# Novel Index Modulation Techniques: A Survey

Tianqi Mao, *Student Member, IEEE*, Qi Wang, *Member, IEEE*, Zhaocheng Wang, *Senior Member, IEEE*, and Sheng Chen, *Fellow, IEEE*

***Abstract*—In fifth generation wireless networks, the escalating teletraffic and energy consumption has necessitated the development of green communication techniques in order to further enhance both the system's spectral efficiency and energy efficiency. In the past few years, the novel index modulation (IM) has emerged as a promising technology that is widely employed in wireless communications. In this paper, we present a survey on IM in order to provide the readers with a better understanding of its principles, advantages, and potential applications. We start with a comprehensive literature review, where the concept of IM is introduced, and various existing IM schemes are classified according to their signal domains, including the frequency domain, spatial domain, time domain and channel domain. Then the principles of different IM-aided systems are detailed, where the transceiver design is illustrated, followed by descriptions of typical systems and corresponding performance evaluation. A range of challenges and open issues on IM are discussed before we conclude this survey.**

***Index Terms*—Index modulation, orthogonal frequency division multiplexing, spatial modulation, media-based modulation, constellation design, maximum-likelihood detection, low-complexity receiver design.**

## GLOSSARY

| | |
|---|---|
| 4G | Fourth Generation |
| 5G | Fifth Generation |
| ABEP | Average Bit Error Probability |
| ACE | Active Constellation Expansion |
| ACO-OFDM | Asymmetrically Clipped Optical OFDM |
| A/D | Analog-to-Digital Converter |
| APM | Amplitude/Phase Modulation |
| AWGN | Additive White Gaussian Noise |
| BER | Bit Error Rate |
| BICM-ID | Bit-Interleaved Coded Modulation with Iterative Decoding |
| BLAST | Bell Labs Layered Space-Time |
| BMST | Block Markov superposition transmission |
| BPSK | Binary Phase Shift Keying |
| CD | Channel Domain |
| CD-IM | Channel-Domain IM |
| CIR | Channel Impulse Response |
| CP | Cyclic Prefix |
| CPEP | Conditioned PEP |
| CSI | Channel State Information |
| CSIT | CSI at the Transmitter |
| D/A | Digital-to-Analog Converter |
| DCO-OFDM | DC-Biased Optical OFDM |
| DMBM | Differential MBM |
| DM-OFDM | Dual-Mode IM-Aided OFDM |
| DM-SCIM | Dual-mode SC-IM |
| ED | Euclidean Distance |
| EGSIM-OFDM | Enhanced GSIM-OFDM |
| EGSIM-OFDM1 | EGSIM-OFDM Version 1 |
| EGSIM-OFDM2 | EGSIM-OFDM Version 2 |
| EGSIM-OFDM-JIQ | In EGSIM-OFDM, IM Conducted on I/Q Components Jointly |
| ESIM-OFDM | Enhanced SIM-OFDM |
| ESM | Enhanced SM |
| FD | Frequency Domain |
| FDE | FD Equalization |
| FD-IM | Frequency-Domain IM |
| FFT | Fast Fourier Transform |
| FSK | Frequency Shift Keying |
| FSO | Free-Space Optical |
| FTN-IM | Faster-than-Nyquist Rate SC-IM |
| GDM-OFDM | Generalized DM-OFDM |
| GPQSM | Generalized Precoded QSM |
| GPSM | Generalized PSM |
| GSIM-OFDM | Generalized ESIM-OFDM |
| GSM | Generalized SM |
| GSM-MBM | GSM Combined with MBM |
| GSSK | Generalized SSK |
| GSTSK | Generalized STSK |
| IAS | Inter-Antenna Synchronization |
| IBI | Inter-Block Interference |
| ICI | Inter-Channel Interference |
| IFFT | Inverse Fast Fourier Transform |
| IM | Index Modulation |
| IM-OFDM | IM-Aided OFDM |
| IoT | Internet of Things |

Manuscript received August 17, 2017; revised January 23, 2018 and May 21, 2018; accepted July 19, 2018. Date of publication July 23, 2018; date of current version February 22, 2019. This work was supported in part by the National Natural Science Foundation of China under Grant 61571267, in part by the Shenzhen Subject Arrangements under Grant JCYJ20160331184124954, and in part by the Shenzhen Fundamental Research under Project JCYJ20170307153116785. *(Corresponding author: Zhaocheng Wang.)*

T. Mao and Z. Wang are with the Beijing National Research Center for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, Beijing 100084, China (e-mail: mtq15@mails.tsinghua.edu.cn; zcwang@tsinghua.edu.cn).

Q. Wang was with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. He is now with Huawei Technologies Company Ltd., Shenzhen 518129, China (e-mail: steven_wq@hotmail.com).

S. Chen is with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K., and also with King Abdulaziz University, Jeddah 21589, Saudi Arabia (e-mail: sqc@ecs.soton.ac.uk).

Digital Object Identifier 10.1109/COMST.2018.2858567

| | |
|---|---|
| I/Q | In-phase and Quadrature |
| ISI | Intersymbol Interference |
| ItM/DD | Intensity Modulation with Direct Detection |
| LDPC | Low-Density Parity Check Code |
| LED | Light Emitting Diode |
| LLR | Log-Likelihood Ratio |
| LMG-SSTSK | Layered Multi-Group Steered STSK |
| LMIMO-MBM | Layered MIMO-MBM |
| LTE | Long-Term Evolution |
| MAP | Mirror Activation Pattern |
| MA-SM | Multiple Active-SM |
| MBM | Media-Based Modulation |
| MI | Mutual Information |
| MIAD | Minimum Intra-Mode Distance |
| MIMO | Multiple-Input Multiple-Output |
| MIMO-OFDM-IM | GSIM-OFDM Combined with MIMO |
| MIRD | Minimum Inter-mode Distance |
| ML | Maximum Likelihood |
| MM-OFDM | Multi-Mode IM-Aided OFDM |
| MMSE | Minimum Mean-Squared Error |
| mmWave | Millimeter-wave |
| MSF-STSK | Multi-Space-Frequency STSK |
| MS-STSK | Multi-Set STSK |
| O-DM-OFDM | Optical DM-OFDM |
| OFDM | Orthogonal Frequency Division Multiplexing |
| O-GSIM-OFDM | Optical GSIM-OFDM |
| OOK | On-Off Keying |
| PAM | Pulse Amplitude Modulation |
| PAM-DMT | PAM Discrete Multi-Tone |
| PAPR | Peak-to-Average Power Ratio |
| PEP | Pairwise Error Probability |
| PLC | Power Line Communications |
| PPM | Pulse-Position Modulation |
| PSK | Phase Shift Keying |
| PSM | Precoded SM |
| PTS | Partial Transmit Sequence |
| QAM | Quadrature Amplitude Modulation |
| QCM | Quadrature Channel Modulation |
| QoS | Quality-of-Service |
| QPSK | Quadrature Phase Shift Keying |
| QSM | Quadrature SM |
| RF | Radio Frequency |
| RX | Receiver |
| SBIM | Source-Based IM |
| SC-FDE | Single-Carrier Frequency-Domain Equalization |
| SC-IM | Single-Carrier-Based IM |
| SD | Spatial Domain |
| SD-IM | Spatial-Domain IM |
| SIC | Successive Interference Cancellation |
| SIMO | Single-Input Multiple-Output |
| SIMO-MBM | MBM Combined with SIMO |
| SIM-OFDM | Subcarrier-index modulated OFDM |
| SISO | Single-Input Single-Output |
| SLM | Selective Mapping |
| SM | Spatial Modulation |
| SM-OFDM | SM Combined with OFDM |
| SNR | Signal-to-Noise Power Ratio |
| S/P | Serial-to-Parallel Converter |
| SSK | Space Shift Keying |
| STBC | Space-Time Block Codes |
| STCM | Space-Time Channel Modulation |
| STFSK | Space-Time-Frequency Shift Keying |
| ST-IM | Space-Time IM |
| STSK | Space-Time Shift Keying |
| TCSM | Trellis Coded Spatial Modulation |
| TD | Time Domain |
| TDD | Time Division Duplexing |
| TD-IM | Time-Domain IM |
| TX | Transmitter |
| U-OFDM | Unipolar OFDM |
| UPEP | Unconditioned PEP |
| VLC | Visible Light Communications |
| ZF | Zero-Forcing |
| ZTM-OFDM | Zero-Padded Tri-Mode IM-Aided OFDM. |

## I. Introduction

DUE TO the fast evolution of smart communication terminals and in order to meet the explosive increase of mobile traffic, high-rate transmission schemes have attracted increasing attention from both the academia and industry, featured by orthogonal frequency division multiplexing (OFDM) [1] and single-carrier frequency-domain equalization (SC-FDE) [2]. OFDM, which has fast become one of the most dominant techniques in 4G communication, is widely employed in multiple wireless standards including Wi-Fi, IEEE 802.11 a/g, 802.16 WiMAX and Long-Term Evolution (LTE) [3]–[6], owing to its distinctive merits of resilience to frequency selective fading channels and facilitating low-complexity hardware implementation with one-tap frequency-domain equalization (FDE) [7]–[9]. However, OFDM symbols suffer from the drawback of high peak-to-average power ratio (PAPR), leading to inevitable nonlinear distortion by the transmitter (TX) power amplifier [10]. To mitigate this issue, the early-invented SC-FDE scheme is still employed in some wireless applications for its lower PAPR, particularly in visible light communication (VLC) systems [11], which are extremely sensitive to the nonlinear characteristics of light emitting diode (LED) emitters. Furthermore, owing to the need for enlarging capacity of wireless communications and ensuring satisfactory quality-of-service (QoS) of mobile devices, multiple-input multiple-output (MIMO) technique [12] has attracted extensive attention, which is capable of improving the system throughput without the sacrifice of additional bandwidth usage or energy consumption [13].

In 5G networks, the explosive growth in mobile data services and the popularisation of smart devices have necessitated the requirement for high spectral and energy efficiency. Following this trend, the novel concept of index modulation (IM) has been promoted in wireless communications to meet the demand for high throughput and low energy consumption. More specifically, unlike conventional communication

schemes, only a fraction of certain indexed resource entities, e.g., subcarriers, antennas, time slots or channel states, are activated for a data transmission, while the others are kept unused by the associated transmission, where additional information bits are conveyed implicitly by the index usage or activation patterns. IM therefore can significantly enhance the spectral and energy efficiency for the reason that, apart from information bits carried by usual amplitude/phase modulation (APM) based transmission, additional index bits are harvested without energy consumption. This makes an IM-aided system capable of employing only part of the resources to achieve the same throughput as its conventional counterpart, leading to a reduced system complexity, while consuming substantially less energy. Besides, due to the flexible structure of IM-aided frame, considerable diversity gains can be achieved from the activated pattern of the indexed entities, types of the constellation alphabets, or the selection of channel states, etc., which contributes to significant performance gain over the non-index-modulated counterparts. Therefore, IM has been applied in various formats to diverse wireless communication applications, including millimeter-wave (mmWave) transmission [14], [15], massive MIMO [16], [17] and network coding [18]. Existing IM schemes can be classified into various signal domain IM schemes, which include frequency-domain IM (FD-IM), spatial-domain IM (SD-IM) and time-domain IM (TD-IM), as well as the channel-domain IM (CD-IM) scheme.

FD-IM is also referred to as IM-aided OFDM (IM-OFDM), where IM is performed on the orthogonal subcarriers [19]. Explicitly, in addition to the binary bits carried by a classical APM scheme, such as pulse amplitude modulation (PAM), phase shift keying (PSK) and quadrature amplitude modulation (QAM), additional information can be implicitly conveyed by the subcarrier activation pattern, that is, the indices of activated subcarriers [20]. Some IM-OFDM systems are capable of achieving higher spectral efficiency than their non-index-modulated counterparts, by harvesting additional index bits. Moreover, the energy efficiency is enhanced, since transmission of index bits is energy-free. IM can be also conducted on the spatial domain (SD) by applying the SD-IM technique, also known as spatial modulation (SM) [21]–[23]. In SM, only a fraction of TX or receiver (RX) antennas are activated for data transmission, while their indices convey additional information bits in an implicit manner. Different from conventional MIMO schemes, such as spatial multiplexing schemes and repetition codes, the antenna-activation pattern of SM is capable of mitigating the inter-channel interference (ICI) as well as the inter-antenna synchronization (IAS) issue. Therefore, SM can achieve significant performance gains, in terms of bit error rate (BER) and energy efficiency, over the existent MIMO systems, such as the Bell Labs layered space-time (BLAST) system and the Alamouti system [24], [25]. IM can also be carried out in the time domain (TD). Specifically, in each data frame, only a fraction of the signaling time slots are activated for data transmission, and the corresponding indices carry information [26], [27], which leads to significant performance gain over its non-IM-aided counterpart. The multi-mode IM philosophy and the faster-than-Nyquist signalling are also applied to TD-IM [28], [29] to further enhance the system performance. Furthermore, TD-IM can be combined with the space-time block code (STBC) system [30], [31], where only subset of the dispersion matrices are selected for data transmission, while their indices carry extra information bits [32], [33].

The aforementioned IM schemes define the so-called source-based IM (SBIM), which embeds additional information bits into the indices of various TX or RX entities, such as antennas, subcarriers or time-slots. Despite of their superior energy efficiency and BER performance, SBIM systems may suffer from severe interferences under the static fading channel. Moreover, the achievable throughput of SBIM increases slowly, only proportional to the logarithm of the number of TX entities. To address these issues whilst maintaining the diversity gain, CD-IM, also known as media-based modulation (MBM) [34], has recently been introduced to radio frequency (RF) communications, which enjoys an inherent diversity against static fading [34], [35]. Specifically, the characteristics of tunable parasitic components external to the TX antennas, such as RF mirrors [36] or electronic switches [37], are intentionally modified to change the RF property consisting of permittivity, permeability and resistivity. Even a small perturbation in a rich scattering environment can be dramatically augmented, leading to differentiable channel realizations. When $m_{\mathrm{rf}}$ RF mirrors are employed, for example, by controlling their on/off status, a total of $2^{m_{\mathrm{rf}}}$ channel states can be generated, which are unknown at the TX but available at the RX using pilots. Therefore, in addition to the information carried by the conventional APM, extra $m_{\mathrm{rf}}$ bits are conveyed by the indices of the selected channel states, leading to enhanced data rate. In comparison with SBIM, whose achievable throughput increases with the logarithm of the number of TX entities, the spectral efficiency of MBM grows linearly with the number of used RF mirrors. Moreover, for SM, the spacing between neighbouring antennas is required to be sufficiently large in order to achieve independent fading. On the other hand, for MBM, RF mirrors can be placed side-by-side [38]. These benefits of MBM have motivated the relevant researches in the past few years.

For clarity, Table I categorizes the existing IM techniques,[1] and Fig. 1 illustrates the general structure of IM-aided systems. Compared to their non-index-modulated counterparts, IM-aided systems convey additional energy-free bits, which enhances the energy efficiency. Due to the flexible system structure of IM-aided systems, superior BER performance can be achieved by optimizing the system parameters. Moreover, the use of SM contributes to low hardware complexity, robustness against ICI and static fading, as well as freeing from IAS issue, since only a fraction of antennas are employed. On the other hand, the introduction of MBM techniques offers the opportunity of dramatically increasing the system throughput as well as provides inherent diversity gains. In

[1]In this paper, we propose a new general MM-OFDM, and MM-OFDM proposed in [49] is a simplified special case of our generic MM-OFDM.

TABLE I
SUMMARY OF REPRESENTATIVE IM-AIDED ARCHITECTURES

| Architecture | Reference | Features |
|---|---|---|
| **FD-IM (IM-OFDM)** | | |
| SIM-OFDM | [39] | Activated subcarriers are determined by the corresponding majority bit-values of an OOK data stream. |
| ESIM-OFDM | [40] | In each subcarrier pair, only one is activated to carry one index bit. |
| GSIM-OFDM | [41] | In each subblock, multiple subcarriers are activated to convey index bits. |
| EGSIM-OFDM | [42], [43] | For each subblock, the number of activated subcarriers is alterable, and IM is also performed on I/Q components. |
| MIMO-OFDM-IM | [44], [45] | GSIM-OFDM is combined with MIMO systems. |
| DM-OFDM | [46] | In each subblock, all subcarriers are modulated by two different constellation alphabets. |
| GDM-OFDM | [47] | In each subblock, the number of subcarriers modulated by either of the two constellation sets is alterable. |
| ZTM-OFDM | [48] | In each subblock, only a fraction of subcarriers are modulated by two different constellation alphabets. |
| MM-OFDM | [49] | In each subblock, all subcarriers are modulated by multiple different constellation alphabets. |
| **SD-IM (SM)** | | |
| Classical SM | [21]–[23] | Information is conveyed by APM and the index of a single activated antenna. |
| SSK | [50] | Information bits are only carried by the index of single activated TX antenna without APM transmission. |
| GSSK | [51] | Information bits are carried by the indices of multiple activated TX antenna without APM transmission. |
| SM-OFDM | [52] | A single TX antenna is activated for OFDM transmission. |
| GSM | [53] | Multiple TX antennas are activated to transmit the same symbol for diversity gain. |
| MA-SM | [54] | Multiple TX antennas are activated to transmit different symbols on different antennas. |
| QSM | [55] | SM is conducted on the I/Q components independently. |
| ESM | [56], [57] | SM is conducted with high flexibility of antenna activation patterns and APM constellation alphabets. |
| PSM | [58], [59] | Information is conveyed by APM and the index of single activated RX antenna. |
| GPSM | [60] | Multiple RX antennas are activated with imperfect CSIT and low rank approximation for large-dimensional MIMO. |
| **TD-IM** | | |
| SC-IM | [26] | Index bits are conveyed by the indices of activated time slots. |
| DM-SCIM | [28] | All the time slots are activated, modulated by two different constellation alphabets. |
| FTN-IM | [29] | Index bits are conveyed by indices of the activated time slots in the FTN-rate based single carrier system. |
| ST-IM | [30] | Index bits are conveyed by indices of the activated antennas and time slots. |
| STSK | [32] | Index bits are conveyed by the single activated dispersion matrix. |
| GSTSK | [33] | Index bits are conveyed by the indices of multiple activated dispersion matrices. |
| **CD-IM (MBM)** | | |
| SIMO-MBM | [34] | Index bits are conveyed by the index of selected channel state in SIMO systems. |
| DMBM | [61] | Differential coding is employed in SIMO type MBM systems. |
| LMIMO-MBM | [35] | Index bits are conveyed by the indices of selected channel states in MIMO systems. |
| GSM-MBM | [36], [62] | Index bits are conveyed by the indices of selected channel states in GSM systems. |
| STCM | [63] | Index bits are conveyed by the indices of selected channel state vectors in Alamouti STBC systems. |

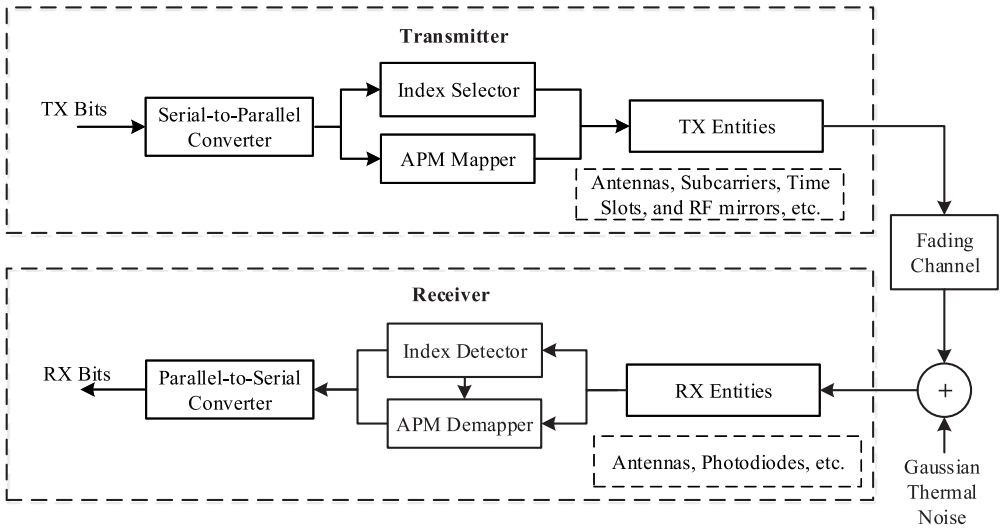



Fig. 1. The transmitter and receiver structure of generic IM-aided system.

summary, some of the distinctive advantages of IM-aided systems include:

- High energy efficiency.
- Low hardware complexity.
- Flexible system structure.
- Superior BER performance.
- Robust against ICI and static fading.
- Free of IAS issue.

To the best of the authors' knowledge, there exist limited literature surveying IM, including [20], [64]–[66].

In [64] and [65], the design framework of SM as well as its pros and cons were comprehensively investigated, where the attention was particularly directed at the transceiver design, multi-dimensional constellation optimization, link adaptation methods, and network protocol design. By discussing many published technical papers, the book [20] reviewed the system models and the detection techniques of the existing SM, IM-OFDM and space-time-domain IM schemes. Some theoretical analysis, such as pairwise error probability (PEP) calculation, and simulation results were provided, while some possible challenges and open issues of IM were also forecasted. The recent paper [66] reviews IM techniques in different signal domains. However, only SM and IM-OFDM are investigated in detail. In particular, for applications and practical issues, only SM and IM-OFDM schemes are discussed in [66].

The organization of this survey paper is given in Fig. 2. Compared with the existing reviews on IM, our new contributions are first summarized below.

1) We present a more comprehensive review and description of FD-IM, by introducing more typical IM-OFDM schemes, such as DM-OFDM, GDM-OFDM, ZTM-OFDM and MM-OFDM, and providing the explicit descriptions of corresponding transceiver structures, constellation designs and performance evaluation, in terms of spectral efficiency, energy efficiency and BER. Moreover, a generic MM-OFDM structure is proposed in the paper for the first time to overcome the disadvantages of the existing MM-OFDM scheme.
2) Our survey includes the novel CD-IM techniques, namely, MBM, by introducing several typical MBM schemes. Limitations of these typical MBM schemes are revealed and possible solutions are suggested.
3) Our survey includes all the important space-time domain IM schemes, such as STSK as well as TD-IM schemes, which were not discussed in [20], and our review, which is much more explicitly than [66], includes the detailed descriptions of transceiver design strategy and comprehensive performance evaluation based on spectral efficiency, energy efficiency and BER.
4) Applications of IM to VLC, not reviewed in most of the existing survey literature, are discussed in our survey paper.
5) Challenges and open issues missed in [20] and [66] are extensively discussed.

Therefore, different from the existing overview on IM, our presentation provides a more comprehensive investigation on the generalized concept of IM, and we review all the existing IM techniques in the frequency domain (FD), SD, TD, and channel domain (CD). State-of-the-art new researches, particularly the novel MBM and dual-mode (multi-mode) IM, are included in our paper, which offers a better illustration of the latest progress in the IM field. More specifically, in this survey, after a comprehensive literature review, we detail the diverse IM frameworks in all the four FD, SD, TD and CD, with an emphasis on the FD-IM. This is followed by the performance evaluation of various typical IM techniques based on a range of performance metrics, including the minimum Euclidean distance (ED), the PEP analysis and the computational complexity analysis. The potential challenges and future work are extensively discussed.

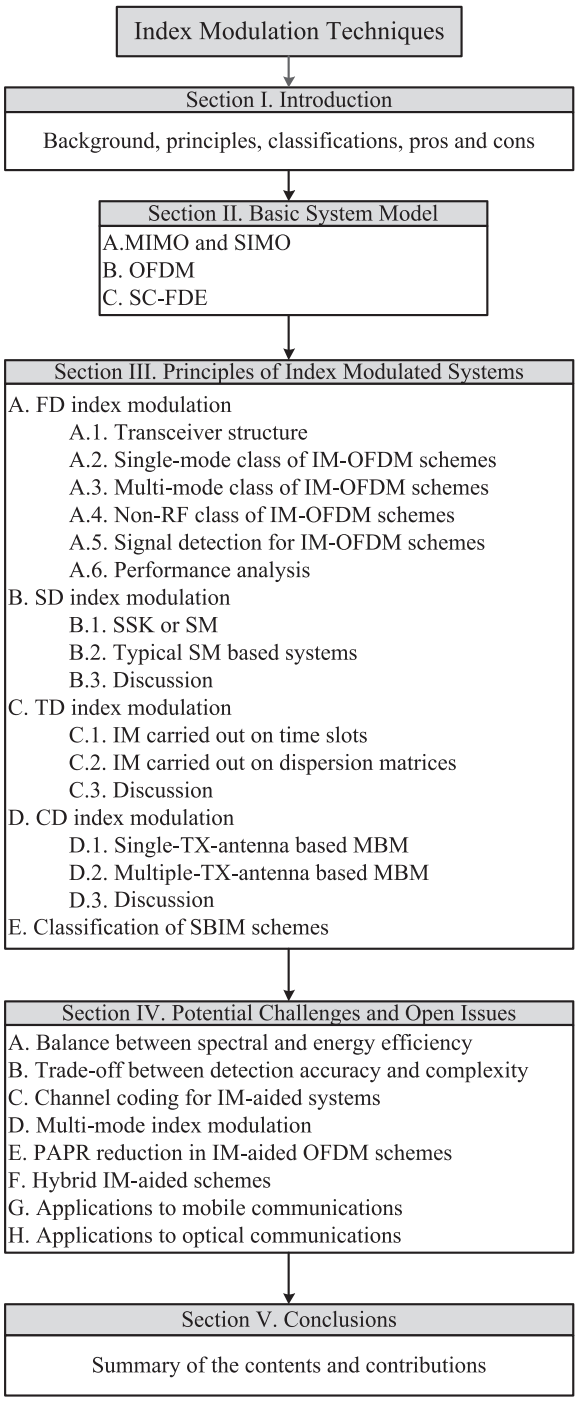



Fig. 2. Outline of the survey paper.

Before we proceed, we point out that placing an emphasis on the FD-IM is because our research is mainly focusing on FD-IM, and we would like to share our expertise and the lessons from our experience to help the readers avoiding pitfalls when embarking their own research in the IM field.

## II. Basic System Model

In this section, we briefly introduce the basic communication system models adopted by various IM techniques,

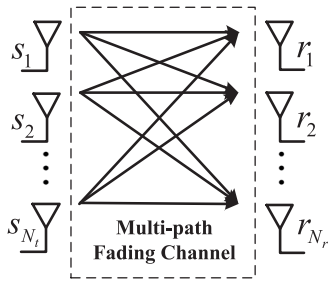



Fig. 3. MIMO communication system diagram.

including MIMO and single-input multiple-output (SIMO) as well as OFDM and SC-FDE.

### A. MIMO and SIMO

Fig. 3 presents the system diagram of spatial multiplexing MIMO equipped with $N_t$ TX antennas and $N_r$ RX antennas. If the input signal vector is denoted as $\boldsymbol{s} = [s_1 \ s_2 \cdots s_{N_t}]^{\mathrm{T}}$, where $(\cdot)^{\mathrm{T}}$ denotes the transpose operator, then the received signal vector $\boldsymbol{r} = [r_1 \ r_2 \cdots r_{N_r}]^{\mathrm{T}}$ can be represented by

$$\boldsymbol{r} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{w}, \tag{1}$$

where $\boldsymbol{w} \in \mathbb{C}^{N_r}$ is the additive white Gaussian noise (AWGN) vector, whose elements obey the circularly symmetric complex Gaussian distribution $\mathcal{CN}(0, \sigma_w^2)$, while $\boldsymbol{H} \in \mathbb{C}^{N_r \times N_t}$ denotes the channel fading matrix, whose (*i, j*)th element $h_{i,j}$ stands for the channel coefficient between the *j*th TX antenna and the *i*th RX antenna, which can be modelled as an independent variable obeying $\mathcal{CN}(0, 1)$ for the static fading channel.

