Lexicographic Codebook Design for OFDM with Index Modulation

Shuping Dang, *Member, IEEE*, Gaojie Chen, *Member, IEEE* and Justin P. Coon, *Senior Member, IEEE*

***Abstract*—In this paper, we propose a novel codebook design scheme for orthogonal frequency-division multiplexing with index modulation (OFDM-IM) to improve system performance. The optimization process can be implemented efficiently by the lexicographic ordering principle. By applying the proposed codebook design, all subcarrier activation patterns with a fixed number of active subcarriers will be explored. Furthermore, as the number of active subcarriers is fixed, the computational complexity for estimation at the receiver is reduced and the zero-active subcarrier dilemma is solved without involving complex higher layer transmission protocols. It is found that the codebook design can potentially provide a trade-off between diversity and transmission rate. We investigate the diversity mechanism and formulate three diversity-rate optimization problems for the proposed OFDM-IM system. Based on the genetic algorithm (GA), the method of solving these formulated optimization problems is provided and verified to be effective. Then, we analyze the average block error rate (BLER) and bit error rate (BER) of OFDM-IM systems applying the codebook design. Finally, all analyses are numerically verified by Monte Carlo simulations. In addition, a series of comparisons are provided, by which the superiority of the codebook design is thereby confirmed.**

***Index Terms*—Orthogonal frequency-division multiplexing with index modulation (OFDM-IM), diversity-rate optimization, genetic algorithm, error performance analysis, diversity gain.**

I. Introduction

ORTHOGONAL frequency-division multiplexing (OFDM) has occupied a crucial position in modern wireless communication networks, since the popularization of fourth generation (4G) networks [1]. Specifically, OFDM utilizes a set of subcarriers with different *orthogonal frequencies* to split a frequency-selective channel into a number of frequency-flat subchannels. Because of the orthogonality of subcarrier frequencies, the subchannels can be regarded as independent and the inter-channel interference (ICI) shall be mitigated or even eliminated [2]. Meanwhile, OFDM has a high spectral efficiency and different subcarriers can overlap in the frequency domain but still maintain orthogonality, as long as certain quantitative relations among their central frequencies can be satisfied. Another attractive feature of OFDM is that it allows independent and different processing on each subcarrier, and provides a higher flexibility for system-level design [3]. From a practical point of view, OFDM is also easy to be modulated and demodulated by inverse fast Fourier transform (IFFT) and fast Fourier transform (FFT) operations, respectively [4]. Based on these merits, OFDM is also believed to play an indispensable role in next generation networks and beyond.

Recently, OFDM with index modulation (OFDM-IM), as a derivative of spatial modulation (SM) for multiple-input and multiple-output (MIMO) systems, has attracted considerable attention in both academia and industry [5]–[8]. The primary principle of OFDM-IM is to extend the modulation dimensions from two (amplitude and phase) to three (amplitude, phase and index of subcarrier). By such an extension, the information conveyed by a transmitted OFDM block is represented by data symbols as well as the indices of active subcarriers, on which these data symbols are carried [5]. Therefore, it has been proven that under certain conditions, OFDM-IM systems will gain a higher reliability and/or transmission rate than conventional OFDM systems [9]–[12].

As a modulation scheme, the superiority of OFDM-IM is highly related to the mapping relation between transmitted bit sequences and subcarrier activation patterns. In this context, how to encode the bit sequence to the data symbols carried on subcarriers as well as the active subcarrier indices becomes the key question. Some early research dedicated to studying the mapping relations is published in [13], [14], which are either with a low spectral efficiency or rely on forward error control techniques. These restrictions make them difficult to use in practical communication networks. A more realistic design of the mapping relation is proposed in [15], in which the information is mapped to the indices regarding a set of a fixed number of subcarriers, instead of the indices of subcarriers *per se*. Although the spectral efficiency of OFDM-IM has been gradually improved by continuous study, the error performance and diversity order of OFDM-IM are awaiting further enhancement.

In this regard, an equiprobable subcarrier activation (ESA) scheme is adopted for OFDM-IM systems to produce a coding gain, but without any diversity gain [16]. In [17], [18], two mapping selection schemes with more flexible mapping relations based on on-off keying (OOK) are presented and are shown to provide frequency diversity gains. However, because the number of active subcarriers is a variable, a zero-active subcarrier dilemma arises and a dual-mode transmission protocol is necessary. Also, because the modulation is per-

This work was supported by the SEN grant (EPSRC grant number EP/N002350/1) and the grant from China Scholarship Council (No. 201508060323).

S. Dang was with Department of Engineering Science, University of Oxford and the R&D Center, Guangxi Huanan Communication Co., Ltd. when completing the major work of this paper, and now with Computer, Electrical and Mathematical Science and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Kingdom of Saudi Arabia (e-mail: shuping.dang@kaust.edu.sa).

G. Chen is with Department of Engineering, University of Leicester, Leicester, U.K., LE1 7RH (email: gaojie.chen@leicester.ac.uk).

J.P. Coon is with Department of Engineering Science, University of Oxford, Oxford, U.K., OX1 3PJ (e-mail: justin.coon@eng.ox.ac.uk).

formed on a variable-length bit sequence, the detection and designs of higher layer protocols at the receiver would be challenging. Another diversity scheme based on source coding and redundancy is given in [19], but the resulting system has a considerable loss of spectral efficiency. Coordinate interleaving is employed in [20]–[22] to enhance the error performance without loss of spectral efficiency, but the diversity order is only improved from unity to two. Spatial diversity provided by relay selections for OFDM-IM systems is also investigated in [23], which renders a higher system complexity and extra signaling overheads due to the coordination among multiple relays.

