} (2^i + 2).$ 

# Algorithm 1 Signal Set Design Method for SLM using Dense Lattices.

**Input:** Generating matrix  $\mathbf{G} \in \mathbb{Z}^{2N_{t} \times 2N_{t}}$ , Maximum power constraint  $P_{\text{max}}$ .

**Output:** Signal set for SLM  $\mathcal{S}^{BW}(N_t, P_{max}) \subset \mathbb{Z}^{2N_t}$ .

- 1: Initialization P = 0.
- 2: for  $P \in [0, 1, 2, \cdots, P_{\max}]$  do
- 3: Define  $2N_t$  dimensional integer row vector  $\mathbf{s} = [s_1, s_2, \dots, s_{2N_t}]$ . Then, exhaustively search s satisfying the power condition P.

 $\tilde{\mathcal{S}}^P := \{ \mathbf{s} \mid \|\mathbf{s}\|_2^2 = P, \ \mathbf{s} \in \mathbb{Z}^{2N_t}, \ |s_i| \ge |s_j|, \forall \ i \ge j \}.$ 

- 4: Let  $\hat{S}^P$  denote the symmetric group, or group of permutations, on  $\tilde{S}_P$ .  $\hat{S}^P := \{Sym\{[s_1, s_2, \dots, s_{2N_*}]\}, \forall s \in \tilde{S}^P\}.$
- 5: Check whether  $\hat{s}$  is in lattice  $\Lambda$ .

$$\mathcal{S}^P := \{ \hat{\mathbf{s}} \mid \hat{\mathbf{s}} \in \hat{\mathcal{S}}_P, \; \hat{\mathbf{s}} \mathbf{G}^{-1} \in \mathbb{Z}^{2N_{\mathrm{t}}} \}.$$

6: end for

Merge all  $\mathcal{S}^P$  sets into  $\mathcal{S}^{BW}$ .

7:  $\mathcal{S}^{\mathrm{BW}}(N_{\mathrm{t}}, P_{\mathrm{max}}) = \bigcup_{P=0}^{P_{\mathrm{max}}} \mathcal{S}^{P}.$ 

By exploiting the Barnes-Wall lattice  $\Lambda_{2^{m+1}}^{BW}$  that is constructed by  $\mathbf{G}_{m+1}^{BW}$  in (17), we propose Algorithm 1, which creates a signal set for SLM using dense lattices. The algorithm finds a signal set  $\mathcal{S}^{BW}$  such that each lattice vector is created by generating matrix  $\mathbf{G} \in \mathbb{Z}^{2N_t \times 2N_t}$  and satisfies the maximum power constraint  $P_{\text{max}}$ . The first step is to exhaustively search all possible integer vectors that satisfy the given power P. For a vector satisfying the power constraint, the symmetric group is selected as a candidate. For example, with  $\Lambda_4^{BW}$  lattice at power condition P = 2,  $[\pm 1, \pm 1, 0, 0]$ ,  $[\pm 1, 0, \pm 1, 0]$ ,  $[\pm 1, 0, 0, \pm 1]$ ,  $[0, \pm 1, \pm 1, 0]$ ,  $[0, \pm 1, 0, \pm 1]$  will be the possible candidates for constellation vectors. After collecting the candidates, we check the condition whether they consist of an integer linear combination of the basis  $\{\mathbf{g}_1, \dots, \mathbf{g}_{2N_t}\}$ . Then, vectors that satisfy this condition are included in our SLM signal set. Algorithm 1 iterates this procedure by increasing P until the power condition P reaches the maximum power constraint  $P_{\text{max}}$ .

**Example 2:** Suppose a four-dimensional Barnes-Wall lattice,  $\Lambda_4^{BW}$ , which is equivalent to the

ITH BARNES-WALL LATTICES $(N_{\rm t} = 2, P_{\rm max} = 6)$					
	$\mathcal{S}^P$	Index	$\mathbf{x}^{\mathrm{T}}$	$\mathcal{S}^P$	
	$\mathcal{S}^0$	46-47	$[0, 0, \pm 2, 0]$	$\mathbf{c}^4$	
		48-49	$[0, 0, 0, \pm 2]$	0	
		50-57	$[\pm 2,\pm 1,\pm 1,0]$		
	$c^2$	58-65	$[\pm 1, \pm 2, \pm 1, 0]$		
	0				

 $[\pm 1, \pm 1, \pm 2, 0]$ 

 $[0, \pm 2, \pm 1, \pm 1]$ 

 $[0, \pm 1, \pm 2, \pm 1]$ 

 $[0, \pm 1, \pm 1, \pm 2]$ 

 $\mathcal{S}^6$ 

TABLE II LOOK-UP TABLE FOR SLM WITH BARNES-WALL LATTICES ( $N_{\rm t}=2,~P_{\rm max}=6$ ).

66-73

122-129

130-137

138-145

 $\mathcal{S}^4$ 

 $\mathbf{x}^{\mathrm{T}}$ 

[0, 0, 0, 0]

 $[\pm 1, \pm 1, 0, 0]$ 

 $[\pm 1, 0, \pm 1, 0]$ 

 $[\pm 1, 0, 0, \pm 1]$ 

 $[0, \pm 1, \pm 1, 0]$ 

 $[0, \pm 1, 0, \pm 1]$ 

 $\frac{[0,0,\pm 1,\pm 1]}{[\pm 1,\pm 1,\pm 1,\pm 1]}$ 

 $[\pm 2, 0, 0, 0]$ 

 $[0, \pm 2, 0, 0]$ 

 $D_4$  lattice. The generating matrix of  ${f \Lambda}_4^{
m BW}$  is

Index

1 2-5

6-9

10-13

14-17

18-21

22-25

26-41

42-43 44-45

$$\mathbf{G}_{4}^{\mathrm{BW}} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$
 (18)

The minimum squared Euclidean distance  $d_{\min}^2 \left( \Lambda_4^{BW} \right) = 2$ , the volume  $\mathcal{V} \left( \Lambda_4^{BW} \right) = |\det \left( \mathbf{G}_4^{BW} \right) | = 2$ , and therefore its nominal coding gain is  $\gamma_c \left( \Lambda_4^{BW} \right) = \sqrt{2} \simeq 1.5$  dB. Using this generating matrix and Algorithm 1, we can generate 145 transmit signal vectors in  $\Lambda_4^{BW}$  with  $P_{\max} = 6$ . The corresponding 145 vectors are listed in Table II<sup>2</sup>.

**Remark 4:** The proposed algorithm is valid when the generating matrix has integer elements. Therefore, except  $A_2 \subset \mathbb{R}^2$ , we are able to construct a signal set for the SLM with any lowdimensional dense lattices as  $BW_{2^{m+1}}$  and Leech lattice  $L_{24}$ . As an extreme case, when the channel matrix  $\mathbf{H} \in \mathbb{R}^{2N_r \times 2N_t}$  is chosen from the set of unitary matrices, the maximum nominal coding gains  $\gamma_c(\mathbf{\Lambda})$  achieved by the known lattices are known (Table III) [26]. For a general MIMO channel, however, these gains are not obtainable because the minimum distance of the signal set seen by the receiver diminishes when the channel matrix is non-unitary.

<sup>&</sup>lt;sup>2</sup> The barnes-Wall lattice vectors of Table II can be indexed by using arithmetic operators without look-up table [25]. This enables to reduce encoding and decoding complexity.

#### TABLE III

THE MAXIMUM NOMINAL CODING GAINS FOR SLM WITH LOW-DIMENSIONAL DENSE LATTICES.

Antennas	Dimensions	Lattice $\Lambda$	Coding gain $\gamma_c(\Lambda)$
1	2	$A_2$	0.62 dB
2	4	$\mathbf{\Lambda}_4^{\mathrm{BW}} (= D_4)$	1.51 dB
4	8	$\mathbf{\Lambda}_8^{\mathrm{BW}}(=E_8)$	3.01 dB
8	16	$\boldsymbol{\Lambda}^{\mathrm{BW}}_{16}(=\Lambda_{16})$	4.52 dB
12	24	$L_{24}$	6.02 dB
16	32	$\Lambda_{32}^{\rm BW} (= \Lambda_{32})$	6.28 dB

# IV. PERFORMANCE ANALYSIS

This section analyzes the average mutual information and average symbol-vector-error-probability (ASVEP) for the MIMO systems when input symbol vectors are uniformly drawn from an arbitrary finite input constellation S. For ease of exposition, we use the complex baseband MIMO model in (1).

# A. Maximum Likelihood (ML) Detection

We assume that the receiver and transmitter share the joint mapping rule between information bits and the SLM signal set. The active dimension and constellation points of a transmit symbol vector are jointly decoded using ML principle

$$[l] = \arg\max_{m} f_{\bar{\mathbf{y}}}(\bar{\mathbf{y}}|\bar{\mathbf{x}}_{m}) = \arg\min_{m} \|\bar{\mathbf{y}} - \bar{\mathbf{H}}\bar{\mathbf{x}}_{m}\|^{2},$$
(19)

for  $\bar{\mathbf{x}}_m \in \bar{\mathcal{S}} = {\{\bar{\mathbf{s}}_1, \bar{\mathbf{s}}_2, \dots, \bar{\mathbf{s}}_{|\bar{\mathcal{S}}|}\}}$ . The conditional probability density function (PDF) of  $\bar{\mathbf{y}}$  given that  $\bar{\mathbf{x}}_m$  is

$$f(\mathbf{\bar{y}}|\mathbf{\bar{x}}_m) = \frac{1}{(\pi\sigma^2)^{N_{\rm r}}} \exp\left(-\frac{\|\mathbf{\bar{y}} - \mathbf{\bar{H}}\mathbf{\bar{x}}_m\|^2}{\sigma^2}\right).$$
 (20)

We use the ML criterion for the performance analysis and for the comparison with other SM schemes.

