



HAL
open science

Variable- N_u Generalized Spatial Modulation for Indoor LOS mmWave Communication: Performance Optimization and Novel Switching Structure

Peng Liu, Marco Di Renzo, Andreas Springer

► **To cite this version:**

Peng Liu, Marco Di Renzo, Andreas Springer. Variable- N_u Generalized Spatial Modulation for Indoor LOS mmWave Communication: Performance Optimization and Novel Switching Structure. IEEE Transactions on Communications, 2017, 65 (6), pp.2625 - 2640. 10.1109/TCOMM.2017.2676818 . hal-01879989

HAL Id: hal-01879989

<https://hal.science/hal-01879989>

Submitted on 7 Jul 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Variable- N_u Generalized Spatial Modulation for Indoor LOS mmWave Communication: Performance Optimization and Novel Switching Structure

Peng Liu, *Member, IEEE*, Marco Di Renzo, *Senior Member, IEEE*, and Andreas Springer, *Member, IEEE*

Abstract—In this paper, we propose to use variable- N_u generalized spatial modulation (VGSM), which is a specific type of generalized spatial modulation (GSM), in indoor line-of-sight (LOS) millimeter-wave (mmWave) communication. As compared to fixed- N_u generalized spatial modulation (FGSM), VGSM needs less number of transmitter (TX) antennas to achieve the same data rate. Reducing the number of antennas is especially important for LOS mmWave MIMO communication since it means reduced array lengths and routing loss in the RF front-end. We first derive and analyze the channel capacity for LOS VGSM, and then propose a novel switching structure for implementation of the VGSM TX at mmWave frequencies and analyze its performance in terms of power added efficiency (PAE) and capacity. Two optimizations of the system performance of VGSM are performed: 1) power allocation optimization for the proposed TX, and 2) antenna separation optimization in order to optimize the capacity in LOS. Our results show that VGSM is a feasible and promising scheme for indoor LOS mmWave communication. For example, we show that an 8×8 VGSM system with practical component values can achieve more than 20 bpcu spectral efficiency within a distance of 1–5 m and more than 17 bpcu within 5–10 m.

Index Terms—Antenna switching, generalized spatial modulation (GSM), indoor millimeter-wave (mmWave) communication, line-of-sight (LOS) MIMO, 60 GHz

I. INTRODUCTION

SPATIAL modulation (SM) and generalized spatial modulation (GSM) are hot topics in wireless communications in recent years [1]. The operating principle of SM/GSM MIMO systems [2]–[4] are radically different from that of conventional spatial multiplexing (SMX) MIMO systems, of which the most notable examples are the D-BLAST (Diagonal-Bell Laboratories Layered Space-Time) [5] and the V-BLAST (Vertical BLAST) [6] schemes. In SMX MIMO, the SMX gain (or increased spectral efficiency) is achieved by simultaneous transmission of multiple data streams at the same frequency.

P. Liu is with NXP Semiconductors Austria GmbH in Gratkorn (e-mail: peng.liu_6@nxp.com), and A. Springer is with the Institute for Communications Engineering and RF-Systems (NTHFS), Johannes Kepler University (JKU) Linz, Austria (e-mail: andreas.springer@jku.at). Their work is in part supported by the Austrian COMET-K2 programme of the Linz Center of Mechatronics (LCM), which is funded by the Austrian federal government and the federal state of Upper Austria.

M. Di Renzo is with the Laboratoire des Signaux et Systèmes, CNRS, CentraleSupélec, Univ Paris Sud, Université Paris-Saclay, 3 rue Joliot Curie, Plateau du Moulon, 91192, Gif-sur-Yvette, France. (e-mail: marco.direnzo@l2s.centralesupelec.fr). His work is partially supported by the European Commission (EC) through the H2020-ETN-5Gwireless research project (grant 641985) and by the Agence Nationale de la Recherche Scientifique (ANR) through the SpatialModulation research project (Société de l'Information et de la Communication – Action Plan 2015).

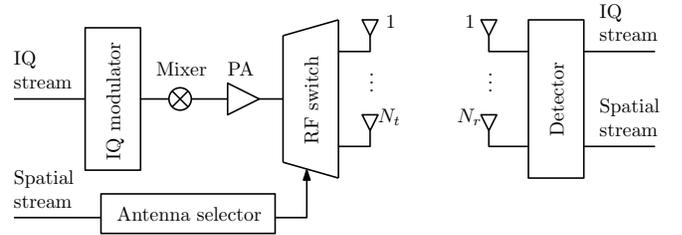


Fig. 1. Conceptual architecture for SM/GSM MIMO.

Actually, SMX MIMO is the most capable MIMO scheme and can achieve the highest spectral efficiency. However, it requires the same number of RF chains as the number of data streams thus suffering from some major drawbacks, e.g., inter-channel interference, inter-antenna synchronization, multiple RF chains [7], [8]. This leads to high hardware and signal processing complexity and low power efficiency in terms of bits/Joule [9]–[12]. SM/GSM MIMO transmits two data streams — one in the conventional in-phase and quadrature (IQ) domain and the other in a so-called spatial domain. As shown in Fig. 1, the single IQ stream requires only a single transmitter (TX) RF chain, while the spatial transmission is based on antenna switching, i.e., only a sub-group of TX antennas are activated within each time slot and the spatial bits are modulated in the indices of the activated antennas. High spectral efficiencies are achieved in SM/GSM MIMO by using a large number of TX antennas. Due to the use of only a single TX RF chain, SM/GSM shows improved bit/Joule power efficiency and reduced detection complexity as compared to SMX MIMO [8]–[12]. So far, SM/GSM has just mainly been studied for low-GHz fading channels [2]–[4].

In the next generation wireless communications (5G), the global spectrum shortage is pushing us to use millimeter-wave (mmWave) frequencies which can provide us with multi-GHz bandwidth and enables transmission rates of multi-Gbps [13]–[15]. Especially, with up to 7 GHz unlicensed bandwidth available worldwide, the 60 GHz band (57 – 64 GHz) has gained significant interest in the industry as evidenced e.g., by the standardization of Wireless HD, IEEE 802.15.3c, and IEEE 802.11ad [16]. In the mmWave literature, multi-stream MIMO communications is mainly based on the conventional SMX principle but additionally with beamforming to combat the much higher free space path loss (FSPL) at mmWave frequencies [17]–[19]. Cost and power consumption of mmWave components are high, therefore it is unlikely to dedicate a separate RF

chain and digital-to-analog converter (DAC)/analog-to-digital converter (ADC) for each antenna, but it is more likely to use fewer RF chains than the number of antennas [18], [20]. In [17], a hybrid beamforming (BF)/SMX architecture, which uses a smaller number of RF chains than the number of antennas, was proposed for line-of-sight (LOS) mmWave communications. In [18], a hybrid RF/baseband precoder architecture, which uses N_t^{RF} TX RF chains to transmit N data streams via N_t TX antennas (where $N \leq N_t^{\text{RF}} \leq N_t$), was proposed for sparsely scattered mmWave channels. However, in both examples, N (or more) RF chains are still needed for transmission of N data streams. This motivates us to consider the simpler and more energy-efficient SM/GSM MIMO also in indoor mmWave communications. More specifically, we focus our study on VGSM, in which the number of used/activated TX antennas N_u is varied, over indoor LOS mmWave channels. Compared to fixed- N_u generalized spatial modulation (FGSM), in which the number of used/activated TX antennas N_u is fixed, VGSM can make more efficient use of the number of TX antennas and achieve a higher data rate.

VGSM has previously been studied, e.g., in [21]–[23] based on single-carrier space modulation and in [24], [25] based on joint space and frequency index modulation. However, none of them has targeted at indoor LOS mmWave communication. To date, only FGSM and space shift keying (SSK) have been studied for indoor mmWave communications [26]–[28]. A major difference between indoor LOS mmWave SM/GSM and low-GHz SM/GSM lies in the different propagation and channel characteristics. At mmWave frequencies, e.g., 60 GHz, the indoor channel is dominated by the LOS component (if not blocked), which is deterministic and determined by the relative geometry of the setup. The NLOS paths, which typically suffer from a loss of 10 dB per reflection and additional path loss due to larger path lengths as compared to the LOS paths [29], have only minor contribution to the received power. The deterministic nature of the dominant LOS component actually gives us opportunities to optimize the performance of LOS SM/GSM [26], [27] and LOS SMX [17], [30]. Optimized operation of LOS SM/GSM and LOS SMX, in terms of symbol error probability (SEP) or capacity, relies on proper separation of the TX and receiver (RX) antennas by an amount that is in general much larger than $\lambda/2$. In contrast, at low-GHz frequencies, the channels are usually rich-scattered and exhibit fading characteristics. Excess separation of the antennas at low-GHz is generally not necessary. Antenna separations down to $\lambda/2$ usually can lead to sufficiently low channel correlation and allow exploitation of diversity and multiplexing gains. To find the optimal antenna separation product (ASP) that maximizes the capacity of VGSM in LOS is one of our objectives.

Another focus of our work is on the practicability of GSM at mmWave frequencies, where we consider a communication bandwidth in the GHz order. Due to the large communication bandwidth, simplicity — simple waveform, simple hardware and simple signal processing — is our major concern. This motivates us to consider a single-carrier and a single IQ stream, instead of multiple sub-carriers and/or multiple IQ streams [22]–[25]. A possible implementation of the TX and RX for

FGSM was proposed in [26]. However, to the best of our knowledge, the switching structure of VGSM has never been addressed and remained a black box.

Against this background, novel contributions of this paper are as follows:

- We derive the channel capacity for LOS VGSM and the main results are summarized in Theorem 1 and Corollary 2 in Section III. Based on the expressions derived, we perform numerical studies in Section IV. The analyses and expressions derived are novel and unique to LOS VGSM. Furthermore, we numerically study the impact of non-line-of-sight (NLOS) components of the channel on the capacity of VGSM in LOS.
- We numerically optimize the ASP of the uniform linear arrays (ULAs) of the TX and RX to maximize the capacity of LOS VGSM. The capacities of LOS FGSM and LOS SMX are also optimized for comparison. It was found in [26] that, to optimize the SEP of LOS SM/GSM, the channel matrix needs to be (column) orthogonal thus requiring considerable ASP (see also Section II-D). However, in this paper, we will show that the capacities of LOS VGSM and FGSM, depending on the signal-to-noise ratio (SNR), can be maximized with a much smaller ASP. This finding suggests compact implementations and the practicality of LOS GSM is greatly enhanced.
- We propose a novel TX switching structure for implementation of VGSM at mmWave frequencies (see Fig. 10) and study its power efficiency. In [26], the authors have proposed: 1) a mmWave TX which, however, supports only FGSM but not VGSM, and 2) a mmWave RX which can be used for both FGSM and VGSM. The implementation of the TX for VGSM remained an open topic. The proposed TX for VGSM in this paper therefore serves the purpose to fill up this gap.
- We optimize the power allocation of VGSM and find that, depending on the IQ constellation, different power allocation strategies have to be used in order to maximize the minimum Euclidean distance of the VGSM symbols. Such optimization is necessary to achieve improved error performance for VGSM. The finding is summarized in Theorem 3.

The rest of the paper is organized as follows. In Section II, we introduce the system and channel models. In Section III, we derive the capacity expressions for VGSM in LOS. Both pure LOS and multipath channels are considered. In Section IV, we present numerical results on capacity. The capacity of VGSM over both pure LOS and multipath channels is studied, and the ASP is optimized in the pure LOS case. In Section V, we introduce the proposed TX for VGSM and optimize the power allocation of the VGSM modulation. The power efficiency of the proposed RF front-end is derived and analyzed. As an example, the achieved capacity is calculated based on practical component values and link parameters. Finally, Section VI concludes the paper.

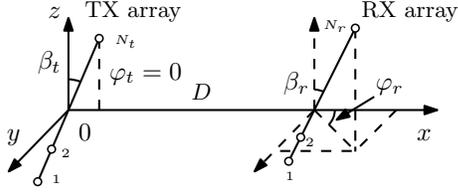


Fig. 2. LOS MIMO using arbitrarily placed ULAs. The coordinate system is chosen such that the center of the TX ULA locates at the origin and the center of the RX ULA lies on the x -axis. Furthermore, the TX ULA lies in the xz -plane, i.e., $\varphi_t = 0$.

II. PRELIMINARIES

A. Channel Model

As shown in Fig. 2, we consider an $N_t \times N_r$ LOS MIMO system using two arbitrarily placed ULAs, where N_t is the number of TX antennas and N_r is the number of RX antennas. The antenna separations in the TX and RX ULAs are respectively s_t and s_r , and the TX-RX distance is D , which is measured between the center points of the TX and RX ULAs. Geometrically, each ULA has two degrees of freedom, namely the azimuth angle φ and the tilt angle β (or equivalently the elevation angle $\theta = \pi/2 - \beta$). Since we are only interested in their relative geometric relation, only three angles are used — the tilt angles β_t and β_r respectively at the TX and RX and the azimuth angle φ_r at the RX. Note that we limit β_t and β_r to $[0, \pi/2]$ yet still being able to cover all cases. We do not consider the mutual coupling between antennas as the antenna separations we are dealing with in this paper are in general much larger than $\lambda/2$.