To provide an easy-to-implement and efficient approach to enhance system performance, we propose a novel codebook design scheme for OFDM-IM systems in this paper. The contributions of this paper are listed as follows:

1) We propose a codebook design scheme for OFDM-IM based on the lexicographic ordering principle. We utilize the principle to generate an optimized codebook specifying the mapping relations between bit sequences and subcarrier activation patterns according to the instantaneous channel state information (CSI) and discard some subcarrier activation patterns that are not appropriate to use. By such an optimization process, we can achieve enhanced system performance without loss of spectral efficiency.
2) We also find that a potential diversity-rate trade-off can be provided by the codebook design scheme. Following this, we investigate the diversity mechanism and formulate three diversity-rate optimization problems suited for different application scenarios with solutions based on the genetic algorithm (GA).
3) We analyze and approximate the average block error rate (BLER) as well as the average bit error rate (BER) of the proposed system in closed form, with the help of the approximation of the Q-function and the union bound.
4) We verify the superiority of the codebook design enhanced OFDM-IM over the conventional OFDM-IM and the OFDM scheme without IM by a series of numerical simulations.

The rest of this paper is organized as follows. In Section II, we introduce the system model as well as the assumptions used in this paper. Then, we detail the algorithm to generate the optimized codebook in Section III. Following the codebook design scheme, we investigate its potential diversity mechanism and formulate three diversity-rate optimization problems with solutions based on the GA in Section IV. After that, we analyze and approximate the average BLER and BER in closed form in Section V, which is subsequently verified by numerical results generated by Monte Carlo simulations in Section VI. Finally, Section VII concludes the paper.

II. System Model

A. System Framework

In this paper, we consider an OFDM-IM system operating over a frequency-selective *slow* Rayleigh fading channel. Assuming a sufficiently long cyclic prefix (CP) is inserted, transmissions over N independent subcarriers without inter-carrier interference (ICI) can be enabled by a N-point IFFT. Then, K out of N subcarriers will be activated to form a *subcarrier activation pattern* according to an equiprobable bit sequence $\mathbf{b}$ with fixed length B[1]. To facilitate the following analysis, we denote the set of all subcarriers as $\mathcal{N}$. It is apparent that the number of all possible patterns can be determined by $L = \binom{N}{K}$, where $\binom{\cdot}{\cdot}$ is the binomial coefficient. However, to modulate a sequence of bits, we have to truncate the number of possible patterns to a power of two. We can denote the truncated set of patterns as $\mathcal{S}(c)$ and the truncated set is termed a *codebook*, where $c \in \mathcal{C}$ is the index of the codebook and $\mathcal{C}$ denotes the set of all possible codebooks. Seemingly, the number of patterns in $\mathcal{S}(c)$ can be derived as

$$S = |\mathcal{S}(c)| = 2^{\lfloor \log_2(L) \rfloor}, \ \forall \ c \in \mathcal{C} \tag{1}$$

where $\lfloor \cdot \rfloor$ represents the floor function. As a consequence, we can determine the number of codebooks to be $C = |\mathcal{C}| = \binom{L}{S}$. Also, by (1), it is straightforward to see that the length of the bit sequence mapped to a subcarrier activation pattern is $B_S = \log_2(S)$. Assume M-ary phase shift keying (M-PSK) is used for amplitude phase modulation (APM) of the symbol carried on each active subcarrier[2]. As there are always K active subcarriers, we can write the transmission rate in bits per channel use (bpcu) (i.e. the length B of the equiprobable bit sequence $\mathbf{b}$) as $B = B_S + B_M$, where $B_M = K \log_2(M)$ represents the length of the segment of the bit sequence mapped to APM constellation symbols. More explicitly, we can express B in terms of N, K and M as

$$B = \left\lfloor \log_2 \binom{N}{K} \right\rfloor + K \log_2(M). \tag{2}$$

From the description above, we can see that the mapping relation between bit sequence and subcarrier activation pattern is exactly the same as stipulated in [15], except for the set truncating process. As a result, the proposed system inherits the merits of having a fixed number of active subcarriers and the zero-active subcarrier dilemma (referring to the case where all subcarriers are switched off in order to represent an all-zero bit sequence) can be prevented in this scenario accordingly [17]. Therefore, the forward error control techniques, dual-mode transmission protocol and always-active control subcarrier employed in [14], [17] and [25] are not required, which will lead to a simpler system framework and unified performance analysis.

B. Signal Transmission

Now, the bit sequence $\mathbf{b}$ with length B can be mapped to the subcarrier activation pattern consisting of K active subcarriers and the K data symbols conveyed on these active subcarriers. To uniquely express the subcarrier activation pattern, we introduce the activation state vector (ASV) of subcarriers by

$$\mathbf{v} = [v(n_1), v(n_2), \ldots, v(n_N)] \in \{0, 1\}^{1 \times N}, \tag{3}$$

[1]The bit sequence $\mathbf{b}$ also holds other information aside from the bits to be mapped to the activation pattern.

[2]It has been proven that PSK, as a constant-envelop APM scheme, outperforms other non-constant-envelop APM schemes, e.g. quadrature amplitude modulation (QAM) [24]. We thereby adopt M-PSK in this paper.

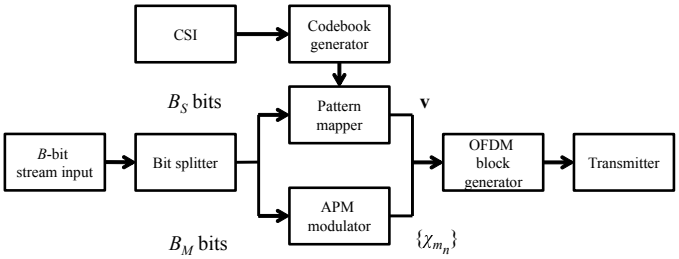


Fig. 1: System block diagram of the transmission part of the proposed OFDM-IM enhanced by the codebook design.