# B. Average Mutual Information Analysis for SLM

When input signal vectors are uniformly drawn from a finite constellation set (e.g., M-ary PAM, M-ary QAM), an exact expression of the mutual information with perfect CSIR was derived in [23] when the channel is fixed. For the completeness of our analysis, we first present

the exact expression of the mutual information. Then, to improve intuition, we also provide a tight approximation expression for the average mutual information.

1) Exact Expression: For a MIMO channel, the mutual information between the discrete input vector  $\bar{\mathbf{x}}$  and the continuous channel output vector  $\bar{\mathbf{y}}$  is

$$\mathcal{I}(\bar{\mathbf{x}}; \bar{\mathbf{y}} | \bar{\mathbf{H}}) = \mathcal{H}(\bar{\mathbf{y}} | \bar{\mathbf{H}}) - \mathcal{H}(\bar{\mathbf{y}} | \bar{\mathbf{x}}, \bar{\mathbf{H}})$$
$$= \mathcal{H}(\bar{\mathbf{y}} | \bar{\mathbf{H}}) - \mathcal{H}(\bar{\mathbf{v}})$$
$$= \mathcal{H}(\bar{\mathbf{y}} | \bar{\mathbf{H}}) - N_{\mathrm{r}} \log_2(\pi e \sigma^2).$$
(21)

Because the received signal vector  $\bar{\mathbf{y}}$  follows a Gaussian distribution for given  $\bar{\mathbf{H}}$  and  $\bar{\mathbf{x}}$ , the PDF of  $\bar{\mathbf{y}}$  is

$$f(\bar{\mathbf{y}}|\bar{\mathbf{H}}) = \frac{1}{|\mathcal{S}|} \sum_{\bar{\mathbf{x}}_i \in \mathcal{S}} \frac{1}{(\pi \sigma^2)^{N_{\mathrm{r}}}} \exp\left(-\frac{\|\bar{\mathbf{y}} - \bar{\mathbf{H}}\bar{\mathbf{x}}_i\|^2}{\sigma^2}\right).$$
(22)

Using  $f(\bar{\mathbf{y}}|\bar{\mathbf{x}}_i, \bar{\mathbf{H}})$  and  $f(\bar{\mathbf{y}}|\bar{\mathbf{H}})$ , the mutual information for given  $\bar{\mathbf{H}}$  is

$$\mathcal{I}(\bar{\mathbf{x}}; \bar{\mathbf{y}} | \bar{\mathbf{H}}) = \sum_{\bar{\mathbf{x}}_i \in \mathcal{S}} \frac{1}{|\mathcal{S}|} \int_{\bar{\mathbf{y}}} f(\bar{\mathbf{y}} | \bar{\mathbf{x}} = \bar{\mathbf{x}}_i, \bar{\mathbf{H}}) \log_2 \frac{f(\bar{\mathbf{y}} | \bar{\mathbf{x}} = \bar{\mathbf{x}}_i, \bar{\mathbf{H}})}{f(\bar{\mathbf{y}} | \bar{\mathbf{H}})} d\bar{\mathbf{y}}.$$
(23)

Using equations (21)-(23), the mutual information for SLM is obtained as

$$\mathcal{I}(\bar{\mathbf{x}}; \bar{\mathbf{y}} | \bar{\mathbf{H}}) = \log |\mathcal{S}| - \frac{1}{|\mathcal{S}|} \sum_{\bar{\mathbf{x}}_i \in \mathcal{S}} \mathbb{E}_{\bar{\mathbf{v}}} \left[ \log_2 \sum_{\bar{\mathbf{x}}_j \in \mathcal{S}} \exp\left(-\frac{\|\bar{\mathbf{H}}(\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j) + \bar{\mathbf{v}}\|^2 - \|\bar{\mathbf{v}}\|^2}{\sigma^2}\right) \right].$$
(24)

2) Approximation of Average Mutual Information: The derived mutual information expression in (24) is not tractable because it involves the multi-dimensional integral with respect to  $\bar{v}$  and it does not capture the randomness in channels. We derive a closed form tight approximation of the average mutual information by exploiting the lower bound of (24). The following proposition closely approximates the average mutual information.

**Proposition 3:** The average mutual information,  $\mathbb{E}_{\mathbf{\bar{H}}} \left[ \mathcal{I} \left( \mathbf{\bar{x}}; \mathbf{\bar{y}} | \mathbf{\bar{H}} \right) \right]$  for SLM, can be approximated as

$$\mathbb{E}_{\bar{\mathbf{H}}} \left[ \mathcal{I} \left( \bar{\mathbf{x}}; \bar{\mathbf{y}} | \bar{\mathbf{H}} \right) \right] \\\approx 2 \log_2 |\mathcal{S}| - \log_2 \sum_{\mathbf{x}_i \in \mathcal{S}} \sum_{\mathbf{x}_j \in \mathcal{S}} \left( 1 + \frac{\| \bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j \|^2}{2\sigma^2} \right)^{-N_r}.$$
(25)

*Proof.* Applying Jensen's inequality yields a lower bound of  $\mathcal{H}(\bar{\mathbf{y}}|\bar{\mathbf{H}})$  as in [23],

$$\mathcal{H}(\bar{\mathbf{y}}|\bar{\mathbf{H}}) \geq -\log_2 \mathbb{E}_{\bar{\mathbf{y}}} \left[ f(\bar{\mathbf{y}}|\bar{\mathbf{H}}) \right] = -\log_2 \int_{\bar{\mathbf{y}}} \left( \frac{1}{|\mathcal{S}|} \sum_{\bar{\mathbf{x}}_i \in \mathcal{S}} \frac{1}{(\pi \sigma^2)^{N_{\mathrm{r}}}} \exp\left(-\frac{\|\bar{\mathbf{y}} - \bar{\mathbf{H}}\bar{\mathbf{x}}_i\|^2}{\sigma^2}\right) \right)^2 d\bar{\mathbf{y}}.$$
(26)

By changing the order of summation and integration in (26),  $\mathcal{H}(\bar{\mathbf{y}}|\bar{\mathbf{H}})$  is lower-bounded by

$$\mathcal{H}(\bar{\mathbf{y}}|\bar{\mathbf{H}}) \geq 2\log_2 |\mathcal{S}| (\pi\sigma^2)^{N_{\mathbf{r}}} - \log_2 \sum_{\bar{\mathbf{x}}_i \in \mathcal{S}} \sum_{\bar{\mathbf{x}}_j \in \mathcal{S}} \int_{\bar{\mathbf{y}}} e^{-\frac{\|\bar{\mathbf{y}}-\bar{\mathbf{H}}\bar{\mathbf{x}}_i\|^2}{\sigma^2}} e^{-\frac{\|\bar{\mathbf{y}}-\bar{\mathbf{H}}\bar{\mathbf{x}}_j\|^2}{\sigma^2}} d\bar{\mathbf{y}}.$$
(27)

The integration in (27) can be calculated as

$$\int_{\bar{\mathbf{y}}} e^{-\frac{\|\bar{\mathbf{y}}-\bar{\mathbf{H}}\bar{\mathbf{x}}_{j}\|^{2}}{\sigma^{2}}} e^{-\frac{\|\bar{\mathbf{y}}-\bar{\mathbf{H}}\bar{\mathbf{x}}_{j}\|^{2}}{\sigma^{2}}} d\bar{\mathbf{y}} = e^{-\frac{\|\bar{\mathbf{H}}\bar{\mathbf{x}}_{i}\|^{2}+\|\bar{\mathbf{H}}\bar{\mathbf{x}}_{j}\|^{2}}{\sigma^{2}}} \times \underbrace{\int_{\bar{\mathbf{y}}} \exp\left(-\frac{2\|\bar{\mathbf{y}}\|^{2}-2\operatorname{Re}\left\{\bar{\mathbf{y}}^{H}(\bar{\mathbf{H}}\bar{\mathbf{x}}_{i}+\bar{\mathbf{H}}\bar{\mathbf{x}}_{j})\right\}}{\sigma^{2}}\right) d\bar{\mathbf{y}}}_{\Theta}.$$
(28)

Because the elements of  $\bar{\mathbf{y}} = [y_1, y_2, \dots, y_{N_r}]^{\top}$  are mutually independent, the multi-dimensional integrations in (28) are computed using element-wise integrations as

$$\Theta = \prod_{k=1}^{N_{\mathbf{r}}} \int_{y_k} \exp\left(-\frac{2\|y_k\|^2 - 2\operatorname{Re}\left\{y_k^*\left\{\bar{\mathbf{H}}(\bar{\mathbf{x}}_i + \bar{\mathbf{x}}_j)\right\}_k\right\}}{\sigma^2}\right) dy_k.$$
(29)