The indoor propagation model is commonly expressed as [17], [29], [30]

$$\mathbf{H}(t) = \mathbf{H}_{\text{LOS}}(t) + \mathbf{H}_{\text{NLOS}}(t), \quad (1)$$

where \mathbf{H}_{LOS} and \mathbf{H}_{NLOS} are respectively the LOS and NLOS components. For convenience, \mathbf{H} is normalized by the FSPL [31] due to the TX-RX distance D .

1) *LOS Component*: The j, i -th element of \mathbf{H}_{LOS} , which represents the LOS path from the i -th TX antenna to the j -th RX antenna, is calculated as

$$h_{ji}(t) = \frac{\lambda}{4\pi d_{ji}} \delta(t - \tau_{ji}) \cdot \frac{4\pi D}{\lambda} \approx \delta(t - \tau_{ji}), \quad (2)$$

where $k = 2\pi/\lambda$ is the wave number, $\delta(\cdot)$ is the Dirac delta function, t is time, $\tau_{ji} = d_{ji}/c$ and d_{ji} are respectively the delay and length of the LOS path between the i -th TX antenna and the j -th RX antenna, c is the speed of light. For the approximation in (2) to hold, we have assumed that the TX and RX array lengths L_t, L_r are much smaller than D , i.e., $L_t, L_r \ll D$.

2) *NLOS components*: We use a similar ray-tracing channel model as the one used by [17], which, as shown in Fig. 3, is modeled in a room of dimensions $5\text{ m} \times 5\text{ m} \times 3\text{ m}$. The floor is assumed to be carpeted such that floor reflections can be ignored due to high reflection loss of carpeted surfaces [32]. The antennas are assumed to have a hemispheric radiation pattern towards the LOS direction. As a result, there exist only

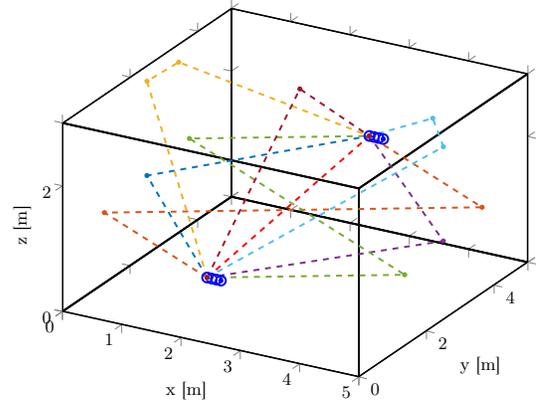


Fig. 3. Indoor propagation environment showing LOS, first-order and second-order paths between one pair of TX and RX antennas.

three first order and four second order reflected paths. The NLOS component of \mathbf{H} is written as

$$\mathbf{H}_{\text{NLOS}}(t) = \sum_{k=1}^3 \mathbf{H}_{\text{R1}}^{(k)}(t) + \sum_{l=1}^4 \mathbf{H}_{\text{R2}}^{(l)}(t), \quad (3)$$

where $\{\mathbf{H}_{\text{R1}}^{(k)}\}$ are the three first-order reflected components off the left and right walls and the ceiling, and $\{\mathbf{H}_{\text{R2}}^{(l)}\}$ are the four second-order reflected components off the left and right walls and the ceiling. The j, i -th element of $\mathbf{H}_{\text{R1}}^{(k)}$ is given by

$$h_{\text{R1},ji}^{(k)}(t) = \Gamma(\theta_{\text{R1},ji}^{(k)}) \frac{D}{d_{\text{R1},ji}^{(k)}} \delta(t - \tau_{\text{R1},ji}^{(k)}), \quad (4)$$

where Γ is the reflection coefficient which is a function of the angle of incidence $\theta_{\text{R1},ji}^{(k)}$, $\tau_{\text{R1},ji}^{(k)} = d_{\text{R1},ji}^{(k)}/c$ and $d_{\text{R1},ji}^{(k)}$ are respectively the delay and length of the k -th first-order NLOS path from the i -TX antenna to the j -th RX antenna. The j, i -th element of $\mathbf{H}_{\text{R2}}^{(l)}$ is given by

$$h_{\text{R2},ji}^{(l)}(t) = \Gamma_1(\theta_{\text{R2},ji}^{(l)}) \Gamma_2(\phi_{\text{R2},ji}^{(l)}) \frac{D}{d_{\text{R2},ji}^{(l)}} \delta(t - \tau_{\text{R2},ji}^{(l)}), \quad (5)$$

where Γ_1 and Γ_2 are respectively the reflection coefficients of the first and second reflections, $\theta_{\text{R1},ji}^{(k)}$ and $\phi_{\text{R1},ji}^{(k)}$ are respectively the incidence angle of the first and second reflections, $\tau_{\text{R2},ji}^{(k)} = d_{\text{R2},ji}^{(k)}/c$ and $d_{\text{R2},ji}^{(l)}$ are respectively the delay and length of the l -th second-order NLOS path from the i -TX antenna to the j -th RX antenna. In the calculation of Γ , Γ_1 and Γ_2 , vertical antenna polarization is assumed and the surface material of the walls and ceiling is assumed to be 9.0-mm-thick plasterboard with a refractive index of $n = 1.74 - j0.023$ (j is the complex unit) [32].

Considering that the LOS component is dominant, we take a similar approach as [17], [26], [29], [30] to first design and optimize our scheme based on the LOS component. The NLOS components will be considered and its impact on channel capacity will be studied in Section IV-C.

B. Signal Model

Consider only the LOS component of the channel¹, the received signal vector \mathbf{y} ($\mathbf{y} \in \mathbb{C}^{N_r \times 1}$) is given by

$$\mathbf{y} = \mathbf{H}_{\text{LOS}} \underbrace{\{\rho_i q_m \mathbf{e}_i\}}_{\triangleq \mathbf{x} \in \mathbb{S}} + \mathbf{n} = \rho_i q_m \sum_{k \in \mathcal{I}_i} \mathbf{h}_k + \mathbf{n}, \quad (6)$$

where

- q_m is an IQ (e.g., M -QAM or M -PSK) symbol, and $m = 1, \dots, M$ is the IQ symbol index;
- \mathbf{e}_i is a spatial symbol which is an $N_t \times 1$ vector containing $N_u^{(i)}$ ones (denoting activated TX antennas) and $N_t - N_u^{(i)}$ zeros (denoting inactivated TX antennas), and $i = 1, \dots, 2^{N_t} - 1$ is the spatial symbol index;
- ρ_i is a power allocation factor that is used to optimize the VGSM constellation (see Theorem 3);
- $\mathbb{S} \triangleq \{\rho_i q_m \mathbf{e}_i \mid i = 1, \dots, 2^{N_t} - 1 \text{ and } m = 1, \dots, M\}$ is the VGSM modulation alphabet;
- $\mathbf{x} \in \mathbb{S}$ is the transmitted VGSM symbol;
- $\mathcal{I}_i = \{k \mid k \in \{1, \dots, N_t\} \text{ and } \mathbf{e}_i(k) = 1\}$ is the set of indices of the activated TX antennas corresponding to the i -th spatial symbol, and we have $|\mathcal{I}_i| = N_u^{(i)}$, where $|\cdot|$ for a set denotes its cardinality, and $N_u^{(i)}$ is the number of activated antennas corresponding to i -th spatial symbol;
- \mathbf{h}_k is the k -th column of \mathbf{H}_{LOS} ;
- $\mathbf{n} \in \mathbb{C}^{N_r \times 1}$ is an independent and identically distributed (i.i.d.) additive white Gaussian noise (AWGN) vector with covariance matrix $\sigma_n^2 \mathbf{I}_{N_r}$, and \mathbf{I}_{N_r} is an $N_r \times N_r$ identity matrix.

Due to the transmission of only a single IQ signal and TX antenna switching, the received signal vector at the RX contains a combination of $N_u^{(i)}$ columns of the channel matrix that are multiplied by the same IQ symbol q_m . This is different from the conventional SMX-MIMO, in which the received signal typically contains a combination of multiple columns of the channel matrix that are multiplied by different IQ symbols.

In VGSM, as the name suggests, N_u is varied and there are 2^{N_t} possible activation patterns of the TX antennas. However, due to the transmission of an IQ stream, it is necessary to guarantee at least one TX antenna to be always activated, thus ruling out the all-OFF pattern. Therefore, practically there are only $2^{N_t} - 1$ usable activation patterns² of the TX antennas thus achieving a spatial data rate of $\log_2(2^{N_t} - 1)$ bpcu and a total data rate of $\log_2(2^{N_t} - 1) + \log_2 M$ bpcu. In comparison, in FGSM, there are $\binom{N_t}{N_u}$ number of possible activation patterns thus achieving a spatial data rate of $\log_2 \binom{N_t}{N_u}$ bpcu and a total data rate of $\log_2 \binom{N_t}{N_u} + \log_2 M$ bpcu. The data rate of FGSM is maximized when $N_u = \lfloor N_t/2 \rfloor$ (denoted as FGSM- $\lfloor N_t/2 \rfloor$) [26]. Since $2^{N_t} - 1$ is larger than $\binom{N_t}{N_u}$ for any N_u ($1 \leq N_u \leq$

N_t), VGSM makes more efficient use of the number of TX antennas than FGSM.

C. Optimal Detection

The optimal detector, which is based on the maximum likelihood (ML) principle as given in [36], can be written for VGSM as

$$\begin{aligned} (\hat{i}, \hat{m}) &= \arg \max_{i,m} p_{\mathbf{y}}(\mathbf{y} \mid s_{i,m}, \mathbf{H}_{\text{LOS}}) \\ &= \arg \min_{i,m} \|\mathbf{y} - q_m \mathbf{H}_{\text{LOS}} \rho_i \mathbf{e}_i\|, \end{aligned} \quad (7)$$

where $p_{\mathbf{y}}(\cdot)$ denotes the probability density function (PDF) of \mathbf{y} , and $\|\cdot\|$ denotes the Euclidean distance. It is worthwhile to note that one of the main advantages of index modulation schemes, such as SM and GSM, is the absence of inter-channel interference at the RX, which renders an ML decoding complexity that does not grow with the IQ constellation size [37], [38].

D. Orthogonal Channel Condition

In [26], it was found that the (column) orthogonal channel condition

$$\langle \mathbf{h}_k, \mathbf{h}_{k'} \rangle_{k \neq k'} = 0 \quad \forall k, k' \in \{1, \dots, N_t\} \quad (8)$$

can maximize the minimum Euclidean distance of the RX-side FGSM symbols thus optimizing the SEPs for LOS FGSM with an IQ constellation of M -QAM or M ($M \leq 8$)-PSK. It can be numerically verified that (8) can also optimize the SEP for LOS VGSM with M -QAM or M ($M \leq 8$)-PSK.

To achieve (8), the minimum ASP of the TX and RX ULAs $s_t s_r$ needs to be

$$s_t s_r \approx D \lambda \zeta / N_r, \quad (9)$$

where $\zeta \triangleq 1/(\cos \beta_t \cos \beta_r)$ is called the tilt factor which measures the degree of tilt of the ULAs. In addition we need $N_r \geq N_t \geq 2$, where $N_r \geq N_t$ is necessary for having N_t orthogonal sub-channels and $N_t = 2$ is the minimum number of TX antennas needed for SM/GSM. Note that condition (9) was derived in [26] under the assumptions of much smaller array lengths than the TX-RX distance, i.e., $L_t, L_r \ll D$ and not too large array tilts, i.e., $\beta_t, \beta_r \ll \pi/2$ (the broadside of the TX and RX arrays should face to each other).

Perfect orthogonality will be lost if the actual ASP differs from the optimal value. To measure this, we use an ASP deviation factor η that was previously defined in [39] as

$$\eta = \frac{\text{Minimum ASP for orthogonality}}{\text{Actual ASP}} = \frac{D \lambda \zeta / N_r}{s_t s_r}. \quad (10)$$

Note that the $s_t s_r$ in the denominator of (10) denotes the actual ASP and should not be confused with the optimal ASP (to achieve channel orthogonality) given in (9). If the actual ASP is larger than the optimal ASP, we have $\eta < 1$; if the actual ASP is smaller than the optimal ASP, we have $\eta > 1$. In what follows, we use η as an (inverted) normalized ASP.

It has to be noted that condition (8) and consequently (9), however, do not optimize the capacities of LOS FGSM and LOS VGSM. Therefore, in Section IV-B1 we will try to find the ASP that maximizes the capacity for LOS FGSM and LOS VGSM.

¹Note that when only the LOS component of the channel is considered, we can ignore the time delay and write the j, i -th element of \mathbf{H}_{LOS} equivalently into $h_{ji} \approx \exp(-jkd_{ji})$.

²The number of TX antenna activation patterns is in general not an integer power of two, thus resulting in fractional information bits being carried per spatial symbol. Fractional bit modulation can be achieved using the methods given in [33] [34] [35].

III. CAPACITY DERIVATION

In this section, we first consider only the dominant LOS component of the channel and derive the capacity of VGSM. In the capacity derivation, we assume no channel state information at TX (CSIT) but perfect channel state information at RX (CSIR). At the end, calculation of the capacity of VGSM in multipath channels will be discussed.

A. Pure LOS Channel

As a discrete constellation in the IQ domain is in general not optimal in maximizing the channel capacity, we rewrite the system model (6) using a continuous complex IQ signal s as

$$\mathbf{y} = \mathbf{H}_{\text{LOS}} s \mathbf{e}_i + \mathbf{n} = s \sum_{k \in \mathcal{I}_i} \mathbf{h}_k + \mathbf{n}. \quad (11)$$

Deriving the VGSM channel capacity involves finding the optimal signaling scheme for s and i , i.e., we should try to find the optimal PDF of s and the probability mass function (PMF) of i . For this reason, ρ_i , whose purpose in (6) was to optimize the VGSM constellation in case of a discrete IQ constellation, is dropped in (11).