At the receiver, the zero-forcing (ZF) detection is often employed for simplicity, which multiplies the received signal vector $\boldsymbol{r}$ by the pseudo-inverse of $\boldsymbol{H}$, formulated as

$$\widehat{\boldsymbol{s}}_{\mathrm{ZF}} = \left(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{r} = \boldsymbol{s} + \left(\boldsymbol{H}^{\mathrm{H}}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{w}, \tag{2}$$

where $(\cdot)^{\mathrm{H}}$ denotes the conjugate transpose operator. It can be seen from (2) that the effect of the channel AWGN is amplified, which may cause performance degradation. To address this issue, several enhanced algorithms can be invoked, which include the minimum mean-squared error (MMSE) detection [67], [68] and the successive interference cancellation (SIC) techniques [69], [70].

SIMO is a special case of MIMO with $N_t = 1$. Apart from spatial multiplexing MIMO, MIMO can also be designed to attain spatial diversity gains in order to enhance the achievable BER performance at the expense of reduced throughput.

To evaluate the performance of SIMO and MIMO systems, capacity analysis is often employed. For a memoryless SIMO system with $N_r$ RX antennas, its capacity is given by [12]

$$C_{\mathrm{SIMO}} = \log_2\left(1 + \gamma \sum_{i=1}^{N_r} \left|h_{i,1}\right|^2\right) \ [\mathrm{bit/s/Hz}], \tag{3}$$

where $|\cdot|$ is the modulo operator and $\gamma$ denotes the system's signal-to-noise power ratio (SNR). As for MIMO systems, the capacity can be calculated according to [12], [71]

$$C_{\mathrm{MIMO}} = \log_2\left(\det\left(\boldsymbol{I}_{N_r} + \frac{\gamma}{N_t}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}\right)\right) \ [\mathrm{bit/s/Hz}], \tag{4}$$

where $\boldsymbol{I}_N$ denotes the $N \times N$ identity matrix and det is the matrix determinant. It can be seen from (3) and (4) that the capacities defined above are random variables, and are not suitable for performance evaluation. Hence, the ergodic capacity (the expectation of the system capacity) [71] and the capacity outage (the capacity supported 90% or 99% of the time period) [72], [73] are typically employed for performance investigation.

### B. OFDM

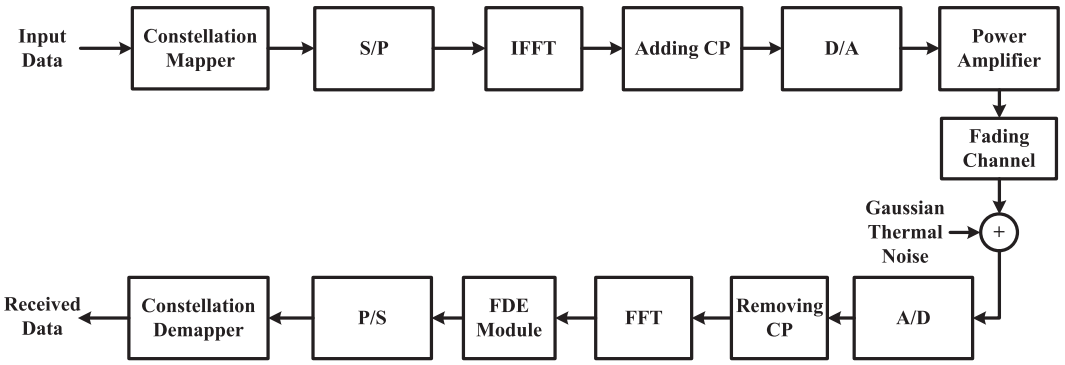
As shown in Fig. 4(a), binary bits are modulated by an APM mapper, such as PAM, PSK or QAM, to generate the complex-valued FD signal block $\mathbf{X} = [X(0) \ X(1) \cdots X(N-1)]^{\mathrm{T}}$, which is then passed through a serial-to-parallel converter (S/P) before fed into the *N*-point inverse fast Fourier transform (IFFT) module to yield the TD signal block $\boldsymbol{x} = [x(0) \ x(1) \cdots x(N-1)]^{\mathrm{T}}$. Next cyclic prefix (CP) of length $L_{\mathrm{CP}}$ is appended to $\boldsymbol{x}$ to combat the inter-block interference (IBI). This is followed by the digital-to-analog conversion (D/A) and power amplification before transmission through the single-input single-output (SISO) fading channel, whose channel impulse response (CIR) is defined by the channel coefficient vector $\boldsymbol{h} = [h_0 \ h_1 \cdots h_{L_h-1}]^{\mathrm{T}}$, where $L_h$ is the length of the CIR. The channel coefficients $h_i$ for $0 \le i \le L_h - 1$ are assumed to follow $\mathcal{CN}(0, 1/L_h)$. To remove completely the IBI, we must have $L_{\mathrm{CP}} \ge L_h$.

At the receiver, after the inverse operations, including the analog-to-digital conversion (A/D), the CP removal and the *N*-point fast Fourier transform (FFT), the received FD signal block $\mathbf{X}_r = [X_r(0) \ X_r(1) \cdots X_r(N-1)]^{\mathrm{T}}$ is generated as

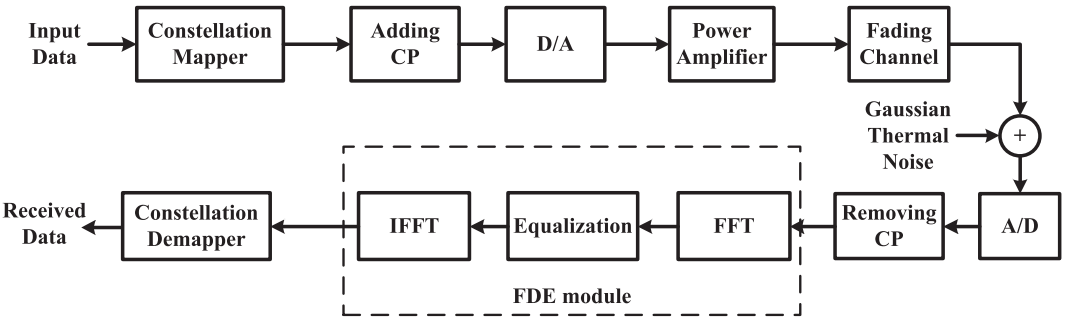
$$\mathbf{X}_r = \mathrm{diag}\{\mathbf{X}\}\mathbf{H} + \mathbf{W}, \tag{5}$$

where $\mathrm{diag}\{\mathbf{X}\}$ stands for the diagonal matrix whose diagonal elements are the elements of $\mathbf{X}$ and $\mathbf{H} = [H(0) \ H(1) \cdots H(N-1)]^{\mathrm{T}}$ is the FD channel coefficient vector obtained by the *N*-point FFT of $\boldsymbol{h}$, while $\mathbf{W} = [W(0) \ W(1) \cdots W(N-1)]^{\mathrm{T}}$ is the equivalent FD AWGN vector. Using channel estimation, $\boldsymbol{h}$ and hence $\mathbf{H}$ can be estimated to construct the one-tap FDE, and the equalized signals are fed into the corresponding constellation demapper for demodulation.

Owing to the IFFT at TX, OFDM signals suffer from high PAPR, which results in severe nonlinear distortions caused by the limited dynamic range and nonlinearities in the power amplifier for RF communications [74] or in the LED emitter for VLC [75]. Therefore, numerous PAPR reduction techniques have been proposed, including hard clipping [76], companding [77], selective mapping (SLM) [78], partial transmit sequence (PTS) [79], and active constellation expansion (ACE) [80], etc., which present a trade-off between the PAPR and the spectral efficiency as well as the received SNR. Alternatively, the digital predistorter [81], [82] or the nonlinear detection scheme [83], [84] can be employed to address




(a) OFDM system diagram.




(b) SC-FDE system diagram.

Fig. 4. Transceiver models of OFDM and SC-FDE systems.

the high PAPR issue, at the expense of increased TX or RX complexity.

### C. SC-FDE

Fig. 4(b) illustrates the transceiver structure of SC-FDE, which transmits a single carrier modulated APM signal. Effective one-tap FDE is implemented at RX with the aid of FFT and IFFT operations [85], [86]. It has been shown that SC-FDE achieves similar BER performance, spectral and energy efficiencies as OFDM. Comparing Fig. 4(b) with Fig. 4(a), it is also clear that the total system complexity of SC-FDE is similar to that of OFDM but the TX complexity of SC-FDE is lower than that of the OFDM TX. This makes SC-FDE attractive for uplink transmission. Moreover, due to the use of single carrier, the PAPR of SC-FDE is much lower, which mitigates the need of power back-off [2] and helps to reduce the nonlinear distortion of power amplifiers. For this aforementioned merit, SC-FDE has also been introduced to VLC [11], where the nonlinearities of LED are severe, and it outperforms other conventional optical modulation schemes, in terms of SNR gains.

## III. Principles of IM-Aided Systems

IM can be performed in the SD on antennas, in the FD on subcarriers, in the TD on time slots, or in the CD on channel state variations. In this section, based on this classification, the frameworks of various IM-aided systems are illustrated, which are exemplified by the corresponding typical schemes.

### A. FD Index Modulation

FD-IM, also known as IM-OFDM, is capable of embedding additional information bits on the indices of the activated OFDM subcarriers without extra energy consumption, which enhances the energy efficiency. Below the IM-OFDM transceiver design is first illustrated. Then different IM-OFDM schemes are classified into single-mode, multi-mode, and non-RF classes, respectively. Afterwards, the signal detection algorithms for typical IM-OFDM schemes are introduced, which is followed by the performance analysis.

*1) Transceiver Structure:* The TX of IM-OFDM up to the creation of the FD signal block $\mathbf{X}$ is presented in Fig. 5. A total of $m$ information bits are divided equally into the $G$ groups of $p$ bits each, i.e., $m = G \cdot p$. Each group is further split into the two subgroups of $p_1$ and $p_2$ bits, respectively, called *index bits* and *symbol bits*. Correspondingly, the $N$ OFDM subcarriers $\mathbf{X}$ are also divided into the $G$ subblocks

$$\mathbf{X}^{(g)} = [X((g-1)L+1)\ X((g-1)L+2) \cdots X(gL)]^{\mathrm{T}}$$

for $1 \leq g \leq G$ with the subblock length of $L = N/G$. For the $g$th OFDM subblock, the indices of the activated subcarriers $\mathbf{i}_g$, referred to as *index pattern*, are determined by the corresponding index bits with the aid of an index selector, and the data symbols $\mathbf{s}_g$ are generated by the constellation mapper according to the symbol bits. Afterwards, by using an OFDM block creator, all the $G$ OFDM subblocks are concatenated together and fed into the subsequent stages which are the same as in a conventional OFDM transmitter.

At the RX, after the FFT operation, information bits can be demodulated by ML detection subblock by subblock, where

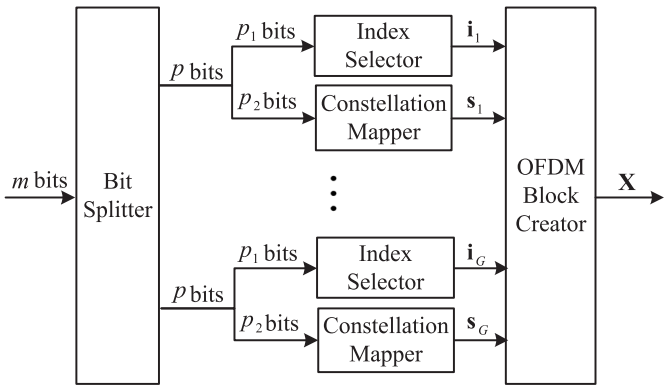



Fig. 5. Transmitter diagram of IM-OFDM.

all the possible realizations of the OFDM subblock, denoted as $\mathcal{X} = \{\widetilde{\mathbf{X}}_1, \widetilde{\mathbf{X}}_2, \cdots, \widetilde{\mathbf{X}}_{2^{p_1}}\}$, are considered to minimize the ED between the received subblock and its estimate. However, since exponentially increasing number of subblock realizations have to be taken into account, the complexity of the optimal ML detection may be too large for practical implementation. Therefore, sub-optimal strategies, such as log-likelihood ratio (LLR) detection, are proposed to reduce the computational overhead at the expense of a slightly degradation in performance. Explicitly, the indices of the activated subcarriers are firstly detected with the aid of LLR calculation subcarrier by subcarrier, and each data symbol is then demodulated by the corresponding demapper according to the estimated index pattern. Detection algorithms for typical IM-OFDM systems will be detailed in Section III-A5.

*2) Single-Mode Class:* In [39], subcarrier-index modulated OFDM (SIM-OFDM) was proposed, where part of the subcarriers are modulated by QAM, and the indices of these activated subcarriers in each OFDM symbol or block are determined by the corresponding majority bit-values of an on-off-keying (OOK) data stream. Since subcarrier grouping is not applied, the transmitter of SIM-OFDM is equivalent to Fig. 5 with $G = N$. The data rate of SIM-OFDM, however, is unstable due to the randomness of the OOK data stream, and this may impose potential error propagation, leading to bursts of errors. To address this problem, Tsonev *et al.* [40] proposed an enhanced SIM-OFDM (ESIM-OFDM) by partitioning the subcarriers into pairs. For each pair, only one subcarrier is activated, whilst the other one remains empty, where the index of the activated subcarrier carries one bit information. Although the BER performance of ESIM-OFDM is better, it suffers the loss of half of the index bits, compared with SIM-OFDM. Moreover, the system's frequency resources are under-utilized, since only one half of the subcarriers are activated for data transmission. In [41], a generalized ESIM-OFDM (GSIM-OFDM)[2] was proposed, where subcarriers are split into OFDM subblocks, and the subblock size is not constrained to be 2 as in ESIM-OFDM. Explicitly, in each OFDM subblock, $k$ out of $L$ subcarriers ($k < L$) are mapped by the constellation alphabet $\mathcal{M} = \{S_1, S_2, \ldots, S_M\}$, where $M$ is the order of the APM constellation, and the others are kept empty. These operations are exemplified by Table II, where $S_{i_1}$ and $S_{i_2}$ are arbitrary elements of $\mathcal{M}$. It is seen that for each OFDM subblock, the indices of the modulated subcarriers are determined by the two index bits, with the bit streams [0,0], [0,1], [1,0] and [1,1] corresponding to the index patterns [1,2], [2,3], [3,4] and [1,4], respectively. Hence, for GSIM-OFDM, $p_1$ and $p_2$ are formulated respectively as

TABLE II
A LOOK-UP TABLE OF GSIM-OFDM WITH $p_1 = 2, L = 4$ AND $k = 2$

| Index bits | Indices | Subblocks |
|---|---|---|
| $[0,0]$ | $[1,2]$ | $[S_{i_1}, S_{i_2}, 0, 0]$ |
| $[0,1]$ | $[2,3]$ | $[0, S_{i_1}, S_{i_2}, 0]$ |
| $[1,0]$ | $[3,4]$ | $[0, 0, S_{i_1}, S_{i_2}]$ |
| $[1,1]$ | $[1,4]$ | $[S_{i_1}, 0, 0, S_{i_2}]$ |

$$p_1 = \left\lfloor \log_2 \binom{L}{k} \right\rfloor = \left\lfloor \log_2 \left( \frac{L!}{(L-k)!k!} \right) \right\rfloor, \tag{6}$$

$$p_2 = k \log_2(M), \tag{7}$$

where $\lfloor \cdot \rfloor$ denotes the integer floor operator. Hence the spectral efficiencies of ESIM-OFDM and GSIM-OFDM are given by

$$\mathrm{SE}_{\mathrm{ESIM-OFDM}} = \frac{N(\log_2(M) + 1)}{2(N + L_{\mathrm{CP}})}, \tag{8}$$

and

$$\mathrm{SE}_{\mathrm{GSIM-OFDM}} = \frac{N\left(k \log_2(M) + \left\lfloor \log_2 \binom{L}{k} \right\rfloor\right)}{L(N + L_{\mathrm{CP}})}, \tag{9}$$

respectively. From (9), it is seen that $k$ and $L$ in GSIM-OFDM can be flexibly adjusted to obtain a high spectral efficiency, which has been investigated in [87]–[89]. Specifically, in [87], the optimal subcarrier grouping strategy was proposed according to the analytical expression of energy efficiency for GSIM-OFDM, while the optimal number of activated subcarriers per subblock was deduced by maximizing the energy efficiency [88]. Furthermore, a closed-form approximated BER upper bound for GSIM-OFDM was derived [89].

Recall that SIM-OFDM carries more index information than ESIM-OFDM due to using the variable number of activated subcarriers. Exploiting this feature, an enhanced GSIM-OFDM (EGSIM-OFDM) version 1 (EGSIM-OFDM1)[3] was proposed by altering the number of activated subcarriers in each OFDM subblock [42]. Explicitly, the number of activated subcarriers is selected from the set $\mathcal{K} = \{k_1, k_2, \ldots, k_T\}$, where $T$ is the size of $\mathcal{K}$. Therefore, the number of the index bits $p_1$ for EGSIM-OFDM1 can be derived as

$$p_1 = \left\lfloor \log_2 \left( \sum\nolimits_{i=1}^{T} \binom{L}{k_i} \right) \right\rfloor, \tag{10}$$

and its spectral efficiency is formulated as

$$\mathrm{SE}_{\mathrm{EGSIM-OFDM1}} = \frac{N\left\lfloor \log_2 \left( \sum_{i=1}^{T} \binom{L}{k_i} M^{k_i} \right) \right\rfloor}{L(N + L_{\mathrm{CP}})}, \tag{11}$$

which is enhanced significantly in comparison with GSIM-OFDM. To further improve the spectral efficiency,

[2]In [41], GSIM-OFDM is referred to as OFDM with IM (OFDM-IM).

[3]In [42], EGSIM-OFDM1 is referred to as OFDM-GIM1.

Fan *et al.* [42] also proposed an EGSIM-OFDM version 2 (EGSIM-OFDM2)[4] by conducting IM on the in-phase and quadrature (I/Q) components separately. Specifically, in each OFDM subblock, a fraction of the I/Q parts of the subcarriers are selected for modulation using PAM separately. Let the numbers of activated subcarriers for the I/Q components be $k_I$ and $k_Q$, which are modulated by $M_I$- and $M_Q$-PAM, respectively. The spectral efficiency of EGSIM-OFDM2 is given by

$$\mathrm{SE}_{\mathrm{EGSIM-OFDM2}} = \frac{N}{L(N+L_{\mathrm{CP}})} \times \left( k_I \log_2(M_I) + k_Q \log_2\left(M_Q\right) + \left\lfloor \log_2 \binom{L}{k_I} \right\rfloor + \left\lfloor \log_2 \binom{L}{k_Q} \right\rfloor \right). \tag{12}$$

Rather than conducting IM on I/Q components separately, Fan *et al.* [90] further improved EGSIM-OFDM2 by performing IM jointly on I/Q components, achieving an even higher spectral efficiency due to an enlarged number of index patterns. The spectral efficiency of this scheme, which we refer to as EGSIM-OFDM-JIQ, is formulated as [90]

$$\mathrm{SE}_{\mathrm{EGSIM-OFDM-JIQ}} = \frac{N\left\lfloor 2\left(\log_2\left(M^k\binom{L}{k}\right)\right)\right\rfloor}{L(N+L_{\mathrm{CP}})}, \tag{13}$$

where the numbers of activated subcarriers are both equal to $k$ for I/Q components, and both I/Q components are modulated by $M$-PAM. A similar scheme, called OFDM-HIQ-IM, was proposed in [91], where linear constellation precoding technique is utilized to attain extra diversity gain. Furthermore, in addition to altering the number of activated subcarriers, different constellation alphabets can be employed for different OFDM subblocks [43].

Moreover, GSIM-OFDM can be combined with MIMO to attain performance gain over conventional MIMO-OFDM, leading to MIMO-OFDM-IM [44], [45]. Additionally, in [92] and [93], a subcarrier-level interleaving technique was incorporated into GSIM-OFDM and MIMO-OFDM-IM, respectively, to enhance the system performance by increasing the minimum ED between different symbols. Similarly, a coordinate interleaved GSIM-OFDM was proposed in [94] by combining GSIM-OFDM with space-time codes and coordinate interleaving technique, in order to attain additional diversity gain. Besides, exploiting the sparsity property of the GSIM-OFDM frame structure, the compressive sensing technique was adopted in GSIM-OFDM [95], to increase both the energy and spectral efficiencies.

*3) Multi-Mode Class:* In the aforementioned single-mode schemes, such as GSIM-OFDM, only a fraction of subcarriers are modulated and, therefore, the precious frequency resources are unavoidably wasted. Moreover, as stated in [96], although GSIM-OFDM exhibits superior performance over conventional OFDM at low data rate, the gain disappears when the system's spectral efficiency is higher than 2 bit/s/Hz. This is because the index bits embedded in the subcarrier indices are only proportional to the logarithm of the subblock size $L$, which are far less than the symbol bits embedded in high-order APM constellation alphabet. This under-utilization of subcarriers may even lead to performance degradation for high throughput systems, compared to conventional OFDM.

To overcome this performance limitation of single-mode based schemes, dual-mode IM aided OFDM (DM-OFDM) was developed in [46], which fully utilizes the valuable spectral resource while maintaining the diversity gain of IM. Explicitly, the subcarriers are partitioned into subblocks, and all the subblocks are utilized for data transmission. Within each OFDM subblock, subcarriers are divided into two groups, A and B, modulated by two distinguishable constellation alphabets $\mathcal{M}_A$ and $\mathcal{M}_B$ with the sizes of $M_A$ and $M_B$, respectively. Additional index information can be conveyed by the indices of the subcarriers in Group A, for example. Without loss of generality, therefore, the indices of the subcarriers in Group A define *index pattern*, which are represented by $\mathcal{I}_{A,i}$ for the $i$th subblock, with $i = 1, 2, \ldots, G$. If $k$ subcarriers are modulated by $\mathcal{M}_A$, and the other $L-k$ subcarriers are mapped by $\mathcal{M}_B$, the spectral efficiency of DM-OFDM is expressed by

$$\mathrm{SE}_{\mathrm{DM-OFDM}} = \frac{N}{L(N+L_{\mathrm{CP}})} \times \left( k \log_2(M_A) + (L-k)\log_2(M_B) + \left\lfloor \log_2 \binom{L}{k} \right\rfloor \right). \tag{14}$$

By comparing (9) with (14), it is readily seen that a significant throughput gain is achieved by DM-OFDM over GSIM-OFDM. Since the index pattern of each OFDM subblock needs to be estimated at the RX, $\mathcal{M}_A$ and $\mathcal{M}_B$ must be differentiable, i.e., $\mathcal{M}_A \wedge \mathcal{M}_B = \varnothing$. Moreover, in order to achieve good BER performance, the minimum ED between two constellation points of the two alphabets should be equal to that of two constellation points within each alphabet. Therefore, a feasible constellation design strategy is to divide a $(M_A + M_B)$-QAM alphabet into two subsets, employed as $\mathcal{M}_A$ and $\mathcal{M}_B$. Fig. 6 illustrates an example of the two constellation alphabets in DM-OFDM, where the two quadrature phase shift keying (QPSK) sets $\mathcal{M}_A = \{-1-\mathrm{j}, 1-\mathrm{j}, 1+\mathrm{j}, -1+\mathrm{j}\}$ and $\mathcal{M}_B = \{1+\sqrt{3}, (1+\sqrt{3})\mathrm{j}, -1-\sqrt{3}, -(1+\sqrt{3})\mathrm{j}\}$ are obtained from an 8-QAM constellation.

By fully utilizing all the subcarriers for data transmission whilst maintaining the advantages of IM, DM-OFDM can achieve significant performance gain over other IM-OFDM schemes [46]. This motivates various works to extend DM-OFDM. By combining the ideas of EGSIM-OFDM1 and DM-OFDM, a generalized DM-OFDM (GDM-OFDM) [47] was proposed to further enhance the spectral efficiency, where the size of $\mathcal{I}_{A,i}$ is no longer constant but chosen from $\mathcal{K} = \{k_1, k_2, \ldots, k_T\}$. With this enhancement, additional diversity gain is achieved at the cost of higher detection complexity. Similar to EGSIM-OFDM1, the spectral efficiency of

[4] In [42], EGSIM-OFDM2 is referred to as OFDM-GIM2.

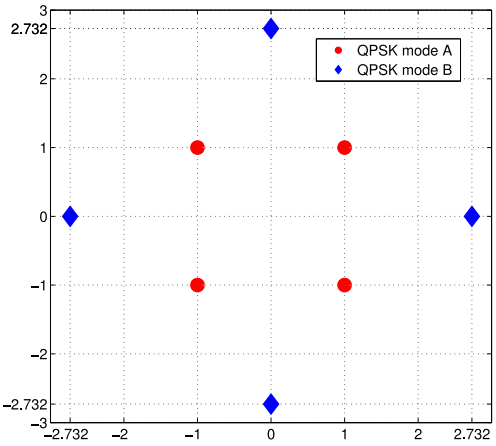



Fig. 6. An example of $\mathcal{M}_{\mathcal{A}}$ and $\mathcal{M}_{\mathcal{B}}$ in DM-OFDM with $M_A = M_B = 4$ [46].

GDM-OFDM can be derived as

$$\mathrm{SE}_{\mathrm{GDM-OFDM}} = \frac{N\left\lfloor \log_2\left(\sum_{j=1}^{T}\binom{L}{k_j} M_A^{k_j} M_B^{L-k_j}\right)\right\rfloor}{L(N+L_{\mathrm{CP}})}. \tag{15}$$

GSIM-OFDM and DM-OFDM designs trade-off between energy efficiency and spectral efficiency. DM-OFDM, by utilizing all the subcarriers for data transmission, attains higher throughput at the cost of additional energy consumption. GSIM-OFDM, by keeping some subcarriers unmodulated, sacrifices spectral efficiency for higher energy efficiency and, consequently, it may attain better BER performance. Based on these observations, a zero-padded tri-mode IM aided OFDM (ZTM-OFDM) [48] was developed, which employs two differentiable constellation alphabets to modulate only a fraction of the subcarriers, leading to less energy consumption than DM-OFDM. The name 'zero-padded tri-mode' comes from the fact that the subcarriers of each subblock are modulated by $\mathcal{M}_{\mathcal{A}}$ or $\mathcal{M}_{\mathcal{B}}$ or unmodulated, i.e., zero-padded. More specifically, in each OFDM subblock, only $k$ ($k < L$) subcarriers are selected for data transmission, which are partitioned into two subgroups of $k_1$ and $(k-k_1)$, subcarriers, modulated by $\mathcal{M}_{\mathcal{A}}$ and $\mathcal{M}_{\mathcal{B}}$, respectively. By choosing $k > k_1 > L/2$, ZTM-OFDM can have more index pattern realizations than DM-OFDM, i.e., more index bits, but its spectral efficiency, which is given by

$$\mathrm{SE}_{\mathrm{ZTM-OFDM}} = \frac{\log_2\left(M_A^{k_1} M_B^{k-k_1}\right) + \left\lfloor \log_2\left(\binom{L}{k} \times \binom{k}{k_1}\right)\right\rfloor}{L(1+L_{\mathrm{CP}}/N)}, \tag{16}$$

is generally lower than that of DM-OFDM.

Furthermore, the constellation modes can be enlarged to attain even higher spectral efficiency, leading to a multiple-mode IM-aided OFDM (MM-OFDM), which we propose here for the first time. Note that the MM-OFDM scheme given in [49] is only a very special and simplified case of our generic MM-OFDM.[5] In our MM-OFDM, all the subcarriers of each OFDM subblock are split into $n_{\mathrm{MM}}$ subsets modulated by $n_{\mathrm{MM}}$ distinguishable constellation alphabets. Assuming that the $i$th subset consists of $k_i$ subcarriers and modulated by $\mathcal{M}_i$ with the size of $M_i$ for $i = 1, 2, \ldots, n_{\mathrm{MM}}$, the spectral efficiency of MM-OFDM is given by

$$\mathrm{SE}_{\mathrm{MM-OFDM}} = \frac{\log_2 \prod_{i=1}^{n_{\mathrm{MM}}} M_i^{k_i} + \left\lfloor \log_2 \prod_{i=1}^{n_{\mathrm{MM}}} \binom{L-\sum_{j=0}^{i-1} k_j}{k_i}\right\rfloor}{L(1+L_{\mathrm{CP}}/N)}, \tag{17}$$

where $k_0 = 0$. From (17), it can be seen that MM-OFDM achieves considerable throughput gain over other IM-OFDM schemes, while its energy efficiency is also enhanced significantly, which can be attributed to using multiple constellation modes to convey more energy-free index bits. Owing to its superior spectrum and energy efficiency, MM-OFDM outperforms the other IM-OFDM counterparts, and can be widely applied to high-rate and low-power-consumption communication systems in 5G networks, VLC and Internet of Things (IoT). However, because of the enlarged number of alphabets, accurate detection of the index pattern for each OFDM subblock becomes difficult, and the multi-mode constellation design is very challenging.