The integration in (29) is conducted separately for real and imaginary parts to yield

$$\int_{y_{k,\mathrm{Re}}} \exp\left(-\frac{2y_{k,\mathrm{Re}}^2 - 2y_{k,\mathrm{Re}}\left\{\bar{\mathbf{H}}(\bar{\mathbf{x}}_i + \bar{\mathbf{x}}_j)\right\}_{k,\mathrm{Re}}}{\sigma^2}\right) dy_{k,\mathrm{Re}}$$
$$\times \int_{y_{k,\mathrm{Im}}} \exp\left(-\frac{2y_{k,\mathrm{Im}}^2 - 2y_{k,\mathrm{Im}}\left\{\bar{\mathbf{H}}(\bar{\mathbf{x}}_i + \bar{\mathbf{x}}_j)\right\}_{k,\mathrm{Im}}}{\sigma^2}\right) dy_{k,\mathrm{Im}}.$$
(30)

The integrals in (30) are computed using the following identity

$$\int_{-\infty}^{\infty} \exp\left(-\frac{2y^2 - 2xy}{\sigma^2}\right) dy = \sqrt{\frac{\pi\sigma^2}{2}} \exp\left(\frac{x^2}{2\sigma^2}\right).$$
(31)

Thus,  $\Theta=\Theta_{\rm Re}\times\Theta_{\rm Im},$  where  $\Theta_{\rm Re}$  and  $\Theta_{\rm Im}$  are given by

$$\Theta_{\text{Re}} = \prod_{k=1}^{N_{\text{r}}} \sqrt{\frac{\pi \sigma^2}{2}} \exp\left(\frac{\left\{\bar{\mathbf{H}}(\bar{\mathbf{x}}_i + \bar{\mathbf{x}}_j)\right\}_{k,\text{Re}}^2}{\sigma^2}\right),$$
  
$$\Theta_{\text{Im}} = \prod_{k=1}^{N_{\text{r}}} \sqrt{\frac{\pi \sigma^2}{2}} \exp\left(\frac{\left\{\bar{\mathbf{H}}(\bar{\mathbf{x}}_i + \bar{\mathbf{x}}_j)\right\}_{k,\text{Im}}^2}{\sigma^2}\right).$$
(32)

Using the results in (32), we obtain the simple

$$\Theta = \prod_{k=1}^{N_{\rm r}} \left(\frac{\pi\sigma^2}{2}\right) \exp\left(\frac{\left\|\left\{\bar{\mathbf{H}}(\bar{\mathbf{x}}_i + \bar{\mathbf{x}}_j)\right\}_k\right\|^2}{2\sigma^2}\right)$$
$$= \left(\frac{\pi\sigma^2}{2}\right)^{N_{\rm r}} \exp\left(\frac{\left\|\bar{\mathbf{H}}(\bar{\mathbf{x}}_i + \bar{\mathbf{x}}_j)\right\|^2}{2\sigma^2}\right).$$
(33)

Using the above equations (27)-(33), the lower bound of  $\mathcal{H}(\bar{\mathbf{y}}|\bar{\mathbf{H}})$  is

$$\mathcal{H}(\bar{\mathbf{y}}|\bar{\mathbf{H}}) \ge 2\log_2 |\mathcal{S}| + \log_2(\pi\sigma^2)^{N_{\mathrm{r}}} + N_{\mathrm{r}}\log_2 2$$
$$-\log_2 \sum_{\bar{\mathbf{x}}_i \in \mathcal{S}} \sum_{\bar{\mathbf{x}}_j \in \mathcal{S}} \exp\left(-\frac{\|\bar{\mathbf{H}}(\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j)\|^2}{2\sigma^2}\right).$$
(34)

Plugging (34) into (21), a lower bound of  $\mathcal{I}(\bar{\mathbf{x}}; \bar{\mathbf{y}} | \bar{\mathbf{H}})$  is obtained as

$$\mathcal{I}^{\text{Low}}(\bar{\mathbf{x}}; \bar{\mathbf{y}} | \bar{\mathbf{H}}) = 2 \log_2 |\mathcal{S}| + N_r \log_2 \frac{2}{e}$$
$$- \log_2 \sum_{\bar{\mathbf{x}}_i \in \mathcal{S}} \sum_{\bar{\mathbf{x}}_j \in \mathcal{S}} \exp\left(-\frac{\|\bar{\mathbf{H}}(\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j)\|^2}{2\sigma^2}\right)$$
$$= 2 \log_2 |\mathcal{S}| + N_r \log_2 \frac{2}{e}$$
$$- \log_2 \left\{ |\mathcal{S}| + \sum_{\bar{\mathbf{x}}_i \in \mathcal{S}} \sum_{\bar{\mathbf{x}}_j \in \mathcal{S} / \{\bar{\mathbf{x}}_i\}} \exp\left(-\frac{\|\bar{\mathbf{H}}(\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j)\|^2}{2\sigma^2}\right)\right\}.$$
(35)

From (35), by taking the limits to the two extreme SNR regimes, we obtain the limit values:

$$\lim_{\text{SNR}\to 0} \mathcal{I}^{\text{Low}}(\bar{\mathbf{x}}; \bar{\mathbf{y}} | \bar{\mathbf{H}}) = N_{\text{r}} \log_2 \frac{2}{e},$$
$$\lim_{\text{SNR}\to\infty} \mathcal{I}^{\text{Low}}(\bar{\mathbf{x}}; \bar{\mathbf{y}} | \bar{\mathbf{H}}) = N_{\text{r}} \log_2 \frac{2}{e} + \log_2 |\mathcal{S}|.$$
(36)

Intuitively, the mutual information should approach zero when SNR is very low, but should approach  $\log_2 |S|$  when SNR is very high. Using these facts and the two extreme values in (36), we define the offset for the approximation as,

$$\Delta_{\mathcal{I}(\bar{\mathbf{x}};\bar{\mathbf{y}}|\bar{\mathbf{H}})} = \mathcal{I}^{\text{Low}}(\bar{\mathbf{x}};\bar{\mathbf{y}}|\bar{\mathbf{H}}) - \mathcal{I}(\bar{\mathbf{x}};\bar{\mathbf{y}}|\bar{\mathbf{H}}) = N_{\text{r}}\log_2\frac{e}{2}.$$
(37)

Subtracting this offset yields a tight approximation of  $\mathcal{I}(\bar{\mathbf{x}}; \bar{\mathbf{y}} | \bar{\mathbf{H}})$ :

$$\hat{\mathcal{I}}(\bar{\mathbf{x}}; \bar{\mathbf{y}} | \bar{\mathbf{H}}) = 2 \log_2 |\mathcal{S}| - \log_2 \sum_{\bar{\mathbf{x}}_i \in \mathcal{S}} \sum_{\bar{\mathbf{x}}_j \in \mathcal{S}} \exp\left(-\frac{\|\bar{\mathbf{H}}(\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j)\|^2}{2\sigma^2}\right)$$
$$= -\log_2 \frac{1}{|\mathcal{S}|^2} \sum_{\bar{\mathbf{x}}_i \in \mathcal{S}} \sum_{\bar{\mathbf{x}}_j \in \mathcal{S}} \exp\left(-\frac{\|\bar{\mathbf{H}}(\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j)\|^2}{2\sigma^2}\right).$$
(38)

In (38), the elements of  $\overline{\mathbf{H}}$  follow the complex Gaussian distribution. Therefore, for given symbol vectors  $\overline{\mathbf{x}}_i, \overline{\mathbf{x}}_j, \kappa = \frac{\|\overline{\mathbf{H}}(\overline{\mathbf{x}}_i - \overline{\mathbf{x}}_j)\|^2}{2\sigma^2}$  follows the Gamma distribution with PDF  $f_{\kappa}(w) \sim \Gamma\left(N_{\mathrm{r}}, \frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right)$ . The exception of the mutual information is approximated by applying Jensen's inequality to  $\hat{\mathcal{I}}$  as

$$\mathbb{E}_{\bar{\mathbf{H}}} \left[ \mathcal{I} \left( \bar{\mathbf{x}}; \bar{\mathbf{y}} | \bar{\mathbf{H}} \right) \right] \approx \mathbb{E}_{\bar{\mathbf{H}}} \left[ \hat{\mathcal{I}} \left( \bar{\mathbf{x}}; \bar{\mathbf{y}} | \bar{\mathbf{H}} \right) \right]$$
  
$$\geq -\log_2 \frac{1}{|\mathcal{S}|^2} \sum_{\bar{\mathbf{x}}_i \in \mathcal{S}} \sum_{\bar{\mathbf{x}}_j \in \mathcal{S}} \mathbb{E}_{\bar{\mathbf{H}}} \left[ \exp \left( -\frac{\| \bar{\mathbf{H}} (\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j) \|^2}{2\sigma^2} \right) \right].$$
(39)

The lower bound is due to the concavity of the log function. In addition, the expectation in (39) is simply calculated using the moment generating function (MGF) of the Gamma random variable  $\kappa$ , i.e.,  $\mathbb{E}\left[e^{-\kappa}\right] = \left(1 + \frac{\|\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j\|^2}{2\sigma^2}\right)^{-N_r}$ . As a result, the approximate expression of the average mutual information is

$$\mathbb{E}_{\bar{\mathbf{H}}} \left[ \mathcal{I} \left( \bar{\mathbf{x}}; \bar{\mathbf{y}} | \bar{\mathbf{H}} \right) \right] \approx -\log_2 \frac{1}{|\mathcal{S}|^2} \sum_{\bar{\mathbf{x}}_i \in \mathcal{S}} \sum_{\bar{\mathbf{x}}_j \in \mathcal{S}} \left( 1 + \frac{\|\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j\|^2}{2\sigma^2} \right)^{-N_{\mathrm{r}}}$$
$$= 2\log_2 |\mathcal{S}| - \log_2 \sum_{\bar{\mathbf{x}}_i \in \mathcal{S}} \sum_{\bar{\mathbf{x}}_j \in \mathcal{S}} \left( 1 + \frac{\|\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j\|^2}{2\sigma^2} \right)^{-N_{\mathrm{r}}}, \tag{40}$$

which completes the proof.