where $v(n_\xi)$ is either '0' or '1' representing whether the n_ξth ordered subcarrier (more detail of the ordering process will be introduced in Section III) is inactive or active, respectively. Accordingly, we can express the transmitted OFDM block as

$$\mathbf{x}(c,\mathbf{b}) = [x(m_1,1), x(m_2,2), \ldots, x(m_N,N)]^T \in \mathbb{C}^{N\times 1}, \tag{4}$$

where $(\cdot)^T$ denotes the matrix transpose operation. The element of the block $\mathbf{x}(c,\mathbf{b})$ is given by

$$x(m_{n_\xi}, n_\xi) = \begin{cases} \chi_{m_{n_\xi}}, & v(n_\xi) = 1 \\ 0, & v(n_\xi) = 0 \end{cases} \tag{5}$$

where χ_{m_n} is a M-ary constellation symbol; $m_{n_\xi} \in \mathcal{M}$ and $\mathcal{M}$ is the set of indices of constellation symbols when M-PSK is used; without loss of generality, we can normalize the constellation symbol by $\chi_{m_{n_\xi}}\chi^*_{m_{n_\xi}} = 1$, $\forall\ m_n \in \mathcal{M}$. For clarity, we illustrate the system block diagram of the transmission part of our proposed OFDM-IM system in Fig. 1.

C. Signal Propagation and Reception

At the receiver with perfect synchronization, after sampling, discarding the CP, and performing a FFT, we can express the received OFDM block as

$$\begin{aligned}\mathbf{y}(c,\mathbf{b}) &= [y(m_1,1), y(m_2,2), \ldots, y(m_N,N)]^T \\ &= \sqrt{\frac{P_t}{K}}\mathbf{H}\mathbf{x}(c,\mathbf{b}) + \mathbf{w} \in \mathbb{C}^{N\times 1},\end{aligned} \tag{6}$$

where $\mathbf{w} = [w(1), w(2), \ldots, w(N)]^T \in \mathbb{C}^{N\times 1}$ denotes the vector of N independent complex additive white Gaussian noise (AWGN) samples on each subcarrier, whose entries $w(n) \sim \mathcal{CN}(0, N_0)$, and N_0 is the noise power; $\mathbf{H} = \mathrm{diag}\{h(1), h(2), \ldots, h(N)\} \in \mathbb{C}^{N\times N}$ is a $N \times N$ diagonal channel state matrix (CSM) characterizing the channel quality; P_t is the total transmit power that can be used for data transmission at the transmitter and we assume it is uniformly distributed over all K active subcarriers.

In this paper, we assume all wireless channels to be frequency-selective *slow*[3] Rayleigh fading channels and the corresponding channel gains to be exponentially independent and identically distributed (i.i.d.) with mean μ. Therefore, the probability density function (PDF) and the cumulative distribution function (CDF) of the channel gain $G(n) = |h(n)|^2$, $\forall\ n \in \mathcal{N}$ are given by [26]

$$f_g(\epsilon) = \exp(-\epsilon/\mu)/\mu \quad \Leftrightarrow \quad F_g(\epsilon) = 1 - \exp(-\epsilon/\mu). \tag{7}$$

At the receiver, because of the slow fading model adopted in this paper, the optimized codebook and CSI, i.e. $\mathcal{S}(c)$ and $\mathbf{H}$ are assumed to be known. As a result of this *a priori* information, the maximum-likelihood (ML) detection method can be employed to perform block detection according to the criterion below:

$$\hat{\mathbf{x}}(c,\hat{\mathbf{b}}) = \underset{\mathbf{x}(c,\mathbf{b})\in\mathcal{X}(c)}{\arg\min} \left\| \dot{\mathbf{y}}(c,\dot{\mathbf{b}}) - \sqrt{\tfrac{P_t}{K}}\mathbf{H}\mathbf{x}(c,\mathbf{b}) \right\|_F, \tag{8}$$

where $\|\cdot\|_F$ denotes the Frobenius norm of the enclosed argument; $\dot{\mathbf{y}}(c,\dot{\mathbf{b}})$ denotes the received OFDM block contaminated by noise; $\mathbf{x}(c,\mathbf{b})$ denotes the estimation trial; $\mathcal{X}(c)$ is the set of all legitimate $\mathbf{x}(c,\mathbf{b})$ when the codebook $\mathcal{S}(c)$ is in use and we can easily see that $X = |\mathcal{X}(c)| = SM^K$.

Then, we can define the *average BLER* as

$$\bar{P}_e = \mathop{\mathbb{E}}_{\dot{\mathbf{b}}}\left\{ \mathop{\mathbb{E}}_{\mathbf{H}}\left\{ P_e(\dot{\mathbf{x}}(c,\dot{\mathbf{b}})|\mathbf{H}) \right\} \right\}, \tag{9}$$

where $\mathbb{E}\{\cdot\}$ denotes the expectation of the enclosed; the conditional BLER when $\dot{\mathbf{x}}(c,\dot{\mathbf{b}})$ is transmitted is approximated by the well-known union bound to be [27]

$$P_e(\dot{\mathbf{x}}(c,\dot{\mathbf{b}})|\mathbf{H}) \le \sum_{\hat{\mathbf{x}}(c,\hat{\mathbf{b}})\neq\dot{\mathbf{x}}(c,\dot{\mathbf{b}})} P_e(\dot{\mathbf{x}}(c,\dot{\mathbf{b}}) \to \hat{\mathbf{x}}(c,\hat{\mathbf{b}})|\mathbf{H}), \tag{10}$$

and $P_e(\dot{\mathbf{x}}(c,\dot{\mathbf{b}}) \to \hat{\mathbf{x}}(c,\hat{\mathbf{b}})|\mathbf{H})$ represents the conditional probability of the event that the original transmitted block $\dot{\mathbf{x}}(c,\dot{\mathbf{b}})$ is erroneously estimated to be $\hat{\mathbf{x}}(c,\hat{\mathbf{b}})$. The average BLER $\bar{P}_e$ is used as an important metric in this paper to evaluate the performance of the proposed OFDM-IM system.