Wen *et al.* [49] introduced a constellation mode selection strategy based on the *intra-mode distance* and *inter-mode distance*. The former is defined as the distance between any two subblocks with only one symbol error, while the latter is defined as the distance between any two subblocks with two symbol errors. At the high-SNR region, since the erroneous detection is more likely to be caused by a single symbol error between subblocks, the constellation design is to maximize the minimum intra-mode distance (MIAD), which is defined as

$$\begin{aligned} d_{\mathrm{intra}} = \min_{1\le i<j\le 2^{p_1}} & \left\|\widetilde{\mathbf{X}}_i - \widetilde{\mathbf{X}}_j\right\|^2, \\ \text{s.t.} \quad & E\left\{\left\|\widetilde{\mathbf{X}}_i\right\|^2\right\} = n_{\mathrm{MM}}, \\ & \mathrm{rank}\left(\mathrm{diag}\left\{\widetilde{\mathbf{X}}_i\right\} - \mathrm{diag}\left\{\widetilde{\mathbf{X}}_j\right\}\right) = 1, \end{aligned} \tag{18}$$

where $\|\cdot\|$ stands for the ED, $E\{\cdot\}$ denotes the expectation operator, and $\mathrm{rank}(\cdot)$ is the matrix rank. By contrast, at the medium-SNR region, the detection errors are more likely caused by double-symbol errors between subblocks than at the high-SNR case. Hence the minimum inter-mode distance (MIRD), which is defined by

$$\begin{aligned} d_{\mathrm{inter}} = \min_{1\le i<j\le 2^{p_1}} & \left\|\widetilde{\mathbf{X}}_i - \widetilde{\mathbf{X}}_j\right\|^2, \\ \text{s.t.} \quad & E\left\{\left\|\widetilde{\mathbf{X}}_i\right\|^2\right\} = n_{\mathrm{MM}}, \\ & \mathrm{rank}\left(\mathrm{diag}\left\{\widetilde{\mathbf{X}}_i\right\} - \mathrm{diag}\left\{\widetilde{\mathbf{X}}_j\right\}\right) = 2, \end{aligned} \tag{19}$$

should be maximized in order to design the constellation modes with good performance. A multiple-constellation mode design using PSK [49] is illustrated in Fig. 7, where $n_{\mathrm{MM}} = 8$, and $M_i = 2$ for $1 \le i \le 8$. The binary phase shift keying

[5] In [49], the length of subblock is restricted to $L = n_{\mathrm{MM}}$. In addition to convey more index bits, the number of index-pattern realizations for our MM-OFDM is dramatically higher than that of the scheme given in [49].

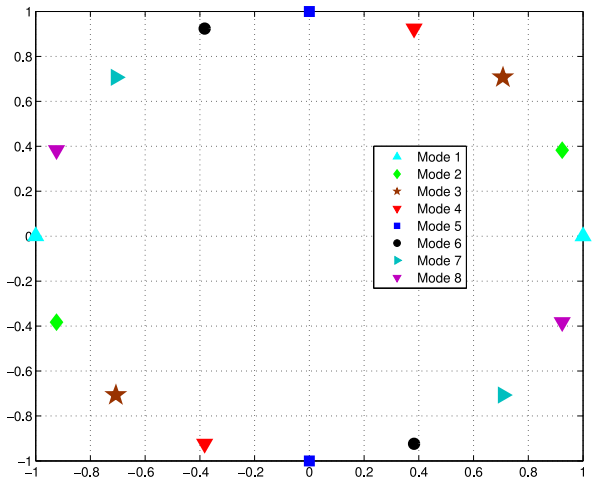



Fig. 7. An example of MM-OFDM constellation design with $M = 4$ and $n_{\mathrm{MM}} = 8$ [49].

(BPSK), a natural choice for the constellation order $M = 2$, is employed as the first constellation mode to generates a set of possible solutions with phase rotation by maximizing the MIAD. This is followed by selecting the rotation angles as $(j-1)\pi/n_{\mathrm{MM}}$ for $2 \le j \le 8$ to maximize the MIRD. The final 8-mode constellations are depicted in Fig. 7.

Note that the design of [49] is only capable of designing constellation modes in certain SNR area with certain modulation constellation. In order to achieve a globally-optimal constellation design with more general APM, there remain numerous challenges to be addressed.

*4) Non-RF Class:* GSIM-OFDM was adopted in underwater acoustic communications in [97], where the negative impacts of ICI on the system performance are illustrated. To address this issue, two adjacent subcarriers in each OFDM subblock are employed for data transmission with opposite polarity in order to combat ICI [97]. Furthermore, in [98], existent literature on ICI self-cancellation were reviewed, and a novel ICI reduction scheme was proposed for the application of IM-OFDM in underwater acoustic communications.

The IM technique was also introduced to VLC. Different from RF communications, intensity modulation with direct detection (ItM/DD) is usually employed in VLC, and thus only real-valued non-negative signals can be used for data transmission [1]. To achieve real-valued signals, Hermitian symmetry is imposed on the FD signal block $\mathbf{X} = [X(0)\ X(1) \cdots X(N-1)]^{\mathrm{T}}$ before IFFT operation as:

$$X(k) = X^*(N-k),\ \ k = 1, 2, \ldots, N/2-1, \qquad (20)$$

where $(\cdot)^*$ is the conjugate operator, and $X(0) = X(N/2) = 0$. In order to generate unipolar signals, several optical OFDM schemes were proposed, including DC-biased optical OFDM (DCO-OFDM) [99], asymmetrically clipped optical OFDM (ACO-OFDM) [100], PAM discrete multi-tone (PAM-DMT) [101], and unipolar OFDM (U-OFDM) [102] or Flip OFDM [103]. In DCO-OFDM, certain DC-bias is imposed on the bipolar signals to make them non-negative. Unlike DCO-OFDM, ACO-OFDM, PAM-DMT and U-OFDM generate unipolar signals without DC-bias by sacrificing half of the data rate. Explicitly, ACO-OFDM only modulates the odd subcarriers, and PAM-DMT only activates the imaginary parts of subcarriers, both leading to anti-symmetric TD signals before clipping at zero. For U-OFDM, the absolute values of the positive and negative signals are transmitted in two OFDM symbols subsequently. These conventional optical OFDM schemes suffer from additional energy consumption or data rate loss. These drawbacks can be effectively addressed by the IM technique.

In [104], an optical GSIM-OFDM (O-GSIM-OFDM) was proposed, which performs IM on the subblocks of DCO-OFDM or ACO-OFDM symbols before Hermitian symmetry operation. More explicitly, in each OFDM subblock, $k$ out of the $L$ subcarriers are modulated by the $M$-ary APM constellation, where the indices of the activated subcarriers convey additional information bits. With the CP length $L_{\mathrm{CP}}$, the spectral efficiencies of O-GSIM-OFDM based on ACO-OFDM and DCO-OFDM can be formulated respectively as

$$\mathrm{SE}_{\mathrm{O-GSIM-OFDM,DCO}} = \frac{(N-2)\left\lfloor \log_2\left(\binom{L}{k} M^k\right)\right\rfloor}{2(N+L_{\mathrm{CP}})L}, \qquad (21)$$

$$\mathrm{SE}_{\mathrm{O-GSIM-OFDM,ACO}} = \frac{N\left\lfloor \log_2\left(\binom{L}{k} M^k\right)\right\rfloor}{4(N+L_{\mathrm{CP}})L}. \qquad (22)$$

Then, Mao *et al.* [105] developed optical DM-OFDM (O-DM-OFDM) by combining DM-OFDM with DCO-OFDM or U-OFDM. In O-DM-OFDM, all the useful subcarriers are partitioned into subblocks, and all the subcarriers within each OFDM subblock are modulated by two different constellation alphabets. Specifically, $k$ subcarriers are mapped by $\mathcal{M}_A$ with size of $M_A$, and the other $(L-k)$ are mapped by $\mathcal{M}_B$ with size of $M_B$. Additional information bits are harvested from the indices of subcarriers modulated by $\mathcal{M}_A$ and $\mathcal{M}_B$. Hence, the spectral efficiencies of the O-DM-OFDM schemes based on DCO-OFDM and U-OFDM can be derived respectively as

$$\mathrm{SE}_{\mathrm{O-DM-OFDM,DCO}} = \frac{(N-2)\left\lfloor \log_2\left(\binom{L}{k} M_A^k M_B^{L-k}\right)\right\rfloor}{2(N+L_{\mathrm{CP}})L}, \qquad (23)$$

$$\mathrm{SE}_{\mathrm{O-DM-OFDM,U}} = \frac{(N-2)\left\lfloor \log_2\left(\binom{L}{k} M_A^k M_B^{L-k}\right)\right\rfloor}{4(N+L_{\mathrm{CP}})L}. \qquad (24)$$

O-DM-OFDM schemes achieve higher spectral efficiency than O-GSIM-OFDM counterparts, which leads to throughput gain. When the nonlinear transfer characteristics of LEDs are taken into account, simulation results have demonstrated the superiority of O-DM-OFDM over conventional optical OFDM schemes, in terms of practical implementation.

*5) Signal Detection for Typical IM-OFDM Schemes:* At the RX, the optimal ML detection can often be employed. Since ESIM-OFDM can be seen as the special case of GSIM-OFDM with $L = 2$ and $k = 1$, while GSIM-OFDM can be considered as the special case of DM-OFDM with either $\mathcal{M}_A$ or $\mathcal{M}_B$ set to $\{0\}$, the ML detectors for these three IM-OFDM

$$\gamma_{i_l}^{(j)} = \ln\left(\frac{\Pr\big(X((\beta-1)L+i_l)\in\mathcal{M}_j|X_r((\beta-1)L+i_l)\big)}{\sum_{q\neq j,q\in\mathcal{Q}^{(l)}}\Pr\big(X((\beta-1)L+i_l)\in\mathcal{M}_q|X_r((\beta-1)L+i_l)\big)}\right),\ \forall\mathcal{M}_j\in\mathcal{M}_{\text{all}}^{(l)} \tag{33}$$

schemes are similar. Specifically, by considering all the possible realizations of OFDM subblock, the data symbols and the corresponding index pattern of the $\beta$th OFDM subblock can be obtained by:

$$\left[\widehat{\mathbf{X}}^{(\beta)},\mathcal{I}^{(\beta)}\right] = \arg\min_{\widehat{\mathbf{X}}^{(\beta)}\in\mathcal{X},\mathcal{I}^{(\beta)}} \sum_{i=1}^{L}\left|X_r^{(\beta)}(i)-H^{(\beta)}(i)\widehat{X}^{(\beta)}(i)\right|^2, \tag{25}$$

where $\mathcal{I}^{(\beta)}$ is the index pattern of $\widehat{\mathbf{X}}^{(\beta)}$, $X_r^{(\beta)}(i)$ is the $i$th element of the $\beta$th received OFDM subblock and $H^{(\beta)}(i) = H(\beta L+i)$. The ML detection for EGSIM-OFDM, GDM-OFDM, ZTM-OFDM and MM-OFDM-IM are more complicated, since larger sets of OFDM subblock realizations have to be considered due to alterable number of activated subcarriers and/or multiple constellation modes.

The ML detector maximizes the minimum ED between the received OFDM subblock with different subblock realizations by exhaustive search, and its computational complexity increases significantly with the sizes of constellation alphabets, the numbers of activated subcarriers per subblock and the number of constellation modes. Therefore, for systems with large sizes of constellation alphabets and/or high number of activated subcarriers, the log-likelihood ratio (LLR) algorithm [46] is invoked for reduced-complexity detection at the cost of slight performance degradation. For the $i$th OFDM subcarrier of DM-OFDM, where $0 \le i \le N-1$, the logarithm of the ratio between the *a posteriori* probabilities for the events whether the subcarrier is modulated by $\mathcal{M}_A$ or $\mathcal{M}_B$ is calculated as

$$\gamma_i = \ln\left(\frac{\sum_{q=1}^{M_A}\Pr\big(X(i)=S_{A,q}|X_r(i)\big)}{\sum_{j=1}^{M_B}\Pr\big(X(i)=S_{B,j}|X_r(i)\big)}\right). \tag{26}$$

From (26), it is seen that the $i$th subcarrier is more likely to be activated by $\mathcal{M}_A$ if the LLR $\gamma_i > 0$, and vice versa. Noting the two *a priori* probabilities $\sum_{q=1}^{M_A}\Pr(X(i)=S_{A,q}) = k/L$ and $\sum_{j=1}^{M_B}\Pr(X(i)=S_{B,j}) = (L-k)/L$ as well as the fact of the channel AWGN having average energy $\sigma_w^2 = \mathrm{N_o}$, (26) can be simplified by applying Bayes' formula as

$$\gamma_i = \ln\frac{M_B k}{M_A(L-k)} + \ln\sum_{q=1}^{M_A}\exp\left(-\frac{\left|X_r(i)-H(i)S_{A,q}\right|^2}{\mathrm{N_o}}\right) - \ln\sum_{j=1}^{M_B}\exp\left(-\frac{\left|X_r(i)-H(i)S_{B,j}\right|^2}{\mathrm{N_o}}\right). \tag{27}$$

After the LLR calculation of the $N$ subcarriers using (27), for each OFDM subblock, there should be $k$ subcarriers with positive LLR values and the other $(L-k)$ subcarriers with negative LLRs, which define a *legitimate index pattern*. However, under extremely noisy conditions, there may exist illegitimate index patterns whose numbers of subcarriers with positive LLRs do not equal to $k$. Therefore, instead of hard decision with zero threshold, in each OFDM subblock, the largest $k$ LLR values are used to determine which $k$ subcarriers are modulated by $\mathcal{M}_A$, whilst the others are considered to be modulated by $\mathcal{M}_B$. After the estimation of index pattern for each OFDM subblock, the data symbols can be demodulated straightforwardly by the corresponding constellation demappers.

LLR detection for GDM-OFDM [47] is more complicated, since the number of activated subcarriers in each OFDM subblock is alterable, which is chosen from $\mathcal{K} = \{k_1, k_2, \ldots, k_T\}$. Therefore, the computational complexity of the LLR detection is $T$ times of the complexity required by DM-OFDM. Specifically, for each OFDM subblock, LLR detection is conducted assuming that the number of subcarriers modulated by $\mathcal{M}_A$ is $k_i$ for $1 \le i \le T$. Then the data symbols are demodulated according to the corresponding estimated index patterns, yielding the $T$ possible solutions $\widehat{\mathbf{X}}^{(\beta,k_i)} = [\widehat{X}^{(\beta,k_i)}(1)\ \ \widehat{X}^{(\beta,k_i)}(2)\cdots\widehat{X}^{(\beta,k_i)}(L)]$ for $1 \le i \le T$. Afterwards, the ML estimate of the data symbols and corresponding index pattern for the $i$th OFDM subblock can be determined based on these $T$ potential solutions as

$$\left[\mathcal{I}^{\left(\beta,\widehat{k}\right)},\widehat{\mathbf{X}}^{\left(\beta,\widehat{k}\right)}\right] = \arg\min_{\widehat{k}\in\mathcal{K}} \sum_{i=1}^{L}\left|X_r^{(\beta)}(i)-H^{(\beta)}(i)\widehat{X}^{\left(\beta,\widehat{k}\right)}(i)\right|^2. \tag{28}$$

LLR detection for ZTM-OFDM [48] can be partitioned into two steps: 1) perform LLR detection to find the indices of activated subcarriers, denoted as $\mathcal{I}_C = \mathcal{I}_A \vee \mathcal{I}_B$; 2) perform LLR detection again on the subcarriers corresponding to $\mathcal{I}_C$ to distinguish the subcarriers modulated by $\mathcal{M}_A$ from those modulated by $\mathcal{M}_B$.

Since the LLR detection algorithm of [49] cannot be applied to the MM-OFDM presented in this paper, we detail our LLR detection for this MM-OFDM system. Denote the initial detection constellation set and its corresponding index set by

$$\mathcal{M}_{\text{all}}^{(0)} = \{\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_{n_{\text{MM}}}\}, \tag{29}$$

$$\mathcal{Q}^{(0)} = \{1, 2, \ldots, n_{\text{MM}}\}, \tag{30}$$

respectively. The size of $\mathcal{Q}^{(0)}$ is denoted by $|\mathcal{Q}^{(0)}|$. For the $\beta$th subblock, the corresponding FD channel coefficients $[H((\beta-1)L+1)\ \ H((\beta-1)L+2)\cdots H(\beta L)]^{\mathrm{T}}$ are first sorted in descending order according to their moduli, yielding, $|H((\beta-1)L+i_1)| \ge |H((\beta-1)L+i_2)| \ge \cdots \ge |H((\beta-1)L+i_L)|$. An $L$-stage LLR calculation procedure is invoked to determine by which constellation mode the $i_l$th subcarrier is modulated for $l = 1, 2, \ldots, L$. Since the number of subcarriers modulated by $\mathcal{M}_j$ is $k_j$ for $1 \le j \le n_{\text{MM}}$ in each subblock, during the

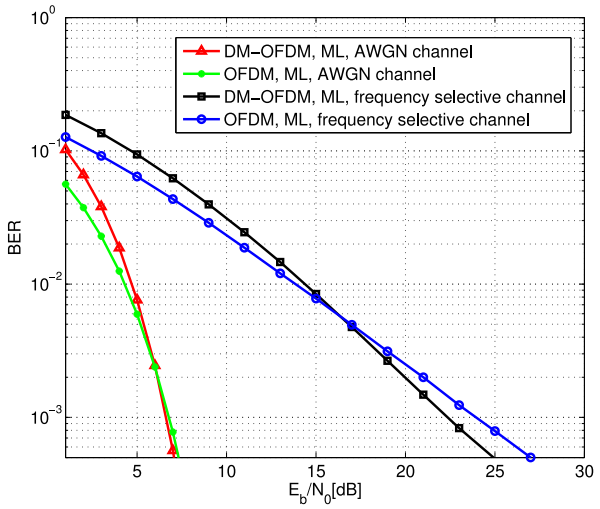



Fig. 8. BER performance comparison between DM-OFDM of 1.33 bits/s/Hz and conventional OFDM of 0.89 bits/s/Hz [46].

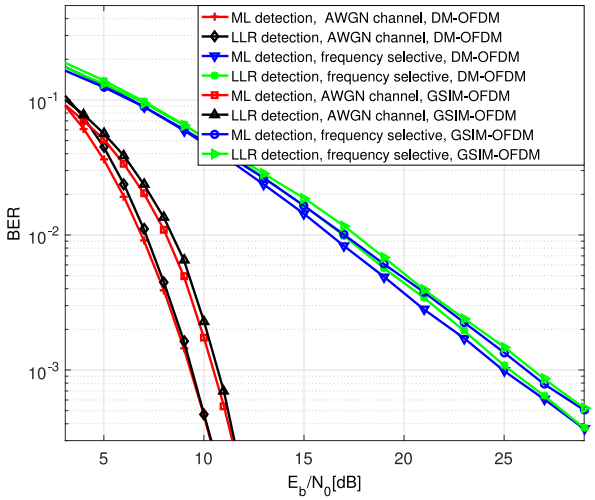



Fig. 9. BER performance comparison between the ML and LLR detectors for EGSIM-OFDM and DM-OFDM under both AWGN and frequency-selective Rayleigh fading channel conditions. The system's spectral efficiency is 2.22 bit/s/Hz.

iteration procedure, if the number of the estimated subcarriers modulated by $\mathcal{M}_{\widehat{j}}$ reach $k_{\widehat{j}}$, the constellation mode $\mathcal{M}_{\widehat{j}}$ should be removed from further consideration. Therefore, we define the constellation set and its corresponding index set at the $l$th iteration respectively as

$$\mathcal{M}_{\text{all}}^{(l)} = \begin{cases} \mathcal{M}_{\text{all}}^{(l-1)} \setminus \mathcal{M}_{\widehat{j}}, & k_{\widehat{j}} \text{ reached at } (l-1)\text{th iteration}, \\ \mathcal{M}_{\text{all}}^{(l-1)}, & \text{otherwise}, \end{cases} \tag{31}$$

$$\mathcal{Q}^{(l)} = \begin{cases} \mathcal{Q}^{(l-1)} \setminus \left\{\widehat{j}\right\}, & k_{\widehat{j}} \text{ reached at } (l-1)\text{th iteration}, \\ \mathcal{Q}^{(l-1)}, & \text{otherwise}. \end{cases} \tag{32}$$

At the $l$th iteration, where $1 \leq l \leq L$, the $|\mathcal{Q}^{(l)}|$ LLRs are calculated for the $i_l$th subcarriers, as given in (33) as shown at the top the previous page. Assume that the index of the largest LLR value among these $|\mathcal{Q}^{(l)}|$ LLRs is $\widehat{j}_l$. Then the $i_l$th subcarrier is considered to be modulated by $\mathcal{M}_{\widehat{j}_l}$. The iterative procedure is continued until $l = L$, and all the $L$ subcarriers are classified. After obtaining the index pattern, the index bits can be recovered with the aid of look-up table, and the symbol bits for each subcarrier are easily demodulated using ML detection.

*6) Performance Analysis:* With their superior spectrum and energy efficiency, IM-OFDM schemes, especially, multi-mode IM-OFDM schemes, are capable of achieving better BER performance over the conventional OFDM of the same throughput. In other words, IM-OFDM is capable of attaining a spectral efficiency gain over the conventional OFDM, while achieving a similar BER performance as the latter. Taking DM-OFDM for instance, Fig. 8 compares the BER performance of the DM-OFDM at 1.33 bit/s/Hz with those of the conventional OFDM at 0.89 bit/s/Hz under both the AWGN and the frequency-selective Rayleigh fading channel environments, where the ML detection is employed [46]. The system's SNR is defined as $\text{SNR} = \text{E}_\text{b}/\text{N}_0$, where $\text{E}_\text{b}$ denotes the average energy per bit and $\text{N}_0 = \sigma_w^2$ is the power spectral density of the channel's AWGN. In addition to the spectral efficiency gain of 0.44 bit/s/Hz, it can be observed that under the frequency-selective Rayleigh fading channel and given SNR $\geq$ 17 dB, the DM-OFDM outperforms the conventional OFDM. More specifically, at the BER level of $5 \times 10^{-4}$, the DM-OFDM attains around 2 dB SNR gain over the conventional OFDM. Under the AWGN channel with SNR $>$ 5 dB, 0.44 bit/s/Hz spectral efficiency gain is attained by the DM-OFDM over the conventional OFDM without suffering from a BER performance degradation.

The detection complexities per subblock of various IM-OFDM schemes using ML detection are compared in Table III, in terms of the number of complex multiplications required, where the operator $\lfloor\zeta\rfloor_2$ outputs the largest power-of-2 integer smaller than $\zeta$, and $\mathsf{O}(M)$ stands for on the order of $M$. From Table III, it can be seen that except for SIM-OFDM and ESIM-OFDM which suffer from the drawback of low spectral efficiency, all the other IM-OFDM schemes impose high computational complexity when employing ML detection. Specifically, the computational complexities of these schemes increase significantly with the systems' parameters, such as the subblock size, the number of activated subcarriers, and/or the constellation size. Therefore, the sub-optimal LLR detectors play an important role in realizing low-complexity IM-OFDM systems. For example, for the LLR-based GSIM-OFDM and DM-OFDM systems, the detection computational complexities, in terms of the number of complex multiplications required, are on the orders of $\mathsf{O}(LM)$ and $\mathsf{O}(L(M_A + M_B))$, respectively, which are dramatically lower than those imposed by ML detection. Moreover, as demonstrated by the simulation results given in [41] and [46], the LLR-based detectors for GSIM-OFDM and DM-OFDM only suffer from slight performance loss, compared to their optimal ML counterparts, which are also illustrated in Fig. 9. Specifically, given the same system's spectral efficiency of 2.22 bit/s/Hz, Fig. 9 compares the BER performance attained by the ML and LLR detectors for GSIM-OFDM and DM-OFDM under both AWGN and frequency-selective Rayleigh fading channel conditions. It can be seen that for both GSIM-OFDM and DM-OFDM, there is only marginal performance gap between the LLR detection

TABLE III
PERFORMANCE COMPARISON OF EXISTING IM-OFDM SCHEMES

| Scheme | Spectral efficiency (bit/s/Hz) | Energy cost | Detection complexity |
|---|---|---|---|
| **Single-Mode Class** | | | |
| SIM-OFDM [39] | unstable | low | $\mathsf{O}(M)$ |
| ESIM-OFDM [40] | $\frac{N}{2(N+L_{\mathrm{CP}})}\left(\log_2(M)+1\right)$ | low | $\mathsf{O}(2M)$ |
| GSIM-OFDM [41] | $\frac{N}{L(N+L_{\mathrm{CP}})}\left(k\log_2(M)+\left\lfloor\log_2\binom{L}{k}\right\rfloor\right)$ | low | $\mathsf{O}\left(\left\lfloor\binom{L}{k}M^k\right\rfloor_2\right)$ |
| EGSIM-OFDM1 [42] | $\frac{N}{L(N+L_{\mathrm{CP}})}\left\lfloor\log_2\left(\sum_{i=1}^{T}\binom{L}{k_i}M^{k_i}\right)\right\rfloor$ | medium | $\mathsf{O}\left(\left\lfloor\sum_{i=1}^{T}\binom{L}{k_i}M^{k}\right\rfloor_2\right)$ |
| EGSIM-OFDM2 [42] | $\frac{N}{L(N+L_{\mathrm{CP}})}\left(\log_2\left(M_I^{k_I}M_Q^{k_Q}\right)+\left\lfloor\log_2\binom{L}{k_I}\right\rfloor+\left\lfloor\log_2\binom{L}{k_Q}\right\rfloor\right)$ | medium | $\mathsf{O}\left(\left\lfloor\binom{L}{k_I}\binom{L}{k_Q}M_I^{k_I}M_Q^{k_Q}\right\rfloor_2\right)$ |
| EGSIM-OFDM-JIQ [90] | $\frac{N}{L(N+L_{\mathrm{CP}})}\left\lfloor 2\log_2\left(M^k\binom{L}{k}\right)\right\rfloor$ | medium | $\mathsf{O}\left(\left\lfloor\left(M^k\binom{L}{k}\right)^2\right\rfloor_2\right)$ |
| **Multi-Mode Class** | | | |
| DM-OFDM [46] | $\frac{N}{L(N+L_{\mathrm{CP}})}\left(\log_2(M_A^kM_B^{L-k})+\left\lfloor\log_2\binom{L}{k}\right\rfloor\right)$ | high | $\mathsf{O}\left(\left\lfloor\binom{L}{k}M_A^kM_B^{L-k}\right\rfloor_2\right)$ |
| GDM-OFDM [47] | $\frac{N}{L(N+L_{\mathrm{CP}})}\left\lfloor\log_2\left(\sum_{j=1}^{T}\binom{L}{k_j}M_A^{k_j}M_B^{L-k_j}\right)\right\rfloor$ | high | $\mathsf{O}\left(\left\lfloor\sum_{j=1}^{T}\binom{L}{k_j}M_A^{k_j}M_B^{L-k_j}\right\rfloor_2\right)$ |
| ZTM-OFDM [48] | $\frac{N}{L(N+L_{\mathrm{CP}})}\left(\log_2\left(M_A^{k_1}M_B^{L-k_1}\right)+\left\lfloor\log_2\left(\binom{L}{k}\binom{k}{k_1}\right)\right\rfloor\right)$ | medium | $\mathsf{O}\left(\left\lfloor\binom{L}{k}\binom{k}{k_1}M_A^{k_1}M_B^{k-k_1}\right\rfloor_2\right)$ |
| MM-OFDM (this paper) | $\frac{\log_2\left(\prod_{i=1}^{n_{\mathrm{MM}}}M_i^{k_i}\right)+\left\lfloor\log_2\left(\prod_{i=1}^{n_{\mathrm{MM}}}\binom{L-\sum_{j=0}^{i-1}k_j}{k_i}\right)\right\rfloor}{L(1+L_{\mathrm{CP}}/N)}$ | high | $\mathsf{O}\left(\left\lfloor\prod_{i=1}^{n_{\mathrm{MM}}}\binom{L-\sum_{j=0}^{i-1}k_j}{k_i}M_i^{k_i}\right\rfloor_2\right)$ |
| **Non-RF Class** | | | |
| O-GSIM-OFDM,DCO [104] | $\frac{(N-2)}{2(N+L_{\mathrm{CP}})L}\left\lfloor\log_2\left(\binom{L}{k}M^k\right)\right\rfloor$ | low | $\mathsf{O}\left(\left\lfloor\binom{L}{k}M^k\right\rfloor_2\right)$ |
| O-GSIM-OFDM,ACO [104] | $\frac{N}{4(N+L_{\mathrm{CP}})L}\left\lfloor\log_2\left(\binom{L}{k}M^k\right)\right\rfloor$ | low | $\mathsf{O}\left(\left\lfloor\binom{L}{k}M^k\right\rfloor_2\right)$ |
| O-DM-OFDM,DCO [105] | $\frac{(N-2)}{2(N+L_{\mathrm{CP}})L}\left\lfloor\log_2\left(\binom{L}{k}M_A^kM_B^{L-k}\right)\right\rfloor$ | high | $\mathsf{O}\left(\left\lfloor\binom{L}{k}M_A^kM_B^{L-k}\right\rfloor_2\right)$ |
| O-DM-OFDM,U [105] | $\frac{(N-2)}{4(N+L_{\mathrm{CP}})L}\left\lfloor\log_2\left(\binom{L}{k}M_A^kM_B^{L-k}\right)\right\rfloor$ | high | $\mathsf{O}\left(\left\lfloor\binom{L}{k}M_A^kM_B^{L-k}\right\rfloor_2\right)$ |

and the ML detection, and this gap becomes negligible in high-SNR region.