Because the proposed SLM uses a set of lattice points with symmetric minimum distance, i.e.,  $\|\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j\|^2 \ge d_{\min}^2(\mathbf{\Lambda})$ ,  $(\bar{\mathbf{x}}_i \neq \bar{\mathbf{x}}_j)$ , we can further simplify the average mutual information in the following corollary:

**Corollary 1:** A lower bound of  $\mathbb{E}_{\bar{\mathbf{H}}}\left[\hat{\mathcal{I}}\left(\bar{\mathbf{x}}; \bar{\mathbf{y}} | \bar{\mathbf{H}}\right)\right]$  for SLM using lattice  $\Lambda$  is

$$\mathbb{E}_{\bar{\mathbf{H}}} \left[ \hat{\mathcal{I}} \left( \bar{\mathbf{x}}; \bar{\mathbf{y}} | \bar{\mathbf{H}} \right) \right]$$
  

$$\geq \log_2 |\mathcal{S}| - \log_2 \left( 1 + (|\mathcal{S}| - 1) \left( 1 + \frac{d_{\min}^2(\mathbf{\Lambda})}{2\sigma^2} \right)^{-N_{\mathrm{r}}} \right). \tag{41}$$

The result in Corollary 1 clearly shows that a higher average mutual information is achievable when dense lattices (or equivalently a large  $d_{\min}^2(\Lambda)$ ) are used and the number of receive antennas increases.

Fig. 2 shows that the proposed approximation of the average mutual information derived in (40) agrees well with simulations at all SNR region for MIMO systems with antenna configurations  $(N_{\rm t}, N_{\rm r}, M) = (2, 4, 2), (2, 4, 4),$  and (4, 4, 2), and various modulation sizes.



Fig. 2. Simulated (lines) and analytic (symbols) average mutual information results when using the proposed SLM with cubic lattices in  $N_t \times 4$  ( $N_t = 2, 4$ ) MIMO systems.

# C. Average Symbol-Vector-Error Probability

This subsection calculates an upper bound of ASVEP for MIMO systems when the proposed SLM is used.

Proposition 4: An upper bound of ASVEP is

$$P_{\text{ASVEP}} \leq \sum_{\bar{\mathbf{x}}_i \in \mathcal{S}} \sum_{\bar{\mathbf{x}}_j \neq \bar{\mathbf{x}}_i} \frac{\mu_{ij}^{N_{\text{r}}} \sum_{k=0}^{N_{\text{r}}-1} {N_{\text{r}}-1+k \choose k} (1-\mu_{ij})^k}{|\mathcal{S}|}.$$
(42)

*Proof.* Let  $P(\bar{\mathbf{x}}_i \to \bar{\mathbf{x}}_j)$  be the pairwise symbol vector error probability (PSVEP) when deciding  $\bar{\mathbf{x}}_j$  given that  $\bar{\mathbf{x}}_i$  is transmitted. Using an union bounding technique [27], ASVEP is upper bounded by the sum of all possible PSVEPs as

$$P_{\text{ASVEP}} \leq \mathbb{E}_{\bar{\mathbf{x}}} \left[ \sum_{j} P(\bar{\mathbf{x}}_{i} \to \bar{\mathbf{x}}_{j}) \right]$$
$$= \frac{1}{|\mathcal{S}|} \sum_{\bar{\mathbf{x}}_{i} \in \mathcal{S}} \sum_{\bar{\mathbf{x}}_{j} \neq \bar{\mathbf{x}}_{i}} P(\bar{\mathbf{x}}_{i} \to \bar{\mathbf{x}}_{j}).$$
(43)

By using the ML detection principle in (19), PSVEP conditioned on  $\overline{H}$  is given by

$$P(\bar{\mathbf{x}}_{i} \to \bar{\mathbf{x}}_{j} | \bar{\mathbf{H}}) = P\left( \| \bar{\mathbf{v}} + \bar{\mathbf{H}}(\bar{\mathbf{x}}_{i} - \bar{\mathbf{x}}_{j}) \|^{2} < \| \bar{\mathbf{v}} \|^{2} \right| \bar{\mathbf{H}} \right)$$
  
$$= P\left( \operatorname{Re}\left\{ \bar{\mathbf{v}}^{H} \bar{\mathbf{H}}(\bar{\mathbf{x}}_{i} - \bar{\mathbf{x}}_{j}) \right\} < -\frac{1}{2} \| \bar{\mathbf{H}}(\bar{\mathbf{x}}_{i} - \bar{\mathbf{x}}_{j}) \|^{2} \right| \bar{\mathbf{H}} \right)$$
  
$$= Q\left(\sqrt{\kappa}\right), \tag{44}$$

where  $Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$  denotes the tail probability of a standard Gaussian distribution. Thus, the PSVEP is obtained by computing the integral, i.e.,

$$P(\bar{\mathbf{x}}_{i} \to \bar{\mathbf{x}}_{j}) = \mathbb{E}_{\bar{\mathbf{H}}} \left[ P(\bar{\mathbf{x}}_{i} \to \bar{\mathbf{x}}_{j} \middle| \bar{\mathbf{H}}) \right]$$
$$= \int_{w=0}^{\infty} Q\left(\sqrt{w}\right) f_{\kappa}(w) dw$$
$$= \mu_{ij}^{N_{\mathrm{r}}} \sum_{k=0}^{N_{\mathrm{r}}-1} \binom{N_{\mathrm{r}}-1+k}{k} (1-\mu_{ij})^{k}, \tag{45}$$

where  $\mu_{ij} = \frac{1}{2} \left( 1 - \sqrt{\frac{\|\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j\|^2}{4\sigma^2 + \|\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j\|^2}} \right)$  and the third equality follows from the closed form expression given in [28]. Substituting (45) into (43) yields

$$P_{\text{ASVEP}} \leq \sum_{\bar{\mathbf{x}}_i \in \mathcal{S}} \sum_{\bar{\mathbf{x}}_j \neq \bar{\mathbf{x}}_i} \frac{\mu_{ij}^{N_{\text{r}}} \sum_{k=0}^{N_{\text{r}}-1} \binom{N_{\text{r}}-1+k}{k} (1-\mu_{ij})^k}{|\mathcal{S}|},\tag{46}$$

which completes the proof.

We can further simplify ASVEP in Proposition 4 by using the well known upper bound of the Gaussian Q-function. Using following inequality  $Q(x) \leq \frac{1}{2} \exp\left(-\frac{x^2}{2}\right)$ , the SVPEP in (44) is upper bounded by

$$P(\bar{\mathbf{x}}_i \to \bar{\mathbf{x}}_j | \bar{\mathbf{H}}) = Q\left(\sqrt{\kappa}\right) \le \frac{1}{2} \exp\left(-\frac{\kappa}{2}\right).$$
(47)

By marginalizing with respect to  $\kappa$ 

$$P(\bar{\mathbf{x}}_i \to \bar{\mathbf{x}}_j) \le \frac{1}{2} \mathbb{E}_{\kappa} \left[ \exp\left(-\frac{\kappa}{2}\right) \right].$$
 (48)

Using the MGF of the Gamma distribution yields

$$P(\bar{\mathbf{x}}_i \to \bar{\mathbf{x}}_j) \le \frac{1}{2} \left( 1 + \frac{\|\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j\|^2}{4\sigma^2} \right)^{-N_{\mathrm{r}}}.$$
(49)

Consequently, from (43), we obtain the upper bound as

$$P_{\text{ASVEP}} \leq \frac{1}{2|\mathcal{S}|} \sum_{\bar{\mathbf{x}}_i \in \mathcal{S}} \sum_{\bar{\mathbf{x}}_j \neq \bar{\mathbf{x}}_i} \left( 1 + \frac{\|\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j\|^2}{4\sigma^2} \right)^{-N_{\text{r}}},\tag{50}$$

which is a similar form of (40).

From the fact that  $\|\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j\|^2 \ge d_{\min}^2(\mathbf{\Lambda})(\bar{\mathbf{x}}_i \neq \bar{\mathbf{x}}_j)$ , we further simplify (50) in the following corollary.

**Corollary 2:** A lower bound of ASVEP for SLM using lattice  $\Lambda$  is

$$P_{\text{ASVEP}} \leq \frac{1}{2|\mathcal{S}|} \sum_{\bar{\mathbf{x}}_i \in \mathcal{S}} \sum_{\bar{\mathbf{x}}_j \neq \bar{\mathbf{x}}_i} \left( 1 + \frac{d_{\min}^2(\mathbf{\Lambda})}{4\sigma^2} \right)^{-N_{\text{r}}} = \frac{(|\mathcal{S}| - 1)}{2} \left( 1 + \frac{d_{\min}^2(\mathbf{\Lambda})}{4\sigma^2} \right)^{-N_{\text{r}}}.$$
(51)

Similar to Corollary 1, the result in Corollary 2 shows that ASVEP decreases as the minimum distance of the lattice  $d_{\min}^2(\mathbf{\Lambda})$  increases or the number of receive antennas  $N_r$  increases.