The mutual information between the input and output of the system can be calculated, according to the chain rule for information [40], as

$$I(i, s; \mathbf{y}) = I(s; \mathbf{y}|i) + I(i; \mathbf{y}). \quad (12)$$

The channel capacity is found by maximizing $I(i, s; \mathbf{y})$ with respect to $p(i)$ (the PMF of i) and $p(s)$ (the PDF of s) as

$$C = \max_{p(i), p(s)} [I(s; \mathbf{y}|i) + I(i; \mathbf{y})] = C_1 + C_2, \quad (13)$$

where C_1 and C_2 are the rates achieved in the IQ and spatial domains, respectively. The system can be viewed as containing $2^{N_t} - 1$ continuous-input continuous-output IQ channels and one discrete-input continuous-output spatial channel. Due to antenna selection, one IQ channel is selected within each time slot for transmission of the IQ signal. Because the channel input contains both discrete and continuous signals, finding a solution for the optimal signaling of i and s is not easily possible. In this paper we use a simple signaling scheme that has been used in [26], [40] — zero-mean complex Gaussian (ZMCG) signaling in the IQ domain, i.e., $s \sim \mathcal{CN}(0, \sigma_s^2)$, and uniform signaling in the spatial domain, i.e., $p(i) = 1/(2^{N_t} - 1)$.

Theorem 1. *The channel capacity attained by VGSM, with ZMCG signaling in the IQ domain and uniform signaling in the spatial domain, over LOS channels is given by (14) on Page 6, where σ_s^2 is the variance of the Gaussian-distributed IQ signal,*

$$p(\mathbf{y}|i) = \frac{1}{|\pi \mathbf{\Gamma}_i|} \exp \left[-\mathbf{y}^H \mathbf{\Gamma}_i^{-1} \mathbf{y} \right], \quad (15)$$

and

$$\mathbf{\Gamma}_i = \sigma_s^2 \left[\sum_{k \in \mathcal{I}_i} \mathbf{h}_k \right] \left[\sum_{k \in \mathcal{I}_i} \mathbf{h}_k^H \right] + \sigma_n^2 \mathbf{I}_{N_r}. \quad (16)$$

Proof: Due to ZMCG signaling in the IQ domain and uniform signaling in the spatial domain, C_1 can be calculated as

$$C_1 = \mathbb{E}_i \left[\log_2 \left(1 + \frac{\sigma_s^2}{\sigma_n^2} \|\mathbf{H}_{\text{LOS}} \mathbf{e}_i\|^2 \right) \right], \quad (17)$$

where maximal ratio combining (MRC) [41, ch. 5] has been used in the IQ domain to maximize the received SNR. It is straightforward to further write (17) into the form in (14). For C_2 , we have

$$C_2 = H(i) - H(i|\mathbf{y}), \quad (18)$$

where $H(i) = 2^{N_t} - 1$. $H(i|\mathbf{y})$ can be derived under the same framework as that for FGSM in [26, Appendix B]. The derivation is omitted here due to space limit. Thus we have C_2 as given in (14). ■

To relate the capacity to SNR, we define the SNR γ as the average TX power to the RX noise power ratio, which is given by

$$\gamma = \frac{1}{2^{N_t} - 1} \sum_{i=1}^{2^{N_t}-1} N_u^{(i)} \cdot \frac{\sigma_s^2}{\sigma_n^2} = \frac{2^{N_t}}{2^{N_t} - 1} \cdot \frac{N_t}{2} \cdot \frac{\sigma_s^2}{\sigma_n^2}, \quad (19)$$

where we have used the binomial identity given in (39).

Corollary 2. *The channel capacity attained by VGSM, with ZMCG signaling in the IQ domain and uniform signaling in the spatial domain, over column-orthogonal LOS channels is given by (20) on the following page, where $\mathbb{E}_{\tilde{\mathbf{y}}|N_u}$ means taking the average over $\tilde{\mathbf{y}}$ given any i that satisfies $|\mathcal{I}_i| = N_u$,*

$$\tilde{\Psi}_{ii'} = \frac{\sum_{i'} \left[\exp \left[-\tilde{\mathbf{y}}^H \left(\tilde{\mathbf{\Gamma}}_{i'}^{-1} \right) \tilde{\mathbf{y}} \right] / |\pi \mathbf{\Gamma}_{i'}| \right]}{\exp \left[-\tilde{\mathbf{y}}^H \left(\tilde{\mathbf{\Gamma}}_i^{-1} \right) \tilde{\mathbf{y}} \right] / |\pi \mathbf{\Gamma}_i|}, \quad (21)$$

and

$$\tilde{\mathbf{\Gamma}}_i = N_r \sigma_s^2 \mathbf{e}_i \mathbf{e}_i^T + \sigma_n^2 \mathbf{I}_{N_t}. \quad (22)$$

Proof: Utilizing the column-orthogonal channel condition (8) and the fact that $\|\mathbf{h}_k\|^2 = N_r$, it is trivial to write C_1 of (14) into the form in (20). Also under condition (8), we can write C_2 of (14) into the form as in (20) based on the equivalent channel model given in [26, (28)]. The derivation is similar to that for FGSM [26, Section IV-B] and thus is omitted here. ■

In Section IV-B1, we will numerically show that the signaling, i.e., ZMCG signaling in the IQ domain and uniform signaling in the spatial domain, given in Corollary 2 for VGSM, is near-optimal in achieving the channel capacity in the case of orthogonal LOS channels.

In general, the C_2 in (14) and (20) cannot be further simplified and we have to resort to numerical methods, e.g., Monte Carlo averaging, for evaluation.

B. Multipath Channel

In the room channel model given in Fig. 3, there exists a delay spread of up to 50 ns between the LOS path and the second-order reflected paths depending on the positions of the ULAs. The multipath channel in general is time-dispersive and frequency-selective and the capacity has to be calculated

$$C = \underbrace{\frac{1}{2^{N_t} - 1} \sum_{i=1}^{2^{N_t} - 1} \left[\log_2 \left(1 + \frac{\sigma_s^2}{\sigma_n^2} \left\| \sum_{k \in \mathcal{I}_i} \mathbf{h}_k \right\|^2 \right) \right]}_{C_1} + \underbrace{\log_2(2^{N_t} - 1) - \frac{1}{2^{N_t} - 1} \sum_i \mathbb{E}_{\mathbf{y}|i} \left[\log_2 \frac{\sum_{i'} p(\mathbf{y}|i')}{p(\mathbf{y}|i)} \right]}_{C_2} \quad (14)$$

$$C = \underbrace{\frac{1}{2^{N_t} - 1} \sum_{N_u=1}^{N_t} \left[\binom{N_t}{N_u} \log_2 \left(1 + \frac{\sigma_s^2}{\sigma_n^2} N_u N_r \right) \right]}_{C_1} + \underbrace{\log_2(2^{N_t} - 1) - \frac{1}{2^{N_t} - 1} \sum_{N_u=1}^{N_t} \left\{ \binom{N_t}{N_u} \mathbb{E}_{\tilde{\mathbf{y}}|N_u} \left[\log_2 \tilde{\Psi}_{ii'} \right] \right\}}_{C_2} \quad (20)$$

block-wise over frequency³ [45, ch. 4.4], where the channel coefficients are constant over each block. The procedure for the calculation is as follows: 1) transform $\mathbf{H}(t)$ as given in (1) into the frequency domain to obtain $\mathbf{H}(f)$; 2) sample $\mathbf{H}(f)$ uniformly over a certain bandwidth and evaluate the capacity using (14) at each of the samples; and 3) take the average of the capacity over frequency.

The frequency-average capacity suggests an orthogonal frequency-division multiplexing (OFDM)-style waveform, i.e., each sub-carrier is treated as an individual VGSM channel with its own IQ stream and space modulation. Implementation of such an OFDM-style VGSM would require digital domain switching at baseband (as opposed to RF switching as shown in Section V-A) and N_t full TX chains from baseband to RF. This leads to a full-complexity MIMO RF front-end and the motivation to employ a simple scheme is lost. Therefore, the frequency-average capacity serves only as a theoretical reference.

IV. NUMERICAL RESULTS ON CAPACITY

In this section, we present numerical results to: 1) validate the capacity expressions derived in the previous section, and 2) compare the performance of VGSM with SMX-MIMO, BF-MIMO, and especially FGSM. Another major task is to numerically optimize the capacity of both VGSM and FGSM, since, as stated earlier, the channel orthogonality condition (8) is not optimal in maximizing the capacity of VGSM and FGSM. Furthermore, we study the impact of NLOS components of the channel on the capacity of VGSM.

A. System Setup

In the numerical studies, we assume a center frequency of 60 GHz, equal antenna separations at the TX and RX, i.e.,

³As SM/GSM schemes are usually applied in single carrier communications, it might appear natural to consider the time-domain representation of the channel instead of its frequency domain representation for computing the attainable capacity, preferably with the (optimal) ML sequence detection (MLSD) [42]. However, MLSD has very high processing complexity and the involvement of intersymbol interference (ISI) makes the problem intractable. Practically, one may use equalization schemes such as minimum mean square error (MMSE) [43], zero forcing (ZF) [43, (17) but with $N_0 = 0$], partial interference cancellation receiver with successive interference cancellation (PIC-R-SIC) [44], etc. However, MMSE and PIC-R-SIC equalized channels are not ISI-free and remain time-dispersive, and we are still unable to compute its GSM capacity directly in time domain. Only the ZF equalizer perfectly removes ISI and enables calculation of the capacity of the equalized channel. The performance of ZF equalizer, however, is known to be inferior to that of MMSE and PIC-R-SIC equalizers, let alone MLSD. As a result, we have to study the frequency-average capacity.

$s_t = s_r$, and parallel and aligned ULAs, i.e., $\zeta = 1$. This configuration actually represents the general case of arbitrarily placed ULAs with their antenna separations along the z -axis (see Fig. 2) being s_t, s_r . The reason is as follows. Let the antenna separations of an arbitrarily placed ULAs be s'_t, s'_r . From (9) and (10), it is apparent that the system with arbitrarily placed ULAs is equivalent to a system with parallel and aligned ULAs (i.e., $\eta = 1$), whose TX and RX antennas separations are $s'_t \cos \varphi_t$ and $s'_r \cos \varphi_r$, respectively. For VGSM, FGSM and BF-MIMO, all of which use a single IQ modulation chain, we assume the same numbers of TX and RX antennas. For FGSM, we only consider the case that achieves the maximum spatial rate, i.e., $N_u = \lfloor N_t/2 \rfloor$ (denoted as FGSM- $\lfloor N_t/2 \rfloor$). For SMX-MIMO, we consider a case that two data streams are transmitted using two TX antennas (denoted as SMX-2), and the same number of RX antennas as the other schemes under consideration. For SMX-2, VGSM and FGSM- $\lfloor N_t/2 \rfloor$ they all transmit two data streams. However, SMX-2 needs two TX and RX RF chains, while VGSM and FGSM- $\lfloor N_t/2 \rfloor$ need only a single TX RF chain and possibly a single RX RF chain. Actually, a major motivation for SM/GSM communications lies in the fact that we can use a large number of TX antennas efficiently by using only a single TX RF chain [1].

It is worthwhile to note that there is a major difference between the physical dimensions of BF-MIMO and the other schemes we consider. In BF-MIMO, the TX and RX antennas are usually densely packed with separations down to $\lambda/2$ [46], [47]. However, to maximize the spatial degrees of freedom, the optimal strategy is proven in [17] to be the channel orthogonality condition (8). Therefore, to attain the maximum channel capacity, SMX-MIMO needs the antenna separations as given in (9), which are in general much larger than $\lambda/2$. As stated in Section II-D, (8) and consequently (9) are needed to optimize the SEPs of both VGSM and FGSM. According to the best of our knowledge, the optimal ASP to maximize the capacity for VGSM and FGSM has not been reported.

B. Pure LOS Capacity

1) *Orthogonal Channels*: Fig. 4 plots the capacities of VGSM, FGSM- $\lfloor N_t/2 \rfloor$, SMX-MIMO (all for orthogonal LOS channels), and BF-MIMO (for LOS channels with densely packed TX and RX antennas). Similar figures but without VGSM have been shown and analyzed in [26]. In what follows, we describe our novel findings related to VGSM.

VGSM attains higher capacity than FGSM- $\lfloor N_t/2 \rfloor$, and the gap is larger for larger N_t . To prove this, we plot C_1 of VGSM

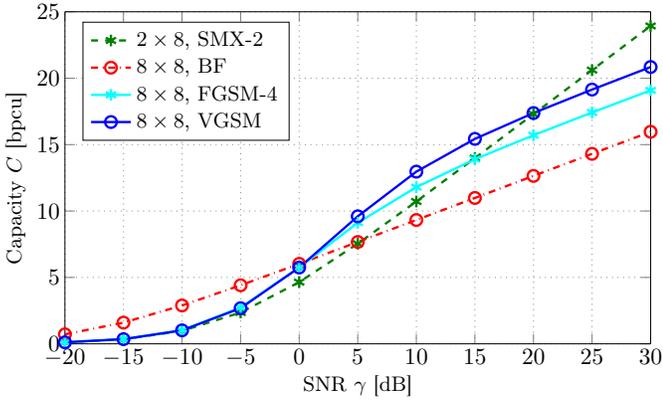


Fig. 4. Capacity comparison of VGSM, FGSM- $\lfloor N_t/2 \rfloor$, SMX-MIMO and BF-MIMO. In the legend, the size of MIMO is indicated in the format $N_t \times N_r$. The curve for VGSM is plotted using (20). And the curves for other schemes are plotted using the same expressions that were used to plot [26, Fig. 8].