Table III also compares the spectral efficiencies and power consumptions of different IM-OFDM schemes. As expected, two trade-offs are witnessed. It can be seen that spectrum more efficient DM-OFDM, GDM-OFDM and MM-OFDM are capable of enhancing the achievable system's throughput significantly in comparison with GSIM-OFDM and its generalized versions. However, the former schemes tend to consume more energy than the latter schemes. This is the well understood trade-off between spectral efficiency and energy efficiency. Furthermore, the systems, which fully exploit the diversities in I/Q components of subcarriers, the number of constellation modes, and/or the number of activated subcarriers in each OFDM subblock, can attain considerable performance gain but inevitably suffer from complexity increase. This is the well-known trade-off between system performance and computational complexity. A well engineered communication system requires careful consideration of these two trad-offs.

The BER performance, an effective QoS indicator for communication systems, is commonly adopted to investigate IM-OFDM systems based on Monte Carlo simulation or theoretical analysis. Using Monte Carlo simulations, the BER performance of various IM-OFDM schemes can be evaluated in a straight-forward manner. Fig. 10 compares the BER performance of three IM-OFDM systems, MM-OFDM, DM-OFDM and GSIM-OFDM, with that of the conventional OFDM under the Rayleigh fading channel condition based on Monte Carlo simulations [49]. The system parameters are set as follows to provide similar spectral efficiency for different systems. Conventional OFDM: 32-QAM is used; GSIM-OFDM: 3 out of 4 subcarriers in each subblock are

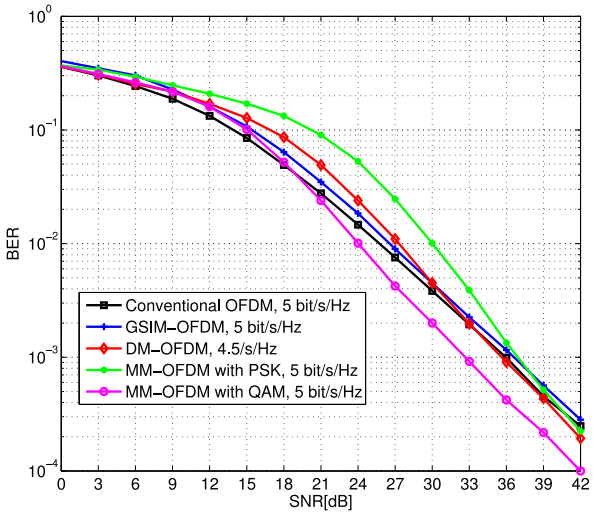



Fig. 10. BER performance comparison between MM-OFDM, GSIM-OFDM, DM-OFDM and conventional OFDM under the frequency-selective Rayleigh fading channel [49].

modulated by 64-QAM; DM-OFDM: 2 out of 4 subcarriers in each subblock are modulated by 16-PSK; MM-OFDM: 4 subcarriers in each subblock are modulated by 4 different PSK or QAM constellation alphabets. Firstly, from Fig. 10, it is seen that IM-OFDM schemes are capable of attaining the SNR gain over the conventional OFDM, at the BER level of $10^{-3}$. Secondly, MM-OFDM is capable of achieving significant performance gain over GSIM-OFDM and DM-OFDM, due to its superior spectral efficiency. Thirdly, it can be seen that the MM-OFDM with PSK constellations suffers from serious performance loss compared to its QAM counterpart, which indicates that appropriate constellation design is particularly crucial for MM-OFDM to achieve superior BER performance.

For IM-OFDM schemes with ML detection in a subblock-by-subblock manner, their BER performances are mainly dependant on the minimum ED $d_{\min}$ between the different possible realizations of OFDM subblock, normalized by the square root of $\mathrm{E_b}$, given by

$$d_{\min} = \min_{1\le i<j\le 2^{p_1}} \sqrt{\frac{1}{\mathrm{E_b}}}\left\|\widetilde{\mathbf{X}}_i - \widetilde{\mathbf{X}}_j\right\|, \tag{34}$$

where the average bit energy $\mathrm{E_b}$ can be calculated as

$$\mathrm{E_b} = \frac{(N+L_{\mathrm{CP}})\mathrm{E_s}}{pG}, \tag{35}$$

in which $\mathrm{E_s}$ denotes the average symbol energy. Since a reduced-complexity LLR detector suffers from negligible performance loss compared to its optimal ML counterpart at high SNRs, the minimum ED metric $d_{\min}$ is also capable of characterizing the performance of LLR detection.

Furthermore, PEP is also widely used to evaluate IM-OFDM systems, which is defined as the probability that one transmitted OFDM subblock is mistakenly detected as another at RX. To estimate the BER, it is sufficient to investigate only a single subblock, since the PEP events are identical for different OFDM subblocks [41], [46]. The conditional PEP (CPEP) of $\mathbf{X}^{(\beta)}$ being wrongly detected as $\widehat{\mathbf{X}}^{(\beta)}$ given the knowledge of the channel state information (CSI) $\mathbf{H}^{(\beta)}$ is defined by

$$\Pr\left(\mathbf{X}^{(\beta)} \to \widehat{\mathbf{X}}^{(\beta)}|\mathbf{H}^{(\beta)}\right) = Q\left(\sqrt{\frac{\mathrm{E_s}\rho}{2\mathrm{N_o}}}\right), \tag{36}$$

where $Q(\cdot)$ is the standard Gaussian Q-function, $\mathbf{X}^{(\beta)}$ and $\widehat{\mathbf{X}}^{(\beta)}$ are arbitrary elements of $\mathcal{X}$, while $\rho$ is calculated by

$$\rho = \left\|\mathrm{diag}\left\{\mathbf{X}^{(\beta)} - \widehat{\mathbf{X}}^{(\beta)}\right\}\mathbf{H}^{(\beta)}\right\|. \tag{37}$$

Since there is no closed-form solution for the CPEP in (36), several schemes were proposed to approximate the Q-function [106]–[110]. In [106], a tight but complicated approximation of the Gaussian Q-function is formulated as

$$Q(x) \approx \frac{1}{\sqrt{2\pi}\left(0.661x + 0.339\sqrt{x^2+5.510}\right)} e^{-x^2/2}, \tag{38}$$

for $x > 0$. To simplify the expression, the works [107], [108] propose an exponential bound of the Q-function less accurate than (38) in general, which is expressed as

$$Q(x) \approx \frac{1}{12}e^{-x^2/2} + \frac{1}{4}e^{-2x^2/3}. \tag{39}$$

The approximation (39) is considered to be a tight upper bound for $x > 0.5$, and it is widely adopted in IM-OFDM schemes for its simplicity [41], [46]. To enhance the approximation accuracy, two more expressions of the Q-function were proposed in [109] based on Prony approximation to the sum of exponential terms, which are given respectively by

$$Q(x) \approx 0.208e^{-0.971x^2} + 0.147e^{-0.525x^2}, \tag{40}$$

$$Q(x) \approx 0.168e^{-0.876x^2} + 0.144e^{-0.525x^2} + 0.002e^{-0.603x^2}, \tag{41}$$

where $x > 0$. In addition, to enhance the tightness of the approximation in the region of small $x$, a novel approximation of the erfc function, which is the derivative of the Q-function, was proposed in [110] as

$$\mathrm{erfc}(x) = \frac{1-e^{-1.98x}}{1.135}\frac{e^{-x^2}}{\sqrt{\pi}x}, \tag{42}$$

for $x \in [0, 20]$. For a comprehensive literature review on Q-function approximation, the reader is referred to [111].

Assuming that $\Pr(\mathbf{X}^{(\beta)} \to \widehat{\mathbf{X}}^{(\beta)}|\mathbf{H}^{(\beta)})$ is approximated as $f(x|\mathbf{H}^{(\beta)})$, we can calculate the unconditional PEP (UPEP) according to

$$\Pr\left(\mathbf{X}^{(\beta)} \to \widehat{\mathbf{X}}^{(\beta)}\right) = E_{\mathbf{H}^{\beta)}}\left\{f\left(x|\mathbf{H}^{(\beta)}\right)\right\}, \tag{43}$$

where $E_{\mathbf{H}}\{\cdot\}$ stands for the expectation with respect to $\mathbf{H}$. After obtaining the UPEP of any two possible realizations of OFDM subblock, the estimated BER, which is referred to as average bit error probability (ABEP), can be derived as [46]

$$\mathrm{ABEP} = \frac{1}{p\cdot 2^p}\sum_{\mathbf{X}^{(\beta)}\neq\widehat{\mathbf{X}}^{(\beta)}} \Pr\left(\mathbf{X}^{(\beta)} \to \widehat{\mathbf{X}}^{(\beta)}\right)\mathcal{E}\left(\mathbf{X}^{(\beta)}, \widehat{\mathbf{X}}^{(\beta)}\right), \tag{44}$$

where $\mathcal{E}(\mathbf{X}^{(\beta)}, \widehat{\mathbf{X}}^{(\beta)})$ denotes the number of bit errors for the event that $\mathbf{X}^{(\beta)}$ is mistakenly detected as $\widehat{\mathbf{X}}^{(\beta)}$.

*Remark:* Empirically, there exist inevitable and non-negligible estimation errors between the ABEP (44) and the Monte Carlo simulated BER at low SNRs, because of the approximation in deriving this ABEP. However, at the high-SNR region, the ABEP is a tight upper bound of the actual BER. In fact, for IM-OFDM systems with lower spectral efficiency, tighter BER bound is more likely to be achieved using PEP calculation, for the reason that there are less possible realizations of OFDM subblock, yielding reduced operations for the UPEP calculation with less approximations. This empirical observation is also demonstrated by the simulation results given in [41] and [46].

To validate the accuracy of our PEP analysis, we consider the ZTM-OFDM [48]. Fig. 11 compares its simulated BER performance with the corresponding ABEP results under the frequency-selective Rayleigh fading channel at the spectral efficiency of 2.22 bit/s/Hz and 1.33 bit/s/Hz, respectively. It is observed that, at low SNRs, there inevitably exists noticeable difference between the simulated results and analytical ABEP results, because of the approximation in the ABEP deviations. However, at high-SNR region, the difference between the ABEP results and the simulated BER disappears, which demonstrates the accuracy of the PEP analysis.

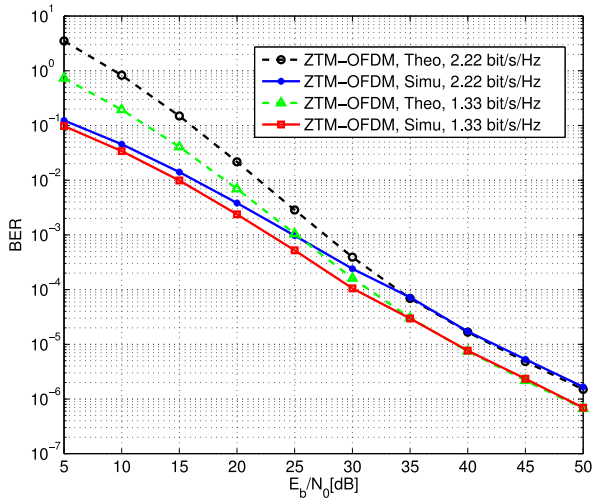



Fig. 11. Comparison between the analytical ABEP results and the simulated BER performance for the ZTM-OFDM at 2.22 bit/s/Hz and 1.33 bit/s/Hz, respectively, under the frequency-selective Rayleigh fading channel [48].

Before we close the discussion on FD based IM techniques, we point out IM-OFDM schemes have been applied in practical communication systems, owing to their advantages of low energy consumption and additional spectral efficiency gain as well as additional diversity gain in the subcarrier-index domain, over the conventional OFDM. For example, IM-OFDM was invoked in vehicle-to-vehicle and vehicle-to-infrastructure communications, where the interleaving technique was incorporated [112], and independent IM was performed on I/Q components to increase the index bits [113]. Physical-layer security techniques can be applied to FD-IM systems to enhance the systems' data secrecy [114], [115]. Specifically, for time division duplexing (TDD) protocol based systems, the CSI of the legitimate link can be utilized as the secret key by exploiting the channel reciprocity property. With the help of this legitimate CSI based secret key, the study [114] adopted joint optical subcarrier index selection and adaptive interleaving to enhance the secrecy, while maintaining reliable data transmission. The work [115] applied randomized mapping rules between the data symbol bits/index bits and the symbol/index pattern, resulting in a high BER at the eavesdropper.

### B. SD Index Modulation

SD-IM, also known as SM, is capable of providing high-rate and reliable data transmission with much lower system complexity compared to other MIMO systems. In SM, only a single TX antenna is employed for data transmission for each signaling time slot, where the activated antenna index carries additional $\lfloor \log_2 N_t \rfloor$ bits information. Since only one antenna is activated all the time, ICI is avoided, leading to reduced detection complexity. Furthermore, unlike standard spatial multiplexing MIMO using multiple antennas for simultaneous data transmission, which may require multiple expensive RF chains [65], only a single RF chain is required in SM, leading to low-hardware-complexity implementation. There already exist comprehensive literature reviewing SM, including [20], [64], [65]. Therefore, this paper will focus on typical SM schemes to help the reader better understand the principles and potentials of SM.

*1) SSK or SM:* The SM structure was first introduced in [50], in which the authors proposed the space shift keying (SSK) scheme to convey information bits by the index of single activated TX antenna. For example, when two antennas are employed at the TX, bits "0" and "1" can be transmitted by activating the 1st and 2nd antennas, respectively. Afterwards, generalized SSK (GSSK) was proposed in [51], where multiple antennas are activated to convey information bits, leading to considerable spectral efficiency gain but also causing severe IAS issue. Since then the interests in SSK grew, and many studies were conducted, including channel-coded and feedback-based design, performance evaluation via theoretical analysis, and system parameter optimization [37], [116]–[122].

At first, the difference between SSK and SM was ambiguous until 2009 when the concepts of SSK and SM were clearly classified [123]. It is now generally understood that in SSK, information bits are only conveyed by the index of single activated antenna and the actual transmitted signal does not carry information bits. By contrast, in SM, in addition to information bits conveyed by the index of single activated antenna, the actual transmitted signal is modulated by APM and, therefore, it carries information bits. According to this classification, SSK is viewed as a special case of SM. Unlike SSK, since both APM and IM are employed for data transmission in SM, it is more suitable for high-rate data transmission and has attracted great attentions from the research community. For instance, in [52], the combined scheme of SM and OFDM (SM-OFDM), was proposed, in which only one TX antenna is activated during each OFDM symbol transmission with extra information bits conveyed by the index of activated antenna.

*2) Typical SM Based Systems:* Compared with BLAST-type OFDM and Alamouti-type OFDM, SM-OFDM achieves dramatic reduction in computational complexity while attaining significant BER performance gain [52]. However, there exists a trade-off by only activating single TX antenna, very similar to the case of GSIM-OFDM. Specifically, although SM-OFDM achieve considerable reduction in system complexity and is free from ICI, the additional index bits conveyed by the index of activated antenna cannot compensate for the throughput loss due to empty TX antennas, particularly for high spectral-efficiency systems. In general, the design of an SM-based system is a careful trade-off between the system's complexity and achievable throughput or BER performance.

Two improved SM schemes, generalized SM (GSM) [53] and multiple active-spatial modulation (MA-SM) [54], both employing multiple $n_{\mathrm{rf}}$ TX antennas for data transmission, which enlarges the possible realizations of antenna activation pattern and attains higher diversity 'gain' over the classical SM.[6] By activating $n_{\mathrm{rf}}$ TX antennas simultaneously, rather than a single TX antenna, the antenna activation pattern of

[6]The concept of diversity here is not a classical one. SM systems have inherently capability of spatial diversity against the spatial channel fading.

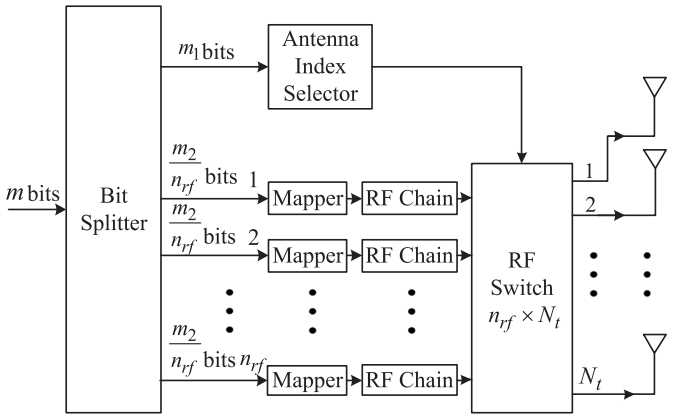



Fig. 12. Transmitter structure of MA-SM.

GSM and MA-SM is capable of conveying $\log_2 \binom{N_t}{n_{\rm rf}}$ index bits, compared to $\log_2 N_t$ bits for classical SM. The difference between GSM [53] and MA-SM [54] is that for GSM, all the activated antennas are modulated by the same symbol, whilst for MA-SM, different symbols are transmitted by different activated antennas. Fig. 12 presents the TX of MA-SM, where $m$ binary bits are firstly divided into $m_1$ index bits and $m_2$ symbol bits with the aid of a bit splitter. The $m_1$ index bits are fed into an antenna index selector to activate $n_{\rm rf}$ antennas for data transmission, while the $m_2$ symbol bits are divided equally into $n_{\rm rf}$ parallel streams of $\frac{m_2}{n_{\rm rf}}$ bits each, which are modulated by $n_{\rm rf}$ $M$-ary constellation mappers, respectively, to generate a $n_{\rm rf}$-dimensional transmit symbol vector. The elements of this transmit symbol vector are then allocated to the $n_{\rm rf}$ activated antennas according to the antenna activation pattern. Clearly, the index bits $m_1$ and the number of activated TX antennas $n_{\rm rf}$ is related by the relationship $m_1 = \lfloor \log_2 \binom{N_t}{n_{rf}} \rfloor$. The TX of GSM is similar to that of MA-SM except that the $m_2$ symbol bits are modulated by a single $M$-ary constellation mapper and the resulting transmit symbol is allocated to the $n_{\rm rf}$ activated antennas according to the antenna activation pattern. Thus, the data rates of GSM and MA-SM, in terms of bits per channel use (bpcu), are given respectively by

$$R_{\rm GSM} = \log_2 M + \left\lfloor \log_2 \binom{N_t}{n_{\rm rf}} \right\rfloor \ [\text{bpcu}], \tag{45}$$

$$R_{\rm MA-SM} = \log_2 M^{n_{\rm rf}} + \left\lfloor \log_2 \binom{N_t}{n_{\rm rf}} \right\rfloor \ [\text{bpcu}]. \tag{46}$$

Since in general $2^{m_1} < \binom{N_t}{n_{rf}}$, only a subset of $2^{m_1}$ activated antenna combinations, out of the total of $\binom{N_t}{n_{rf}}$ combinations, are used for data transmission. A feasible solution of selecting $2^{m_1}$ antenna activation patterns was proposed in [53], where the BER value is approximated using the union bounding technique [124], and the selection of activated antenna combinations is attained by minimizing the BER expression.

Attracted by its distinctive merits, MA-SM was applied to STBC systems [125] in order to attain extra spectral efficiency gain as well as to optimize the diversity and coding gains. In 2015, quadrature spatial modulation (QSM) was proposed to fully exploit both the in-phase and quadrature dimensions [55], in which SM is conducted on both the in-phase and quadrature components separately. More explicitly, for each signaling time instant, one $M$-QAM symbol is generated as usual but its real and imaginary parts are transmitted by two different activated antennas. By this arrangement, QSM is capable of doubling the diversity gain. Moreover, the ICI is completely avoided since the signals on two activated antennas are orthogonal. The data rate of QSM is formulated as

$$R_{\rm QSM} = 2\log_2 N_t + \log_2 M \ [\text{bpcu}]. \tag{47}$$

It can be seen that the idea of QSM is very similar to the aforementioned EGSIM-OFDM2 [42], where IM is performed on the I/Q components of each subcarrier separately. It is also worth pointing out that the idea of the improved EGSIM-OFDM2, namely, EGSIM-OFDM-JIQ [90], can also be applied to QSM in order to further enhance its data rate.

An interesting SM system design called precoded or preprocessed SM (PSM) was proposed in [58] and [59]. Unlike the aforementioned SM schemes which perform IM on TX antennas, RX antennas are considered in PSM, whose indices carry additional information bits. Specifically, a preprocessing matrix $\boldsymbol{P} = [\boldsymbol{p}_1\ \mathbf{p}_2 \cdots \boldsymbol{p}_{M_p}]$ containing the $M_p$ different preprocessing vectors is pre-stored at the TX and RX. For each signaling time instant, an $M$-ary transmitted symbol $x_t$ is multiplied by the chosen vector $\boldsymbol{p}_i$ from the preprocessing matrix, where the index of the chosen preprocessing vector conveys $\log_2 M_p$ binary bits. The optimization of $\boldsymbol{P}$ can be carried out by maximizing the throughput whilst minimizing the interferences on the other antennas [58], which is very complicated. One feasible and simple design of PSM is to apply precoding targeting at only single RX antenna, where $M_p = \lfloor \log_2 N_r \rfloor$, so that the received vector only has one non-zero element for each signaling time instance, and its index carries additional information. This is equivalent to performing IM on the RX antennas. This simplified system design is adopted in our description. Inspired by the idea of MA-SM, generalized PSM (GPSM) was proposed in [60] by activating multiple $n_{\rm rf}$ RX antennas for data transmission, where two practical issues, negative effects of imperfect CSI at the TX (CSIT) and the low-rank approximation employed for large-dimensional MIMO systems, were investigated, and a reinforcement matrix was applied to attain further performance gain. Clearly, PSM can be viewed as a special case of GPSM with $n_{\rm rf} = 1$. The TX structure of GPSM is shown in Fig. 13. Similar to MA-SM, at each signaling instant, $m$ binary bits are mapped onto $n_{\rm rf}$ $M$-ary symbols and the corresponding index pattern. The $n_{\rm rf}$ resultant symbols are preprocessed to produce an $N_r \times 1$ signal vector with the aid of the $n_{\rm rf} \times N_r$ RF switch, where the indices of non-zero elements correspond to the input index pattern, and additional information bits are conveyed by the activation pattern of the RX antennas. After an $N_r \times N_t$ precoding operation, the $N_t \times 1$ vector is fed to the TX antennas for data transmission. If $M$-ary constellation is employed for modulation, the data rate of GPSM is

$$R_{\rm GPSM} = \left\lfloor \log_2 \binom{N_r}{n_{\rm rf}} \right\rfloor + n_{\rm rf} \log_2 M \ [\text{bpcu}]. \tag{48}$$

Moreover, inspired by QSM, generalized precoded QSM (GPQSM) was proposed in [126], where both the I/Q components of the RX antennas are employed for separate

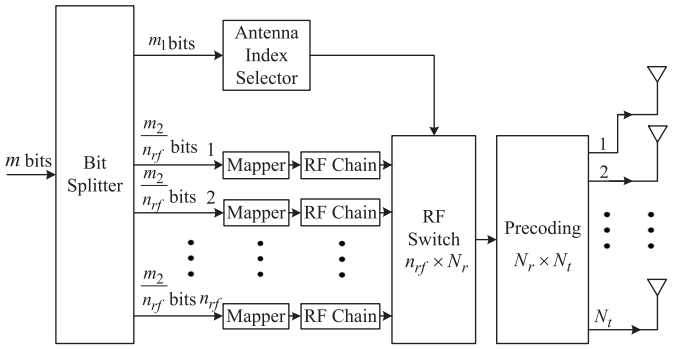



Fig. 13. Transmitter structure of GPSM.

IM. GPQSM doubles the index bits, compared to GPSM, and its data rate is formulated as

$$R_{\mathrm{GPQSM}} = 2\left\lfloor \log_2 \binom{N_r}{n_{\mathrm{rf}}} \right\rfloor + n_{\mathrm{rf}} \log_2 M \ \ [\mathrm{bpcu}]. \quad (49)$$

*Remark:* In comparison with the other SM schemes which perform IM on the TX antenna domain, the PSM-based schemes convey index information in the RX antenna domain. IM on the RX antenna domain is totally independent from IM on the TX antenna domain. Therefore, there is considerable scope for further research to perform IM simultaneously on the TX and RX antennas, which is capable of further enhance the achievable system's performance significantly, in terms of both throughput and diversity enhancements.

Although the data rates of the aforementioned modified SM systems have been enhanced in comparison with the standard SM, there still exists a throughput gap between these schemes and classical spatial multiplexing MIMO. For instance, given $N_t = N_r = 4$ and 16-QAM modulation, the data rate of spatial multiplexing MIMO is 16 [bpcu], whereas for MA-SM and GPSM with $n_{\mathrm{rf}} = 2$, the data rate equals to 10 [bpcu]. Even combined with QSM, the data rate of GPQSM can only be raised to 12 [bpcu]. It can be seen that this is similar to the case of GSIM-OFDM. Specifically, when constellations of lower order are utilized, the SM-based schemes are capable of achieving throughput gain over spatial multiplexing MIMO. However, for high-order constellations, spatial multiplexing MIMO provides higher-rate data transmission than its SM counterparts. This is because for high spectral efficiency systems, extra information bits conveyed by antenna activation index are insufficient to compensate for the throughput loss due to empty antennas.