III. Codebook Generating Algorithm

A. Mathematical Principle of the Codebook Generating Algorithm

As we specified above, there are $L = \binom{N}{K}$ available patterns in total, from which only $S = 2^{\lfloor \log_2(L) \rfloor}$ patterns will be selected to form a codebook and used for OFDM-IM. In other words, there are

$$\Delta(N,K) = L - S = \binom{N}{K} - 2^{\lfloor \log_2 \binom{N}{K} \rfloor} \tag{11}$$

patterns, which should be discarded. Specifically, denoting the set of all patterns as $\mathcal{L}$, we need to find $\mathcal{S}(c) \subseteq \mathcal{L}$, by which the corresponding average BLER can be reduced. However, as we derived above, $C = |\mathcal{C}| = \binom{L}{S}$ will be a large number for intermediate N and K, which makes the codebook design demanding[4]. Here, to perform the optimization efficiently, we propose a codebook generating algorithm based on the lexicographic ordering principle [28]. First, according to the

[3]The slow fading/quasi-static fading model adopted in this paper means that the CSM will remain constant for a sufficiently long period of time, which justifies the codebook design mechanism, as the signaling overhead and computational resource for performing such an optimization can be maintained below a reasonable level.

[4]For example, when $N = 8$ and $K = 4$, the numbers of available and selected patterns are determined by $L = \binom{8}{4} = 70$ and $S = 2^{\lfloor \log_2 \binom{8}{4} \rfloor} = 64$. Therefore, $C = \binom{70}{64} = 131115985$.

obtained CSI, we rank $G(1), G(2), \ldots, G(N)$ in terms of their values in ascending order. Then, we have the ordered channel gains as

$$G_{\langle 1\rangle}(n_1) < G_{\langle 2\rangle}(n_2) < \cdots < G_{\langle N\rangle}(n_N), \tag{12}$$

where n_ξ denotes the original index of the channel gain ranked in the ξth place in ascending order. For clarity, we term n_ξ and ξ the *index* and the *order* in this paper, respectively.

We can generate a list consisting of L ASVs by a certain order, and each of them represents a unique subcarrier activation pattern with K active subcarriers. Since multiple subcarriers are considered in OFDM-IM systems, which all have impacts on the performance, it would be tough to define a general criterion to evaluate and order all L patterns in terms of their quality. Therefore, to facilitate this ordering process and ease optimality analysis, we only consider the asymptotic region when the transmit power becomes large. In the asymptotic region, it has been proven that the error performance of OFDM systems is dominated by the worst subarrier (i.e. the one with the lowest channel gain) [29]. In this regard, the patterns in which the worst active subcarriers have larger gains must be superior to those whose worst active subcarriers have smaller gains. This provides a hint for the pattern ordering and we might resort to the *reverse lexicographic order* [28]. As a result, we can treat an ASV as a binary sequence where the most significant digit is on the left side and convert it to a decimal number, which can be used as an indicator to characterize the performance of the corresponding pattern in the asymptotic region. Mathematically, for two arbitrary ASVs $\mathbf{v}_i$ and $\mathbf{v}_j$, we define the lexicographic ordering relation as [30]

$$\mathbf{v}_i <_{\mathsf{L}} \mathbf{v}_j \;\Leftrightarrow\; \omega_i < \omega_j, \tag{13}$$

where $\omega_i = \mathrm{b2d}_l(\mathbf{v}_i)$ is called an ordering indicator and $\mathrm{b2d}_l(\cdot)$ converts a binary vector to a decimal number given the most significant digit on the left side. With the help of ω_i, we can easily enumerate all L ASVs in the lexicographically ascending order as $\mathbf{v}_1 <_{\mathsf{L}} \mathbf{v}_2 <_{\mathsf{L}} \cdots <_{\mathsf{L}} \mathbf{v}_L$, since the reverse lexicographic order is a *well-order relation* [31]. As we have ordered the channel gains in ascending order, patterns with smaller ordering indicators are superior, and we can simply select the S lexicographically smallest ASVs to form the optimized codebook $\mathcal{S}(c)$ by

$$\mathcal{S}(c) = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_S\}. \tag{14}$$

B. An Simplistic Example of the Codebook Generating Process

To be clear, we demonstrate a simplistic example of the codebook generating process by reverse lexicographic order when $N = 4$ and $K = 2$ as follows.

1) Initialization: For $N = 4$ and $K = 2$, it can be calculated that $L = \binom{N}{K} = \binom{4}{2} = 6$ and $S = 2^{\lfloor \log_2(L)\rfloor} = 2^{\lfloor \log_2(6)\rfloor} = 4$. Also, the length of the segment of bit sequence mapped to subcarrier activation patterns can be determined by $B_S = \log_2(S) = \log_2(4) = 2$. Then, we initialize an empty set with size S to represent $\mathcal{S}(c)$.

TABLE I: Mapping relation between the B_S-bit sequence segment and subcarrier activation patterns for a simplistic example when $N = 4$ and $K = 2$.

B_S bits	ASVs	Orders	Indices	$\mathbf{x}(c, \mathbf{b})$
$[0, 0]$	$\mathbf{v}_1 = [0, 0, 1, 1]$	$\{3, 4\}$	$\{2, 3\}$	$[0, \chi_{m_2}, \chi_{m_3}, 0]$
$[0, 1]$	$\mathbf{v}_2 = [0, 1, 0, 1]$	$\{2, 4\}$	$\{2, 4\}$	$[0, \chi_{m_2}, 0, \chi_{m_4}]$
$[1, 0]$	$\mathbf{v}_3 = [0, 1, 1, 0]$	$\{2, 3\}$	$\{3, 4\}$	$[0, 0, \chi_{m_3}, \chi_{m_4}]$
$[1, 1]$	$\mathbf{v}_4 = [1, 0, 0, 1]$	$\{1, 4\}$	$\{1, 2\}$	$[\chi_{m_1}, \chi_{m_2}, 0, 0]$

TABLE II: Relation among orders, indices of subcarriers and ASVs when $N = 4$ and $K = 2$ (the entries of ASVs corresponding to the worst active subcarriers in a given activation pattern are in boldface).