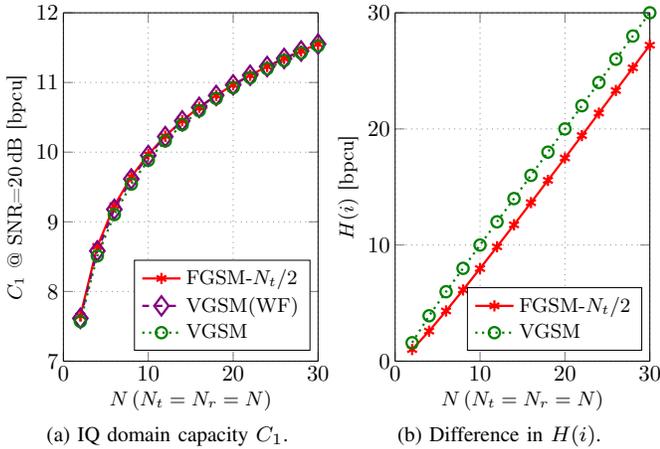


Fig. 5. Achievable rate: VGSM vs. FGSM- $\lfloor N_t/2 \rfloor$.

and FGSM- $\lfloor N_t/2 \rfloor$ for $N_t = 2$ to 30 (and $N_r = N_t$) at SNR $\gamma = 20$ dB in Fig. 5(a). Due to the fluctuations in the transmit power caused by the activation of a variable number of TX antennas, VGSM achieves smaller C_1 than FGSM- $\lfloor N_t/2 \rfloor$. However, the gap is negligible (approximately 0.14 bpcu). In Fig. 5(b), we plot $\log_2(2^{N_t}) - \log_2(2^{N_t} - 1)$, which is the difference between the upper bound of C_2 of VGSM, i.e., $H(i)$, and that of FGSM- $\lfloor N_t/2 \rfloor$. We see that the gap increases from 0.5 bpcu towards 3 bpcu with increasing N_t . Since it is evident from Fig. 4 that $H(i)$ can be achieved at high SNR, the claim now becomes obvious.

VGSM and FGSM- $\lfloor N_t/2 \rfloor$ can surpass BF-MIMO and SMX-2 MIMO at moderate and high SNR, and the gap is larger with a larger number of antennas. At high SNR, the capacity of VGSM has linear growth with respect to N_t (note that for GSM we assume $N_r \geq N_t$). The capacity of FGSM- $\lfloor N_t/2 \rfloor$ is lower than that of VGSM but the gap is within 0.5 – 3 bpcu. However, the capacity for BF-MIMO and SMX-2 MIMO, which are $\log_2(1 + N_t N_r \cdot \text{SNR})$ and $2 \log_2(1 + N_r/2 \cdot \text{SNR})$ respectively, only have logarithmic growth.

At high SNR, where the full spatial rate $\log_2(2^{N_t} - 1)$ is attained, the signaling i.e., ZMCG signaling in the IQ domain

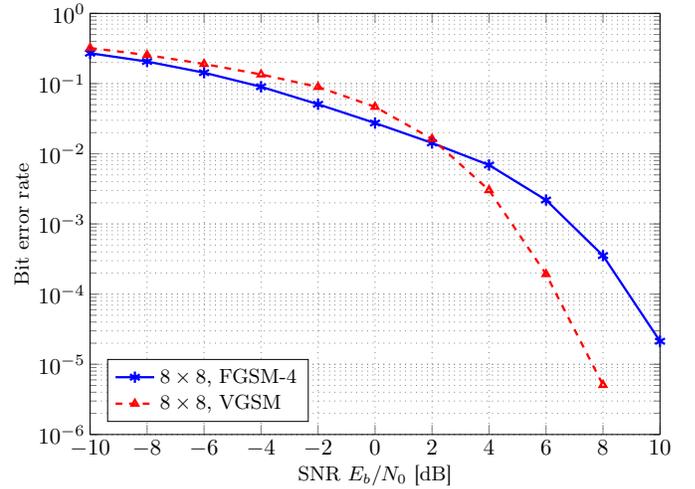


Fig. 6. Simulated bit error rate (BER): VGSM vs FGSM- $\lfloor N_t/2 \rfloor$. An 8×8 ($N_t \times N_r$) MIMO with orthogonal setup (i.e., $\eta = 1$) is assumed and the modulation rate is $R = 10$ bpcu.

and uniform signaling in the spatial domain, is near-optimal in attaining the capacity of VGSM over orthogonal LOS channels. At high SNR, the sub-optimality of the signaling lies in the use of a constant IQ signal variance σ_s^2 in the IQ domain for varying N_u . If only the IQ capacity is concerned, the optimal signaling for s is still ZMCG, but the variance has to be changed with respect to N_u according to water-filling [41]. It is shown in Fig. 5(a) that the improvement in C_1 due to water-filling, however, is marginal.

VGSM has shown to have an advantage over FGSM- $\lfloor N_t/2 \rfloor$ in terms of channel capacity. To alleviate the concern if the advantage can be translated into a gain in bit error rate (BER), especially when fractional bit modulation is involved. We further performed BER simulations using the fractional bit modulation method given in [35]⁴. And it is confirmed that VGSM indeed has an advantage over FGSM- $\lfloor N_t/2 \rfloor$ in BER. As an example, we plot the simulation results with an 8×8 ($N_t \times N_r$) orthogonal MIMO setup in Fig. 6. It is shown that although 8×8 VGSM is slightly worse than 8×8 FGSM-4 at low SNR in BER, it shows a clear advantage of about 3 dB at the higher SNR of around 8 dB.

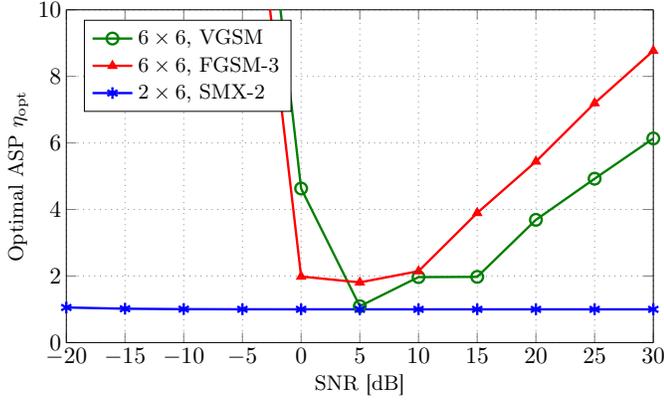
2) *Capacity-Optimized Channel*: We formulate the optimization problem as follows

$$C_{\text{opt}} = \max_{\eta} C, \quad (23)$$

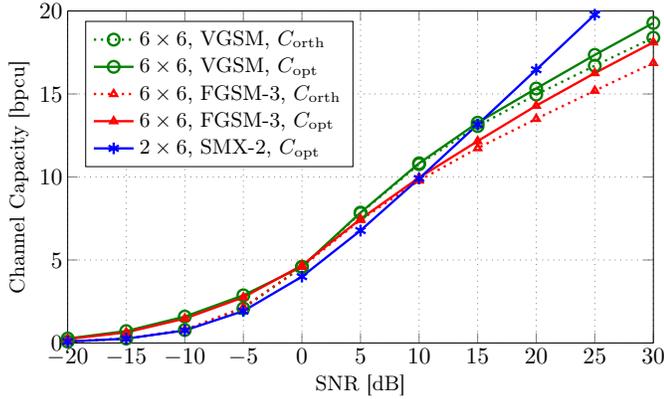
where η (inverted ASP) is the parameter to vary. The expressions for the capacity of VGSM, FGSM and SMX are respectively given by (14), [26, (30)] and [26, (46)]. A set of numerical results for the optimized capacity and ASP is shown in Fig. 7. For SMX-2, the results simply confirm that the channel orthogonality condition (8) is needed to maximize the spatial degrees of freedom [17]. In what follows, we report the novel findings concerning VGSM and FGSM- $\lfloor N_t/2 \rfloor$.

At low SNR, the beamforming gain is more important. This is already evident in Fig. 4, where the single-stream BF-MIMO

⁴We prefer the method of [35] because it eliminates the error propagation issue of [33] and the bit sliding issue of [34].



(a) Optimal ASP

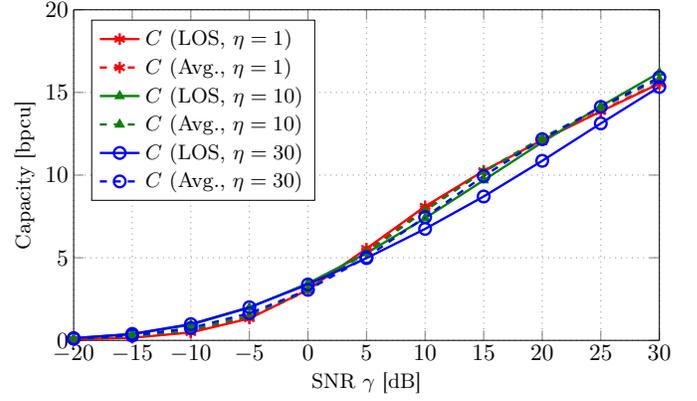


(b) Channel capacity

Fig. 7. Channel capacity optimization for VGSM, FGSM- $\lfloor N_t/2 \rfloor$ and SMX-2.

achieves higher capacity than VGSM, FGSM- $\lfloor N_t/2 \rfloor$ and SMX-2 at low SNR. In Fig. 7(a), for VGSM and FGSM- $\lfloor N_t/2 \rfloor$ the optimization algorithm effectively “recognized” the fact that, at low SNR, the antenna separations should be made small, i.e., η should be made large, such that beamforming gains can be exploited. However, for SMX-2, the optimization algorithm converged to $\eta = 1$ instead of a large η . This is due to the fact that without CSIT, SMX cannot exploit TX array gain. At moderate SNR, the channel orthogonality condition (8) tends to be needed in order to maximize the capacities of VGSM and FGSM- $\lfloor N_t/2 \rfloor$. This is evident from the convex shape of η_{opt} for both VGSM and FGSM- $\lfloor N_t/2 \rfloor$ in Fig. 7(a). At high SNR, for VGSM and FGSM- $\lfloor N_t/2 \rfloor$, η_{opt} increases with increasing SNR, which means that higher capacity is attained even with smaller ASP than which is given by (9). This is because at high SNR, C_2 can reach its upper limit, i.e., $\log_2(2^{N_t} - 1)$, even with a large η . Meanwhile, a large η means higher beamforming gain and thus a large C_1 . As a result, $C = C_1 + C_2$ is maximized at high SNR with large η . The above finding of a large η_{opt} , i.e., a small optimal ASP, for maximum capacity at both low and high SNR leads to a small physical size of the TX and RX ULAs, and therefore is of great practical importance.

The optimal ASP varies as a function of SNR in Fig. 7(a). In practical systems, however, ASP is usually fixed by design and the SNR varies depending on, e.g., the TX-RX distance.

Fig. 8. Capacity of VGSM: multipath channel vs pure LOS channel in a 4×4 system.

The optimal ASP may not always be realizable in practical scenarios. Taking into consideration the fact shown in Fig. 7(b) that only a small improvement of C_{opt} as compared to C_{orth} (the capacity corresponding to orthogonal channels) at low and high SNRs can be seen for VGSM and FGSM- $\lfloor N_t/2 \rfloor$, and the findings from Fig. 4, we make the following conclusions:

- GSM and SMX MIMO schemes are suitable to be operated at moderate and high SNRs, while in the low SNR regime, BF MIMO should be deployed;
- For GSM MIMO, it is beneficial to drive the RX into the moderate and high SNR regimes, and the system benefits from a marginally increased capacity but significantly reduced ASP.

C. Impact of NLOS Components on Capacity

So far, only the LOS component of the channel is considered in the capacity analysis. In this subsection, we investigate the impact of the NLOS components on the capacity of VGSM.

The system setup is as follows. We consider a 4×4 system and the indoor propagation environment as shown in Fig. 3. Both the TX and RX ULAs are placed horizontally and parallel to the x -axis. The midpoint of the TX ULA is fixed at $(2.5, 0.1, 1)$ and the midpoint of the RX ULA is (uniformly) randomly placed in a 0.25×0.25 square that is centered at $(2.5, 4.65, 1.5)$. The capacity is averaged both over frequency and over space. Over frequency, the frequency response of the multipath channel is sampled with a sampling interval of 2 MHz over a bandwidth of 1 GHz around 60 GHz. Over space, ten random locations of the RX ULA are generated. The capacity is averaged first over frequency according to the procedure given in Section III-B and then over the ten locations. Three ASP values ($\eta = 1, 10, 30$) are used and the results are shown in Fig. 8.