To mitigate this inherent throughput loss of SM-based schemes, an enhanced SM (ESM) scheme with multiple antenna patterns and signal constellations was proposed in [56], where information bits are carried by different combination of antenna activation patterns and symbol constellations. Explicitly, for each signaling time instant, the number of activated antennas is variable and different constellation alphabets are employed, which considerably increases the possible realizations of the transmitted signal vector, leading to an enhanced throughput. For instance, for systems with $N_t = 2$ TX antennas, a possible antenna and constellation combination, referred to as a subblock, is designed by employing a primary constellation alphabet $\mathcal{M}_0$ when one TX antenna is activated, and two secondary alphabets $\mathcal{M}_1$ and $\mathcal{M}_2$ when both TX antennas are activated, where the size of $\mathcal{M}_0$ is *M* while the size of $\mathcal{M}_1$ and $\mathcal{M}_2$ is *M/2* so that information bits carried by different subblocks are equal. Table IV illustrates an ESM system with $N_t = 2$ (ESM-2TX), where QPSK primary constellation, $\mathcal{M}_0 = \{1+\mathrm{j}, 1-\mathrm{j}, -1-\mathrm{j}, -1+\mathrm{j}\}$, and two different BPSK secondary constellations, $\mathcal{M}_1 = \{1, -1\}$ and $\mathcal{M}_2 = \{\mathrm{j}, -\mathrm{j}\}$, are employed. The four possible antenna patterns are listed in the left column of Table IV, which convey two index bits, while the corresponding transmitted signal vectors are given in the right column of Table IV. More specifically, when only one antenna is activated for data transmission, QPSK is employed with data symbol $S_4 \in \mathcal{M}_0$ to generate two activation patterns. By contrast, when both antennas are used, two BPSK constellations are used, respectively, with $S_2^{(1)} \in \mathcal{M}_1$ and $S_2^{(2)} \in \mathcal{M}_2$ to produce two more activation patterns. In general, for ESM-2TX with the size of primary constellation equal to *M*, its data rate is given by

TABLE IV
A LOOK-UP TABLE OF ESM-2TX WITH QPSK PRIMARY CONSTELLATION AND TWO DIFFERENT BPSK SECONDARY CONSTELLATIONS

| Index bits | Transmitted vector |
|---|---|
| $[0,0]$ | $[S_4, 0]$ |
| $[0,1]$ | $[0, S_4]$ |
| $[1,0]$ | $[S_2^{(1)}, S_2^{(2)}]$ |
| $[1,1]$ | $[S_2^{(2)}, S_2^{(1)}]$ |

$$R_{\mathrm{ESM-2TX}} = \log_2 M + 2 \ [\mathrm{bpcu}]. \quad (50)$$

Clearly, the three constellations, $\mathcal{M}_0$, $\mathcal{M}_1$ and $\mathcal{M}_2$, must be mutually distinguishable in order to facilitate accurate demodulation at the RX.

The ESM technique can further be generalized by employing more TX antennas and more constellations, which is detailed in [56]. Moreover, in order to achieve higher throughput and better BER performance, novel design strategies were proposed to ESM in [57]. Since ESM signals are transmitted in a vector manner, the system BER is determined by the normalized minimum ED between the different possible realizations of the transmitted vector similar to (34). To attain SNR gain, therefore, one needs to enlarge the normalized minimum ED. To achieve this objective, a constellation design needs to ensure that the minimum ED between different transmitted vectors is no smaller than the minimum ED between the constellation points within each employed alphabet, and at the same time, to enable that more realizations of transmitted vector are generated to convey more index bits, which is equivalent to enlarge the normalized minimum ED. In the work [57], three designs were developed, namely, ESM-Type 1 and ESM-Type 2 as well as their combination called ESM-Type 3. In ESM-Type 1, a 4-TX-antenna ESM system is considered, where one *M*-QAM primary constellation and a distinguishable *M*/2-size secondary constellation are used for modulation. For each signaling time instant, only $n_{\mathrm{rf}} = 2$ antennas are activated, which are modulated by one primary

constellation symbol and one secondary constellation point, respectively. With this arrangement, each transmitted vector is capable of doubling index bits in comparison with GSM with $n_{\mathrm{rf}} = 2$. ESM-Type 2 attains even higher throughput gain by expanding the transmitted signal space. Specifically, for systems with $N_t = 4$, and $n_{\mathrm{rf}} = 2$, a union of 4 subspaces

$$\mathcal{L} = \{\mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3, \mathcal{L}_4\} \tag{51}$$

is designed from which the transmitted vector $\boldsymbol{s} \in \mathcal{L}$ is selected [57]. Each of the subspaces $\mathcal{L}_1$, $\mathcal{L}_2$ and $\mathcal{L}_3$ consists of $4$ different activation patterns. For each of these activation patterns, the two activated antennas are modulated by two differentiable constellation constellations $\mathcal{P}_{M/2}$ and $\mathcal{S}_{M/2}$, both having the size of *M/2*. Thus $\log_2 M$ symbol bits are conveyed by each of these three subspaces, and additional 2 prefix bits are conveyed by the index of selected subspace. As for the subspace $\mathcal{L}_4$, it contains 8 activation patterns, and two different constellations used for modulating $\mathcal{L}_4$ should also convey $\log_2 M$ symbol bits. Moreover, the activation patterns in $\mathcal{L}_4$ should be specially designed, so that the minimum ED between transmitted vectors is unchanged. It can be seen that the design of this 'expanded' subspace $\mathcal{L}_4$ is complicated, which can be found in [57].

Before concluding our presentation for typical SM-based systems, we briefly comment on their RX designs. Since the signals are transmitted in a vector form for almost all the SM-based systems, the optimal ML detection can be directly applied, similar to the ML detector for IM-OFDM schemes. Nevertheless, due to the significantly increased detection complexity, the optimal ML detector may not be suitable for practical implementation, and several reduced-complexity sub-optimal detectors were proposed as a trade-off [54], [125].

*3) Discussion:* Because of its promising performance as well as high energy efficiency, SM has been introduced to different applications, including wireless networking [127], [128] and VLC [129]–[133], etc. In various existing SM-based systems, IM is performed on different entities, such as the indices of TX and RX antennas, the I/Q components of transmitted signals, and the selection of used precoding vectors, etc. All these IM designs achieve significant diversity gains over classical spatial multiplexing MIMO. However, due to intentional empty TX antennas, the spatial resources are under-utilized in SM based systems, which may lead to inevitable throughput loss. To address this issue, considerable research efforts have been directed towards designing ESM systems, which attain considerable throughput gain at the cost of increasing design complexity and detection complexity. This in turn demands more effective system design and low-complexity detector design. In this survey, we have concentrated on a few typical SM based systems, with the aim of promoting the understanding of SM principle. In the literature, there exist numerous researches focusing on various aspects of SM based systems, including enhanced system design, performance evaluation and system optimization. The reader is referred to the existing survey papers [20], [64], [65] and the references therein for more details.

Data confidentiality is clearly an important issue for practical SM systems. Several works have been proposed to enhance the secrecy of SM systems [134]–[136]. Explicitly, jamming signals were used in [134] and [135], while the randomized mapping rule between the antenna activation pattern and index bits was utilized in [136], both approaches utilizing the CSI of the legitimate link as the secret key.

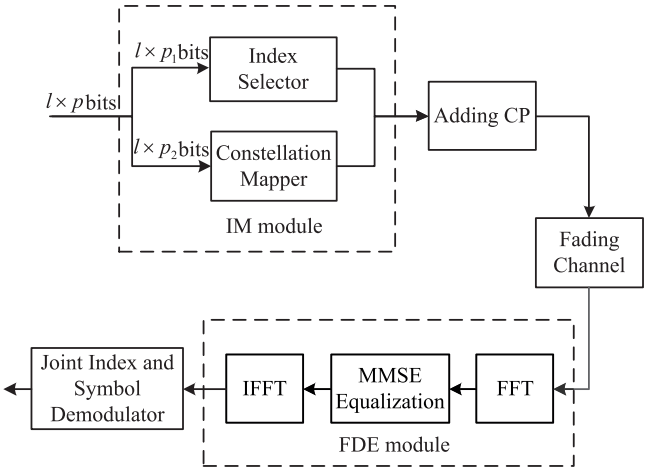



Fig. 14. Transceiver structure of FDE-aided SC-IM.

### C. TD Index Modulation

Various TD-IM schemes can be classified into two classes: schemes performing IM on time slots of signaling frame and schemes conducting IM on dispersion matrices of STBC system. According to this classification, below typical TD-IM schemes are briefly illustrated, in terms of their principles as well as the achievable system performance.

*1) IM Carried Out on Time Slots:* As mentioned previously, in the uplink of many current wireless standards, SC-FDE is usually employed for data transmission over SISO dispersive channel for its low PAPR and low TX complexity [2], [137], [138]. To enhance the energy efficiency of conventional SC-FDE, a pure TD-IM scheme called single-carrier IM (SC-IM) was proposed by performing IM on time slots of each transmit frame [26], where only part of the time slots are occupied for data transmission, and the others are kept empty. With this arrangement, in addition to the conventional APM bits, extra information bits are conveyed by the indices of the activated time slots with no energy consumption. The transceiver structure of the FDE-aided SC-IM system is depicted in Fig. 14. A total of $l \times p$ binary bits are partitioned into $l \times p_1$ bits for the index selector and $l \times p_2$ bits fed to the *M*-ary constellation mapper to generate the $l$ subframes $\boldsymbol{s}^{(i)} = [s_1^{(i)}\ s_2^{(i)} \cdots s_{l_s}^{(i)}]^{\mathrm{T}}$ for $1 \le i \le l$. Since only a fraction of the $l_s$ time slots in each subframe are utilized for data transmission, $s_j^{(i)}$ of the *j*th time slot can either be an *M*-ary symbol or 0, where $1 \le j \le l_s$. The overall transmission frame of length $N_s = l \times l_s$ is obtained by concatenating the $l$ subframes as $\boldsymbol{s} = [(\boldsymbol{s}^{(1)})^{\mathrm{T}}\ (\boldsymbol{s}^{(2)})^{\mathrm{T}} \cdots (\boldsymbol{s}^{(l)})^{\mathrm{T}}]^{\mathrm{T}}$. After adding the CP of length $L_{\mathrm{CP}}$, the resultant signals are transmitted through the fading channel, and the received signals are equalized by the one-tap FDE based on the MMSE criterion. Finally, the data symbols and corresponding index pattern are jointly estimated with ML detection. Assuming that *k* time

slots are activated in each subframe, the data rate of SC-IM is formulated by

$$R_{\mathrm{SC-IM}} = \frac{N_s\left(k\log_2 M + \left\lfloor \log_2 \binom{l_s}{k} \right\rfloor\right)}{(N_s + L_{\mathrm{CP}})l_s} \ [\mathrm{bit/frame}]. \quad (52)$$

It can be seen that TD-IM is equivalent to FD-IM, where time slots are analogous to subcarriers.

The dual-mode IM technique used in DM-OFDM can also be applied to SC-IM, yielding the dual-mode SC-IM (DM-SCIM) scheme [28], where the $l_s$ time slots of each subframe are activated by two distinguishable constellation alphabets. By assuming that $k$ time slots of each subframe are modulated by $M_A$-ary constellation, and the others are modulated by $M_B$-ary constellation, then its data rate can be calculated as

$$R_{\mathrm{DM-SCIM}} = \frac{N_s\left(\log_2 M_A^k M_B^{l_s-k} + \left\lfloor \log_2 \binom{l_s}{k} \right\rfloor\right)}{(N_s + L_{\mathrm{CP}})l_s} \ [\mathrm{bits/frame}]. \quad (53)$$

Additionally, the faster-than-Nyquist (FTN) rate sampling technique is incorporated to SC-IM and DM-SCIM, respectively, in [28] and [29] to enhance the data rate over the standard SC-IM and DM-SCIM, at the cost of higher intersymbol interference (ISI). For instance, unlike the conventional time-orthogonal transmission scheme based on Nyquist criterion, where the minimum symbol interval is $T_0 = \frac{1}{2}W$ with $W$ denoting the signal bandwidth, the FTN-IM of [29] shortens the symbol interval as $T = \alpha T_0$ in which $\alpha < 1$. By using a root raised cosine filter with a roll-off factor of $\beta$, the data rate of FTN-IM can be derived as

$$R_{\mathrm{FTN-IM}} = \frac{N_s\left(k\log_2 M + \left\lfloor \log_2 \binom{l_s}{k} \right\rfloor\right)}{\alpha(1+\beta)(N_s + L_{\mathrm{CP}})l_s} \ [\mathrm{bit/frame}]. \quad (54)$$

It is worth pointing out that the concept of TD-IM can be traced back to a much earlier modulation scheme called pulse-position modulation (PPM) [139], which is widely adopted in optical communications. Specifically, in the $N_p$-PPM, each of the symbol duration is partitioned into the $N_p$ time slots, where only one slot is employed to transmit a pulse, and the others are kept unused. Thus, instead of using the signal intensity to convey information, $\log_2 N_p$ information bits are conveyed by the position of the pulse. In other words, $\log_2 N_p$ bits are embedded in the index of the activated time slot.

SM and TD-IM can be combined, resulting in space-time IM (ST-IM) for MIMO [30]. More specifically, based on a standard SM system, each signaling frame is partitioned into $N_{\mathrm{ns}}$ time slots, and only $k$ slots are activated for each RF chain, and the others remain empty. Therefore, in ST-IM, in addition to the information bits conveyed by data symbol and spatial indexing of SM, extra information bits can be carried by the activation index of time slots. If we assume that $M$-ary constellation is used for modulation and $L_{\mathrm{CP}}$ CP symbols are added to each data stream at the TX, then the data rate of ST-IM can be represented by

$$R_{\mathrm{ST-IM}} = \frac{N_{\mathrm{ns}}\left(k\log_2 N_t + \left\lfloor \log_2 \binom{N_{\mathrm{ns}}}{k} \right\rfloor + k\log_2 M\right)}{N_{\mathrm{ns}} + L_{\mathrm{CP}}} \ [\mathrm{bpcu}], \quad (55)$$

in which $k\log_2 N_t$ is the number of the bits transmitted by SM, $\lfloor \log_2 \binom{N_{\mathrm{ns}}}{k} \rfloor$ is the number of the TD index bits, and $k\log_2 M$ bits are transmitted by $k$ $M$-ary symbols. At the RX, the optimal ML detector with high complexity can be used for joint detection. To reduce the detection complexity, a suboptimal 2-stage ST-IM detector was proposed in [30], which is quite similar to the two-stage LLR detector for GDM-OFDM systems. At the first stage, a MMSE estimator is employed to obtain the indices of the activated antennas, yielding the estimated index bits conveyed by SM. Afterwards, a message passing algorithm is applied in an iterative manner to find the *a posteriori* probabilities of the activated time slots, which are utilized to recover the data symbols of activated time slots and their corresponding TD index pattern. Jacob *et al.* [30] further proposed a 3-stage detector with message passing over a bipartite graph to improve the system reliability.

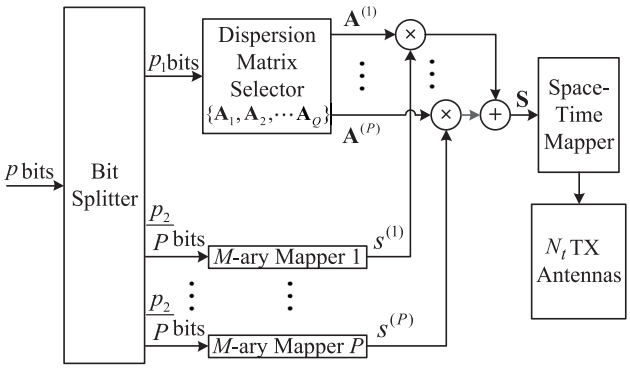



Fig. 15. Transmitter diagram of GSTSK.

*2) IM Carried Out on Dispersion Matrices:* Instead of using time slots, IM on the TD can be carried out in a very different but equivalent manner, leading to space-time shift keying (STSK) for MIMO [32]. Unlike SM, STSK is capable of striking a flexible trade-off of diversity gain and multiplexing gain. Moreover, like SM, the ICI is completely avoided in STSK based systems. Explicitly, for each signaling block, one out of the $Q$ dispersion matrices is selected for data transmission, whose index carries additional information bits. To further enhance the achievable system throughput, generalized STSK (GSTSK) [33] was developed by activating $P > 1$ dispersion matrices simultaneously, at the expense of increased detection complexity. Since STSK is a special case of GSTSK with $P = 1$, here only GSTSK is described in detail.

The TX structure of GSTSK employing $N_t$ TX antennas and $N_r$ RX antennas is illustrated in Fig. 15. For each block transmission, which is lasting for the duration of $N_{\mathrm{ns}}$ time slots, the $p$ information bits are partitioned into the $p_1$ index bits and $p_2$ symbol bits. The index bits are employed to determine the activation pattern of the dispersion matrices, namely, selecting the $P$ dispersion matrices $\{\boldsymbol{A}^{(1)}, \boldsymbol{A}^{(2)}, \ldots, \boldsymbol{A}^{(P)}\}$

from the set of the $Q$ pre-designed dispersion matrices

$$\mathcal{A} = \{\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_Q\}, \tag{56}$$

where $\boldsymbol{A}_i$ for $1 \leq i \leq Q$ are complex-valued $N_t \times N_{\rm ns}$ matrices. On the other hand, the $p_2$ symbol bits are further split equally into the $P$ sub-groups, which are modulated by the $P$ $M$-ary constellation mappers to generate the symbol sequence $\{s^{(1)}, s^{(2)}, \ldots, s^{(P)}\}$. The resultant symbols are multiplied by the corresponding activated dispersion matrices to generate the coded signal matrix $\boldsymbol{S} \in \mathbb{C}^{N_t \times N_{\rm ns}}$, which is given by

$$\boldsymbol{S} = \sum\nolimits_{i=1}^{P} s^{(i)} \boldsymbol{A}^{(i)}. \tag{57}$$

The signal matrix $\boldsymbol{S}$ is transmitted through the $N_t$ TX antennas over the block duration of $N_{\rm ns}$ time slots, and the received signal matrix $\boldsymbol{Y} \in \mathbb{C}^{N_r \times N_{\rm ns}}$ is represented by

$$\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{S} + \boldsymbol{W}, \tag{58}$$

where $\boldsymbol{H} \in \mathbb{C}^{N_r \times N_t}$ is the MIMO channel matrix and $\boldsymbol{W} \in \mathbb{C}^{N_r \times N_{\rm ns}}$ is the channel AWGN matrix. Given the estimated $\boldsymbol{H}$, ML detector can be employed to detect all the $p$ transmitted bits by considering all the possible combinations of the activated dispersion matrices and data symbols [33]. From the aforementioned description, $p_1$ and $p_2$ are given by

$$p_1 = \left\lfloor \log_2 \binom{Q}{P} \right\rfloor, \tag{59}$$

$$p_2 = P \log_2 M. \tag{60}$$

Then the data rate for GSTSK is derived as

$$R_{\rm GSTSK} = \frac{p_1 + p_2}{N_{\rm ns}} = \frac{\left\lfloor \log_2 \binom{Q}{P} \right\rfloor + P \log_2 M}{N_{\rm ns}} \ [\text{bpcu}]. \tag{61}$$

The structure of GSTSK is highly flexible and it includes all the existing MIMO structures, such as STBC, BLAST, SSK and SM, as its special cases [33], [140], [141]. Because of its unveiling advantages, the STSK technique has drawn much attention from the research community. In [142], a soft-decision-aided detector was proposed for STSK and GSTSK, which significantly reduces the detection complexity with only slight performance degradation in comparison with the optimal maximum *a posteriori* detector. STSK was further extended to include the FD, leading to the space-time-frequency shift keying (STFSK) scheme [143], where information bits are also modulated with frequency shift keying (FSK) on the FD. Unlike other SD-IM based schemes, which can only be applied to the narrow-bank MIMO channel of (1), STFSK is designed for the frequency-selective MIMO channel. In [32], the differential STSK was also proposed, where the signals are differentially encoded at TX with the aid of Cayley unitary transform, and hence the CSI is not required at the RX. As a continuous work, in [144], OFDM was amalgamated with differential STSK, and the iterative soft multiple-symbol differential sphere decoder was employed to obtain the SNR gain over the conventional soft-decision-aided differential detection method. More recently, Hemadeh *et al.* [145] generalized the concept of STSK to the multi-user MIMO scenario, and proposed the layered multi-group steered STSK (LMG-SSTSK) for MIMO downlink, which combines the multi-user TX beamforming, STSK and OFDM techniques for simultaneous data transmission to multiple users. Furthermore, multi-set STSK (MS-STSK) was proposed in [146]. In MS-STSK, in addition to the bits transmitted by the STSK codeword, at each subcarrier, a unique selection of TX antennas are activated for data transmission, and their indices carry extra information bits. Then, MS-STSK was extended to include the FD, resulting in multi-space-frequency STSK (MSF-STSK) [146]. Both MS-STSK and MSF-STSK achieve throughput and performance gains over the conventional STSK.

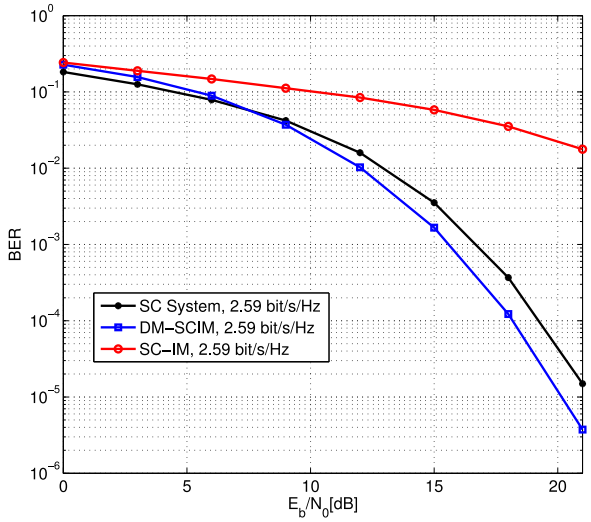



Fig. 16. BER performance comparison between the conventional SC system, SC-IM and DM-SCIM with two RX antennas under the frequency-selective Rayleigh fading channel [28].

*3) Discussion:* IM can be conducted alone in the TD, and a pure TD-IM scheme is SC-IM. Inspired by the dual-mode IM philosophy of FD-IM, DM-SCIM is proposed to enhance the achievable throughput of SC-IM. FTN signaling can also be incorporated with SC-IM for further performance gain. Fig. 16 compares the BER performance of the conventional SC system with those of the two typical TD-IM schemes, SC-IM and DM-SCIM, under the frequency-selective Rayleigh fading channel at the spectral efficiency of 2.59 bit/s/Hz [28]. To achieve this high spectral efficiency, the SC system adopts 8-PSK. For the SC-IM and DM-SCIM, the subblock length is 2, where a single subcarrier is activated by 32-ary constellation in SC-IM, while the two subcarriers are modulated by QPSK and 8-PSK in each subblock of DM-SCIM. Since a fraction of time slots in SC-IM are intentionally unused to convey index bits information, higher-order constellation alphabet has to be adopted to reach the same data rate as the conventional SC system and DM-SCIM, leading to inevitable high performance penalty. Consequently, SC-IM suffers from severe performance loss in comparison with the other two schemes, as can be seen from Fig. 16. Similar to GSIM-OFDM, SC-IM can outperform its non-IM-aided counterpart only at low data rate. By contrast, DM-SCIM can outperform the conventional SC system at high data rate, as can be observed from Fig. 16. Specifically, at the

TABLE V
PERFORMANCE COMPARISON OF TYPICAL TD-IM SCHEMES

| Scheme | Data rate | Detection complexity |
|---|---|---|
| SC-IM [26] | $\frac{N_s\left(k\log_2 M+\left\lfloor\log_2\binom{l_s}{k}\right\rfloor\right)}{\left(N_s+L_{\mathrm{CP}}\right)l_s}$ [bit/frame] | $\mathrm{O}\left(M^k\binom{l_s}{k}\right)$ |
| DM-SCIM [28] | $\frac{N_s\left(k\log_2 M_A+(l_s-k)\log_2 M_B+\left\lfloor\log_2\binom{l_s}{k}\right\rfloor\right)}{\left(N_s+L_{\mathrm{CP}}\right)l_s}$ [bit/frame] | $\mathrm{O}\left(M_A^k M_B^{l_s-k}\binom{l_s}{k}\right)$ |
| FTN-IM [29] | $\frac{N_s\left(k\log_2 M+\left\lfloor\log_2\binom{l_s}{k}\right\rfloor\right)}{\alpha(1+\beta)\left(N_s+L_{\mathrm{CP}}\right)l_s}$ [bit/frame] | $\mathrm{O}\left(M^k\binom{l_s}{k}\right)$ |
| ST-IM [30] | $\frac{N_{\mathrm{ns}}\left(k\log_2 N_t+\left\lfloor\log_2\binom{N_{\mathrm{ns}}}{k}\right\rfloor+k\log_2 M\right)}{N_{\mathrm{ns}}+L_{\mathrm{CP}}}$ [bpcu] | $\mathrm{O}\left(M^k N_t\binom{N_{\mathrm{ns}}}{k}\right)$ |
| GSTSK [33] | $\frac{\left\lfloor\log_2\binom{Q}{P}+P\log_2 M\right\rfloor}{N_{\mathrm{ns}}}$ [bpcu] | $\mathrm{O}\left(2N_tN_rN_{\mathrm{ns}}^2Q+(2N_rN_{\mathrm{ns}}P+N_rN_{\mathrm{ns}})\binom{Q}{P}N_t^P\right)$ |

BER level of $10^{-3}$, DM-SCIM attains 1 dB SNR gain over the conventional SC system.

IM can also be performed in both the SD and TD, and two examples are ST-IM and STSK or its generalized version GSTSK. However, the IM operations are carried out very differently by ST-IM and GSTSK. ST-IM can be viewed as the combined SM and SC-IM, where IM operations in the TD and in the SD are independent. This enables multi-stage detectors to reduce the detection complexity. By contrast, GSTSK signals are transmitted in a space-time block manner, where the indices of the activated dispersion matrices convey additional information, and IM operations in the SD and in the TD are correlated. This gives the high flexibility to GSTSK systems. For these three typical TD-IM schemes, Table V summarizes their achieved data rates and the ML detection complexities per frame or block, in terms of multiplications required. From the aforementioned descriptions, it is observed that STSK technique can be extended from single carrier system to multicarrier system by including IM on the FD. Apart from throughput and performance gains, these STFSK designs are particularly applicable to frequency-selective MIMO channel environments. For a comprehensive review of these STFSK based schemes, the reader is referred to the survey paper [147] and the references therein.

### D. CD Index Modulation

The concept of CD-IM, also known as MBM, was firstly proposed in [34]. In MBM, IM operations are not performed on the TX or RX entities, such as antennas, subcarriers or time slots, but on the variable channel states for conveying additional information [34]. Specifically, by placing special parasitic components, such as RF mirror [36] and electronic switches [37], close to the TX to perturb the rich scattering environment, numerous independent channel realizations are produced. These channel states, although unknown to the TX, can be made available to the RX through training. Therefore, the index of the used channel state can be utilize to convey additional information bits.

Compared to its SBIM counterparts, MBM can significantly enhances spectral and energy efficiency [34], [38]. For example, for SM based systems, the index bits increases with the logarithm of the TX antenna number, whilst for a MBM based system, the extra information bits that can be conveyed increase linearly with the number of RF mirrors. Since no energy is consumed for transmitting larger number of CD index bits, the system's energy efficiency is also naturally improved. Moreover, the RF mirrors can be closely placed and thus no space constraint is imposed on CD-IM. By contrast, for SD-IM schemes, such as SM and STSK, the distance between neighbouring antennas has to be sufficiently large.

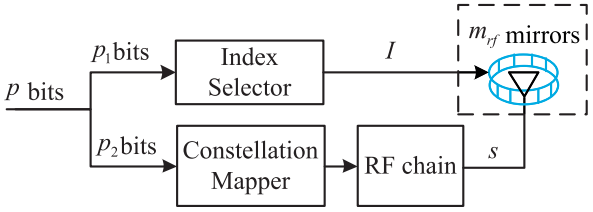



Fig. 17. Transmitter diagram of SIMO-MBM.

There also exist several challenges in realizing MBM [34]. An obvious difficulty is that since the number of possible channel realizations should be large in order to attain high spectral and energy efficiency, training overhead for acquiring these channel states at the RX can be huge. Other difficulties include the possible outage due to unavailability of CSIT and the randomness of the states, although MBM is reasonably robust to channel fading as deep fading does not have permanent effect on variable channel states [35].

Because the advantages clearly outweigh the disadvantages, MBM has become a promising novel technique to enhance the BER performance and the system throughput [148]. We classify various CD-IM systems into two groups: those employ single TX antenna, i.e., SIMO based systems, and those utilize multiple TX antennas, i.e., MIMO based systems.

*1) Single-TX-Antenna Based MBM:* We use SIMO-MBM [34] and differential MBM (DMBM) [61] to illustrate the basic concept of single-TX-antenna based MBM.