Orders ξ	1	2	3	4
Indices n_ξ	1	4	3	2
$\mathbf{v}_1$	[0	0	**1**	1]
$\mathbf{v}_2$	[0	**1**	0	1]
$\mathbf{v}_3$	[0	**1**	1	0]
$\mathbf{v}_4$	[**1**	0	0	1]
$\mathbf{v}_5$	[**1**	0	1	0]
$\mathbf{v}_6$	[**1**	1	0	0]

2) Channel ordering: Assume all channel gains can be perfectly estimated and we know that $G(1) = 0.2$, $G(2) = 2.6$, $G(3) = 2.5$ and $G(4) = 0.4$. Then, we can rank them in terms of the value in ascending order and obtain $G_{\langle 1\rangle}(1) < G_{\langle 2\rangle}(4) < G_{\langle 3\rangle}(3) < G_{\langle 4\rangle}(2)$.

3) Generating $\mathcal{S}(c)$: It can be easily shown that we have the following six ASVs in reverse lexicographic order: $\mathbf{v}_1 = [0, 0, 1, 1]$ ($\omega_1 = 3$), $\mathbf{v}_2 = [0, 1, 0, 1]$ ($\omega_2 = 5$), $\mathbf{v}_3 = [0, 1, 1, 0]$ ($\omega_3 = 6$), $\mathbf{v}_4 = [1, 0, 0, 1]$ ($\omega_4 = 9$), $\mathbf{v}_5 = [1, 0, 1, 0]$ ($\omega_5 = 10$) and $\mathbf{v}_6 = [1, 1, 0, 0]$ ($\omega_6 = 12$). Then, we can simply select $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ and $\mathbf{v}_4$ to form the optimized codebook $\mathcal{S}(c) = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$. According to the ascending order $G_{\langle 1\rangle}(1) < G_{\langle 2\rangle}(4) < G_{\langle 2\rangle}(3) < G_{\langle 4\rangle}(2)$ and $\mathcal{S}(c)$ as well as the definition of ASV given in (3), we have the relation among orders, indices and ASVs presented in Table I. Following the relation given in Table I, we have the mapping relation between the B_S-bit stream segment and subcarrier activation patterns in Table II, by which it is obvious that the activation frequency of a subcarrier in the optimized codebook $\mathcal{S}(c)$ is related to its channel quality[5]. The lower ordered subcarriers are not used as much but the higher ordered subcarriers are more likely to be employed (e.g. subcarrier 1 (the worst subcarrier) is only used once in all selected patterns, but subcarrier 2 (the best subcarrier) is used three times).

IV. Diversity Mechanism and Diversity-Rate Optimization

A. Diversity Mechanism

The diversity gain of multi-carrier systems can be expressed by the number of independent sets of subcarrier activation patterns covering all legitimate $\mathbf{x}(c, \mathbf{b})$ [29], [32]. Following this rationale, it can be deduced that for multi-carrier systems, the error performance is dominated by the deepest faded

[5]We intend to list the discarded ASVs $\mathbf{v}_5$ and $\mathbf{v}_6$ in order to illustrate how the worst subcarrier changes after performing the reverse lexicographic ordering process.

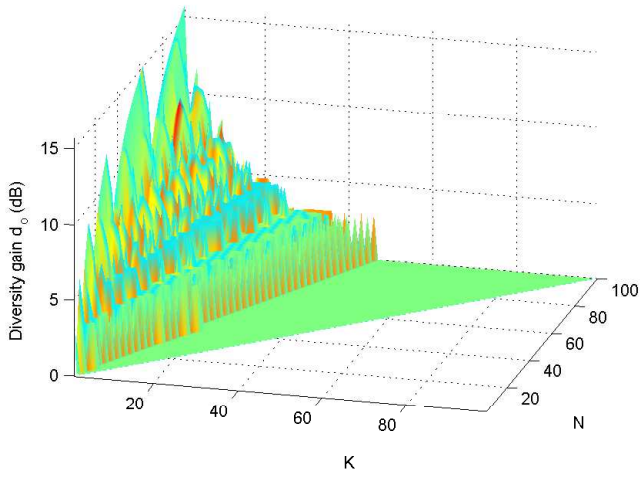


Fig. 2: Relation among diversity gain d_o, N and K.

subcarrier (i.e. the subcarrier with the lowest channel gain) when transmit power becomes large [33]. Therefore, if we denote the lowest order of the subcarrier taken in $\mathcal{S}(c)$ as $\xi_{\min}$, the diversity gain can be expressed as $d_o = \xi_{\min}$. As we order all subcarriers to generate the codebook by the lexicographic ordering principle, there is the potential to provide a diversity gain by properly choosing N and K, so that some inappropriate subcarriers can be completely eliminated from the generated codebook. From the viewpoint of ordered subcarrier combinations, we can also express the number of subcarrier activation patterns L by

$$L = \Omega(N, K, N-K+1) = \sum_{\xi=1}^{N-K+1} \binom{N-\xi}{K-1}, \tag{15}$$

where $\Omega(N, K, \nu) = \sum_{\xi=1}^{\nu} \binom{N-\xi}{K-1}$ representing the number of combinations out of $\binom{N}{K}$ with ones in the ν most significant bits. As a result, we can explicitly give the expression of d_o in terms of N and K as follows:

$$d_o = \begin{cases} 1, & \Delta(N,K) < \Omega(N,K,1) \\ \nu, & \Omega(N,K,\nu-1) \le \Delta(N,K) < \Omega(N,K,\nu) \end{cases} \tag{16}$$

which can be plotted in Fig. 2. It can be observed from Fig. 2 that the relation among diversity gain d_o, N and K is non-monotone and highly complex.