At low SNR, BF gain is more important [26] and that is why $\eta = 30$ corresponds to the best capacity for $\gamma < 0$ dB. In what follows, we only consider the moderate and high SNR region where $\gamma > 5$ dB. The pure LOS capacity (in solid curves) shows the trend that has been discovered in Section IV-B2. That is, $\eta = 1$ corresponds to the maximized capacity for VGSM at moderate SNR and an appropriately reduced ASP

(or increased η of, e.g., 10) achieves the maximum capacity at high SNR. However, an overly reduced ASP (or increased η of e.g., 30) leads to compromised capacity at moderate SNR. For $\eta = 1$ and $\eta = 10$, the multipath channel shows no clear improvement in capacity as compared to the pure LOS channel.

For $\eta = 30$, the NLOS components seem to remarkably improve the capacity as compared to the pure LOS case. A further investigation showed that this is because the NLOS components are able to randomize the channel and improve the average spatial capacity (but not the average IQ capacity). This finding is surprising because it was not expected that the NLOS components, which are weak in the received power, could notably improve the spatial domain capacity. Recall that in the multipath channel model, the antennas are assumed to have a hemispheric radiation pattern. In practice, however, more focused antennas with moderate gains of at least a few dBi may be needed in order to compensate for the high path loss at mmWave frequencies and enable GSM communication at high SNR [26]. In that case, the NLOS rays may be attenuated by antennas and the impact of the NLOS components may be reduced. Another practical issue, which was stated in the end of Section III-B, is that it might not be desirable to implement a full-complexity TX and RX in order to handle the time dispersion (or frequency selectivity) and exploit the NLOS components.

To conclude, when good LOS capacity (in regard to C_{opt}) is achieved as a result of sufficient antenna separation, the contribution of NLOS components to VGSM capacity is marginal. However, when the LOS capacity is poor (in regard to C_{opt}) due to insufficient antenna separation, NLOS components might introduce enough randomization to the channel and significantly improve the VGSM capacity. Nevertheless, the improved VGSM capacity in the latter case does not seem to exceed the capacity of the former case.

V. PROPOSED TX ARCHITECTURE

In this section, we propose a practical switching structure for the implementation of VGSM at mmWave frequencies. The feasibility of VGSM remained an un-touched topic and this is a first attempt to address this issue, particularly for mmWave frequencies. Practicability and power efficiency are the major concerns that will be addressed in what follows. In addition, a power allocation factor will be considered and used to optimize the VGSM constellation. Like in any other schemes, e.g., QAM, an optimized constellation is the basis to achieve good system performance.

A. Switching Structure

The variable N_u -to- N_t switching at symbol rate with high power efficiency needed by VGSM poses a major challenge in implementation, especially at mmWave frequencies. The switching structure in Fig. 9 solved the problem of fixed N_u -to- N_t switching (please refer to [26] for its working principle). However, as will be shown in Section V-C, it inherently suffers from low power efficiency due to the insertion loss introduced by the switches and routing section. Furthermore, in Fig. 9, each PA has to drive a fixed number of N_u/N_p TX antennas in order

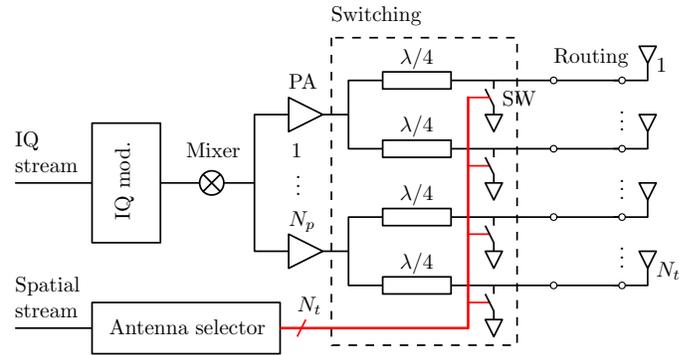


Fig. 9. TX architecture for FGSM [26, Fig. 11(a)].

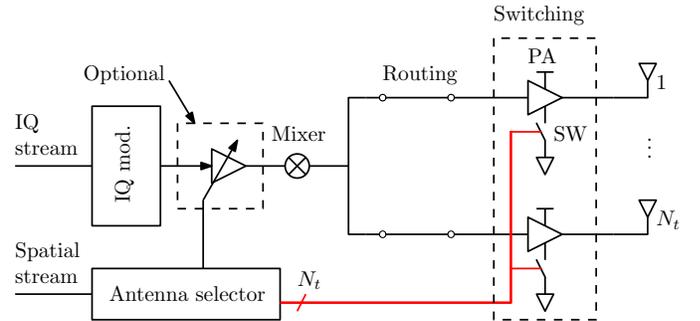


Fig. 10. Proposed TX architecture for VGSM.

to achieve impedance matching. Impedance matching would be totally destroyed if variable N_u switching were intended to be applied. Therefore, the structure in Fig. 9 is not suitable for VGSM.

To address the challenges, as shown in Fig. 10, we propose a novel switching structure for VGSM. The proposed TX supports transmission of a single IQ stream and a spatial stream. The modulated IQ signal is split into N_t branches after up-conversion to RF. Switching and power amplification are combined in a single stage using the so-called switching PAs [48]–[51]. The operational principle of the switching PAs is illustratively shown in the figure — the DC power supply of the PAs can be switched OFF. From a signal point of view, a switching PA functions the same as a series combination of a single-pole single-throw (SPST) switch and a conventional power amplifier (PA). However, from a circuit point of view, it further has the following two advantages: 1) since the PAs are placed after the power splitter, impedance matching can be achieved by making the input impedance of the switching PAs fixed over ON and OFF states and matched at the junction; and 2) when switched OFF, a switching PA does not consume any DC power thus providing improved power efficiency. Variable N_u switching is achieved by turning ON/OFF each branch individually. All switching patterns of the TX antennas are thus supported. The power allocation factor ρ_i , which will be optimized in the next subsection, is modeled as a variable gain stage as shown in the dashed box between the IQ modulator and the mixer. This switching mechanism of the proposed TX for VGSM is rather different from the TX for FGSM shown in Fig. 9. In the latter, switching OFF/ON of a branch

is achieved by short-circuiting/un-short-circuiting the branch. For further details of latter TX, the reader is referred to [26]. In comparison, the proposed TX is expected to achieve higher data rates at the same DC power consumption due to 1) the higher data rate achieved by VGSM, and 2) placement of the routing sections before the PAs. We will present a quantitative comparison in Section V-D based on numerical results.

Since the antennas of the TX and RX ULAs need to be sufficiently separated in order to achieve optimized SEP and capacity, a routing section is necessary at each branch. We propose to place the (switching) PAs as close as possible to the feed of the TX antennas so as to reduce the routing loss, which is much higher at mmWave frequencies [52], [53]. For example, as shown in the two references, the routing loss at 50 – 60 GHz can be 0.017 – 0.04 dB/mm depending on the substrate technology. According to (9), to make an 8×8 MIMO channel (with $\zeta = 1$, $s_t = s_r$ and $D = 1$ m) orthogonal in order to optimize the SEP, the TX array length needs to be $L_t = 175$ mm which translates into a maximum routing loss of 1.5 – 3.5 dB. And the routing loss will increase for larger D , which for room communications can be a few meters. Thanks to our finding that was presented in Section IV-B2, for maximized capacity we do not need to make the channel orthogonal. However, even though, the routing loss will be noticeable.

B. Power Allocation Optimization

The purpose of power allocation optimization is to optimize the VGSM constellation, which is needed to have a good error performance. More precisely, we seek to maximize the minimum Euclidean distance of the VGSM symbols, since at high SNR the error probability is determined by the minimum symbol distances [54]. The optimization problem is formulated as

$$\underset{\{\rho_i\}}{\text{maximize}} \quad \min_{\substack{i,i',m,m' \\ (i,m) \neq (i',m')}} \underbrace{\|q_m \rho_i \mathbf{e}_i - q_{m'} \rho_{i'} \mathbf{e}_{i'}\|^2}_{\triangleq \tilde{\Phi}} \quad (24a)$$

$$\text{subject to} \quad \rho_i \geq 0 \quad \forall i = 1, \dots, 2^{N_t} - 1, \quad \text{and} \quad (24b)$$

$$\sum_i \sum_m \| \rho_i q_m \mathbf{e}_i \|^2 = P \quad (24c)$$

where $\tilde{\Phi}$ is a simplified notation of $\tilde{\Phi}_{i,m,i',m'}$ denoting the pairwise Euclidean distance between the TX-side VGSM symbols (i, m) and (i', m') . Constraint (24b) is to allow only positive power allocations, and (24c) gives the total power budget of the VGSM constellation. Deriving a closed-form (optimal) solution to such kind of problem is challenging and remains an open issue [55]. Therefore, in this paper we do not pursue the optimal solution to (24) but instead try to find possible sub-optimal solutions analytically. The main findings are summarized below in Theorem 3.

Theorem 3. Consider the VGSM constellation \mathbb{S} defined in (6). An optimized, though not necessarily optimal, power allocation strategy, in the sense of maximizing the minimum Euclidean distance of the VGSM symbols, is given as follows:

1) When the IQ constellation is square/rectangular M ($M \geq 4$)-QAM or M ($2 \leq M \leq 6$)-PSK, the VGSM symbol power should be allocated, for $i = 1, \dots, 2^{N_t} - 1$, according to

$$\rho_i = \sqrt{\frac{P}{N_t 2^{N_t-1} \sum_m |q_m|^2}}; \quad (25)$$

2) When the IQ constellation is M ($M > 6$)-PSK, the VGSM symbol power should be allocated, for $i = 1, \dots, 2^{N_t} - 1$, according to

$$\rho_i = \begin{cases} k_* \sqrt{1/N_u^{(i)}} & \text{for } N_u^{(i)} \leq N_u^* \\ 2k_* \sin\left(\frac{\pi}{M}\right) & \text{for } N_u^{(i)} > N_u^* \end{cases}, \quad (26)$$

where

$$N_u^* \triangleq \left\lceil \left[4 \sin^2\left(\frac{\pi}{M}\right) \right]^{-1} \right\rceil,$$

$$k_* \triangleq \sqrt{\frac{P}{\left[|\mathcal{I}_1| + 4 \sin^2\left(\frac{\pi}{M}\right) \sum_{i \in \mathcal{I}_2} N_u^{(i)} \right] \sum_m |q_m|^2}},$$

$$\mathcal{I}_1 \triangleq \left\{ i \mid \forall i = 1, \dots, 2^{N_t} - 1, \text{ and } N_u^{(i)} \leq N_u^* \right\},$$

$$\mathcal{I}_2 \triangleq \left\{ i \mid \forall i = 1, \dots, 2^{N_t} - 1, \text{ and } N_u^{(i)} > N_u^* \right\}.$$

Proof: The proof is provided in Appendix A. ■

To demonstrate the effect of the power allocation scaling, we apply the scaling strategies given in (25) and (26) to VGSM with different IQ constellations (8-QAM, 8-PSK, 16-QAM, and 16-PSK) and various MIMO sizes (N_t, N_r varying from 2 to 10) and the results are plotted in Fig. 11. Note that, except for the lower bounds (in magenta), which are calculated from closed form expressions, all the other curves are numerical results obtained by exhaustive search. It is evident that the strategies given in Theorem 3, depending on the MIMO size and IQ constellation, can indeed significantly improve the minimum Euclidean distance of the VGSM constellation. The derivation in Appendix A is validated by the fact that the theoretical lower bounds given in (41) and (47) indeed match the accurate numerical curves in the figure. Additionally, we see that, when optimized (by applying Theorem 3), 16-PSK shows a clear advantage over 16-QAM in terms of the minimum Euclidean distance of the VGSM symbols $\tilde{\Phi}_{\min}$, e.g., for the case with $N_t = N_r = 10$, the $\tilde{\Phi}_{\min}$ of VGSM with 16-PSK is 4 dB larger than that of VGSM with 16-QAM. This finding is in accordance with the conclusion of [56], which was made for single-antenna activating SM, that: “In many cases multi-ring star-QAM is a suitable constellation for SM. However, when the number of transmit antennas is large, the numerically derived (optimized) constellation converges to the conventional PSK constellation.”

Knowing Theorem 3, it is immediately clear that in the proposed TX architecture in Fig. 10, for square/rectangular M ($M \geq 4$)-QAM and M ($2 \leq M \leq 6$)-PSK constellations the gain control stage depicted in the dashed box is not necessary and can be omitted. However, it is needed in other cases of an M -PSK constellation in which non-constant ρ_i 's are needed.