*a) SIMO-MBM:* The TX of SIMO-MBM system is characterized in Fig. 17, where a single RF chain is presented. A block of $p$ binary bits are divided into two sub-groups of $p_1$ and $p_2$ bits, respectively. The block of $p_1$ information bits are fed into the index selector to generate the index pattern $I$,

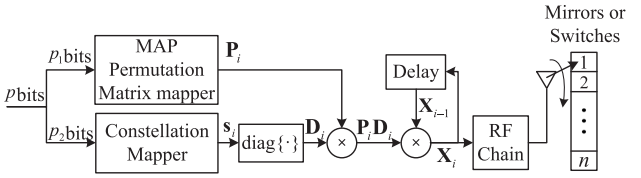



Fig. 18. Transmitter diagram for DMBM.

which serves as the control signal to determine the on-off status of the $m_{\rm rf}$ RF mirrors. This selects one used channel state out of the total of $2^{m_{\rm rf}}$ possible realizations. The block of $p_2$ information bits are modulated as usual by an $M$-ary constellation mapper, and the resultant signal $s$ is then sent to the TX antenna for data transmission.

Denote the set of all the $2^{m_{\rm rf}}$ possible channel states by

$$\mathcal{H}^{(c)} = \left\{ \boldsymbol{h}_1^{(c)}, \boldsymbol{h}_2^{(c)}, \ldots, \boldsymbol{h}_{2^{m_{\rm rf}}}^{(c)} \right\}, \tag{62}$$

where $\boldsymbol{h}_i^{(c)} \in \mathbb{C}^{N_r \times 1}$ is the $i$th SIMO channel realization, for $i \in \{1, 2, \ldots, 2^{m_{\rm rf}}\}$. Further assume that the index pattern corresponds to the $n$th channel state $\boldsymbol{h}_n^{(c)}$. Then the received signal vector $\boldsymbol{y} \in \mathbb{C}^{N_r \times 1}$ can be expressed as

$$\boldsymbol{y} = \boldsymbol{h}_n^{(c)} s + \boldsymbol{w}, \tag{63}$$

where $\boldsymbol{w} \in \mathbb{C}^{N_r \times 1}$ denotes the channel AWGN vector. In order to demodulate the MBM signal, including the index bits and symbol bits, the RX must have the CSI. That is, it must know which channel state is 'activated' or used for the current data transmission. In other words, the RX must have the knowledge of the channel state set $\mathcal{H}^{(c)}$. Therefore, pilots or training sequences are utilized for channel estimation. It can be seen that since the size of $\mathcal{H}^{(c)}$ increases exponentially with the number of RF mirrors $m_{\rm rf}$, training overhead imposes a practical limitation on $m_{\rm rf}$, i.e., $m_{\rm rf}$ cannot be too large. The data rate of SIMO-MBM is clearly given by

$$R_{\rm SIMO-MBM} = m_{\rm rf} + \log_2 M \ \ [\text{bpcu}]. \tag{64}$$

It was demonstrated in [38] that SIMO-MBM is capable of converting the SIMO channel into multiple parallel AWGN channels and, therefore, it is capable of approaching the capacity of random coding over an AWGN channel.

*b) DMBM:* It is seen that for the aforementioned SIMO-MBM, the CSI is required at the RX for demodulation, which is usually estimated by training. However, perfect CSI is impossible to acquire, since estimation errors inevitably occur during training. Moreover, for MBM systems employing a large number of RF mirrors, training overhead becomes impractically large. Inspired by differential SM [149], differential MBM (DMBM) was proposed in [61], so that CSI may not be required for demodulation at the RX.

Fig. 18 illustrates the TX of DMBM system equipped with single TX antenna and a number of RF mirrors or electronic switches.[7] The operations of DMBM is very different from other MBM schemes, such as SIMO-MBM. Specifically, differentially encoded DMBM signals are transmitted block-by-block with the block size of $n$, i.e., each block consists of $n$ time slots. Moreover, RF mirrors are repeatedly switched from state 1 to $n$ sequentially, so that the channel matrix experienced over the $i$th block can be represented by[8]

$$\boldsymbol{H}^{(i)} = \left[ \boldsymbol{h}_1^{(i)} \ \boldsymbol{h}_2^{(i)} \cdots \boldsymbol{h}_n^{(i)} \right] \in \mathbb{C}^{N_r \times n}. \tag{65}$$

The $i$th coded transmission block, represented by an $n \times n$ matrix, is generated as follows. First, the block of $p_1$ information bits are utilized by the mirror activation pattern (MAP) generator to produce the $n \times n$ permutation matrix $\boldsymbol{P}_i$ which has only one non-zero value in each column and in each row.[9] There are total of $n!$ permutations of the MAP $\boldsymbol{P}_i$, which represent the $n!$ possible channel states that may be experienced. On the other hand, the block of $p_2$ information bits are mapped onto the $n$ $M$-PSK symbols, denoted by $\boldsymbol{s}_i = [s_i(1) \ s_i(2) \cdots s_i(n)]^{\rm T} \in \mathbb{C}^{n \times 1}$. Then the $i$th uncoded signal block is represented by the $n \times n$ diagonal matrix $\boldsymbol{D}_i = \text{diag}\{\boldsymbol{s}_i\}$. Thus, the differentially encoded $i$th signal block for transmission over the $n$ time slots is given by

$$\boldsymbol{X}_i = \begin{cases} \boldsymbol{I}_n, & i = 0, \\ \boldsymbol{X}_{i-1} \boldsymbol{P}_i \boldsymbol{D}_i, & i \geq 1, \end{cases} \tag{66}$$

where $\boldsymbol{I}_n$ denotes the $n \times n$ identity matrix. Clearly, $\boldsymbol{X}_i$ has only one non-zero value in each column and in each row, and there are total of the $n!$ possible permutations of $\boldsymbol{X}_i$ in this sense. Since $p_1 = \lfloor \log_2(n!) \rfloor$ index bits and $p_2 = n \log_2 M$ symbol bits are transmitted over the $n$ time slots, the data rate of DMBM is given by

$$R_{\rm DMBM} = \log_2 M + \frac{1}{n} \lfloor \log_2(n!) \rfloor \ \ [\text{bpcu}]. \tag{67}$$

At the RX, the received signals are differentially demodulated block by block, without the need of CSI. Explicitly, two subsequent $N_r \times n$ received blocks $\boldsymbol{Y}_{i-1}$ and $\boldsymbol{Y}_i$ can be written respectively as

$$\boldsymbol{Y}_{i-1} = \boldsymbol{H}^{(i-1)} \boldsymbol{X}_{i-1} + \boldsymbol{W}_{i-1}, \tag{68}$$

$$\boldsymbol{Y}_i = \boldsymbol{H}^{(i)} \boldsymbol{X}_i + \boldsymbol{W}_i, \tag{69}$$

where $\boldsymbol{W}_i \in \mathbb{C}^{N_r \times n}$ denotes the corresponding channel AWGN matrix. Under the same classical assumption for conventional differential demodulation, namely, if the fading is sufficiently slow, we may have

$$\boldsymbol{H}^{(i-1)} = \boldsymbol{H}^{(i)}. \tag{70}$$

Consequently, we have

$$\boldsymbol{Y}_i = \boldsymbol{Y}_{i-1} \boldsymbol{P}_i \boldsymbol{D}_i + \widetilde{\boldsymbol{W}}_i, \tag{71}$$

where $\widetilde{\boldsymbol{W}}_i = \boldsymbol{W}_i - \boldsymbol{W}_{i-1} \boldsymbol{P}_i \boldsymbol{D}_i$ also follows a white Gaussian distribution. Therefore, the ML criterion can be

[7] The TX structure of the DMBM system presented in [61] is incorrect. Moreover, [61] did not correctly describe the operations of DMBM, specifically, how to generate the channel matrix over the transmission block.

[8] At the $l$th time slot of every block, although the RF mirrors are placed in the same $l$th state, the SIMO channel vectors experienced in the $l$th slot of two different blocks will generally be different owing to channel fading. Therefore, the block index should be indicated in the channel matrix $\boldsymbol{H}$.

[9] The term 'mirror activation pattern' originated from [61] is somewhat misleading, since the MAP $\boldsymbol{P}_i$ has nothing to do with the generation of the $i$th channel matrix $\boldsymbol{H}^{(i)}$. In fact, it is the single no-zero value of each column of $\boldsymbol{P}_i$, in other words, the index pattern, that determines the single channel coefficient experience in the corresponding time slot.

applied to estimate the transmitted data symbols and the corresponding index pattern, which is formulated as

$$\left[\widehat{\boldsymbol{P}}_i, \widehat{\boldsymbol{D}}_i\right] = \arg \min_{\forall \boldsymbol{P}_i, \boldsymbol{D}_i} \| \boldsymbol{Y}_i - \boldsymbol{Y}_{i-1} \boldsymbol{P}_i \boldsymbol{D}_i \|_F, \quad (72)$$

where $\| \cdot \|_F$ stands for the matrix Frobenius norm. With $\widehat{\boldsymbol{P}}_i$, the index bits for the *i*th block are recovered with the aid of look-up table, while the symbol bits can be demodulated with the demapper given $\widehat{\boldsymbol{D}}_i$. Furthermore, to reduce the computational overhead, a low-complexity ML detector based on the trace of matrix was also derived in [61].

It is worth pointing out that if fading is fast, the assumption (70) is no longer valid, and the ML detector (72) will suffer from serious performance degradation, leading to BER floor, just as classical differential detection systems. We further point out that the optimal multiple-symbol differential detection and the reduced-complexity multiple-symbol differential sphere detection [150]–[152] can be adopted to overcome the difficulty caused by fast fading channel, at the expense of increased detection complexity.

*2) Multiple-TX-Antenna Based MBM:* After the initial proposal of SIMO-MBM in [34] and [38], it has drawn considerable attention from researchers. It can be seen that to attain high throughput and diversity gains, the number of RF mirrors $m_{\mathrm{rf}}$ must be large. However, even under the static channel environment, training overhead is huge, since the possible channel states $2^{m_{\mathrm{rf}}}$ is huge. Moreover, huge storage space is required in RX to store these CSIs for signal detection.

*a) LMIMO-MBM:* To mitigate the aforementioned difficulties, MIMO, i.e., multiple TX antennas, can be applied to split the $m_{\mathrm{rf}}$ mirrors equally into the $N_t$ groups for each TX antenna. Thus, given the same data rate, the training task can be divided into the $N_t$ smaller tasks, reducing the overall training overhead significantly, and moreover, the detection complexity may be reduced considerably. Explicitly exploiting the desirable merits of combining the spatial multiplexing MIMO and the idea of MBM, layered MIMO-MBM (LMIMO-MBM) was proposed in [35].

Let the number of RF mirrors per TX antenna be $m_{\mathrm{rf}}^a$, and denote the set of all the $2^{m_{\mathrm{rf}}^a}$ possible SIMO channel states related to the *i*th TX antenna by

$$\mathcal{H}^{(i)} = \left\{ \boldsymbol{h}_1^{(i)}, \boldsymbol{h}_2^{(i)}, \ldots, \boldsymbol{h}_{2^{m_{\mathrm{rf}}^a}}^{(i)} \right\}, \quad (73)$$

where $\boldsymbol{h}_l^{(i)} \in \mathcal{C}^{N_r \times 1}$ denotes the *l*th realization of the *i*th SIMO channel vector. Further denote $s_i$ for $1 \leq i \leq N_t$ as the *M*-ary symbol transmitted through the *i*th TX antenna. Then the received signal vector $\boldsymbol{y} \in \mathcal{C}^{N_r \times 1}$ can be expressed as

$$\boldsymbol{y} = \sum\nolimits_{i=1}^{N_t} s_i \boldsymbol{h}_c^{(i)} + \boldsymbol{w}, \quad (74)$$

where $\boldsymbol{h}_c^{(i)} \in \mathcal{H}^{(i)}$ for $1 \leq i \leq N_t$ denote the SIMO CSIs actually used for the current transmission, i.e., they correspond to the current index pattern, while $\boldsymbol{w} \in \mathcal{C}^{N_r \times 1}$ is the MIMO channel AWGN vector. It can be seen that an $N_t \times N_r$ LMIMO-MBM system can be viewed as the superposition of the $N_t$ $1 \times N_r$ SIMO-MBM systems. Clearly, the data rate for LMIMO-MBM is given by

$$R_{\mathrm{LMIMO-MBM}} = N_t \log_2 M + N_t m_{\mathrm{rf}}^a \ [\mathrm{bpcu}]. \quad (75)$$

At the RX, ML detection can naturally be employed for signal demodulation. By denoting $\boldsymbol{s} = [s_1 \ s_2 \cdots s_{N_t}]^{\mathrm{T}}$ and $\boldsymbol{H}_c = [\boldsymbol{h}_c^{(1)} \ \boldsymbol{h}_c^{(2)} \cdots \boldsymbol{h}_c^{(N_t)}]$, the received signal vector (74) can be written as

$$\boldsymbol{y} = \boldsymbol{H}_c \boldsymbol{s} + \boldsymbol{w}, \quad (76)$$

and the optimal ML detection is carried out according to

$$\left[\widehat{\boldsymbol{H}}_c, \widehat{\boldsymbol{s}}\right] = \arg \min_{\forall \boldsymbol{H}_c, \boldsymbol{s}} \|\boldsymbol{y} - \boldsymbol{H}_c \boldsymbol{s}\|^2. \quad (77)$$

Afterwards, the index bits can be recovered based on $\widehat{\boldsymbol{H}}_c$ with the aid of look-up table, and the symbol bits are demodulated from $\widehat{\boldsymbol{s}}$ using the corresponding demapper. The computational complexity of the optimal ML detector (77) increases significantly with the system's parameters $N_t$, $M$ and $m_{\mathrm{rf}}^a$. An suboptimal detector based on greedy search was also given in [35], which largely reduces the detection complexity but suffers from inevitable performance loss. Moreover, from the information security perspective, artificial noise can be added to the transmitted signals to impose the interference only to the eavesdropper, leading to secrecy rate enhancement [153].

*b) STCM:* By combining Alamouti-STBC and MBM, space-time channel modulation (STCM) was proposed in [63], which can be viewed as a special case of LMIMO-MBM. Specifically, as in classical Alamouti-STBC, two complex symbols, denoted by $s_1$ and $s_2$, are transmitted through 2 TX antennas in two time slots, and the Alamouti $2 \times 2$ transmission matrix is given by

$$\boldsymbol{S} = \begin{bmatrix} s_1 & -s_2^* \\ s_2 & s_1^* \end{bmatrix}, \quad (78)$$

where $(\cdot)^*$ denotes the conjugate operation. Both TX antennas are equipped with $m_{\mathrm{rf}}^a$ RF mirrors, and different channel states can be selected in each time slot. Hence, additional index bits are conveyed by the corresponding MAPs. Given *M*-QAM symbol constellation, the data rate of STCM is given by

$$R_{\mathrm{STCM}} = \log_2 M + m_{\mathrm{rf}}^a \ [\mathrm{bpcu}]. \quad (79)$$

At the RX, similar to the more general case of LMIMO-MBM, the transmitted symbols and the corresponding index pattern can be demodulated using the ML detection [63].

*c) GSM-MBM:* As discussed previously, SM based schemes have low hardware overhead and lower energy consumption as well as are robust to ICI. In [36] and [62], the GSM scheme is combined with the MBM, leading to the GSM-MBM, where IM is performed on both the SD and CD. To achieve even higher performance gain in MIMO-MBM schemes, MAP selection based on ED metric, feedback-based phase compensation and constellation rotation were applied.

Fig. 19 depicts the TX of an $N_t \times N_r$ GSM-MBM system with $m_{\mathrm{rf}}^a$ RF mirrors per TX antenna. The block of $m$ information bits is split into three subblocks of $m_1$, $m_2$ and $m_3$ bits, respectively. The first subblock of $m_1$ bits is further divided

TABLE VI
PERFORMANCE COMPARISON OF TYPICAL MBM SCHEMES

| Scheme | Data rate | Detection complexity |
|---|---|---|
| SIMO-MBM [34] | $m_{\rm rf}+\log_2 M$ [bpcu] | $\mathsf{O}\left(M2^{m_{\rm rf}}\right)$ |
| DMBM [61] | $\log_2 M+\frac{1}{n}\lfloor\log_2(n!)\rfloor$ [bpcu] | $\mathsf{O}\left(2^{\lfloor\log_2(n!)\rfloor}M^n\right)$ |
| LMIMO-MBM [35] | $N_t\log_2 M+N_t m_{\rm rf}^a$ [bpcu] | $\mathsf{O}\left(M^{N_t}2^{N_t m_{\rm rf}^a}\right)$ |
| STCM [63] | $m_{\rm rf}^a+\log_2 M$ [bpcu[ | $\mathsf{O}\left(2^{2m_{\rm rf}^a+1}M\right)$ |
| GSM-MBM [62] | $n_{\rm rf}m_{\rm rf}^a+\left\lfloor\log_2\binom{N_t}{n_{\rm rf}}\right\rfloor+n_{\rm rf}\log_2 M$ [bpcu] | $\mathsf{O}\left(2^{n_{\rm rf}m_{\rm rf}^a}\left\lfloor\binom{N_t}{n_{\rm rf}}\right\rfloor_2 M^{n_{\rm rf}}\right)$ |

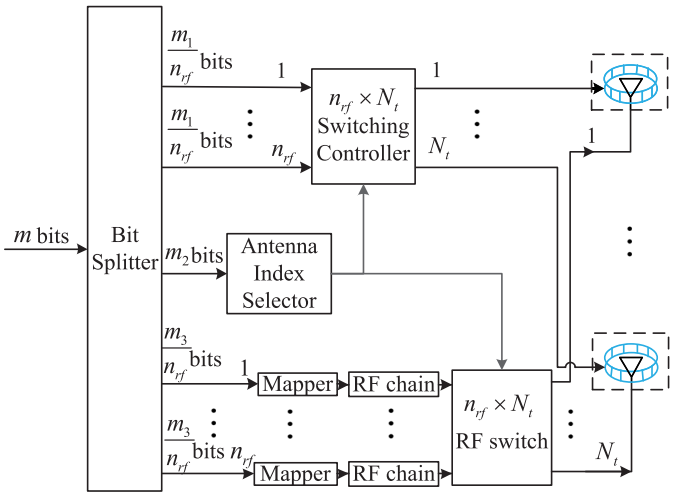



Fig. 19. Transmitter diagram for GSM-MBM.

into $n_{\rm rf}=m_1/m_{\rm rf}^a$ parallel streams to determine the index pattern of the RF mirrors for each TX antenna, with the aid of an $n_{\rm rf}\times N_t$ switching controller. The second subblock of $m_2$ bits is fed into an index selector to generate the antenna activation pattern, which determines which $n_{\rm rf}$ TX antennas, out of the total of $N_t$ TX antennas, are activated for data transmission. The third subblock of $m_3$ bits is partitioned equally into $n_{\rm rf}$ parts, and each part is modulated by an $M$-ary constellation before loaded onto the activated antennas. Clearly, $m_1$, $m_2$ and $m_3$ are given respectively by

$$m_1=n_{\rm rf}m_{\rm rf}^a, \tag{80}$$

$$m_2=\left\lfloor\log_2\binom{N_t}{n_{\rm rf}}\right\rfloor, \tag{81}$$

$$m_3=n_{\rm rf}\log_2 M. \tag{82}$$

Therefore, the data rate of GSM-MBM is given by

$$R_{\rm GSM-MBM}=n_{\rm rf}m_{\rm rf}^a+\left\lfloor\log_2\binom{N_t}{n_{\rm rf}}\right\rfloor+n_{\rm rf}\log_2 M\ [\text{bpcu}]. \tag{83}$$

The received signal $\boldsymbol{y}\in\mathbb{C}^{N_r\times 1}$ can be modeled by

$$\boldsymbol{y}=\sum\nolimits_{j=1}^{N_t}s_j\boldsymbol{H}^{(j)}\boldsymbol{e}_{m_j}+\boldsymbol{w}, \tag{84}$$

where $\boldsymbol{s}=[s_1\ s_2\cdots s_{N_t}]^{\rm T}$ consists of only $n_{\rm rf}$ non-zero elements which are $M$-ary modulated symbols, $\boldsymbol{H}^{(j)}=[\boldsymbol{h}_1^{(j)}\ \boldsymbol{h}_2^{(j)}\cdots\boldsymbol{h}_{2^{m_{\rm rf}^a}}^{(j)}]$ denotes the matrix containing all the possible channel state realizations for the $j$th TX antenna, and $\boldsymbol{e}_{m_j}$ stands for the $2^{m_{\rm rf}^a}\times 1$ vector with all zero elements except for the $m_j$th element which is equal to 1. Note that $\boldsymbol{e}_{m_j}$ picks up the $m_j$th channel realization from all the possible realizations of $\boldsymbol{H}^{(j)}$, which is the currently used CSI as indicated by the MAP. On the other hand, the indices or positions of the $n_{\rm rf}$ nonzero elements in $\boldsymbol{s}$ are specified by the antenna activation pattern. Due to the similar structures of LMIMO-MBM and GSM-MBM, the ML detection (77) can be also applied to GSM-MBM to estimate the antenna and RF mirror index patterns as well as the $M$-ary symbols.

Similar to the derivation of GSM-MBM, by combining QSM and MBM technique, the quadrature channel modulation (QCM) scheme was formed [154], [155], which is capable of improving the performance over both QSM and MBM alone. Here, we only detail the GSM-MBM to demonstrate how to combine SM techniques with MBM. For the QCM, the reader is referred to [154], [155] for detail.

*3) Discussion:* For the aforementioned five MBM schemes, Table VI summarize their data rates and ML-based detection complexities, in terms of complex multiplications required. Compared to SBIM, MBM is capable of enhancing the achievable system's throughput significantly, since the index bits increase linearly with the number of RF mirrors. Except for the DMBM system, however, the *a priori* MBM constellation alphabet or the set of all the channel realizations must be available at the RX of a MBM system. This set of all the channel realizations is typically acquired by training before data transmission, which imposes huge training overhead, particularly when a large number of RF mirrors are employed. Moreover, during the data transmission, when the MAP indicates a particular channel state, the actual channel may be different from this channel state acquired in training, unless the channel is static. Thus, channel fading seriously affects the performance of MBM systems. This does not seem to be recognized fully in the existing literature, which often choose to emphasize the advantage of throughput enhancing capability of MBM. Thus, DMBM is a promising technique for the reason that no CSI is required at the RX for demodulation, and it only suffers from 2 to 4 dB SNR loss, compared to SIMO-MBM with unrealistic assumption of perfect knowledge of the MBM channel state set at the RX [61]. It is worth emphasizing again that even for DMBM, channel fading must be sufficiently slow so that the channels in two consecutive frames remain unchanged.

TABLE VII
CLASSIFICATION OF TYPICAL SBIM-AIDED SCHEMES

| Classification | FD-IM | SD-IM | TD-IM |
|---|---|---|---|
| Single-mode class | SIM-OFDM [39], ESIM-OFDM [40], GSIM-OFDM [41], EGSIM-OFDM [42], [43], MIMO-OFDM-IM [44], [45] | Classical SM [21]–[23], SSK [50], GSSK [51], SM-OFDM [52], GSM [53], MA-SM [54], QSM [55], PSM [58], [59], GPSM [60] | SC-IM [26], FTN-IM [29], ST-IM [30], STSK [32], GSTSK [33] |
| Multi-mode class | DM-OFDM [46], GDM-OFDM [47], ZTM-OFDM [48], MM-OFDM [49] | ESM [56], [57] | DM-SCIM [28] |

Before concluding the discussion on CD-IM systems, we discuss the issue of MBM constellation design or MAP selection. It is well known that the channel gains can be modelled as circularly symmetric complex Gaussian variables. Thus, the $N_r$-dimensional MBM constellation points or channel states are randomly located, and the distances between any two constellation points are irregular. In order to improve the system performance, it is highly desired to only use 'good' channel states or MBM constellation points. An effective approach to do so is to employ a larger number of $M_{\rm rf}$ ($> m_{\rm rf}$) RF mirrors, and to choose a subset of $m_{\rm sub} = 2^{m_{\rm rf}}$ good channel states from the full set of $m_{\rm all} = 2^{M_{\rm rf}}$ possible channel states. Two selection strategies were considered in [62], namely, mutual-information (MI) based MAP selection and ED based MAP selection. Let the full set of all the $m_{\rm all}$ possible channel realizations be denoted by $\mathcal{H}_{\rm all} = \{\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_{m_{\rm all}}\}$. By arranging all these MBM constellation points in the descending order of their energy

$$\left\|\boldsymbol{h}_{l_1}\right\| \geq \left\|\boldsymbol{h}_{l_2}\right\| \geq \cdots \geq \left\|\boldsymbol{h}_{l_{m_{\rm sub}}}\right\| \geq \cdots \geq \left\|\boldsymbol{h}_{l_{m_{\rm all}}}\right\|, \quad (85)$$

the MI-based MAP selection, originally proposed in [38], chooses the subset $\mathcal{H}_{\rm sub} = \{\boldsymbol{h}_{l_1}, \boldsymbol{h}_{L_2}, \ldots, \boldsymbol{h}_{l_{m_{\rm sub}}}\}$, consisting of the $m_{\rm sub}$ largest energy MBM constellation points in order to attain the largest MI. By contrast, the ED-based MAP selection is based on maximizing the minimum ED between different constellation points, and it is capable of achieving the optimal BER performance. Specifically, let $\boldsymbol{H}_{\rm sub}$ denote the channel matrix consisting of $m_{\rm sub}$ elements of $\mathcal{H}_{\rm all}$, and further denote an arbitrary transmitted signal vector by $\boldsymbol{s}$. Then the optimally selected subset channel matrix is given by

$$\boldsymbol{H}_{\rm sub}^{\rm opt} = \arg \max_{\forall \boldsymbol{H}_{\rm sub}} \left\{ \min_{\boldsymbol{s} \neq \widetilde{\boldsymbol{s}}} \|\boldsymbol{H}_{\rm sub}(\boldsymbol{s} - \widetilde{\boldsymbol{s}})\|^2 \right\}. \quad (86)$$

This optimal ED-based MAP selection scheme is based on exhaustive search, and it suffers from the drawback of extremely high computational complexity. On the other hand, the MI-based MAP selection scheme enjoys the advantage of very low complexity, but suffers from severe performance loss compared to the optimal ED-based counterpart [62]. In [156], a greedy iterative algorithm was implemented to find a sub-optimal solution of the optimization problem (86), which significantly reduces the computational complexity, while only suffering from negligible performance loss.

### *E. Classification of Typical SBIM-Aided Schemes*

From the detailed presentations of Sections III-A–C, it can be seen that transmission resources for SBIM systems are multiple subcarriers (FD-IM), antennas (SD-IM) or time-slots (TD-IM). IM can activate a single subset of the transmission resources (subcarriers, antennas or time-slots) for transmission, leading to 'single-mode' class of IM-aided schemes, or IM can activate multiple subsets of the transmission resources (subcarriers, antennas or time-slots) for transmission, resulting in 'multi-mode' class of IM-aided schemes. According to the modes of IM, most of the IM-aided schemes in FD, SD and TD can be classified into the single-mode class and multi-mode class, which is illustrated in Table VII. Note that the transmission resources (subcarriers, antennas or time-slots) are all available physically for data transmission, and by using or activating all the available transmit entities, we have the corresponding traditional non-IM systems.

In single-mode class, only a fraction of the transmit entities, i.e., subcarriers, TX antennas, time slots, or dispersion matrices, are activated for data transmission, which are modulated by a single constellation alphabet, whilst the other subcarriers remain empty. With the single-mode IM arrangement, index bits are conveyed by the indices of the activated transmit entities or resources. By contrast, in multi-mode class, multiple subsets of the transmit entities are activated using multiple distinguishable constellation alphabets, and additional information bits are also conveyed by the index pattern.

CD-IM systems are physically very different from SBIM systems discussed above. True, IM can be performed by 'selecting' or 'activating' a specific channel state from the set of potential channel states, and hence it conveys extra bits. But the transmission entity or resource is a single channel state and, therefore, the single-mode/multi-mode classification does not apply. Physically, there exists no non-IM counterpart of a CD-IM system.

## IV. POTENTIAL CHALLENGES AND OPEN ISSUES

Based on our comprehensive review of the existing literature for IM, we now outline the potential challenges and open issues that require solutions and further research.