### V. LATTICE SPHERE DECODING

In this section, we present a low complexity detection method for SLM, referred to as *lattice sphere decoding (LSD)*. One drawback of the SLM introduced in Section III is that the ML detection complexity increases exponentially with the number of transmit antennas or the modulation size, i.e.,  $\mathcal{O}((M+1)^{2N_t})$ . Therefore, for massive MIMO systems with a large number of transmit antennas, SLM is not appropriate due to the forbidding detection complexity. The proposed LSD overcomes this drawback by reducing the effective search space to the closest lattice vectors from an initially estimated lattice vector.

## A. Lattice Sphere Decoding

This subsection explains the key idea and the algorithm of lattice sphere decoding (LSD). The core idea of LSD is to search only the closest lattice vectors from an initially estimated lattice vector.

1) Initial estimate: The first step of LSD is to find an initial estimate  $\hat{x}$  by using linear detection methods such as a minimum mean-square error (MMSE) detection method, namely,

$$\hat{\mathbf{x}} = \left(\mathbf{H}^{\top}\mathbf{H} + \frac{\sigma^2}{E_s/N_t}\mathbf{I}\right)^{-1}\mathbf{H}^{\top}\mathbf{y}.$$
(52)

2) Vector quantization: The estimated  $\hat{\mathbf{x}}$  is quantized to the closest lattice vector by using the vector quantization function<sup>3</sup>  $Q : \mathbb{R}^{2N_t} \to \Lambda^{BW}$  in [29], i.e.,

$$\bar{\mathbf{x}} = Q(\hat{\mathbf{x}}). \tag{53}$$

<sup>&</sup>lt;sup>3</sup>For more details on the lattice vector quantization in [29], see Appendix A.

Since the transmitted signal is a lattice vector that satisfies the maximum power constraint  $P_{\text{max}}$ by the construction of SLM, we need to rescale the initial estimate  $\hat{\mathbf{x}}$  if  $||Q(\hat{\mathbf{x}})||^2 > P_{\text{max}}$ , i.e.,

$$\bar{\mathbf{x}} = Q\left(\frac{\hat{\mathbf{x}}}{\|\hat{\mathbf{x}}\|}\sqrt{P_{\max}}\right).$$
(54)

3) Reduced lattice set construction inside a sphere: We construct a subset  $\hat{\mathcal{X}}^{BW} \subset \Lambda^{BW}$  whose elements lie in the sphere centered at the quantization output vector  $\bar{\mathbf{x}}$  with radius of d. In particular, we set the radius to the smallest Euclidean norm, i.e.,  $d = d_{\min}^2(\Lambda^{BW})$ . This sphere radius guarantees that the cardinality of  $\hat{\mathcal{X}}$  is equal to the kissing number of a lattice plus one, i.e,  $|\hat{\mathcal{X}}^{BW}| = K_{\min}(\Lambda^{BW}) + 1$ .

Using the fact that a lattice is closed under addition, we find the elements of the subset  $\hat{\mathcal{X}}^{BW}$ in a systematic manner. Let define a set  $\mathcal{D}_{\min}$  that has elements with the smallest Euclidean norm  $d_{\min}^2(\Lambda)$ , i.e., the set with the closest lattice vectors from origin. By the lattice property, the cardinality of  $\mathcal{D}_{\min}$  is equal to the kissing number of a lattice, i.e,  $|\mathcal{D}_{\min}| = K_{\min}(\Lambda)$ . Using  $\mathcal{D}_{\min}$ , it is possible to find the elements of  $\hat{\mathcal{X}}^{BW}$  by adding the initial estimate lattice point  $\bar{\mathbf{x}}$  to the elements in  $\mathcal{D}_{\min}$ , namely,

$$\hat{\mathcal{X}}^{BW} = \{ \hat{\mathbf{x}} | \hat{\mathbf{x}} = \bar{\mathbf{x}} + \mathbf{d}, \mathbf{d} \in \mathcal{D}_{\min} \cup \{ \mathbf{0} \} \}.$$
(55)

Under the maximum power constraint  $P_{\text{max}}$ , the set  $\hat{\mathcal{X}}^{\text{BW}}$  is further reduced to a set  $\mathcal{X}^{\text{BW}}$  by only merging the vectors which satisfy this power constraint.

$$\mathcal{X}^{\mathrm{BW}} = \{ \mathbf{x} | \mathbf{x} \in \hat{\mathcal{X}}^{\mathrm{BW}}, ||\mathbf{x}||^2 \le P_{\max} \}.$$
(56)

4) Decoding over the reduced set: Using the reduced lattice set in the sphere,  $\mathcal{X}^{BW}$ , the receiver performs ML detection over the reduced set, i.e.,

$$\hat{\mathbf{s}} = \arg\min_{\mathbf{x}\in\mathcal{X}^{BW}} ||\mathbf{y} - \mathbf{H}\mathbf{x}||^2.$$
(57)

As seen in Fig. 3, the blue x-mark represents the initial estimate  $\hat{\mathbf{x}}$ . This blue x-mark is quantized to the lattice vector (the blue circle). Because the power of the blue circle exceeds  $P_{\text{max}}$ , the blue x-mark are rescaled and re-quantized. Then, we obtain the lattice set inside the sphere (the red stripes region), which includes the quantized lattice vector. Finally, we calculate ML metrics over the reduced lattice set  $\mathcal{X}^{\text{BW}}$ . The proposed algorithm is summarized in Algorithm 2.

Throughout this section, we mainly focus on LSD with Barnes-Wall lattices. With a similar manner, we are able to extend the LSD algorithm for cubic lattices. In the case of cubic lattices,

# Algorithm 2 Lattice Sphere Decoding.

**Input:** Received signal y, Channel matrix H, SLM signal set  $S^{BW}(N_t, P_{max})$ .

**Output:** Estimated input signal point  $\hat{\mathbf{s}} \in \mathcal{S}^{BW}(N_t, P_{max})$ .

1: Using MMSE estimator, calculate the initial estimate  $\hat{\mathbf{x}}$ .

 $\hat{\mathbf{x}} = [\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_{2N_t}] = (\mathbf{H}^{\mathrm{T}}\mathbf{H} + \frac{\sigma^2}{E_{\mathrm{s}}/N_{\mathrm{t}}}\mathbf{I})^{\mathrm{T}}\mathbf{y}.$ 

If the power of  $Q(\hat{\mathbf{x}})$  exceeds  $P_{\max}$ , rescale the power of  $\hat{\mathbf{x}}$  to  $P_{\max}$  and then quantize it.

2: if 
$$||Q(\hat{\mathbf{x}})||^2 > P_{\max}$$
 then

3: 
$$\mathbf{\hat{x}} = \frac{\mathbf{x}}{||\hat{\mathbf{x}}||} \cdot \sqrt{P_{\max}}.$$

- 4: **end if**
- 5:  $\bar{\mathbf{x}} = Q(\hat{\mathbf{x}}).$

Obtain reduced lattice sphere set  $\hat{\mathcal{X}}^{BW}$ .

6:  $\hat{\mathcal{X}}^{\mathrm{BW}} = \{\hat{\mathbf{x}} | \hat{\mathbf{x}} = \bar{\mathbf{x}} + \mathbf{d}, \mathbf{d} \in \mathcal{D}_{\min} \cup \{\mathbf{0}\}\}.$ 

Merge the elements that satisfy the maximum power constraint  $P_{\text{max}}$  into  $\mathcal{X}^{\text{BW}}$ .

7: 
$$\mathcal{X}^{\mathrm{BW}} = \{\mathbf{x} | \mathbf{x} \in \mathcal{X}^{\mathrm{BW}}, ||\mathbf{x}||^2 \le P_{\mathrm{max}}\}.$$

8:  $\hat{\mathbf{s}} = \arg\min_{\mathbf{x} \in \mathcal{X}^{BW}} ||\mathbf{y} - \mathbf{H}\mathbf{x}||^2$ .



Fig. 3. The illustration of the proposed lattice sphere decoding (LSD) for SLM.

we use a simple element-wise quantization function in which every elements of the estimated lattice vector is sliced to the closest integer of  $\bar{C} = \left\{\frac{-M}{2}, \frac{-M}{2} + 1, \dots, 0, \dots, \frac{M}{2} - 1, \frac{M}{2}\right\}$ . Also, the power constraint is changed into the *M*-ary constellation constraint.

To provide a better understanding of the proposed LSD, it is instructive to consider an example.