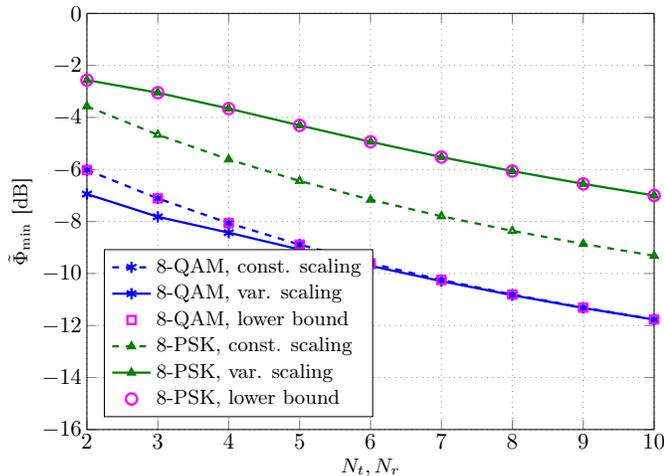
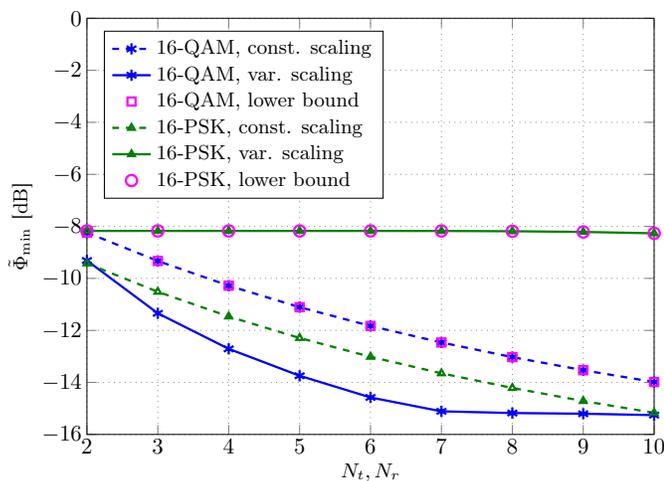
(a) Size of IQ constellation: $M = 8$ (b) Size of IQ constellation: $M = 16$

Fig. 11. Effect of power allocation scaling on the minimum Euclidean distance of the VGSM symbols Φ_{\min} . It is assumed that $N_t = N_r$ and, for unit average VGSM symbol power, $P = 2^M(2^{N_t} - 1)$. In the legend, “const. scaling” and “var. scaling” refer to the scaling strategies given in (25) and (26), respectively, and “lower bound” refers to the lower bounds given by (41) and (47) for “const. scaling” and “var. scaling”, respectively.

C. Power-Added Efficiency

We are interested in the efficiency of the section of the RF front-end between the mixer and the TX antennas. Its power added efficiency (PAE) is defined as

$$\eta_{\text{TX}} = \frac{\text{Total transmitted power} - \text{Input power}}{\text{Total supplied DC power}}. \quad (27)$$

The PAE of PAs, i.e.,

$$\eta_{\text{PA}} = \frac{\text{Output power} - \text{Input power}}{\text{Supplied DC power}}, \quad (28)$$

influences the power efficiency of the front-end and thus will be contained in η_{TX} .

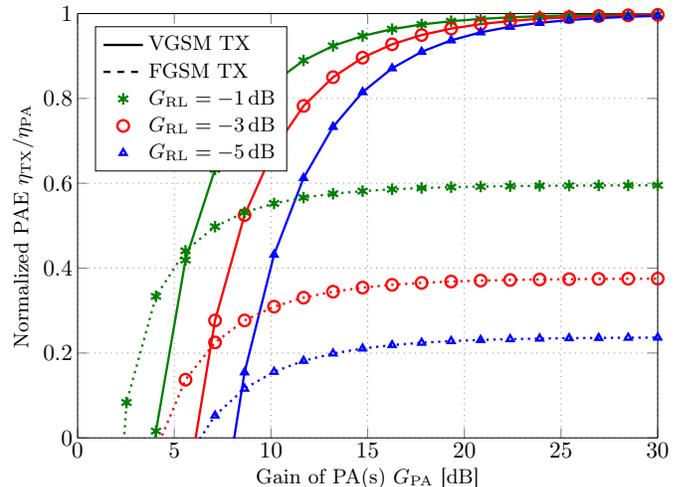


Fig. 12. PAE of the VGSM TX in Fig. 10 and the FGSM TX in Fig. 9.

Lemma 4. The PAE for the TX in Fig. 10 is given by

$$\eta_{\text{TX}} = \eta_{\text{PA}} \left(1 - \frac{\epsilon/G_{\text{RL}} - 1}{G_{\text{PA}} - 1} \right) \quad (29a)$$

$$\geq \eta_{\text{PA}} \left(1 - \frac{2.4/G_{\text{RL}} - 1}{G_{\text{PA}} - 1} \right), \quad (29b)$$

where η_{PA} is the PAE of the switching PAs in their ON state, G_{RL} is the routing loss of the RF front-end from the mixer to the TX antennas, G_{PA} is the ON-gain of the switching PAs, $\epsilon = N_t^2/(2^{N_t} - 1) {}_3F_2(1, 1, 1 - N_t; 2, 2; -1)$ is a parameter that accounts for the power waste due to the use of the fixed 1-to- N_t power splitter and ${}_3F_2(\dots; \dots; \cdot)$ is the generalized hyper-geometric function [57]; The PAE for the TX in Fig. 9 is given by

$$\eta_{\text{TX}} = \eta_{\text{PA}} \frac{G_{\text{RL}} G_{\text{SW}} G_{\text{PA}} - 1}{G_{\text{PA}} - 1}. \quad (30)$$

where G_{RL} is the routing loss from the PA to the feed of the TX antennas, G_{SW} is the insertion loss of the switch.

Proof: The derivation of (29a) is given in Appendix B. A plot of ϵ over $2 \leq N_t \leq 200$ shows that it increases from 1.67 to 2.4 over $2 \leq N_t \leq 6$ but decreases asymptotically to 2 over $6 \leq N_t \leq 200$. Thus we have (29b) as a lower bound for (29a) which does not have dependence on N_t . The derivation for (30) is similar to the derivation for (29a) and is omitted here. ■

Equation (29b) indicates that the PAE of the RF front-end of the proposed TX for VGSM can reach its maximum of η_{PA} when the PAs are operating in a high gain mode (such that $(2.4/G_{\text{RL}} - 1)/(G_{\text{PA}} - 1) \approx 0$). This is in contrast to the TX for FGSM, in which the PAE as given in (30), regardless of the PAs' gain, can never reach η_{PA} as it is limited by the switching and routing losses. Therefore, the proposed TX for VGSM inherently has a better PAE performance than the TX for FGSM. In the following, we give a quantitative analysis with numerical results.

The PAE, normalized by η_{PA} , is plotted in Fig. 12 for the two TXs. For the VGSM TX, the lower bound PAE given

TABLE I
LINK BUDGET PARAMETERS[†]

	TX #1 (Fig. 10)	TX #2 (Fig. 9)
Center frequency	GHz	60
Bandwidth	GHz	1
PA output power [47]	dBm	$P_{PA} = 11$
Switch insertion loss [59], [60]	dB	–
Routing loss (0.04 dB/mm [52])	dB	$3.5/N_p$
TX antenna gain [61]	dB	7
FSPL	dB	$G_{path} = 20 \log_{10} D + 68$
RX antenna gain [61]	dB	7
Noise power (290 K)	dBm	–84
RX noise figure [62], [63]	dB	6
SNR (TX #1)	dB	$4P_{PA} + G_{path} + 92$
SNR (TX #2)	dB	$N_p P_{PA} - 3.5/N_p + G_{path} + 89$

[†]Both systems are assumed to be 8×8 and with parallel and aligned ULAs, i.e., $\zeta = 1$. The TX and RX antenna separations are $s_t = s_r = 2.5$ cm, which achieve channel orthogonality at $D = 1$ m.

in (29b) is used. For the FGSM TX, the exact PAE given in (30) is used and we have assumed an optimistic switching loss of $G_{SW} = 1.25$ dB, which was achieved in [58] for a 40 – 70 GHz single-pole double-throw (SPDT) switch. For multi-throw switches, the insert loss is generally larger. It is shown that a higher PA gain leads to a higher PAE for both TXs. It is also shown that there is a cross point, below which the corresponding PA gain is lower and the PAE of the FGSM TX is larger. However, we are more concerned with the PAE in the high power regime. In the region above the cross point where a larger PA gain is used, the VGSM TX shows a much higher PAE. Furthermore, considering that a multi-dB routing loss is very likely due to the antenna separations needed to achieve high channel capacity, the PAE performance of the FGSM TX may be highly limited and much worse than that of the VGSM TX.

D. Capacity Achieved

The main purpose of this subsection is to demonstrate the capacity that is attained by the proposed TX for VGSM. In addition, we show the capacity attained by the FGSM TX for comparison. Since the proposed VGSM TX is intended for single-carrier transmission and its operation relies on the dominant LOS component, here we focus on the pure LOS capacity. The link budget parameters are listed in Table I and the attained capacities are plotted in Fig. 13. We see that with an 8×8 MIMO system, which has a TX and RX array length of 17.5 cm, more than 20 bpcu channel capacity can be achieved by VGSM within a distance of 1 – 5 m and more than 17 bpcu can be achieved within 5 – 10 m. Although eight PAs are used by the VGSM TX, due to switching, on average there are only four of them being activated. For FGSM-4 with four PAs (i.e., $N_p = 4$) and the same per-PA output power (i.e., $P_{PA} = 11$ dBm), about 3 bpcu less capacity is attained within 1 – 5 m, and over 2 bpcu less capacity is attained within 5 – 10 m. The performance of the FGSM TX seems to be limited by the insertion loss introduced by switching and routing. This

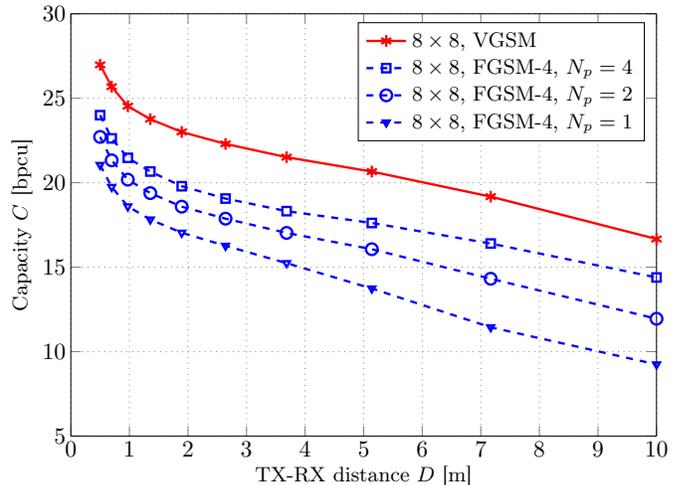


Fig. 13. Comparison of the capacities attained by the TX for VGSM in Fig. 10 and the TX for FGSM in Fig. 9.

is evident from the fact that increasing N_p of the FGSM TX increases the capacity attained. However, for fixed total transmit power, increasing N_p of the FGSM TX would decrease the capacity attained [26]. In summary, VGSM with the proposed TX in Fig. 10 can achieve remarkably higher capacity than FGSM with the TX in Fig. 9.

VI. CONCLUSIONS

In this paper, we studied VGSM for indoor LOS mmWave communications. VGSM is a known scheme but only for low-GHz fading channels. We focused on its application to indoor mmWave broadband communication and the capacity optimization of VGSM in LOS. The most important findings of this paper are summarized as follows:

- VGSM can achieve higher capacity than FGSM by making more efficient use of the number of TX antennas. For example, at moderate SNR, 8×8 VGSM (with a single TX chain) achieves higher capacity than SMX-MIMO (with two TX chains) and much higher capacity than BF-MIMO (with a single TX chain). The simplicity and good capacity exhibited makes VGSM a very attractive scheme for indoor LOS mmWave communication.
- We found that the optimal ASP to maximize the capacities of VGSM and FGSM in LOS is significantly smaller than the ASP needed to achieve channel orthogonality, which was found in [26] to be needed to maximize the SEP of SM/GSM in general. The findings of capacity optimization are necessary supplements to the findings of SEP optimization given in [26] and enriches our understanding about GSM in LOS. More importantly, a small ASP suggests compact system setups and makes the schemes more practical.
- We proposed and optimized a novel TX solution for the implementation of VGSM at mmWave frequencies. This is a first attempt to address the practicability issue of VGSM, which, in the context of mmWave communications, is of particular importance. The TX for VGSM proposed in

this work shows inherently much higher PAE than the TX for FGSM proposed in [26].

- Furthermore, the impact of the NLOS components on capacity was found to be marginal when the antennas are sufficiently separated and good LOS capacity is achieved. The finding that the NLOS components may significantly improve the spatial capacity for under-separated systems shows both opportunity and practical challenges, and this may trigger research into the exploitation of the NLOS components in indoor mmWave GSM communications.

We have shown that VGSM is an attractive scheme for indoor LOS mmWave communication. The results are promising and motivate future work on performance evaluation using real-world measured channel models which include the antenna characteristics.