### *A. Balance Between Spectral and Energy Efficiency*

As underpinned by Shannon's fundamental capacity, in designing an IM-aided system, there exists a trade-off between the spectral efficiency and energy efficiency. For instance, GSIM-OFDM is more energy efficient than DM-OFDM for the reason that only a fraction of all the subcarriers are employed in GSIM-OFDM, leading to a reduced power consumption. However, since all the subcarriers are utilized in DM-OFDM,

it utilizes the frequency resources more efficiently than GSIM-OFDM. Because the system's BER performance is mainly determined by the received SNR, to optimize the system's performance, it is essential to maximize the received SNR by adjusting the system's critical parameters, such as the number of activated subcarriers in each subblock, the subblock size, and the employed constellation alphabets, etc. This optimization is equivalent to balancing the system's spectral efficiency and energy efficiency.

From the viewpoint of balancing the system's spectral efficiency and energy efficiency, the 'optimal' FD-IM system may take the form between DM-OFDM and GSIM-OFDM, which can simply be exemplified by combining some DM-OFDM and GSIM-OFDM subblocks together for data transmission. The optimization involved, however, is extremely complicated, since the received SNR is an implicit function of many system's critical parameters, where joint optimization is required. Several existing works, including [87], [88], considered to optimize a few key parameters for GSIM-OFDM, where the number of activated subcarriers or the constellations are constrained to be the same for each OFDM subblock. In the existing literature, there certainly exists no research addressing the above-mentioned joint optimization design, which is challenging but the gain will be very significant and, therefore, it warrants further research efforts.

### *B. Trade-off Between Detection Accuracy and Complexity*

For most IM-aided systems, ML detection is often employed to detect the transmit symbol vector, such as subblock for IM-OFDM and transmit frame for TD-IM, which considers all the possible realizations of the transmit signal vector and minimizes the ED between the candidates and received signals. Although attaining the optimal detection accuracy, the computational complexity of the optimal ML detector increases significantly with the system 's critical parameters, such as the OFDM subblock length and the size of MBM constellation alphabet. Therefore, ML detection is unsuitable for large-scale systems. For IM-OFDM and TD-IM schemes, reduced-complexity LLR detection has been employed at the cost of only marginal performance loss at low SNRs. For SM and TD-IM systems, however, ML detection is mainly adopted, and there exists very limited literature discussing low-complexity RX design. Moreover, several greedy-based low-complexity detectors for MBM systems suffer from inevitable performance degradation. Therefore, it is necessary to develop low-complexity detector with reliable performance, where better or even optimal trade-off between the detection accuracy and computational overhead can be made.

As mentioned previously, most of MBM based designs suffer from the drawback of high training overhead and can only be applied under a static channel environment. DMBM by contrast is a promising design for practical implementation of MBM, since it does not require the CSI at the RX. However, DMBM still requires that the channel fading is sufficiently slow, in order to ensure that the CSIs at two consecutive frames remain the same. In order to extend the application to the fast fading channel environment where the assumption of the unchanged CSI during two consecutive frames is invalid, further reseach efforts are warranted to investigate how to apply multiple-symbol differential detection and the reduced-complexity multiple-symbol differential sphere detection [150]–[152] to differentially encoded MBM systems. In such a multiple-symbol differential detection based MBM system, the detection is based on multiple observed frames. The higher the number of frames considered, the better the achievable system performance and more robust to the fast fading channel, but the detection complexity may become unacceptably high. It can be seen that this is a classical trade-off optimization between detection accuracy and complexity.

### *C. Channel Coding for IM-Aided Systems*

Channel coding is a critical technique to achieve a reliable communication over hostile environment. Representative channel coding methods include Hamming codes [157], turbo codes [158], low-density parity check (LDPC) codes [159], block Markov superposition transmission (BMST) codes [160], and polar codes [161], which can be applied to both the APM symbol bits and the index bits in IM-aided schemes. Although channel-coded systems suffers from inevitable data rate loss and additional computational complexity due to the need of channel encoding and decoding, the system performance becomes significantly enhanced by coding protection of both the symbol bits and index bits, which is a necessary requirement for practical implementation. However, channel coding is not employed in most of the researches on IM. Some preliminary work has been conducted for SD-IM [162]–[164]. In [162], polar codes are implemented to protect both the symbol bits and index bits in SM. The works [163] and [165] apply turbo codes to the trellis coded spatial modulation (TCSM) system, while the study [164] utilizes BMST codes to enhance the reliability of SM. Moreover, there have been studies on channel coding for FD-IM [166]–[168], where LDPC and BMST codes are applied to GSIM-OFDM for different applications, including VLC, highly mobile scenarios and power line communications (PLC), leading to superior BER performance. Recently, LDPC codes have also been employed to protect both the symbol bits and index bits of GDM-OFDM [47], which finds that the calculation of soft information for each bit is very complicated, and totally different from conventional OFDM. Since IM is performed on FD signals, conventional methods cannot be applied to obtain the bit soft information, and the LLR for each subcarrier indicating the index pattern has to be taken into account. Therefore, the design of channel-coded IM-aided systems is challenging but necessary for real-world application. To reduce the complexity, a feasible trade-off can be made by only encoding the symbol bits while leaving the index bits uncoded or vice versa.

More importantly, turbo detection/decoding, in which soft detection and soft channel decoding operations are carried out iteratively, is the enabling technology to operate near-capacity communication systems with affordable complexity [169]–[171]. These iterative detection/decoding designs

also do not require accurate CSI and, consequently, can significantly reduce training overhead for large-scale MIMO systems, since errors caused by the CSI estimation error can also be corrected by the iterative detection/decoding procedure. Future research is therefore warranted to investigate how to design turbo detection/decoding for IM-aided systems, in order to fully realize their performance potential with affordable computational complexity. One example is to combine bit-interleaved coded modulation with iterative decoding (BICM-ID) [172] and IM-OFDM by exchanging soft information between the index pattern demodulator and the channel decoder.

### D. Multi-Mode Index Modulation

For IM-OFDM, the multi-mode schemes, including DM-OFDM [46], ZTM-OFDM [48] and MM-OFDM [49], achieve superior performance over the single-mode counterparts. This is because when multiple distinguishable constellations are employed for IM, the number of possible realizations of OFDM subblock becomes much larger, leading to spectral efficiency enhancement. Furthermore, additional diversity is harvested by the use of different constellation sets, resulting in significant performance gain over the single-mode schemes. In this paper, we have proposed the generic MM-OFDM scheme, and have pointed out that the MM-OFDM scheme presented in [49] is only a very special case of our MM-OFDM scheme. Because of its distinctive merits, there is a scope to extend this multi-mode IM approach in the FD to other IM-aided systems in other domains, such as SM and TD-IM.

Multiple constellation design is crucial to ensure good performance for multi-mode IM based systems. In [49], the multi-mode constellation design is carried out based on the MIAD and MIRD criteria for a given spectral efficiency value. It is claimed in [49] that the MIAD criterion ensures good performance of the constellations at high SNR, whilst the MIRD criterion filters the solutions produced by the MIAD criterion in order to guarantee the medium-SNR performance. However, the resultant solutions may not be optimal for the following reason. It is well known that for different constellation orders, the required SNR values are different for reaching certain level of BER performance, e.g., $10^{-3}$. Therefore, both the constellation orders and the received SNR values should jointly be taken into account in the multiple constellation design. Moreover, the scheme of [49] is a very special and simplified case of the generic MM-OFDM proposed in this paper. The design strategy of [49] may not be directly applicable to our generic MM-OFDM, which includes several additional critical factors to consider, e.g., the types of constellation and the number of constellation modes. For example, in Fig. 10, we have compared two MM-OFDM systems with the same design except that one adopts PSK constellations and the other employs QAM constellations. The MM-OFDM with PSK constellations is observed to suffer from serious performance loss, compared to the MM-OFDM with QAM constellations. Therefore, effective multiple constellation design for the generic MM-OFDM is an urgent issue that have to be addressed.

### E. PAPR Reduction in IM-Aided OFDM Schemes

Conventional OFDM signals are notoriously known to have high PAPR, which causes severe nonlinear distortions due to the nonlinear transfer characteristics of TX devices, such as power amplifiers and LEDs. IM-OFDM schemes also suffer from the high PAPR problem. Although ESIM-OFDM and GSIM-OFDM have slightly lower PAPRs, compared to conventional OFDM, their PAPRs are still much higher than other non-OFDM systems, and will cause serious performance degradation. For DM-OFDM and MM-OFDM, where all the subcarriers are modulated, the system's PAPR is almost the same as conventional OFDM. Therefore, PAPR reduction techniques must be adopted to IM-OFDM based systems to mitigate the detrimental effects of high PAPR. Some classical PAPR reduction methods, such as clipping [76], SLM [78] and PTS [79], can be directly applied to IM-OFDM based systems. But other PAPR reduction techniques need modifications by taking the signal property of IM-OFDM into consideration in order to achieve better overall performance. For instance, when the ACE algorithm [80] is adopted, the constellation extension strategy is correlated to the high-dimensional constellation design of IM-OFDM, which requires joint optimization in order to harvest maximum performance gain.

All the above-mentioned PAPR reduction techniques trade off the system's PAPR with the spectral efficiency and/or the received SNR. It is highly desired to combat the high PAPR issue without the need to make such a compromise. For conventional OFDM downlink, the digital predistorter [81], [82] can alternatively be adopted to compensate the nonlinear distortions of the TX power amplifier, while for conventional OFDM uplink, the nonlinear detection scheme [83], [84] can be employed to address the nonlinear distortions at TX. For optical wireless, the work [173] presented a RX-side predistortion scheme to compensate for the nonlinear distortions of LEDs at TX. These novel techniques are attractive, because they are capable of overcoming the problem caused by high PAPR, without sacrificing the system's spectral efficiency or received SNR. Rather they achieve the desired goal at the expense of increased TX or RX complexity. There exists no work yet to apply these techniques to IM-OFDM based systems, which is worth considering in future research.

### F. Hybrid IM-Aided Schemes

As discussed previously, IM can be performed on different domains, including FD, SD, TD and CD. Moreover, in the SD based IM systems, IM operation can be carried out on TX antennas or RX antennas. By performing IM operation simultaneously on several domains, the system performance can be enhanced significantly, in comparison with performing IM on single domain. For instance, in the ST-IM system, a fraction of the TX antennas are activated for data transmission, where binary bits can be conveyed by their indices, while on each activated antenna, symbols are transmitted on part of the time slots, whose indices carry additional information bits. This leads to significant throughput gain over the pure SM and TD-IM schemes. It can be

seen that by combining multiple IM operations together, considerable diversity gain can be achieved, which enhances the system's throughput and performance. Currently, there exist a few works [30], [36], [62], [174], which consider hybrid IM-aided systems. However, the design strategy to combine three types or even four types of IM-aided schemes has not been properly investigated yet. Because of the enlarged multiple different signal sets in such a hybrid IM-aided system, jointly optimal constellation design is required and, moreover, low-complexity detection algorithm is also a potential issue for the hybrid architecture.

### G. Applications to Mobile Communications

Rapid proliferation of smart mobile devices and explosive increase in mobile data traffic have imposed the urgent need of speedily upgrading the current 3G and 4G mobile networks to the future-generation mobile network known as 5G. Andrews *et al.* [175] list the 'big three' key technologies for 5G networks: ultra-densification, mmWave, and massive MIMO. In a nutshell, 5G networks must be capable of offering much higher data rate, lower latency and less cost as well as consuming less energy than the current 3G and 4G communications. The IM technique provides an ideal enabling technology to address these engineering challenges. To be more specific, as discussed in this survey paper, some IM-aided schemes, such as multi-mode IM-aided systems, are capable of significantly enhancing the data rate by conveying additional energy-free bits, which leads to considerable energy efficiency gain in comparison with non-index-modulated counterparts. Moreover, some IM-aided schemes can be utilized to simplify the transceiver structures. This contributes to lower cost and reduces computational complexity, which in turn leads to lower latency. For instance, SM is capable of reducing the number of required RF chains and the detection complexity. Due to the aforementioned merits, preliminary work has been conducted to adopt the IM technique in mmWave transmissions [14], [15], massive MIMO [16], [17], and network coding [18]. It is believed that the IM technique can be incorporated with different critical technologies in 5G networks, where further research is needed.

### H. Applications to Optical Communications

Due to the shortage of frequency resources and the ever-increasing demand for high-rate communication, optical communication has become as a promising complementary technique to RF communication, which provides additional advantages of unlicensed spectrum, robustness to the electromagnetic interferences, information security and safety to human health [176]. Optical communications can be generally classified into VLC [177] and coherent optical communications [178]–[180]. At present, there exists limited literature applying the IM technique to optical communicating systems.

Indeed, enormous benefits can be obtained by applying IM to optical communications. In VLC, due to Hermitian symmetry and the unique frame structures, as in ACO-OFDM and PAM-DMT systems, optical OFDM suffers from inevitable data rate loss, which can be recovered at least partially by IM. By applying the zero-padded IM technique, energy consumption is reduced, which compensates for the low energy efficiency caused by the addition of DC-bias, as in DCO-OFDM system. Furthermore, since illumination and communication are the dual purposes of many LED based lighting systems, dimming control is essential for VLC system design, which may be also realized with the aid of IM by adjusting the number of activated TX entities.

Additionally, coherent optical communications, originated in optical fiber transmission, plays an important role in free-space optical (FSO) systems [181]. FSO signals suffer from severe attenuation and interferences due to atmosphere turbulence or complex climate conditions such as fog or haze [181]. Although IM offers an attractive and effective means of enhancing the energy efficiency and the received SNR for VLC systems, it is challenging to apply IM to coherent optical communications. This is because in coherent optical communications, due to the use of coherent detection, both the amplitude and phase of the incoming light can be detected, which means that transmitted signals are complex-valued. Moreover, two independent data streams with orthogonal polarization can be transmitted simultaneously, harvesting additional diversities. Therefore, strategies of employing IM in VLC and RF architectures cannot be applied in coherent optical communications. Further research is warranted to address this critical issue of how to apply IM to coherent optical communications.

## V. Conclusion

Index modulation, which is capable of enhancing the system's energy efficiency and providing high-rate data transmission, has become a promising technique for wireless communications. In IM-aided systems, in addition to the symbol bits carried by conventional APM, extra information bits are conveyed by the indices of activated transmit entities, without extra energy consumption. Transmit entities that can be exploited for IM include the subcarriers of OFDM, TX antennas of MIMO, time slots of single-carrier systems, and channel states of MBM, etc. With the aid of IM, considerable diversity gains can be achieved, leading to significant performance enhancement over the non-index-modulated counterparts. Thus, IM technique has attracted the full attention from the academia and industry, and numerous relevant researches have been proposed in the past few years.

This paper has provided a comprehensive survey on IM-aided systems by classifying various IM based schemes in terms of indexing domains, which include frequency domain, spatial domain, time domain and channel domain. This general classification yields the four categories of IM schemes: IM-OFDM, SM, TD-IM and MBM. For each category of IM based systems, we have provided the corresponding system model followed by the transceiver design of the typical IM-aided systems, to help the reader to better understand IM principles. Performance evaluation, in terms of spectral efficiency/throughput, energy efficiency, BER and detection complexity, have been presented, along with a brief introduction to performance analysis methods, including the minimum ED calculation and PEP approximation. In this survey, a

special emphasis has been placed on IM-OFDM schemes. We believe this is helpful to the understanding of other IM-aided systems, since their fundamental principles are the same as the IM-OFDM. Based on the survey for the current development of IM, the potential challenges and open issues have been summarized to shed light on the future research. By addressing these challenges, we believe full benefits of the IM technique can be realized, and this attractive novel technique will play an important role in 5G communications.

## REFERENCES


[1] G. Zhang, M. D. Leenheer, A. Morea, and B. Mukherjee, "A survey on OFDM-based elastic core optical networking," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 65–87, 1st Quart., 2013.

[2] D. Falconer, S. L. Ariyavisitakul, A. Benyamin-Seeyar, and B. Eidson, "Frequency domain equalization for single-carrier broadband wireless systems," *IEEE Commun. Mag.*, vol. 40, no. 4, pp. 58–66, Apr. 2002.

[3] J.-S. Lee, Y.-W. Su, and C.-C. Shen, "A comparative study of wireless protocols: Bluetooth, UWB, ZigBee, and Wi-Fi," in *Proc. IECON*, Taipei, Taiwan, Nov. 2007, pp. 46–51.

[4] B. P. Crow, I. Widjaja, J. G. Kim, and P. T. Sakai, "IEEE 802.11 wireless local area networks," *IEEE Commun. Mag.*, vol. 35, no. 9, pp. 116–126, Sep. 1997.

[5] A. Ghosh, D. R. Wolter, J. G. Andrews, and R. Chen, "Broadband wireless access with WiMax/802.16: Current performance benchmarks and future potential," *IEEE Commun. Mag.*, vol. 43, no. 2, pp. 129–136, Feb. 2005.

[6] A. Ghosh, R. Ratasuk, B. Mondal, N. Mangalvedhe, and T. Thomas, "LTE-advanced: Next-generation wireless broadband technology [invited paper]," *IEEE Wireless Commun.*, vol. 17, no. 3, pp. 10–22, Jun. 2010.

[7] W. Y. Zou and Y. Wu, "COFDM: An overview," *IEEE Trans. Broadcast.*, vol. 41, no. 1, pp. 1–8, Mar. 1995.

[8] P. H. Moose, D. Roderick, R. North, and M. Geile, "A COFDM-based radio for HDR LOS networked communications," in *Proc. ICC*, Vancouver, BC, Canada, Jun. 1999, pp. 187–192.

[9] I. Koffman and V. Roman, "Broadband wireless access solutions based on OFDM access in IEEE 802.16," *IEEE Commun. Mag.*, vol. 40, no. 4, pp. 96–103, Apr. 2002.

[10] Y. Rahmatallah and S. Mohan, "Peak-to-average power ratio reduction in OFDM systems: A survey and taxonomy," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 4, pp. 1567–1592, 4th Quart., 2013.

[11] A. Nuwanpriya, S.-W. Ho, J. A. Zhang, A. J. Grant, and L. Luo, "PAM-SCFDE for optical wireless communications," *J. Lightw. Technol.*, vol. 33, no. 14, pp. 2938–2949, Jul. 15, 2015.

[12] D. Gesbert, M. Shafi, D.-S. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: An overview of MIMO space-time coded wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, pp. 281–302, Apr. 2003.

[13] K. Zheng *et al.*, "Survey of large-scale MIMO systems," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 3, pp. 1738–1760, 3rd Quart., 2015.

[14] C. Sacchi, T. F. Rahman, I. A. Hemadeh, and M. El-Hajjar, "Millimeter-wave transmission for small-cell backhaul in dense urban environment: A solution based on MIMO-OFDM and space-time shift keying (STSK)," *IEEE Access*, vol. 5, pp. 4000–4017, 2017.

[15] T. F. Rahman, A. Habib, C. Sacchi, and M. El-Hajjar, "Mm-wave STSK-aided single carrier block transmission for broadband networking," in *Proc. ISCC*, Heraklion, Greece, Jul. 2017, pp. 507–514.

[16] T. L. Narasimhan, P. Raviteja, and A. Chockalingam, "Generalized spatial modulation in large-scale multiuser MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3764–3779, Jul. 2015.

[17] Y. Cui and X. Fang, "Performance analysis of massive spatial modulation MIMO in high-speed railway," *IEEE Trans. Veh. Technol.*, vol. 65, no. 11, pp. 8925–8932, Nov. 2016.

[18] D. U. Sudhakaran and B. S. Rajan, "Index coded PSK modulation for prioritized receivers," *IEEE Trans. Veh. Technol.*, vol. 66, no. 12, pp. 11151–11165, Dec. 2017.

[19] E. Başar, "Index modulation techniques for 5G wireless networks," *IEEE Commun. Mag.*, vol. 54, no. 7, pp. 168–175, Jul. 2016.

[20] M. Wen, X. Cheng, and L. Yang, *Index Modulation for 5G Wireless Communications*. Cham, Switzerland: Springer, 2017.

[21] S. Ganesan, R. Mesleh, H. Haas, C. W. Ahn, and S. Yun, "On the performance of spatial modulation OFDM," in *Proc. ACSSC*, Pacific Grove, CA, USA, Oct./Nov. 2006, pp. 1825–1829.

[22] R. Mesleh, H. Haas, C. W. Ahn, and S. Yun, "Spatial modulation—A new low complexity spectral efficiency enhancing technique," in *Proc. ChinaCom*, Beijing, China, Oct. 2006, pp. 1–5.

[23] R. Mesleh, H. Haas, C. W. Ahn, and S. Yun, "Spatial modulation-OFDM," in *Proc. 11th Int. OFDM Workshop*, Hamburg, Germany, Aug. 2006, pp. 288–292.

[24] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. URSI Int. Symp. Signals Syst. Electron.*, Pisa, Italy, Sep./Oct. 1998, pp. 295–300.

[25] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 17–24, Oct. 1998.

[26] M. Nakao, T. Ishihara, and S. Sugiura, "Single-carrier frequency-domain equalization with index modulation," *IEEE Commun. Lett.*, vol. 21, no. 2, pp. 298–301, Feb. 2017.

[27] S. Sugiura, T. Ishihara, and M. Nakao, "State-of-the-art design of index modulation in the space, time, and frequency domains: Benefits and fundamental limitations," *IEEE Access*, vol. 5, pp. 21774–21790, 2017.

[28] M. Nakao, T. Ishihara, and S. Sugiura, "Dual-mode time-domain index modulation for Nyquist-criterion and faster-than-Nyquist single-carrier transmissions," *IEEE Access*, vol. 5, pp. 27659–27667, 2017.

[29] T. Ishihara and S. Sugiura, "Faster-than-Nyquist signaling with index modulation," *IEEE Wireless Commun. Lett.*, vol. 6, no. 5, pp. 630–633, Oct. 2017.

[30] S. Jacob, T. L. Narasimhan, and A. Chockalingam, "Space-time index modulation," in *Proc. WCNC*, San Francisco, CA, USA, Mar. 2017, pp. 1–6.

[31] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

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

[33] S. Sugiura, S. Chen, and L. Hanzo, "Generalized space-time shift keying designed for flexible diversity-, multiplexing- and complexity-tradeoffs," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1144–1153, Apr. 2011.

[34] A. K. Khandani, "Media-based modulation: A new approach to wireless transmission," in *Proc. ISIT*, Istanbul, Turkey, Jul. 2013, pp. 3050–3054.

[35] E. Seifi, M. Atamanesh, and A. K. Khandani, "Media-based MIMO: Outperforming known limits in wireless," in *Proc. ICC*, Kuala Lumpur, Malaysia, May 2016, pp. 1–7.

[36] Y. Naresh and A. Chockalingam, "On media-based modulation using RF mirrors," in *Proc. ITA*, La Jolla, CA, USA, Jan./Feb. 2016, pp. 1–10.

[37] Z. Bouida, H. El-Sallabi, A. Ghrayeb, and K. A. Qaraqe, "Reconfigurable antenna-based space-shift keying (SSK) for MIMO Rician channels," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 446–457, Jan. 2016.

[38] A. K. Khandani, "Media-based modulation: Converting static Rayleigh fading to AWGN," in *Proc. ISIT*, Honolulu, HI, USA, Jun./Jul. 2014, pp. 1549–1553.

[39] R. Abu-Alhiga and H. Haas, "Subcarrier-index modulation OFDM," in *Proc. PIMRC*, Tokyo, Japan, Sep. 2009, pp. 177–181.

[40] D. Tsonev, S. Sinanovic, and H. Haas, "Enhanced subcarrier index modulation (SIM) OFDM," in *Proc. GLOBECOM Workshops*, Houston, TX, USA, Dec. 2011, pp. 728–732.

[41] E. Başar, Ü. Aygölü, E. Panayırcı, and H. V. Poor, "Orthogonal frequency division multiplexing with index modulation," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5536–5549, Nov. 2013.

[42] 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, Oct. 2015.

[43] X. Yang, Z. Zhang, P. Fu, and J. Zhang, "Spectrum-efficient index modulation with improved constellation mapping," in *Proc. HMWC*, Xi'an, China, Oct. 2015, pp. 91–95.

[44] E. Başar, "Multiple-input multiple-output OFDM with index modulation," *IEEE Signal Process. Lett.*, vol. 22, no. 12, pp. 2259–2263, Dec. 2015.

[45] E. Başar, "On multiple-input multiple-output OFDM with index modulation for next generation wireless networks," *IEEE Trans. Signal Process.*, vol. 64, no. 15, pp. 3868–3878, Aug. 2016.

[46] T. Mao, Z. Wang, Q. Wang, S. Chen, and L. Hanzo, "Dual-mode index modulation aided OFDM," *IEEE Access*, vol. 5, pp. 51–60, 2017.

[47] T. Mao, Q. Wang, and Z. Wang, "Generalized dual-mode index modulation aided OFDM," *IEEE Commun. Lett.*, vol. 21, no. 4, pp. 761–764, Apr. 2017.

[48] T. Mao, Q. Wang, J. Quan, and Z. Wang, "Zero-padded tri-mode index modulation aided OFDM," in *Proc. GLOBECOM*, Singapore, Dec. 2017, pp. 1–5.

[49] M. Wen, E. Başar, Q. Li, B. Zheng, and M. Zhang, "Multiple-mode orthogonal frequency division multiplexing with index modulation," *IEEE Trans. Commun.*, vol. 65, no. 9, pp. 3892–3906, Sep. 2017.

[50] Y. A. Chau and S.-H. Yu, "Space modulation on wireless fading channels," in *Proc. VTC Fall*, Atlantic City, NJ, USA, Oct. 2001, pp. 1668–1671.

[51] J. Jeganathan, A. Ghrayeb, and L. Szczecinski, "Generalized space shift keying modulation for MIMO channels," in *Proc. PIMRC*, Cannes, France, Sep. 2008, pp. 1–5.

[52] R. Y. 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.

[53] A. Younis, N. Serafimovski, R. Mesleh, and H. Haas, "Generalised spatial modulation," in *Proc. ASILOMAR*, Pacific Grove, CA, USA, Nov. 2010, pp. 1498–1502.

[54] 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.

[55] 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.

[56] 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.

[57] 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.

[58] L.-L. Yang, "Transmitter preprocessing aided spatial modulation for multiple-input multiple-output systems," in *Proc. VTC Spring*, Yokohama, Japan, May 2011, pp. 1–5.

[59] A. Stavridis, S. Sinanovic, M. D. Renzo, and H. Haas, "Transmit precoding for receive spatial modulation using imperfect channel knowledge," in *Proc. VTC Spring*, Yokohama, Japan, May 2012, pp. 1–5.

[60] R. Zhang, L.-L. Yang, and L. Hanzo, "Generalised pre-coding aided spatial modulation," *IEEE Trans. Wireless Commun.*, vol. 12, no. 11, pp. 5434–5443, Nov. 2013.

[61] Y. Naresh and A. Chockalingam, "A low-complexity maximum-likelihood detector for differential media-based modulation," *IEEE Commun. Lett.*, vol. 21, no. 10, pp. 2158–2161, Oct. 2017.

[62] Y. Naresh and A. Chockalingam, "On media-based modulation using RF mirrors," *IEEE Trans. Veh. Technol.*, vol. 66, no. 6, pp. 4967–4983, Jun. 2017.

[63] E. Başar and I. Altunbas, "Space-time channel modulation," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, pp. 7609–7614, Aug. 2017.

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

[65] M. D. Renzo, H. Haas, and P. M. Grant, "Spatial modulation for multiple-antenna wireless systems: A survey," *IEEE Commun. Mag.*, vol. 49, no. 12, pp. 182–191, Dec. 2011.

[66] E. Başar *et al.*, "Index modulation techniques for next-generation wireless networks," *IEEE Access*, vol. 5, pp. 16693–16746, 2017.

[67] A. Sezgin, E. A. Jorswieck, and E. Costa, "LDC in MIMO Ricean channels: Optimal transmit strategy with MMSE detection," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 313–328, Jan. 2008.

[68] P. Liu and I.-M. Kim, "Exact and closed-form error performance analysis for hard MMSE-SIC detection in MIMO systems," *IEEE Trans. Commun.*, vol. 59, no. 9, pp. 2463–2477, Sep. 2011.