**Example 3:** Consider a system in which  $N_t = 2$  and the signal points in Table II are used for SLM-BW. In other words, SLM-BW is designed with the maximum power constraint of 6,

i.e.,  $P_{\text{max}} = 6$ . Following by Algorithm 2, we firstly obtain an initial estimate  $\hat{\mathbf{x}}$ . For example,

$$\hat{\mathbf{x}} = [+1.32, -2.51, -0.41, +2.70]^{\mathrm{T}}.$$
 (58)

Then,  $\hat{\mathbf{x}}$  is quantized to the closest lattice vector  $Q(\hat{\mathbf{x}})^4$ .

$$Q(\hat{\mathbf{x}}) = [+1, -2, 0, +3]^{\mathrm{T}}.$$
(59)

The power of  $Q(\hat{\mathbf{x}})$  equals 14. Therefore, we rescale the power of  $\hat{\mathbf{x}}$  to  $P_{\max}$ .

$$\frac{\hat{\mathbf{x}}}{\|\hat{\mathbf{x}}\|} \sqrt{P_{\max}} = [+0.82, -1.57, -0.25, +1.68]^{\mathrm{T}}.$$
$$\bar{\mathbf{x}} = Q\left(\frac{\hat{\mathbf{x}}}{\|\hat{\mathbf{x}}\|} \sqrt{P_{\max}}\right) = [+1, -1, 0, +2]^{\mathrm{T}}.$$
(60)

After obtaining the quantized output  $\bar{\mathbf{x}}$ , the lattice sphere set  $\mathcal{X}^{BW}$  is generated.  $\mathcal{X}^{BW}$  of SLM-BW is obtained by adding  $\bar{\mathbf{x}}$  to  $\mathcal{D}_{\min}(=\mathcal{S}^2$  in Table II) and removing the elements that violate the power constraint as follows:

$$\mathcal{X}^{BW} = \left\{ \begin{bmatrix} +1 \\ -1 \\ 0 \\ +2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ +2 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ -1 \\ -1 \\ +2 \end{bmatrix}, \cdots, \begin{bmatrix} +1 \\ -2 \\ 0 \\ +1 \end{bmatrix}, \begin{bmatrix} +1 \\ -1 \\ -1 \\ -1 \\ +1 \end{bmatrix} \right\}.$$
 (61)

Then, the receiver calculates the ML metrics over the twelve symbol vectors in  $\mathcal{X}^{BW}$  and decides a symbol vector that has the minimum value of  $\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2$ . Similarly, the receiver performs the same procedure for SLM-CB.

#### **B.** Detection Complexity Analysis

In this subsection, the detection complexities of the proposed LSD and ML detectors for SLM are analyzed. In addition, these are compared to the complexities of detectors for the existing spatial modulation and the spatial multiplexing method. Each real addition, multiplication, and rounding operation are counted as one floating operation (flop).

1) MLD: Let  $N_{\rm a}$  be the number of activated dimensions in an  $2N_{\rm t}$ -dimensional real space. Therefore,  $N_{\rm a}$  of SM equals 2,  $N_{\rm a}$  of spatial multiplexing equals  $2N_{\rm t}$ , and  $N_{\rm a}$  of SLM has values between 0 and  $2N_{\rm t}$ . The complexity of ML detector in (19) can be computed as follows:

- The operator Hx requires  $N_{\rm r}N_{\rm a}$  real multiplications and  $N_{\rm r}(N_{\rm a}-1)$  real additions.

<sup>&</sup>lt;sup>4</sup>In Appendix A, we provide the quantization procedures of  $\hat{\mathbf{x}}$  as an example.

#### TABLE IV

### DETECTION COMPLEXITIES OF VARIOUS MODULATION METHODS.

Spatial Lattice Modulation		Spatial Modulation		Spatial Multiplexing	
Detection	# of real-valued flops	Detection	# of real-valued flops	Detection	# of real-valued flops
ML	$\sum_{N_{\rm a}=0}^{2N_{\rm t}} \binom{2N_{\rm t}}{N_{\rm a}} M^{N_{\rm a}} (2N_{\rm r}(N_{\rm a}+1)-1)$	ML	$(4N_{\rm r}-1)\log_2(N_{\rm t}M^2)$	ML	$(2N_{\rm r}(N_{\rm t}+1)-1)\log_2(M^{2N_{\rm t}})$
Proposed LSD	$4N_{\rm t}^3 + 12N_{\rm t}^2N_{\rm r} + 7N_{\rm t}^2 + 6N_{\rm t}N_{\rm r} - 4N_{\rm t}$	MRRC [1] $8N_{\rm t}N_{\rm r} + 2$	$8N.N \pm 2$	MMSE	$4N_{\rm t}^3 + 12N_{\rm t}^2N_{\rm r} + 7N_{\rm t}^2$
(worst case)	$ \mathcal{D}_{\min} (4N_tN_r+4N_t+2N_r-1)+\mathcal{O}(2N_t)$		$o_1 v_1 v_r + 2$		$+6N_{\rm t}N_{\rm r}+2N_{\rm t}$

• The operator  $\mathbf{y} - \mathbf{H}\mathbf{x}$  requires  $N_{\mathrm{r}}$  real additions.

• The operator  $\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2$  requires  $N_r$  real multiplications and  $N_r - 1$  real additions.

To compute  $\|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2$ ,  $2N_r(N_a + 1) - 1$  flops are required for each symbol vector  $\mathbf{x} \in S$ . SM and spatial multiplexing requires  $(4N_r - 1)\log_2(N_tM^2)$  flops and  $(2N_r(N_t + 1) - 1)\log_2(M^{2N_t})$ flops, respectively. For SLM using cubic lattices,  $N_a$  varies from 0 to  $2N_t$ ; thereby, the required flops of the SML are given by

$$\sum_{N_{\rm a}=0}^{2N_{\rm t}} |\mathcal{S}^{N_{\rm a}}| (2N_{\rm r}(N_{\rm a}+1)-1), \tag{62}$$

where  $|S^{N_a}| = {\binom{2N_t}{N_a}}M^{N_a}$ . Note that the detection complexity of SLM with barnes-wall lattices is similar to that of SLM using cubic lattices.

2) LSD: As described in Algorithm 2, the effective search space of LSD is determined by the cardinality of  $\mathcal{X}$ . This set size can be changed depending on the lattice structure for SLM and the quantized method. For convenience, we analyze the worst case of LSD, i.e,  $|\mathcal{X}| = |\mathcal{D}_{\min}| + 1$ . The complexity of LSD in Algorithm 2 is computed as follows:

- By following [30], the MMSE estimator (H<sup>T</sup>H + <sup>σ<sup>2</sup></sup>/<sub>E<sub>s</sub>/N<sub>t</sub></sub>I)<sup>T</sup>y requires 2N<sup>3</sup><sub>t</sub> + 6N<sup>2</sup><sub>t</sub>N<sub>r</sub> + 6N<sup>2</sup><sub>t</sub> + 4N<sub>t</sub>N<sub>r</sub> real multiplications and 2N<sup>3</sup><sub>t</sub> + 6N<sup>2</sup><sub>t</sub>N<sub>r</sub> + N<sup>2</sup><sub>t</sub> + 2N<sub>t</sub>N<sub>r</sub> real additions.
- As indicated in [29], the quantization complexity order of SLM-BW with  $N_t$  antennas follows  $\mathcal{O}(2N_t)$ . For example, SLM-BW with  $N_t = 2$  requires  $8N_t 2$  flops.
- The construction of  $\mathcal{X}^{BW}$  requires  $|\mathcal{D}_{min}|(4N_t 1)$  real additions and  $2|\mathcal{D}_{min}|N_t$  real multiplications.
- The operator ||y − Hx||<sup>2</sup> is computed |D<sub>min</sub>| + 1 times. Therefore, the entire flops required are (|D<sub>min</sub>| + 1)(2N<sub>r</sub>(2N<sub>t</sub> + 1) − 1).

We summarize the detection complexities of ML, MMSE, and LSD in Table IV and compared with other modulation and detection methods.

## VI. NUMERICAL RESULTS

This section provides numerical results on the average mutual information, uncoded symbolvector-error-rate (SVER), and coded frame-error-rate (FER) of the proposed SLM methods. These metrics are used to show SNR gains compared to the existing SM and the spatial multiplexing techniques. All results are obtained by using Monte Carlo simulations on independent flat-fading channel realizations for various spectral efficiencies as a function of SNR.

# A. Average Mutual Information

Fig. ?? depicts the achievable spectral efficiencies of various transmission strategies for a MIMO system with  $N_t = 2$ ,  $N_r = 4$ , and the *M*-PAM input constellation set per in-phase and quadrature component (equivalently,  $M^2$ -QAM). The solid-black line shows the capacity of  $2 \times 4$  MIMO channel when using Gaussian input signaling, which serves an upper bound of the spectral efficiencies attained by the other transmission methods. In Fig. ??, the proposed SLM method with cubic lattices (SLM-CB) uses 4-PAM per dimension, which is able to generates  $5^4$  lattice symbol vectors. In case of SLM with Barnes-Wall lattices (SLM-BW), we select  $5^4$  elements from  $S^{BW}(2, 18)$  in ascending order of the power consumption. The proposed SLM-CB and SLM-BW both achieve the spectral efficiency of  $\log_2(5^4) \approx 9.29$  bits/sec/Hz beyond 20 dB SNR. Especially, SLM-BW achieves higher spectral efficiency than that of SLM-CB in the mid SNR because Barnes-Wall lattices provide a higher nominal coding gain than do cubic lattices,  $\gamma(\Lambda_4^{BW}) > \gamma(\Lambda_4^{CB})$ . In contrast, spatial multiplexing and SM achieve the spectral efficiencies of 8 bits/sec/Hz and 5 bits/sec/Hz respectively when SNR is high enough. Similar results are observed when 2-ary PAM input signal is used. Thus, SLM achieves the higher spectral efficiency than do the conventional methods.