APPENDIX A PROOF OF THEOREM 3

We start by finding the minimum Euclidean distance of the VGSM symbol, i.e.,

$$\tilde{\Phi}_{\min} = \min_{\substack{i, i', m, m' \\ (i, m) \neq (i', m')}} \tilde{\Phi}(i, i', m, m'), \quad (31)$$

where $\tilde{\Phi}$ is the pairwise Euclidean distance between the TX-side VGSM symbols defined in (24a). We assume, without loss of generality, that $N_u^{(i)} \geq N_u^{(i')}$, and break the condition $(i, m) \neq (i', m')$ into two cases: 1) For $i = i'$ (and $m \neq m'$), we have

$$\tilde{\Phi} = \|(q_m - q_{m'}) \rho_i \mathbf{e}_i\|^2 = d_{mm'}^2 \rho_i^2 N_u^{(i)} \geq d_{\min}^2 \rho_i^2 N_u^{(i)}, \quad (32)$$

where $N_u^{(i)} = |\mathbf{e}_i|_1$ (with $|\mathbf{e}_i|_1$ being the ℓ^1 -norm of \mathbf{e}_i) is the number of activated antennas corresponding to \mathbf{e}_i , and $d_{mm'} \triangleq |q_m - q_{m'}|^2$ is the Euclidean distance between IQ symbols q_m and $q_{m'}$, and $d_{\min} \triangleq \min_{m, m' (m \neq m')} d_{mm'}$ is the minimum IQ symbol distance. 2) For $i \neq i'$ ($m = m'$ or $m \neq m'$), we have

$$\begin{aligned} \tilde{\Phi} &= r_m^2 \rho_i^2 N_u^{(i)} + r_{m'}^2 \rho_{i'}^2 N_u^{(i')} - 2\Re\{q_m^* q_{m'}\} \rho_i \rho_{i'} \mathbf{e}_i^T \mathbf{e}_{i'} \\ &\geq r_m^2 \rho_i^2 N_u^{(i)} + r_{m'}^2 \rho_{i'}^2 N_u^{(i')} - 2r_m r_{m'} \rho_i \rho_{i'} \mathbf{e}_i^T \mathbf{e}_{i'}, \end{aligned} \quad (33)$$

where $r_m = |q_m|$. To reach the last step we have used the fact that $\Re\{q_m^* q_{m'}\} = r_m r_{m'} \cos(\theta_m - \theta_{m'}) \leq r_m r_{m'}$, where $\theta_m = \angle q_m$. Equality holds when $\theta_m = \theta_{m'}$.

We further break case 2 into two sub-cases: $N_u^{(i)} = N_u^{(i')}$ and $N_u^{(i)} > N_u^{(i')}$. For $N_u^{(i)} = N_u^{(i')}$, by using (33) and the fact that for $i \neq i'$ and $N_u^{(i)} = N_u^{(i')}$ $\mathbf{e}_i^T \mathbf{e}_{i'} \leq N_u^{(i)} - 1$, we have

$$\begin{aligned} \tilde{\Phi} &\geq r_m^2 \rho_i^2 N_u^{(i)} + r_{m'}^2 \rho_{i'}^2 N_u^{(i)} - 2r_m r_{m'} \rho_i^2 (N_u^{(i)} - 1) \\ &= (r_m \rho_i - r_{m'} \rho_{i'})^2 N_u^{(i)} + 2r_m r_{m'} \rho_i \rho_{i'} \\ &\stackrel{(a)}{\geq} 2r_m r_{m'} \rho_i \rho_{i'} \geq 2r_{\min}^2 \rho_i \rho_{i'}, \end{aligned} \quad (34)$$

where $r_{\min} = \min_m |r_m|$ is the minimum magnitude of the IQ symbols. For $N_u^{(i)} > N_u^{(i')}$, from (33) we have

$$\begin{aligned} \tilde{\Phi} &\geq r_m^2 \rho_i^2 N_u^{(i)} + r_{m'}^2 \rho_{i'}^2 N_u^{(i')} - 2r_m r_{m'} \rho_i \rho_{i'} N_u^{(i')} \\ &= r_m^2 \rho_i^2 (N_u^{(i)} - N_u^{(i')}) + (r_m \rho_i - r_{m'} \rho_{i'})^2 N_u^{(i')} \\ &\stackrel{(a)}{\geq} r_m^2 \rho_i^2 \geq r_{\min}^2 \rho_i^2, \end{aligned} \quad (35)$$

where in the first step we have used the fact that for $i \neq i'$ and $N_u^{(i)} > N_u^{(i')}$ there is $\mathbf{e}_i^T \mathbf{e}_{i'} \leq \min(N_u^{(i)}, N_u^{(i')}) = N_u^{(i')}$. By comparing (34) and (35), we see that $\tilde{\Phi}_{\min}$ in case 2 is determined by the lower bound given in (35).

Now it is clear that $\tilde{\Phi}_{\min}$ is generated either by (32) or by (35), i.e.,

$$\tilde{\Phi} \geq \begin{cases} d_{\min}^2 \rho_i^2 N_u^{(i)} & \text{for } i = i' \\ r_{\min}^2 \rho_i^2 & \text{for } i \neq i' \end{cases} \quad (36a)$$

where in (36a) $m \neq m'$ is implied by $i = i'$ because $(i, m) \neq (i', m')$.

For rectangular and square $M(M \geq 4)$ -QAM constellations [54], there is $d_{\min}^2 = 2r_{\min}^2$ and (36a) becomes

$$\tilde{\Phi} \geq 2r_{\min}^2 \rho_i^2 N_u^{(i)}. \quad (37)$$

Since $N_u^{(i)} \geq 1$, it follows that $\tilde{\Phi}_{\min}$ is generated by (36b). Therefore, to maximize $\tilde{\Phi}_{\min}$ we have to maximize $\{\rho_i^2\}$. Considering the constraints on $\{\rho_i\}$ given in (24b) and (24c), it is clear that the $\tilde{\Phi}_{\min}$ is maximized by choosing equal ρ_i 's, thus we have

$$\rho_i = \sqrt{\frac{P}{\sum_i \|\mathbf{e}_i\|^2 \sum_m |q_m|^2}} \quad i = 1, \dots, 2^{N_t} - 1. \quad (38)$$

Using the following binomial identity

$$\sum_{k=1}^n \binom{n}{k} k = \sum_{k=0}^n \binom{n}{k} k = n2^{n-1}, \quad (39)$$

which is obtained by differentiating $(1+x)^n = \sum_{k=0}^n \binom{n}{k} x^k$ with respect to x and then letting $x = 1$, we have

$$\sum_{i=1}^{2^{N_t}-1} \|\mathbf{e}_i\|^2 = \sum_{i=1}^{2^{N_t}-1} N_u^{(i)} = \sum_{N_u=1}^{N_t} \binom{N_t}{N_u} N_u = N_t 2^{N_t-1}. \quad (40)$$

Plugging (40) into (38) leads to (25). As a result, the maximized lower bound of $\tilde{\Phi}$ is given by

$$\tilde{\Phi} \geq r_{\min}^2 \rho_i^2 = \frac{P}{N_t 2^{N_t-1} \sum_m |q_m / r_{\min}|^2}. \quad (41)$$

For an $M(M \geq 2)$ -PSK constellation [54], $r_m = \text{const.} \triangleq r$ and $d_{\min} = 2r \sin(\pi/M)$. Substitute these into (36) and we have

$$\tilde{\Phi} \geq \begin{cases} 4 \sin^2\left(\frac{\pi}{M}\right) r^2 \rho_i^2 N_u^{(i)} & \text{if } N_u^{(i)} \leq N_u^* \\ r^2 \rho_i^2 & \text{otherwise} \end{cases} \quad (42a)$$

$$(42b)$$

where

$$N_u^* \triangleq \left\lfloor \frac{1}{4 \sin^2\left(\frac{\pi}{M}\right)} \right\rfloor. \quad (43)$$

For $2 \leq M \leq 6$, there is $N_u^{(i)} \geq 1 \geq N_u^*$ and $\tilde{\Phi}_{\min}$ is generated by (42b). In this case, similarly as in the square/rectangular M -QAM case, $\tilde{\Phi}_{\min}$ is maximized by allocating power according to (25) and the maximized lower bound of $\tilde{\Phi}$ is as given by (41). For $M > 6$ and large N_t such that $N_t > N_u^*$, $\tilde{\Phi}_{\min}$ can be generated by both (42a) and (42b). To maximize $\tilde{\Phi}_{\min}$

in this case, we divide the indices of the spatial symbols into two sets

$$\mathcal{I}_1 \triangleq \{i | \forall i = 1, \dots, 2^{N_t} - 1, \text{ and } N_u^{(i)} \leq N_u^*\} \quad (44a)$$

$$\mathcal{I}_2 \triangleq \{i | \forall i = 1, \dots, 2^{N_t} - 1, \text{ and } N_u^{(i)} > N_u^*\} \quad (44b)$$

and make

$$\begin{cases} \rho_i^2 N_u^{(i)} = \rho_*^2 N_u^* & \text{for } i \in \mathcal{I}_1 \\ \rho_i^2 = 4 \sin^2\left(\frac{\pi}{M}\right) \rho_*^2 N_u^* & \text{for } i \in \mathcal{I}_2 \end{cases}, \quad (45)$$

where ρ_* is a constant to be found. By uniting (45) and (24c) and solving for ρ_* , we have

$$\rho_* = \sqrt{\frac{P / (N_u^* \sum_m |q_m|^2)}{|\mathcal{I}_1| + 4 \sin^2\left(\frac{\pi}{M}\right) \sum_{i \in \mathcal{I}_2} N_u^{(i)}}} \quad (46a)$$

Thus we have (26), in which $k_* \triangleq \rho_* \sqrt{N_u^*}$. For $M > 6$ but not so large N_t such that $N_u^{(i)} \leq N_t \leq N_u^*$, $\tilde{\Phi}_{\min}$ is generated by (42a) and is maximized when we make $\rho_i^2 N_u^{(i)} = \text{const}$. It turns out that (26) serves also as the solution for this case. For both cases (i.e., $N_t > N_u^*$ and $N_t \leq N_u^*$), the maximized lower bound of $\tilde{\Phi}$ is given by

$$\begin{aligned} \tilde{\Phi} &\geq 4 \sin^2\left(\frac{\pi}{M}\right) r^2 \rho_*^2 N_u^* \\ &= \frac{4 \sin^2\left(\frac{\pi}{M}\right) P / M}{|\mathcal{I}_1| + 4 \sin^2\left(\frac{\pi}{M}\right) \sum_{i \in \mathcal{I}_2} N_u^{(i)}}. \end{aligned} \quad (47)$$

APPENDIX B

DERIVATION OF THE PAE OF THE VGSM TX

The VGSM TX is shown in Fig. 10. For improved PAE, we have assumed that the PAs are placed close to the feed of the TX antennas. For ease of analysis, we assume the wires in the routing section are of equal length. Denote the input (IQ signal) power to the power splitter as P_{in} , and the total transmitted power corresponding to the i -th spatial symbol as $P_{\text{out}}^{(i)}$, thus we have

$$P_{\text{out}}^{(i)} = P_{\text{in}} \cdot \underbrace{\frac{1}{N_t} \cdot G_{\text{RL}} \cdot G_{\text{PA}}}_{P_{\text{in,PA}}} \cdot N_u^{(i)}, \quad (48)$$

where $1/N_t$ is due to N_t -branch power splitting, G_{RL} accounts for the routing loss from the mixer to the input of PAs, $P_{\text{in,PA}}$ is the input power to each of the PAs, G_{PA} is the ON-gain of the PAs, and $N_u^{(i)}$ is the number of ON branches corresponding to the i -th spatial symbol. Denote the total DC power supply of the PAs corresponding to the i -th spatial symbol as $P_{\text{DC}}^{(i)}$, thus we have

$$P_{\text{DC}}^{(i)} = \frac{1}{\eta_{\text{PA}}} (G_{\text{PA}} - 1) P_{\text{in,PA}} \cdot N_u^{(i)}, \quad (49)$$

where η_{PA} is the PAE of the PAs in their ON-state as defined in (28). Note that, due to the use of switching PAs, in (49) we only account for the DC power consumption of the $N_u^{(i)}$ activated PAs, i.e., the inactivated PAs are assumed to consume

no power. The PAE of the TX RF front-end is calculated for the i -th spatial symbol as

$$\eta_{\text{TX}}^{(i)} = \frac{P_{\text{out}}^{(i)} - P_{\text{in}}}{P_{\text{DC}}^{(i)}} = \eta_{\text{PA}} \left[1 - \frac{N_t / N_u - G_{\text{RL}}}{G_{\text{RL}} (G_{\text{PA}} - 1)} \right]. \quad (50)$$

Assume the spatial symbols are equal-probable, then the average PAE over all possible spatial symbols is

$$\eta_{\text{TX}} = \mathbb{E}_i \left[\eta_{\text{TX}}^{(i)} \right] = \eta_{\text{PA}} \left[1 - \frac{\epsilon - G_{\text{RL}}}{G_{\text{RL}} (G_{\text{PA}} - 1)} \right], \quad (51)$$

where

$$\begin{aligned} \epsilon &= \mathbb{E}_i \left[\frac{N_t}{N_u^{(i)}} \right] = \frac{1}{2^{N_t} - 1} \sum_{i=1}^{2^{N_t}-1} \left[\frac{N_t}{N_u^{(i)}} \right] \\ &= \frac{1}{2^{N_t} - 1} \sum_{N_u=1}^{N_t} \left[\binom{N_t}{N_u} \frac{N_t}{N_u} \right]. \end{aligned} \quad (52)$$

The sum in the final step of (52) can be carried out using any computer algebra software, thus we have found η_{TX} as given in (29a).