Remark 1: The above explanations as well as the expression of d_o given in (16) might not be straightforward to understand. To understand them better, we might view it from the perspective of independent fading replicas. According to [27], diversity techniques can be understood as the supply of multiple replicas of the same information-bearing signal. Hence, as long as the best replica (for example with the largest end-to-end channel gain) selected from these multiple replicas can be successfully decoded at the receiver, the information intended to be transmitted is retrieved[6]. It has also been proven that for M-ary orthogonal signals, the error performance is governed by the worst pairwise error event[7], in which the worst subcarrier in mapping codebook is activated. Obviously, without codebook design and subcarrier selection, the worst subcarrier involved in the codebook will be the worst one in the full set of subcarriers $\mathcal{N}$, and thereby the diversity gain $d_o = 1$ for this case. Following this rationale, if we can have multiple independent subcarrier activation patterns[8] to represent the same bit sequence, we can select the best subcarrier activation pattern so as to raise the order of the worst subcarrier from 1 to $\xi_{\min}$. This is the essential principle of the diversity mechanism of the lexicographic codebook design scheme. More intuitively speaking, the diversity gain comes from the prevention of using 'bad' subcarrier(s) in the codebook, which can be regarded as a special case of subcarrier assignment for multi-carrier systems [34]–[36]. Therefore, we resort to the basics of combinatorics to identify the order of the worst subcarrier contained in the mapping codebook generated by the lexicographic ordering principle, which immediately gives (16).

B. Diversity-Rate Optimization

By (2), we can also plot the relation among transmission rate B, N and K with different APM order M in Fig. 3, from which it is also evident that the relation between B and K is non-monotone, especially when M is small. Considering both diversity and data rate are important performance evaluation metrics for wireless communication systems' reliability and efficiency, we can formulate three diversity-rate optimization problems as follows, which will cover a wide range of applications. The first diversity-rate optimization is for diversity-critical systems, in which reliability is of high importance, e.g. Internet of Things, vehicular networks and military wireless sensor networks (WSNs) [37]–[39]. In particular, there is only a minimum quality of service (QoS) requirement for transmission rate, say $\underline{B}$, and the objective is to maximize the diversity gain. We can mathematically formulate this optimization problem infra:

$$\begin{aligned} &\max_{\{N,K\}} \{d_o\} \\ &\text{s.t.} \quad 2 \le N \le \overline{N},\ N \in \mathbb{N}^+ \\ &\qquad 1 \le K \le N-1,\ K \in \mathbb{N}^+ \\ &\qquad B > \underline{B} \end{aligned} \tag{17}$$

where $\overline{N}$ is the maximum allowed number of subcarriers in OFDM-IM systems, which is regulated by the spectral resource assigned. It would be emphasized that the optimization constraint B is a natural number according to (2).

[6]For traditional communication systems, independent fading replicas can be generated either in the time, frequency or space domains, which yields time, frequency or space diversity, respectively. For OFDM-IM systems, independent fading replicas could be generated in the *index domain* [15], and thereby the so-called *index diversity* is attainable, which can be regarded as a special case of frequency diversity.

[7]This phenomenon is also observed and investigated in [33].

[8]The independence among subcarrier activation patterns refers to the exclusive use of different worst subcarriers according to different codebooks.

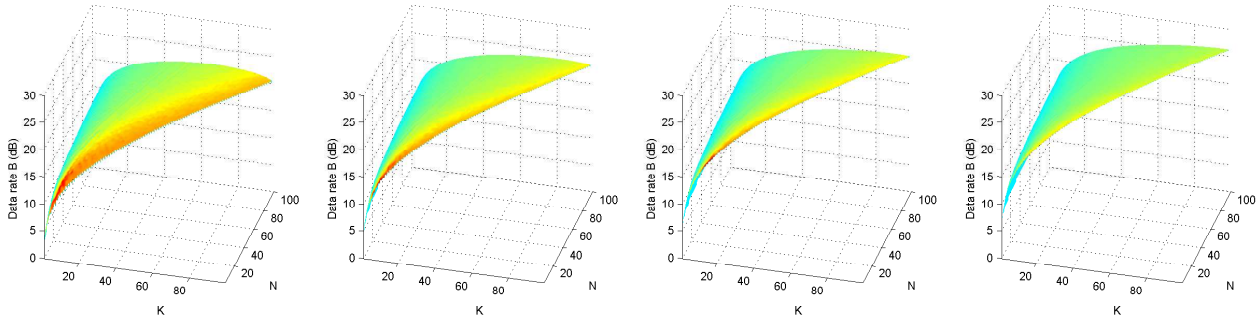


Fig. 3: Relation among transmission rate B, N and K with $M = 2, 4, 8, 16$ (from left to right).

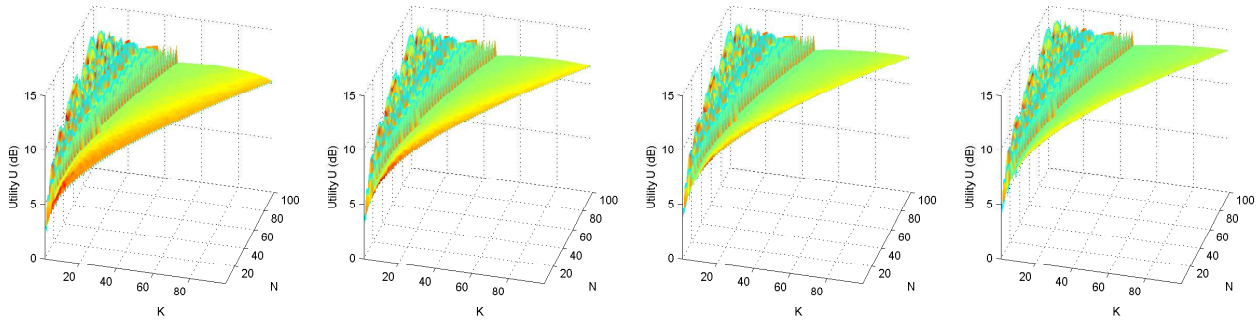


Fig. 4: Relation among utility U, N and K with $M = 2, 4, 8, 16$ (from left to right), given $w_1 = w_2 = 0.5$.