## B. Symbol Vector Error Rate

We compare the SVERs of the proposed SLM methods with that of the existing SM, GSM, GSM-Eisenstein, ESM, and spatial multiplexing methods. For fair comparisons, each modulation scheme is allowed to use a different modulation size M to achieve similar target spectral efficiencies. It should be noted that the results of Fig. 5 and Fig. 6 are obtained under ML detection and the results of Fig. 7 and Fig. 8 are obtained under ML detection and LSD.

Fig. 4 shows the SVERs when the target spectral efficiency is 9 bits/sec/Hz for the  $2 \times 8$  MIMO system. To meet this criterion, we set different modulation size for each method:



Fig. 4. SVERs for the 2×8 MIMO system with the target spectral efficiency of 9 bits/sec/Hz.



Fig. 5. SVERs for the  $4 \times 8$  MIMO system with the target spectral efficiency of 12 bits/sec/Hz.

- SM with 256-QAM, which uses 256-QAM to modulate 8 bits and the antenna index to modulate 1 bit;
- Spatial multiplexing with 16-QAM and 32-QAM, which uses 16-QAM to modulate 4 bits and 32-QAM to modulate 5 bits, and sends two symbols simultaneously via different transmit antennas;

- ESM with 64-QAM, which uses 64-QAM as a primary signal constellation set to modulate 9 bits;
- QSM with 128-QAM, which uses 128-QAM to modulate 7 bits and the antenna indies to modulate 2 bit;
- SLM-CB with 16-QAM, which is constructed by selecting 512 elements from  $S^{CB}(2,4)$  in ascending order of the power consumption to modulate 9 bits;
- SLM-BW, which is constructed by selecting 512 elements from S<sup>BW</sup>(2,14) in ascending order of the power consumption to modulate 9 bits;

Fig. 4 shows that the proposed SLM-CB and SLM-BW achieve lower SVERs than all other transmission methods at all SNR. SLM-CB provides about 1 dB SNR gain over the spatial multiplexing at SVER =  $10^{-3}$ . SLM-BW provides 1 dB SNR gain over SLM-CB. This increase occurs because the Barnes-Wall lattice has a larger nominal coding gain than does cubic lattice. In addition, our analytic expression of SVER obtained from (46) is shown to be tight at high SNR.

In Fig. 5, SVERs are also measured at the target spectral efficiency of 12 bits/sec/Hz for the 4 × 8 MIMO system. For GSM-Eisenstein and ESM, the target spectral efficiency of 12 bits/sec/Hz is not well defined in [15] and [17]. Alternatively, we choose the target spectral efficiency of GSM-Eisenstein at 10.9 bits/sec/Hz and ESM at 11 bits/sec/Hz. The modulation size *M* for each method and the corresponding spectral efficiency are summarized in the legend of Fig. 5. As a result, SLM-BW and SLM-CB achieve a lower SVER than that of the existing methods. In particular, SLM-BW provides about 3 dB SNR gains over spatial multiplexing at SVER =  $10^{-3}$ , which is the 1 dB additional gain compared to the 2 × 8 MIMO system. These gains are obtained because SLM achieves a high spectral efficiency with a low modulation size when the number of transmit antennas is large enough. For example, with  $N_t = 4$  SLM using 2-PAM per dimension (or 4-QAM per two dimensions) is able to achieve the spectral efficiency of 12 bits/sec/Hz, whereas the spatial multiplexing method needs to use 8-PSK to attain 12 bits/sec/Hz. Intuitively, the use of low modulation size leads to the increase of  $d_{min}^2$  under the transmit power constraint.

Fig. 6 demonstrates the detection performance of the proposed LSD. The proposed LSD is compared to ML and MMSE detection methods. For comparison, we consider  $2 \times N_r$  ( $N_r = 4, 8$ ) MIMO systems with target spectral efficiencies of about 7 bits/sec/Hz (Fig. 6(a)) and 9 bits/sec/Hz (Fig. 6(b)). The corresponding simulation setups are listed in each sub-caption. For LSD, we



Fig. 6. Comparison of the SVERs under LSD and ML detection with the  $2 \times N_r$  ( $N_r = 4, 8$ ) MIMO systems. (a): SLM-BW with  $S^{BW}(2, 6)$ , 7.17 bits/sec/Hz, and spatial multiplexing with 8-PSK and 16-QAM, 7 bits/sec/Hz. (b): SLM-BW with  $S^{BW}(2, 14)$ , 9.32 bits/sec/Hz, and spatial multiplexing with 16-QAM and 32-QAM, 9 bits/sec/Hz.

use the set  $S^2$  in Table II as  $D_{\min}$ . Therefore, among 145 (Fig. 6(a)), 601 (Fig. 6(b)) SLM vectors, the maximum 25 lattice vectors are computed for the ML metric. In other words, the effective search space of LSD is reduced by 17%, 4% of ML detection. Even with the significant complexity reduction, remarkably, the performance of LSD (the green marks) is almost same with SLM with ML detection (the red solid lines). Additionally, SLM with LSD achieves a lower SVER than that of spatial multiplexing with ML and MMSE detection.

## C. Frame Error Rate

Fig. 7 shows FERs of different transmission strategies in a  $2 \times 4$  coded MIMO system with the target spectral efficiency of 6 bits/sec/Hz. In the simulation, one frame is assumed to consist of 1200 coded bits with a rate of 1/2 turbo code. Thus, 600 information bits are sent per frame. We assume an independent fading channel model in which channel matrices H vary at every channel use. Rate-1/2 turbo codes based on parallel concatenated codes with feedforward and feedback polynomial (15,13) in octal notation are used for the simulations. Using soft information, turbo decoding iterates six times, which is sufficient to reduce coded FER compared to the uncoded case.



Fig. 7. Comparison of the FERs for the  $2 \times 4$  coded MIMO system.

For the target spectral efficiency of 6 bits/sec/Hz, SM uses 32QAM, spatial multiplexing uses 8PSK, SLM-CB uses 2-PAM, and SLM-BW is designed with  $P_{\text{max}} = 6$ . Among 81 (SLM-CB), 145 (SLM-BW) lattice vectors, we select 64 lattice vectors in ascending order of the power consumption. In other words, the indices from 1 to 64 in Table I and II are used to modulate 6 bits. For bits-to-symbol mapping, the binary switching algorithm <sup>5</sup> in [31] is applied as a lattice indexing algorithm. We uploaded bits-to-symbol mapping tables in [32]. Fig. 7 demonstrates that the proposed SLM methods provides 1 dB SNR gain over the spatial multiplexing, and 3 dB over SM at FER =  $10^{-2}$  when ML detection is applied. The trend of this result is similar to that in Fig. 4, in which the number of transmit antennas is the same. The red and green marks in Fig. 7 represent the FERs of SLM-BW and SLM-CB when LSD is applied. Soft information for LSD is extracted by calculating log likelihood-ratio values of the vectors in  $\mathcal{X}$ . Similar to Fig 6, the performance of LSD achieves near ML performance.

#### VII. CONCLUSION

We have presented SLM, a new spatial modulation technique for MIMO systems. By jointly mapping information bits into a set of  $2N_t$ -dimensional lattice points, we have shown that SLM

<sup>&</sup>lt;sup>5</sup>The binary switching algorithm in [31] provides locally optimal solutions for codebook indexing. Therefore, the performance of SLM in Fig. 7 can be further improved by using an novel indexing algorithm.

is able to achieve the maximum spectral efficiency at high SNR under the PAM input constraint per dimension. We also demonstrated that the SLM that uses dense Barnes-Wall lattices offers a considerable SNR gain. We derived a tight approximation of the average mutual information and the upper bound of average symbol-vector-error-probability and used simulations to validate the effectiveness of our analysis. In addition, lattice sphere decoding for SLM was proposed to diminish the detection complexity at the receiver. Simulations showed that the performance of LSD closely matches that of ML at every SNR region with practical MIMO setups.

A promising direction for future work includes a study of impact on SLM for the case where the number of RF chains is greater than the number of transmit antennas, i.e.,  $N_{\rm t} > N_{\rm RF}$ . Other interesting research directions are to devise SLM techniques combined with index modulation methods [33] for multi-carrier MIMO systems and to devise SLM techniques for the integerforcing framework in [34].

#### APPENDIX

# A. Lattice Vector Quantizer

In this Appendix, we explain the lattice vector quantizers (LVQ) proposed in [29]. The LVQs in [29] can quantize any *n*-dimensional vectors into  $D_n$  ( $D_4 = \Lambda_4^{\text{BW}}$ ) or  $E_n$  ( $E_8 = \Lambda_8^{\text{BW}}$ ) lattices. The main advantages of these LVQs are that they can be implemented with simple arithmetic operations and the quantization complexities follow the order of  $\mathcal{O}(n)$ . Each lattice requires a different quantization procedure. Due to the paper limitation, we only focus on the LVQ of  $D_n$  lattices.