[69] M. Mandloi, M. A. Hussain, and V. Bhatia, "Improved multiple feedback successive interference cancellation algorithms for near-optimal MIMO detection," *IET Commun.*, vol. 11, no. 1, pp. 150–159, Jan. 2017.

[70] P. Li, R. C. de Lamare, and R. Fa, "Multiple feedback successive interference cancellation detection for multiuser MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2434–2439, Aug. 2011.

[71] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, 1999.

[72] P. F. Driessen and G. J. Foschini, "On the capacity formula for multiple input-multiple output wireless channels: A geometric interpretation," *IEEE Trans. Commun.*, vol. 47, no. 2, pp. 173–176, Feb. 1999.

[73] A. Lozano, F. R. Farrokhi, and R. A. Valenzuela, "Lifting the limits on high speed wireless data access using antenna arrays," *IEEE Commun. Mag.*, vol. 39, no. 9, pp. 156–162, Sep. 2001.

[74] H. Ku and J. S. Kenney, "Behavioral modeling of nonlinear RF power amplifiers considering memory effects," *IEEE Trans. Microw. Theory Techn.*, vol. 51, no. 12, pp. 2495–2504, Dec. 2003.

[75] K. Ying *et al.*, "Nonlinear distortion mitigation in visible light communications," *IEEE Wireless Commun.*, vol. 22, no. 2, pp. 36–45, Apr. 2015.

[76] H. Ochiai and H. Imai, "On clipping for peak power reduction of OFDM signals," in *Proc. GLOBECOM*, San Francisco, CA, USA, Nov./Dec. 2000, pp. 731–735.

[77] X. Huang, J. Lu, J. Zheng, J. Chuang, and J. Gu, "Reduction of peak-to-average power ratio of OFDM signals with companding transform," *Electron. Lett.*, vol. 37, no. 8, pp. 506–507, Apr. 2001.

[78] H. Breiling, S. H. Muller-Weinfurtner, and J. B. Huber, "SLM peak-power reduction without explicit side information," *IEEE Commun. Lett.*, vol. 5, no. 6, pp. 239–241, Jun. 2001.

[79] A. D. S. Jayalath and C. Tellambura, "Adaptive PTS approach for reduction of peak-to-average power ratio of OFDM signal," *Electron. Lett.*, vol. 36, no. 14, pp. 1226–1228, Jul. 2000.

[80] B. S. Krongold and D. L. Jones, "PAR reduction in OFDM via active constellation extension," *IEEE Trans. Broadcast.*, vol. 49, no. 3, pp. 258–268, Sep. 2003.

[81] S. Chen, "An efficient predistorter design for compensating nonlinear memory high power amplifier," *IEEE Trans. Broadcast.*, vol. 57, no. 4, pp. 856–865, Dec. 2011.

[82] S. Chen, X. Hong, Y. Gong, and C. J. Harris, "Digital predistorter design using B-spline neural network and inverse of De Boor algorithm," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 60, no. 6, pp. 1584–1594, Jun. 2013.

[83] X. Hong, S. Chen, Y. Gong, and C. J. Harris, "Nonlinear equalization of Hammerstein OFDM systems," *IEEE Trans. Signal Process.*, vol. 62, no. 21, pp. 5629–5639, Nov. 2014.

[84] S. Chen *et al.*, "Single-carrier frequency-domain equalization with hybrid decision feedback equalizer for Hammerstein channels containing nonlinear transmit amplifier," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 3341–3354, May 2017.

[85] H. Sari, G. Karam, and I. Jeanclaud, "Frequency-domain equalization of mobile radio and terrestrial broadcast channels," in *Proc. GLOBECOM*, San Francisco, CA, USA, Nov./Dec. 1994, pp. 1–5.

[86] H. Sari, G. Karam, and I. Jeanclaude, "Transmission techniques for digital terrestrial TV broadcasting," *IEEE Commun. Mag.*, vol. 33, no. 2, pp. 100–109, Feb. 1995.

[87] M. Wen, X. Cheng, and L. Yang, "Optimizing the energy efficiency of OFDM with index modulation," in *Proc. ICCS*, Macau, China, Nov. 2014, pp. 31–35.

[88] W. Li, H. Zhao, C. Zhang, L. Zhao, and R. Wang, "Generalized selecting sub-carrier modulation scheme in OFDM system," in *Proc. ICC Workshops*, Sydney, NSW, Australia, Jun. 2014, pp. 907–911.

[89] Y. Ko, "A tight upper bound on bit error rate of joint OFDM and multi-carrier index keying," *IEEE Commun. Lett.*, vol. 18, no. 10, pp. 1763–1766, Oct. 2014.

[90] R. Fan, Y. J. Yu, and Y. L. Guan, "Improved orthogonal frequency division multiplexing with generalised index modulation," *IET Commun.*, vol. 10, no. 8, pp. 969–974, May 2016.

[91] M. Wen, B. Ye, E. Başar, Q. Li, and F. Ji, "Enhanced orthogonal frequency division multiplexing with index modulation," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4786–4801, Jul. 2017.

[92] Y. Xiao *et al.*, "OFDM with interleaved subcarrier-index modulation," *IEEE Commun. Lett.*, vol. 18, no. 8, pp. 1447–1450, Aug. 2014.

[93] S. Wang, B. Xu, H. Bai, Y. Xiao, and L. Dan, "MIMO-OFDM with interleaved subcarrier-index modulation," in *Proc. WiCOM*, Beijing, China, Sep. 2014, pp. 35–37.

[94] E. Başar, "OFDM with index modulation using coordinate interleaving," *IEEE Wireless Commun. Lett.*, vol. 4, no. 4, pp. 381–384, Aug. 2015.

[95] H. Zhang, L.-L. Yang, and L. Hanzo, "Compressed sensing improves the performance of subcarrier index-modulation-assisted OFDM," *IEEE Access*, vol. 4, pp. 7859–7873, 2016.

[96] N. Ishikawa, S. Sugiura, and L. Hanzo, "Subcarrier-index modulation aided OFDM—Will it work?" *IEEE Access*, vol. 4, pp. 2580–2593, 2016.

[97] M. Wen, Y. Li, X. Cheng, and L. Yang, "Index modulated OFDM with ICI self-cancellation in underwater acoustic communications," in *Proc. 48th IEEE Asilomar Conf. Signals Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2014, pp. 338–342.

[98] M. Wen *et al.*, "Index modulated OFDM for underwater acoustic communications," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 132–137, May 2016.

[99] J. B. Carruthers and J. M. Kahn, "Multiple-subcarrier modulation for nondirected wireless infrared communication," *IEEE J. Sel. Areas Commun.*, vol. 14, no. 3, pp. 538–546, Apr. 1996.

[100] J. Armstrong and A. J. Lowery, "Power efficient optical OFDM," *Electron. Lett.*, vol. 42, no. 6, pp. 370–372, Mar. 2006.

[101] S. C. J. Lee, S. Randel, F. Breyer, and A. M. J. Koonen, "PAM-DMT for intensity-modulated and direct-detection optical communication systems," *IEEE Photon. Technol. Lett.*, vol. 21, no. 23, pp. 1749–1751, Dec. 1, 2009.

[102] D. Tsonev, S. Sinanovic, and H. Haas, "Novel unipolar orthogonal frequency division multiplexing (U-OFDM) for optical wireless," in *Proc. VTC Spring*, Yokohama, Japan, May 2012, pp. 1–5.

[103] N. Fernando, Y. Hong, and E. Viterbo, "Flip-OFDM for unipolar communication systems," *IEEE Trans. Commun.*, vol. 60, no. 12, pp. 3726–3733, Dec. 2012.

[104] E. Başar and E. Panayırcı, "Optical OFDM with index modulation for visible light communications," in *Proc. IWOW*, Istanbul, Turkey, Sep. 2015, pp. 11–15.

[105] T. Mao, R. Jiang, and R. Bai, "Optical dual-mode index modulation aided OFDM for visible light communications," *Opt. Commun.*, vol. 391, pp. 37–41, May 2017.

[106] P. Borjesson and C.-E. Sundberg, "Simple approximations of the error function $Q(x)$ for communications applications," *IEEE Trans. Commun.*, vol. COM-27, no. 3, pp. 639–643, Mar. 1979.

[107] M. Chiani and D. Dardari, "Improved exponential bounds and approximation for the Q-function with application to average error probability computation," in *Proc. GLOBECOM*, Taipei, Taiwan, Nov. 2002, pp. 1399–1402.

[108] M. Chiani, D. Dardari, and M. K. Simon, "New exponential bounds and approximations for the computation of error probability in fading channels," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 840–845, Jul. 2003.

[109] P. Loskot and N. C. Beaulieu, "Prony and polynomial approximations for evaluation of the average probability of error over slow-fading channels," *IEEE Trans Veh. Technol.*, vol. 58, no. 3, pp. 1269–1280, Mar. 2009.

[110] G. K. Karagiannidis and A. S. Lioumpas, "An improved approximation for the Gaussian Q-function," *IEEE Commun. Lett.*, vol. 11, no. 8, pp. 644–646, Aug. 2007.

[111] M. Lopez-Benitez and F. Casadevall, "Versatile, accurate, and analytically tractable approximation for the Gaussian Q-function," *IEEE Trans Commun.*, vol. 59, no. 4, pp. 917–922, Apr. 2011.

[112] X. Cheng, M. Wen, L. Yang, and Y. Li, "Index modulated OFDM with interleaved grouping for V2X communications," in *Proc. ITSC*, Qingdao, China, Oct. 2014, pp. 1097–1104.

[113] M. Zhang, H. Wang, X. Cheng, L.-Q. Yang, and X. Zhou, "Quadrature index modulated OFDM with interleaved grouping for V2X communications," in *Proc. ICNC*, Kauai, HI, USA, Feb. 2016, pp. 1–5.

[114] J. M. Hamamreh, E. Başar, and H. Arslan, "OFDM-subcarrier index selection for enhancing security and reliability of 5G URLLC services," *IEEE Access*, vol. 5, pp. 25863–25875, 2017.

[115] Y. Lee, H. Jo, Y. Ko, and J. Choi, "Secure index and data symbol modulation for OFDM-IM," *IEEE Access*, vol. 5, pp. 24959–24974, 2017.

[116] S. Nagaraj, "Performance analysis of coded SSK modulation on block-fading channels," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6773–6777, Aug. 2016.

[117] P. Maheswaran and M. D. Selvaraj, "Performance analysis of feedback-based dynamic SSK-BPSK system," *IEEE Wireless Commun. Lett.*, vol. 5, no. 1, pp. 96–99, Feb. 2016.

[118] P. Som and A. Chockalingam, "Bit error probability analysis of SSK in DF relaying with threshold-based best relay selection and selection combining," *IEEE Commun. Lett.*, vol. 18, no. 1, pp. 18–21, Jan. 2014.

[119] S. S. Ikki and R. Mesleh, "A general framework for performance analysis of space shift keying (SSK) modulation in the presence of Gaussian imperfect estimations," *IEEE Commun. Lett.*, vol. 16, no. 2, pp. 228–230, Feb. 2012.

[120] M. D. Renzo and H. Haas, "Space shift keying (SSK) modulation with partial channel state information: Optimal detector and performance analysis over fading channels," *IEEE Trans. Commun.*, vol. 58, no. 11, pp. 3196–3210, Nov. 2010.

[121] M. D. Renzo and H. Haas, "A general framework for performance analysis of space shift keying (SSK) modulation for MISO correlated Nakagami-m fading channels," *IEEE Trans. Commun.*, vol. 58, no. 9, pp. 2590–2603, Sep. 2010.

[122] M. D. Renzo and H. Haas, "Improving the performance of space shift keying (SSK) modulation via opportunistic power allocation," *IEEE Commun. Lett.*, vol. 14, no. 6, pp. 500–502, Jun. 2010.

[123] 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.

[124] J. G. Proakis, *Digital Communications*. New York, NY, USA: McGraw-Hill, 1995.

[125] E. Başar, U. Aygolu, E. Panayırcı, and H. V. Poor, "Space-time block coded spatial modulation," *IEEE Trans. Commun.*, vol. 59, no. 3, pp. 823–832, Mar. 2011.

[126] J. Li *et al.*, "Generalized precoding-aided quadrature spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 1881–1886, Feb. 2017.

[127] A. Stavridis, D. Basnayaka, S. Sinanovic, M. D. Renzo, and H. Haas, "A virtual MIMO dual-hop architecture based on hybrid spatial modulation," *IEEE Trans. Commun.*, vol. 62, no. 9, pp. 3161–3179, Sep. 2014.

[128] A. Afana, T. M. N. Ngatched, O. A. Dobre, and S. Ikki, "Cooperative DF cognitive radio networks with spatial modulation with channel estimation errors," in *Proc. WCNC*, San Francisco, CA, USA, Mar. 2017, pp. 1–5.

[129] S. P. Alaka, T. L. Narasimhan, and A. Chockalingam, "Generalized spatial modulation in indoor wireless visible light communication," in *Proc. GLOBECOM*, San Diego, CA, USA, Dec. 2015, pp. 1–7.

[130] Y. Chen and M. Jiang, "Joint colour-and-spatial modulation aided visible light communication system," in *Proc. VTC Spring*, Nanjing, China, May 2016, pp. 1–5.

[131] C. He, T. Q. Wang, and J. Armstrong, "Performance comparison between spatial multiplexing and spatial modulation in indoor MIMO visible light communication systems," in *Proc. ICC*, Kuala Lumpur, Malaysia, May 2016, pp. 1–6.

[132] K. Xu, H. Yu, and Y.-J. Zhu, "Channel-adapted spatial modulation for massive MIMO visible light communications," *IEEE Photon. Technol. Lett.*, vol. 28, no. 23, pp. 2693–2696, Dec. 1, 2016.

[133] C. R. Kumar and R. K. Jeyachitra, "Power efficient generalized spatial modulation MIMO for indoor visible light communications," *IEEE Photon. Technol. Lett.*, vol. 29, no. 11, pp. 921–924, Jun. 1, 2017.

[134] L. Wang, S. Bashar, Y. Wei, and R. Li, "Secrecy enhancement analysis against unknown eavesdropping in spatial modulation," *IEEE Commun. Lett.*, vol. 19, no. 8, pp. 1351–1354, Aug. 2015.

[135] C. Liu, L.-L. Yang, and W. Wang, "Secure spatial modulation with a full-duplex receiver," *IEEE Wireless Commun. Lett.*, vol. 6, no. 6, pp. 838–841, Dec. 2017.

[136] X. Wang, X. Wang, and L. Sun, "Spatial modulation aided physical layer security enhancement for fading wiretap channels," in *Proc. WCSP*, Yangzhou, China, Oct. 2016, pp. 1–5.

[137] H. G. Myung, J. Lim, and D. J. Goodman, "Single carrier FDMA for uplink wireless transmission," *IEEE Veh. Technol. Mag.*, vol. 1, no. 3, pp. 30–38, Sep. 2006.

[138] N. Al-Dhahir, "Single-carrier frequency-domain equalization for space-time block-coded transmissions over frequency-selective fading channels," *IEEE Commun. Lett.*, vol. 5, no. 7, pp. 304–306, Jul. 2001.

[139] G. Lee and G. Schroeder, "Optical pulse position modulation with multiple positions per pulsewidth," *IEEE Trans. Commun.*, vol. COM-25, no. 3, pp. 360–364, Mar. 1977.

[140] S. Sugiura, S. Chen, and L. Hanzo, "A unified MIMO architecture subsuming space shift keying, OSTBC, BLAST and LDC," in *Proc. VTC Fall*, Ottawa, ON, Canada, Sep. 2010, pp. 1–5.

[141] S. Sugiura, S. Chen, and L. Hanzo, "Space-time shift keying: A unified MIMO architecture," in *Proc. Globecom*, Miami, FL, USA, Dec. 2010, pp. 1–5.

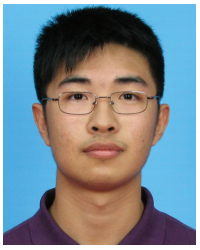
[142] S. Sugiura, C. Xu, S. X. Ng, and L. Hanzo, "Reduced-complexity iterative-detection-aided generalized space-time shift keying," *IEEE Trans. Veh. Technol.*, vol. 61, no. 8, pp. 3656–3664, Oct. 2012.

[143] H. A. Ngo, C. Xu, S. Sugiura, and L. Hanzo, "Space-time-frequency shift keying for dispersive channels," *IEEE Signal Process. Lett.*, vol. 18, no. 3, pp. 177–180, Mar. 2011.

[144] M. I. Kadir, S. Chen, K. Hari, K. Giridhar, and L. Hanzo, "OFDM-aided differential space-time shift keying using iterative soft multiple-symbol differential sphere decoding," *IEEE Trans. Veh. Technol.*, vol. 63, no. 8, pp. 4102–4108, Oct. 2014.

[145] I. A. Hemadeh, M. El-Hajjar, S. Won, and L. Hanzo, "Layered multi-group steered space-time shift-keying for millimeter-wave communications," *IEEE Access*, vol. 4, pp. 3708–3718, 2016.

[146] I. A. Hemadeh, M. El-Hajjar, S. Won, and L. Hanzo, "Multi-set space-time shift keying and space-frequency space-time shift keying for millimeter-wave communications," *IEEE Access*, vol. 5, pp. 8324–8342, 2017.
[147] M. I. Kadir, S. Sugiura, S. Chen, and L. Hanzo, "Unified MIMO-multicarrier designs: A space-time shift keying approach," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 550–579, 2nd Quart., 2015.
[148] N. Pillay and H. Xu, "Uncoded space-time labeling diversity—Application of media-based modulation with RF mirrors," *IEEE Commun. Lett.*, vol. 22, no. 2, pp. 272–275, Feb. 2018.
[149] Y. Bian *et al.*, "Differential spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 64, no. 7, pp. 3262–3268, Jul. 2015.
[150] D. Divsalar and M. K. Simon, "Multiple-symbol differential detection of MPSK," *IEEE Trans. Commun.*, vol. 38, no. 3, pp. 300–308, Mar. 1990.
[151] L. Lampe, R. Schober, V. Pauli, and C. Windpassinger, "Multiple-symbol differential sphere decoding," *IEEE Trans. Commun.*, vol. 53, no. 12, pp. 1981–1985, Dec. 2005.
[152] P. Zhang, S. Chen, and L. Hanzo, "Differential space-time shift keying-aided successive-relaying-assisted decode-and-forward cooperative multiuser CDMA," *IEEE Trans. Veh. Technol.*, vol. 62, no. 5, pp. 2156–2169, Jun. 2013.
[153] C. Xu and M. Zhao, "Mutual information and secrecy enhancement analysis of media-based modulation," in *Proc. WCSP*, Yangzhou, China, Oct. 2016, pp. 1–5.
[154] I. Yildirim, E. Başar, and I. Altunbas, "Quadrature channel modulation," *IEEE Wireless Commun. Lett.*, vol. 6, no. 6, pp. 790–793, Dec. 2017.
[155] N. Pillay and H. Xu, "Quadrature spatial media-based modulation with RF mirrors," *IET Commun.*, vol. 11, no. 16, pp. 2440–2448, Nov. 2017.
[156] L. Yapeng, T. Cheng, L. Liu, and L. Yongzhi, "Novel reduced-complexity channel state selection algorithms for media-based modulation," in *Proc. ICSP*, Chengdu, China, Nov. 2016, pp. 1259–1263.
[157] R. W. Hamming, "Error detecting and error correcting codes," *Bell Syst. Techn. J.*, vol. 29, no. 2, pp. 147–160, Apr. 1950.
[158] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo-codes," *IEEE Trans. Commun.*, vol. 44, no. 10, pp. 1261–1271, Oct. 1996.
[159] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 619–637, Feb. 2001.
[160] C. Liang, K. Huang, X. Ma, and B. Bai, "Block Markov superposition transmission with bit-interleaved coded modulation," *IEEE Commun. Lett.*, vol. 18, no. 3, pp. 397–400, Mar. 2014.
[161] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for Symmetric binary-input memoryless channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051–3073, Jul. 2009.
[162] P. Akuon and H. Xu, "Polar coded spatial modulation," *IET Commun.*, vol. 8, no. 9, pp. 1459–1466, Jun. 2014.
[163] C. Vladeanu, "Turbo trellis-coded spatial modulation," in *Proc. GLOBECOM*, Anaheim, CA, USA, Dec. 2012, pp. 4024–4029.
[164] L. Wang, C. Liang, Z. Yang, and X. Ma, "Two-layer coded spatial modulation with block Markov superposition transmission," *IEEE Trans. Commun.*, vol. 64, no. 2, pp. 643–653, Feb. 2016.
[165] R. Mesleh, M. Di Renzo, H. Haas, and P. M. Grant, "Trellis coded spatial modulation," *IEEE Trans. Wireless Commun.*, vol. 9, no. 7, pp. 2349–2361, Jul. 2010.
[166] S. Alaka, T. L. Narasimhan, and A. Chockalingam, "Coded index modulation for non-DC-biased OFDM in multiple LED visible light communication," in *Proc. VTC Spring*, Nanjing, China, May 2016, pp. 1–5.
[167] L. Wang and X. Ma, "Coded index modulation with block Markov superposition transmission for highly mobile OFDM systems," in *Proc. VTC Spring*, Nanjing, China, May 2016, pp. 1–5.
[168] H. Zhang, L.-L. Yang, and L. Hanzo, "LDPC-coded index-modulation aided OFDM for in-vehicle power line communications," in *Proc. VTC Spring*, Nanjing, China, May 2016, pp. 1–5.
[169] P. Zhang, S. Chen, and L. Hanzo, "Reduced-complexity near-capacity joint channel estimation and three-stage turbo detection for coherent space-time shift keying," *IEEE Trans. Commun.*, vol. 61, no. 5, pp. 1902–1913, May 2013.
[170] P. Zhang, S. Chen, and L. Hanzo, "Embedded iterative semi-blind channel estimation for three-stage-concatenated MIMO-aided QAM turbo transceivers," *IEEE Trans. Veh. Technol.*, vol. 63, no. 1, pp. 439–446, Jan. 2014.
[171] P. Zhang, S. Chen, and L. Hanzo, "Two-tier channel estimation aided near-capacity MIMO transceivers relying on norm-based joint transmit and receive antenna selection," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 122–137, Jan. 2015.
[172] X. Li and J. A. Ritcey, "Bit-interleaved coded modulation with iterative decoding using soft feedback," *Electron. Lett.*, vol. 34, no. 10, pp. 942–943, May 1998.
[173] Q. Wang, Z. Wang, S. Chen, and L. Hanzo, "Enhancing the decoding performance of optical wireless communication systems using receiver-side predistortion," *Opt. Express*, vol. 21, no. 25, pp. 30295–30305, 2013.
[174] B. Shamasundar, S. Jacob, and A. Chockalingam, "Time-indexed media-based modulation," in *Proc. VTC Spring*, Sydney, NSW, Australia, Jun. 2017, pp. 1–5.
[175] J. G. Andrews *et al.*, "What will 5G be?" *J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.
[176] J. M. Kahn and J. R. Barry, "Wireless infrared communications," *Proc. IEEE*, vol. 85, no. 2, pp. 265–298, Feb. 1997.
[177] D. Karunatilaka, F. Zafar, V. Kalavally, and R. Parthiban, "LED based indoor visible light communications: State of the art," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 3, pp. 1649–1678, 3rd Quart., 2015.
[178] T. G. Hodgkinson, D. W. Smith, R. Wyatt, and D. J. Malyon, "Coherent optical communications," in *Proc. Opt. Fiber Commun. Conf.*, San Diego, CA, USA, Feb. 1985.
[179] M. A. Khalighi and M. Uysal, "Survey on free space optical communication: A communication theory perspective," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 2231–2258, 4th Quart., 2014.
[180] J. Jiang, P. Zhang, R. Zhang, S. Chen, and L. Hanzo, "Aperture selection for ACO-OFDM in free-space optical turbulence channel," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6089–6100, Aug. 2016.
[181] V. W. S. Chan, "Free-space optical communications," *J. Lightw. Technol.*, vol. 24, no. 12, pp. 4750–4762, Dec. 2006.



**Tianqi Mao** (S'15) received the B.S. and M.S. degrees (with Highest Hons.) in electronic engineering from Tsinghua University, Beijing, China, in 2015 and 2018, respectively, where he is currently pursuing the Ph.D. degree with the Department of Electronic Engineering at Tsinghua University.

His current research interests include modulation and signal processing for wireless communication and visible light communication. He was a recipient of the Special Scholarship of Tsinghua University, the Outstanding Master Graduate of Tsinghua University, the Outstanding Master Graduate of Beijing City, the Excellent Master Dissertation of Tsinghua University, and National Scholarship.

**Qi Wang** (S'15–M'16) received the B.E. and Ph.D. degrees (with Highest Hons.) in electronic engineering from Tsinghua University, Beijing, China, in 2011 and 2016, respectively. From 2014 to 2015, he was a Visiting Scholar with the Electrical Engineering Division, Centre for Photonic Systems, Department of Engineering, University of Cambridge. From 2016 to 2017, he was a Research Fellow with Southampton Wireless Group, University of Southampton. Since 2017, he has been a Senior Engineer with Huawei Technologies, China.

He has authored over 30 IEEE/OSA journal papers and several conference papers. He has also co-authored a book entitled *Visible Light Communications: Modulation and Signal Processing* (Wiley-IEEE Press). His research interests include modulation and signal processing for wireless communication and visible light communication. He was a recipient of the Excellent Doctoral Dissertation of the Chinese Institute of Electronics, the Outstanding Ph.D. Graduate of Tsinghua University, the Excellent Doctoral Dissertation of Tsinghua University, National Scholarship, and the Academic Star of Electronic Engineering Department in Tsinghua University. He was selected by the IEEE Series on Digital and Mobile Communication. He serves as an Associate Editor for IEEE ACCESS and a TPC member for many IEEE conferences, including Globecom, GlobalSIP, CSNDSP, and IWCMC.

**Zhaocheng Wang** (M'09–SM'11) received the B.S., M.S., and Ph.D. degrees from Tsinghua University, Beijing, China, in 1991, 1993, and 1996, respectively. From 1996 to 1997, he was a Post-Doctoral Fellow with Nanyang Technological University, Singapore. From 1997 to 1999, he was with OKI Techno Centre (Singapore) Pte. Ltd., Singapore, where he was first a Research Engineer and later became a Senior Engineer. From 1999 to 2009, he was with Sony Deutschland GmbH, where he was first a Senior Engineer and later became a Principal Engineer. He is currently a Professor of electronic engineering with Tsinghua University and serves as the Director of Broadband Communication Key Laboratory, Tsinghua National Laboratory for Information Science and Technology.

He has authored or co-authored over 100 journal papers (SCI indexed). He holds 34 granted U.S./EU patents. He has co-authored two books, one of which entitled *Millimeter Wave Communication Systems*, was selected by IEEE Series on Digital and Mobile Communication (Wiley–IEEE Press). His research interests include wireless communications, visible light communications, millimeterwave communications, and digital broadcasting. He currently serves as an Associate Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and the IEEE COMMUNICATIONS LETTERS, and has also served as the technical program committee co-chair of various international conferences. He is a fellow of the Institution of Engineering and Technology.

**Sheng Chen** (M'90–SM'97–F'08) received the B.Eng. degree in control engineering from East China Petroleum Institute, Dongying, China, in 1982, the Ph.D. degree in control engineering from the City University, London, in 1986, and the Higher Doctoral degree (Doctor of Science) from the University of Southampton, Southampton, U.K., in 2005. From 1986 to 1999, he held research and academic appointments with the University of Sheffield, U.K., the University of Edinburgh, U.K., and the University of Portsmouth, U.K. Since 1999, he has been with Electronics and Computer Science, University of Southampton, where he holds the post of a Professor in intelligent systems and signal processing.

He is a Distinguished Adjunct Professor with King Abdulaziz University, Jeddah, Saudi Arabia. His research interests include adaptive signal processing, wireless communications, modeling and identification of nonlinear systems, neural network and machine learning, intelligent control system design, evolutionary computation methods, and optimization. He has published over 600 research papers. He was a recipient of ISI Highly Cited Researcher Award in engineering in 2004. He is a fellow of the United Kingdom Royal Academy of Engineering and a fellow of IET.