In a similar manner, we have the diversity-rate optimization for rate-critical systems, which dedicate to reinforcing the transmission rate, e.g. data networks for multimedia and virtual reality (VR) applications [40], [41]. In these cases, there is only a minimum QoS requirement for reliability, denoted as $\underline{d}_o$, and the optimization problem can be formulated as

$$\begin{aligned} &\max_{\{N,K\}} \{B\} \\ &\text{s.t.} \quad 2 \leq N \leq \overline{N}, \ N \in \mathbb{N}^+ \\ &\qquad 1 \leq K \leq N-1, \ K \in \mathbb{N}^+ \\ &\qquad d_o > \underline{d}_o \end{aligned} \tag{18}$$

Here, it should also be emphasized that the optimization constraint d_o is a natural number according to (16).

Both optimization problems formulated for diversity-critical and rate-critical systems in (17) and (18) belong to the category of single-objective optimization problems. In addition, we can also regard d_o and B as a whole and optimize both together as a multi-objective optimization problem, which provides a flexibility to trade off the QoS requirements between diversity and transmission rate and is thereby more suited for heterogeneous networks (HetNets) [42]. To be specific, as both d_o and B are expected to be maximized, we can simply employ their product with corresponding powers of weights as a single utility function

$$U = d_o^{w_1} B^{w_2}, \tag{19}$$

where $w_1 \in [0, 1]$ and $w_2 \in [0, 1]$ are the weights depending on the significance of both measures and we also have[9] $w_1 + w_2 = 1$. Subsequently, we can formulate the joint diversity-rate optimization problem as

$$\begin{aligned} &\max_{\{N,K\}} \{U\} = \max_{\{N,K\}} \{d_o^{w_1} B^{w_2}\} \\ &\text{s.t.} \quad 2 \leq N \leq \overline{N}, \ N \in \mathbb{N}^+ \\ &\qquad 1 \leq K \leq N-1, \ K \in \mathbb{N}^+ \end{aligned} \tag{20}$$

To provide intuition, the relation among utility U, N and K with different APM order M is plotted in Fig. 4, given $w_1 = w_2 = 0.5$.

C. Solutions to Formulated Optimization Problems by the Genetic Algorithm

In this subsection, let us focus on solving the formulated optimization problems in the last subsection. First, we can characterize the computational complexity by the total number of combinations of $\{N, K\}$ (a.k.a. the size of the search space) for a given $\overline{N}$ as

$$\Xi(\overline{N}) = \sum_{n=1}^{\overline{N}-1} n = \frac{1}{2}\overline{N}(\overline{N}-1), \tag{21}$$

which is obviously with a quadratic time complexity when implementing a brute-force method searching for the optimal solution over the full set of $\{N, K\}$. As given in (2) and (16), we can easily find the optimization problems formulated in (17), (18) and (20) to be nonlinear integer programming problems [43]. Also, as plotted in Fig. 2 and Fig. 3, although B can be proven to a discrete concave function of N and K [44], this is not the case for d_o, owing to its zigzag nature. Therefore, it is not possible to apply convex optimization methods to find out the optimal solutions of $\{N, K\}$ to

[9]The determination of w_1 and w_2 in engineering practice should take the policy and QoS requirements into consideration and requires domain knowledge [40]. As this is out of the scope of this paper, we therefore omit an in-depth discussion regarding relevant issues.

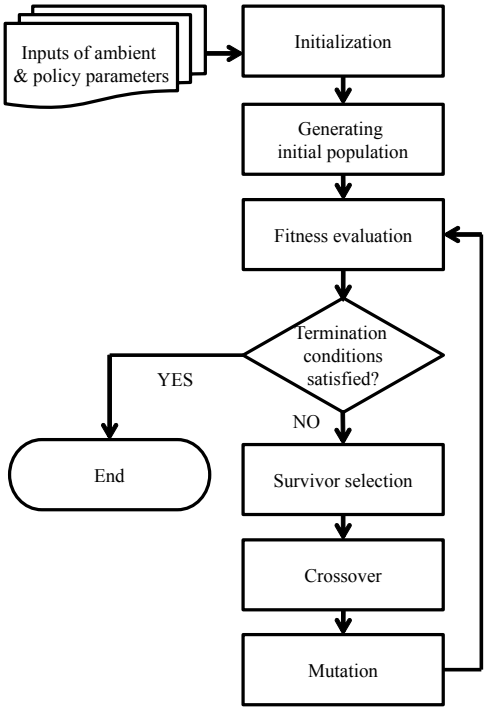


Fig. 5: A generic structure of the GA employed in this paper to solve the diversity-rate optimization problems.

the above formulated optimization problems. Alternatively, in this subsection, we apply the GA to search for the optimal solutions, by which a combination $\{N, K\}$ is represented by a *chromosome* C. Moreover, it has been rigorously proven that as long as a sufficiently large operation time is allowed, the solution generated by the GA will converge to the optimal solution [45]. A generic structure of the GA to solve the diversity-rate optimization problems is shown in Fig. 5. Now, we expatiate each of the functional blocks in this structure and construct the optimization model based on the GA in the following steps.

1) Initialization and input parameters: In this initial stage, all required parameters for all other stages should be specified and relevant operations are stipulated (i.e. the meta-data and meta-rules). For generating the initial population, we have to specify the size of population V_p and the rule of generating the initial V_p chromosomes (i.e. V_p combinations $\mathsf{C} = \{N, K\}$). For fitness evaluation, we need to input the optimization objective and constraints, which serve as criteria to evaluate each chromosome and assign a corresponding fitness metric. Then, we should explicitly regulate the termination condition, i.e. when this optimization process should be terminated and how the best chromosome in the last generation can be selected as the output combination. If the termination conditions are not satisfied, we then carry out the selection process and thereby we should tell the optimizer how to perform the selection when initializing. Having selected new parental chromosomes by a certain rule, we then perform crossover processing to exchange genes between two parental chromosomes and produce two offspring chromosomes for the next generation of V_p chromosomes. The random process of crossover should also be stipulated with relevant parameters, e.g. the crossover probability P_c. To avoid being trapped in a local optimum, we mutate over the new generation of chromosomes to explore a larger search space and the rule as well as the mutation probability P_m are thus required to be defined.