REFERENCES

- [1] M. Di Renzo, H. Haas, A. Ghayeb, 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.
- [2] N. Serafimovski, A. Younis, R. Mesleh, P. Chambers, M. Di Renzo, C.-X. Wang, P. Grant, M. Beach, and H. Haas, "Practical Implementation of Spatial Modulation," *IEEE Trans. Veh. Technol.*, vol. 62, no. 9, pp. 4511–4523, Nov. 2013.
- [3] A. Younis, W. Thompson, M. Di Renzo, C.-X. Wang, M. Beach, H. Haas, and P. Grant, "Performance of Spatial Modulation Using Measured Real-World Channels," in *2013 IEEE 78th Vehicular Technology Conf. (VTC Fall)*, Sep. 2013, pp. 1–5.
- [4] J. Zhang, Y. Wang, L. Ding, and N. Zhang, "Bit Error Probability of Spatial Modulation over Measured Indoor Channels," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1380–1387, Mar. 2014.
- [5] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.
- [6] P. Wolniansky, G. Foschini, G. Golden, and R. Valenzuela, "V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel," in *1998 URSI Int. Symp. on Signals, Systems, and Electronics*, Sep. 1998, pp. 295–300.
- [7] C.-X. Wang, F. Haider, X. Gao, X.-H. You, Y. Yang, D. Yuan, H. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, "Cellular architecture and key technologies for 5g wireless communication networks," *Communications Magazine, IEEE*, vol. 52, no. 2, pp. 122–130, 2014.
- [8] R. Mesleh, H. Haas, C. W. Ahn, and S. Yun, "Spatial modulation — A new low complexity spectral efficiency enhancing technique," in *1st Int. Conf. on Communications and Networking in China*, Oct. 2006, pp. 1–5.
- [9] A. Stavridis, S. Sinanovic, M. Di Renzo, H. Haas, and P. Grant, "An energy saving base station employing spatial modulation," in *2012 IEEE 17th Int. Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, Sep. 2012, pp. 231–235.
- [10] X. Wu, S. Sinanovic, M. Di Renzo, and H. Haas, "Base station energy consumption for transmission optimised spatial modulation (TOSM) in correlated channels," in *2012 IEEE 17th Int. Workshop on Computer Aided Modeling and Design of Communication Links and Networks*, Sep. 2012, pp. 261–265.
- [11] A. Stavridis, S. Sinanovic, M. Di Renzo, and H. Haas, "Energy evaluation of spatial modulation at a multi-antenna base station," in *2013 IEEE 78th Vehicular Technology Conf. (VTC Fall)*, Sep. 2013, pp. 1–5.
- [12] A. Younis, D. A. Basnayaka, and H. Haas, "Performance analysis for generalised spatial modulation," in *Proc. 20th European Wireless Conf.*, May 2014, pp. 1–6.

- [13] S. K. Yong and C.-C. Chong, "An overview of multigigabit wireless through millimeter wave technology: Potentials and technical challenges," *EURASIP J. Wirel. Commun. Netw.*, vol. 2007, no. 1, pp. 50–50, Jan. 2007.
- [14] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101–107, Jun. 2011.
- [15] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5g cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.
- [16] R. Daniels, J. Murdock, T. Rappaport, and R. Heath, "60 GHz wireless: Up close and personal," *IEEE Microw. Mag.*, vol. 11, no. 7, pp. 44–50, Dec. 2010.
- [17] E. Torkildson, U. Madhow, and M. Rodwell, "Indoor millimeter wave MIMO: Feasibility and performance," *IEEE Trans. Wireless Commun.*, vol. 10, no. 12, pp. 4150–4160, Dec. 2011.
- [18] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.
- [19] W. Hong, K.-H. Baek, Y. Lee, Y. Kim, and S.-T. Ko, "Study and prototyping of practically large-scale mmWave antenna systems for 5g cellular devices," *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 63–69, Sep. 2014.
- [20] R. W. Heath, "Millimeter wave MIMO: A signal processing perspective," Apr. 2015. [Online]. Available: <http://users.ece.utexas.edu/~rheath/presentations/2015/MmWaveMIMOaSPerspectiveICASSP2015Heath.pdf>
- [21] A. Younis, "Spatial Modulation: Theory to Practice," Ph.D. dissertation, University of Edinburgh, 2014. [Online]. Available: <http://hdl.handle.net/1842/8990>
- [22] O. Osman, "Variable active antenna spatial modulation," *IET Microwaves, Antennas & Propagation*, vol. 9, no. 15, pp. 1816–1824, Dec. 2015.
- [23] 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.
- [24] T. Datta, H. Eshwariaiah, and A. Chockalingam, "Generalized space and frequency index modulation," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 4911–4924, Jul. 2016.
- [25] R. Fan, Y. J. Yu, and Y. L. Guan, "Generalization of orthogonal frequency division multiplexing with index modulation," *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, pp. 5350–5359, 2015.
- [26] P. Liu, M. Di Renzo, and A. Springer, "Line-of-sight spatial modulation for indoor mmWave communication at 60 GHz," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7373–7389, Nov. 2016.
- [27] P. Liu and A. Springer, "Space shift keying for LOS communication at mmWave frequencies," *IEEE Wireless Commun. Lett.*, vol. 4, no. 2, pp. 121–124, Apr. 2015.
- [28] N. Ishikawa, R. Rajashekar, S. Sugiura, and L. Hanzo, "Generalized spatial modulation based reduced-RF-chain millimeter-wave communications," *IEEE Trans. Veh. Technol.*, vol. PP, no. 99, pp. 1–5, 2016.
- [29] A. Maltsev, V. Erceg, E. Perahia, C. Hansen, R. Maslennikov, A. Lomayev, A. Sevastyanov, A. Khoryaev, G. Morozov, M. Jacob, S. Priebe, T. Kürner, S. Kato, H. Sawada, K. Sato, and H. Harada, "Channel models for 60 GHz WLAN systems," Tech. Rep. IEEE 802.11-09/0334r8, May 2010. [Online]. Available: <https://mentor.ieee.org/802.11/dcn/09/11-09-0334-08-00ad-channel-models-for-60-ghz-wlan-systems.doc>
- [30] F. Bohagen, P. Orten, and G. Oien, "Design of optimal high-rank line-of-sight MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 6, no. 4, pp. 1420–1425, Apr. 2007.
- [31] H. Friis, "A note on a simple transmission formula," *Proceedings of the IRE*, vol. 34, no. 5, pp. 254–256, May 1946.
- [32] K. Sato, H. Kozima, H. Masuzawa, T. Manabe, T. Ihara, Y. Kasashima, and K. Yamaki, "Measurements of reflection characteristics and refractive indices of interior construction materials in millimeter-wave bands," in *1995 IEEE 45th Vehicular Technology Conf.*, vol. 1, Jul. 1995, pp. 449–453 vol.1.
- [33] N. Serafimovski, M. Di Renzo, S. Sinanovic, R. Mesleh, and H. Haas, "Fractional bit encoded spatial modulation (FBE-SM)," *IEEE Commun. Lett.*, vol. 14, no. 5, pp. 429–431, May 2010.
- [34] Y. Yang and S. Aissa, "Bit-padding information guided channel hopping," *IEEE Commun. Lett.*, vol. 15, no. 2, pp. 163–165, Feb. 2011.
- [35] —, "Information guided channel hopping with an arbitrary number of transmit antennas," *IEEE Commun. Lett.*, vol. 16, no. 10, pp. 1552–1555, Oct. 2012.
- [36] J. Jeganathan, A. Ghrayeb, and L. Szczecinski, "Spatial modulation: optimal detection and performance analysis," *IEEE Commun. Lett.*, vol. 12, no. 8, pp. 545–547, Aug. 2008.
- [37] H. Men and M. Jin, "A low-complexity ML detection algorithm for spatial modulation systems with M -PSK constellation," *IEEE Commun. Lett.*, vol. 18, no. 8, pp. 1375–1378, Aug. 2014.
- [38] R. Rajashekar, K. Hari, and L. Hanzo, "Reduced-complexity ML detection and capacity-optimized training for spatial modulation systems," *IEEE Trans. Commun.*, vol. 62, no. 1, pp. 112–125, Jan. 2014.
- [39] F. Bohagen, P. Orten, and G. E. Oien, "Construction and capacity analysis of high-rank line-of-sight MIMO channels," in *2005 IEEE Wireless Communications and Networking Conf.*, vol. 1. IEEE, 2005, pp. 432–437.
- [40] Y. Yang and B. Jiao, "Information-guided channel-hopping for high data rate wireless communication," *IEEE Commun. Lett.*, vol. 12, no. 4, pp. 225–227, Apr. 2008.
- [41] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, UK ; New York: Cambridge University Press, Jul. 2005.
- [42] P. Som and A. Chockalingam, "Spatial modulation and space shift keying in single carrier communication," in *2012 IEEE 23rd Int. Symp. on Personal Indoor and Mobile Radio Communications (PIMRC)*, Sep. 2012, pp. 1962–1967.
- [43] 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. PP, no. 99, pp. 1–1, 2014.
- [44] R. Rajashekar, K. Hari, and L. Hanzo, "Spatial modulation aided zero-padded single carrier transmission for dispersive channels," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2318–2329, Jun. 2013.
- [45] Andrea Goldsmith, *Wireless communications*, 1st ed. Cambridge University Press, Aug. 2005.
- [46] A. Valdes-Garcia, S. Nicolson, J.-W. Lai, A. Natarajan, P.-Y. Chen, S. Reynolds, J.-H. Zhan, D. Kam, D. Liu, and B. Floyd, "A fully integrated 16-element phased-array transmitter in SiGe BiCMOS for 60-GHz communications," *IEEE J. Solid-State Circuits*, vol. 45, no. 12, pp. 2757–2773, Dec. 2010.
- [47] A. Natarajan, S. Reynolds, M.-D. Tsai, S. Nicolson, J.-H. Zhan, D. G. Kam, D. Liu, Y.-L. Huang, A. Valdes-Garcia, and B. Floyd, "A fully-integrated 16-element phased-array receiver in SiGe BiCMOS for 60-GHz Communications," *IEEE J. Solid-State Circuits*, vol. 46, no. 5, pp. 1059–1075, May 2011.
- [48] X. Li, P. Baltus, D. Milosevic, W. Deng, P. van Zeijl, N. Bird, and A. van Roermond, "Wireless wire-the 60 GHz ultra-low power radio system," in *IEEE Radio and Wireless Symp., 2009. RWS '09*, Jan. 2009, pp. 215–218.
- [49] F. Lin, J. Brinkhoff, K. Kang, D. D. Pham, and X. Yuan, "A low power 60GHz OOK transceiver system in 90nm CMOS with innovative on-chip AMC antenna," in *IEEE Asian Solid-State Circuits Conf., 2009*, Nov. 2009, pp. 349–352.
- [50] J. J. Lee and C. S. Park, "60-GHz gigabits-per-second OOK modulator with high output power in 90-nm CMOS," *IEEE Trans. Circuits Syst. II*, vol. 58, no. 5, pp. 249–253, May 2011.
- [51] A. Chakrabarti and H. Krishnaswamy, "Design considerations for stacked Class-E-like mmWave high-speed power DACs in CMOS," in *2013 IEEE MTT-S Int. Microwave Symp. Dig.*, Jun. 2013, pp. 1–4.
- [52] "Taconnic TacLamplus datasheet." [Online]. Available: <http://www.taconnic-add.com/pdf/taclamplus.pdf>
- [53] Y. Uemichi, O. Nukaga, R. Hosono, N. Guan, J. Hirokawa, and M. Ando, "A ultra low-loss silica-based transformer between microstrip line and post-wall waveguide for millimeter-wave antenna-in-package applications," *IEEE MTT-S Int. Microwave Symp. Dig.*, pp. 3–5, 2014.
- [54] J. Proakis and M. Salehi, *Digital Communications, 5th Edition*, 5th ed. Boston: McGraw-Hill Science/Engineering/Math, Nov. 2007.
- [55] P. Yang, Y. Xiao, B. Zhang, and S. Li, "Power Allocation-Aided Spatial Modulation for Limited-Feedback MIMO Systems," *IEEE Trans. Veh. Technol.*, vol. 64, no. 5, pp. 2198–2204, May 2015.
- [56] M. Maleki, H. R. Bahrami, and A. Alizadeh, "Constellation design for spatial modulation," in *2015 IEEE International Conference on Communications (ICC)*, no. 1. IEEE, Jun. 2015, pp. 2739–2743.
- [57] E. W. Weisstein, "Generalized hypergeometric function." [Online]. Available: <http://mathworld.wolfram.com/GeneralizedHypergeometricFunction.html>
- [58] M. Thian and V. Fusco, "Ultrafast low-loss 42–70 GHz differential SPDT switch in 0.35 m SiGe technology," *IEEE Trans. Microw. Theory Techn.*, vol. 60, no. 3, pp. 655–659, Mar. 2012.
- [59] Y. A. Atesal, B. Cetinoneri, and G. M. Rebeiz, "Low-loss 0.13- μ m CMOS 50–70 GHz SPDT and SP4T switches," in *2009 IEEE Radio Freq. Integr. Circuits Symp.*, vol. 1. IEEE, Jun. 2009, pp. 43–46.

- [60] R. L. Schmid, P. Song, and J. D. Cressler, "A compact, transformer-based 60 GHz SPDT RF switch utilizing diode-connected SiGe HBTs," in *2013 IEEE Bipolar/BiCMOS Circuits and Technology Meeting (BCTM)*. IEEE, Sep. 2013, pp. 111–114.
- [61] D. Liu, J. Akkermans, and B. Floyd, "A superstrate patch antenna for 60-GHz applications," in *3rd European Conf. on Antennas and Propagation*, Mar. 2009, pp. 2592–2594.
- [62] Y. Yu, P. G. Baltus, and A. H. van Roermund, *Integrated 60GHz RF Beamforming in CMOS*. Dordrecht: Springer Netherlands, 2011.
- [63] H.-C. Kuo and H.-R. Chuang, "A 60-GHz high-gain, low-power, 3.7-dB noise-figure low-noise amplifier in 90-nm CMOS," in *2013 European Microwave Integrated Circuits Conf.*, Oct. 2013, pp. 584–587.