Let  $f(\hat{x}_i)$  be the closest integer of  $\hat{x}_i$ . Then, for  $\hat{\mathbf{x}} = [\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_n]^T \in \mathbb{R}^n$ , we define  $f(\hat{\mathbf{x}})$  as

$$f(\hat{\mathbf{x}}) = [f(\hat{x}_1), f(\hat{x}_2), \cdots, f(\hat{x}_n)]^{\mathrm{T}}.$$
 (63)

We also define  $\delta(\hat{\mathbf{x}})$  as the difference between  $\hat{\mathbf{x}}$  and  $f(\hat{\mathbf{x}})$ .

$$\delta(\hat{\mathbf{x}}) = [\hat{x}_1 - f(\hat{x}_1), \hat{x}_2 - f(\hat{x}_2), \cdots, \hat{x}_n - f(\hat{x}_n)]^{\mathrm{T}}.$$
(64)

Using  $f(\hat{\mathbf{x}})$  and  $\delta(\hat{\mathbf{x}})$ , we define  $g(\hat{\mathbf{x}})$  as

$$g(\hat{\mathbf{x}}) = [f(\hat{x}_1), \cdots, w(\hat{x}_k), \cdots, f(\hat{x}_n)]^{\mathrm{T}},$$
(65)

$$w(\hat{x}_k) = f(\hat{x}_k) + 1, \text{ if } \delta(\hat{x}_k) \ge 0,$$
  
=  $f(\hat{x}_k) - 1, \text{ if } \delta(\hat{x}_k) < 0.$  (66)

The notable fact is that the sum of all  $D_n$  lattice vectors should be even, i.e.,  $mod(\sum_i \hat{x}_i, 2) = 0$ . Therefore, we need to check whether  $\sum_i f(\hat{x}_i)$  or  $\sum_i g(\hat{x}_i)$  is even. Then, we admit the vector which has even sum as the quantized output  $Q(\hat{\mathbf{x}})$ .

**Example 4:** Suppose  $\hat{\mathbf{x}} = [+1.32, -2.51, -0.41, +2.70]^{\mathrm{T}}$ . Then  $f(\hat{\mathbf{x}})$ ,  $\delta(\hat{\mathbf{x}})$ , and  $g(\hat{\mathbf{x}})$  are computed as

$$f(\hat{\mathbf{x}}) = [+1, -3, 0, +3]^{\mathrm{T}},$$
  

$$\delta(\hat{\mathbf{x}}) = [+0.32, +0.49, -0.41, -0.30]^{\mathrm{T}},$$
  

$$g(\hat{\mathbf{x}}) = [+1, -2, 0, +3]^{\mathrm{T}}.$$
(67)

Considering that  $\sum_i f(\hat{x}_i) = 1$  (odd) and  $\sum_i g(\hat{x}_i) = 2$  (even), we select  $g(\hat{\mathbf{x}})$  as the quantized output of  $\hat{\mathbf{x}}$ , i.e.,  $Q(\hat{\mathbf{x}}) = g(\hat{\mathbf{x}})$ .

#### REFERENCES

- R. Mesleh, H. Haas, S. Sinanovic, C. W. Ahn, and S. Yun, "Spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2228–2241, Jul. 2008.
- [2] M. Di Renzo, H. Haas, and P. M. Grant, "Spatial modulation for multiple-antenna wireless systems: A survey," *IEEE Commun. Mag.*, vol. 49, no. 12, Dec. 2011.
- [3] 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. 56–103, Jan. 2014.
- [4] 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, First Quarter 2015.
- [5] P. Yang, Y. Xiao, Y. L. Guan, K. Hari, A. Chockalingam, S. Sugiura, H. Haas, M. Di Renzo, C. Masouros, Z. Liu 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, Third Quarter 2016.
- [6] J. Jeganathan, A. Ghrayeb, L. Szczecinski, and A. Ceron, "Space shift keying modulation for MIMO channels, *IEEE Trans. Wireless Commun.*, vol. 8 no. 7, pp. 3692–3703, Jul. 2009.
- [7] A. Younis, N. Seramovski, R. Mesleh, and H. Haas, "Generalised spatial modulation," in *Proc. IEEE Signals, Syst. Comput.* (ASILOMAR), Pacific Grove, CA, USA, Nov. 2010, pp 1498–1502.
- [8] J. Jeganathan, A. Ghrayeb, and L. Szczecinski, "Generalized space shift keying modulation for MIMO channels," in *Proc. IEEE Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Cannes, France, Sep. 2008, pp. 1–5.

- [9] J. Wang, S. Jia, and J. Song, "Generalised spatial modulation system with multiple active transmit antennas and low complexity detection scheme," *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1605-1615, Apr. 2012.
- [10] K. M. Humadi, A. I. Sulyman, and A. Alsanie, "Experimental results for generalized spatial modulation scheme with variable active transmit antennas," in *IEEE Int. Conf. on Cognitive Radio Oriented Wireless Networks (CROWNCOM)*, Doha, Qatar, April. 2015, pp. 260–270.
- [11] R. Mesleh, S. S. Ikki, and H. M. Aggoune, "Quadrature spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 64, no. 6, pp. 2738–2742, Jun. 2015.
- [12] A. A. I. Ibrahim, T. Kim, and D. J. Love, "On the achievable rate of generalized spatial modulation using multiplexing under a Gaussian mixture model," *IEEE Trans. Commun.*, vol. 64, no. 4, pp. 1588–1599, Apr. 2016.
- [13] N. Ma, A. Wang, C. Han, and Y. Ji, "Adaptive joint mapping generalised spatial modulation," in *IEEE Int. Conf. Commun. in China (ICCC)*, China, Beijing, Aug. 2012, pp. 520–523.
- [14] S. Guo, H. Zhang, S. Jin, and P. Zhang, "Spatial modulation via 3-D mapping," *IEEE Commun. Lett.*, vol. 20, no. 6, pp. 1096–1099, Jun. 2016.
- [15] C.-C. Cheng, H. Sari, S. Sezginer, and Y. T. Su, "Enhanced spatial modulation with multiple signal constellations," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2237–2248, Jun. 2015.
- [16] C.-C. Cheng, H. Sari, S. Sezginer, and Y. T. Su, "New signal designs for enhanced spatial Modulation," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7766–7777, Nov. 2016.
- [17] J. Freudenberger and S. Shavgulidze, "Signal constellations based on Eisenstein integers for generalized spatial modulation," *IEEE Commun. Lett.*, vol. 21, no. 3, pp. 556–559, Mar. 2017.
- [18] G. D. Forney Jr. and G. Ungerboeck, "Modulation and coding for linear Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2384–2415, Jun. 1998.
- [19] J. Conway and N. Sloane, Sphere packings, lattices and groups, Springer Science & Business Media, 2013.
- [20] G. Nebe, E. M. Rains, and N. J. A. Sloane, "A simple construction of the Barnes-Wall lattices," in *Codes, Graphs, and Systems*, Springer, 2002, pp. 333–342.
- [21] X. Guan, Y. Cai, W. Yang, "On the mutual information and precoding for spatial modulation with finite alphabet," *IEEE Wireless Commun. Lett.*, vol. 2, no. 4, pp. 383–386, Aug. 2013.
- [22] S.-R. Jin, W.-C. Choi, J.-H. Park, and D.-J. Park, "Linear precoding design for mutual information maximization in generalized spatial modulation with finite alphabet inputs," *IEEE Commun. Lett.*, vol. 19, no. 8, pp. 1323–1326, Aug. 2015.
- [23] S. Guo, H. Zhang, J. Zhang, and D. Yuan, "On the mutual information and constellation design criterion of spatial modulation MIMO systems," in *Proc. IEEE Int. Conf. Commun. Syst. (ICCS)*, Macau, China, Nov. 2014, pp. 487–491.
- [24] B. P. Buckles and M. Lybanon, "Generation of a vector from the lexicographical index," *ACM Trans. Math. Softw.*, vol. 3, no. 2, pp. 180-182, Jun. 1977.
- [25] P. Rault and C. Guillemot, "Indexing algorithms for  $Z_n, A_n, D_n$ , and  $D_n^{++}$  lattice vector quantizers." *IEEE Trans. Multimedia*, vol. 3, no. 4, pp. 395–404, Dec. 2001.
- [26] V. Tarokh, A. Vardy, and K. Zeger, "Universal bound on the performance of lattice codes," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 670–681, Mar. 1999.
- [27] D. Tse and P. Viswanath, Fundamentals of wireless communication, Cambridge university press, 2005.
- [28] M.-S. Alouini and A. Goldsmith, "A unified approach for calculating error rates of linearly modulated signals over generalized fading channels," *IEEE Trans. Commun.*, vol. 47, no. 9, pp. 1324–1334, Sept. 1999.
- [29] J. Conway and N. Sloane, "Fast quantizing and decoding and algorithms for lattice quantizers and codes," *IEEE Trans. Inf. Theory*, vol.28, no. 2, pp. 227–232, Mar. 1982.

- [30] L. Xiao et al., "Efficient compressed sensing detectors for generalized spatial modulation systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 1284–1298, Feb. 2017.
- [31] K. Zeger abd A. Gersho, "Pseudo-Gray coding," IEEE Trans. Commun., vol. 38, pp. 2147–2158, Dec. 1990.
- [32] J. Choi, Y. Nam, and N. Lee, "Bits-to-symbol mapping tables for spatial lattice modulation", [Online]. Available: http://wisepostech.com.
- [33] E. Başar, Ü. Aygölü, E. Panayırcı, and H. V. Poor, "Orthogonal frequency division multiplexing with index modulation," *IEEE Trans. Sig. Process.*, vol. 61, no. 22, pp. 5536–5549, Nov. 2013.
- [34] J. Zhan, B. Nazer, U. Erez, and M. Gastpar, "Integer-Forcing Linear Receivers," *IEEE Trans. Inf. Theory*, vol. 60, no. 12, pp. 7661–7685, Dec. 2014.