2) Generating initial population: To start the optimization process by the GA, we have to generate the first population consisting of V_p chromosomes. We can easily do so by two steps. In the first step, we randomly select an integer for N from set $[2, \overline{N}]$ with a uniform probability (allowing repetitions). Then, for an arbitrary integer N, we randomly select an integer K from set $[1, N-1]$ to complete a chromosome structure $\mathsf{C} = \{N, K\}$. Then, we repeat this process V_p times and obtain the initial population of V_p chromosomes.

3) Fitness evaluation: In this stage, we can check the fitnesses of V_p chromosomes by the objective given in (17), (18) or (20). Then, for each $\mathsf{C}_i = \{N_i, K_i\}$, $i = 1, 2, \ldots, \mathsf{V}_p$, we will have a measurement M_i, which could be the diversity, data rate or combined utility. It is worth noting that if any of optimization constraints are violated by a chromosome, its measurement will be set to zero[10]. Then, we normalize the measurement M_i by

$$\mathsf{m}_i = \frac{\mathsf{M}_i}{\mathsf{M}_1 + \mathsf{M}_2 + \cdots + \mathsf{M}_{\mathsf{V}_p}}, \tag{22}$$

which is defined as the *fitness* of the corresponding chromosome and equals the survival probability used for the survivor selection.

4) Termination conditions: After obtaining the fitnesses of all chromosomes in the present generation, we then examine the termination conditions. To simplify the optimization process, we can simply assume a fixed number of generations $\overline{\mathsf{G}}$, which is *sufficiently* large to reach an appropriate average fitness level. Then, if the current generation has not reached $\overline{\mathsf{G}}$ yet, we carry on the optimization process and increase the current generation G by one increment. Otherwise, we terminate the process and give the index of the optimized combination and finally have the optimized combination as

$$\begin{cases} i_{\text{opt}} = \arg\max\limits_{i} \{\mathsf{m}_i\} \\ \mathsf{C}_{i_{\text{opt}}} = \{N_{i_{\text{opt}}}, K_{i_{\text{opt}}}\} \end{cases}. \tag{23}$$

5) Survivor selection: As the fitness we defined above has been a normalized value ranging between 0 and 1, we can directly apply them to the standard roulette wheel selection process as the *survival probabilities* [46]. Then, we conduct the standard roulette wheel selection process V_p times[11] to select V_p parental chromosomes prepared for mating to reproduce the next generation chromosomes.

6) Crossover: Crossover is an efficient approach to explore the search space. In terms of the parity of V_p, we define two modes to perform crossover: 1) if V_p is even, the crossover

[10]In the case where the measurements of all chromosomes are set to zero, the chromosome selection cannot be performed. To mitigate this problem, we stipulate that the measurements are equal to $1/\mathsf{V}_p$ when the optimization constraints have been violated by all V_p chromosomes.

[11]This is because in this paper we need to guarantee the same sample numbers of chromosomes in all generations. Although a shrinking/variable number of chromosomes in the next generation is also possible, we do not adopt this setting for simplicity.

process might be performed between two adjacent parental chromosomes with probability P_c, as long as both offspring chromosomes do not violate the optimization constraints; 2) if V_p is odd, the crossover process might be performed between two adjacent parental chromosomes with probability P_c, as long as both offspring chromosomes do not violate the optimization constraints, but the last single parental chromosome will remain the same without crossover. After performing the crossover process, we will have V_p offspring chromosomes.

7) Mutation: To avoid being trapped in a local optimum, we resort to mutation to provide a larger exploration of the search space. Considering the chromosome structure $\mathsf{C} = \{N, K\}$ in this optimization process, we design a joint mutation process here. In the joint mutation process, both N and K will be mutated together with probability P_m or remain the same with probability $1 - \mathsf{P}_m$. Specifically, Once mutation occurs, N in the chromosome $\mathsf{C} = \{N, K\}$ is mutated to $\hat{N}$, which is randomly chosen from the set $[2, \overline{N}]\backslash\{N\}$. Then, K in the chromosome $\mathsf{C} = \{N, K\}$ is mutated to $\hat{K}$, which is randomly chosen from the set $[1, \hat{N}-1]\backslash\{K\}$. In this way, we can avoid violating the constraint between N and K after performing mutation, so that the convergence rate of GA will be improved.

By the detailed descriptions of all functional blocks, we have the pseudocode of the GA-based optimization in **Algorithm 1**. From **Algorithm 1**, it can be found that the computational complexity of the proposed optimization algorithm based on the GA is $\mathcal{O}(\bar{\mathsf{G}}\mathsf{V}_p)$ regardless of the search space, while the computational complexity of the brute-force method is $\mathcal{O}(\bar{N}^2)$ as the search space is given by (21), where $\mathcal{O}(\cdot)$ is the big O notation[12].

V. Error Performance Analysis

In this section, we analyze the error performance of OFDM-IM systems assuming the optimal values of N and K are attained via the GA described above. In this paper, we utilize average BLER to characterize the error performance. To calculate the average BLER, we first need to investigate the order statistics of the channel gain, as the channel ordering algorithm is involved in the reverse lexicographic ordering process. According to (7), we can write the PDF of the ξth order statistic of the channel gain among N subcarriers as [47]

$$\phi_{\langle\xi\rangle}(\epsilon) = \frac{N!(F_g(\epsilon))^{\xi-1}(1-F_g(\epsilon))^{N-\xi}f_g(\epsilon)}{(\xi-1)!(N-\xi)!}, \tag{24}$$

which can be used to average the conditional BLER after performing the codebook design over all channel states. For now, we temporarily neglect the codebook design process. We focus on the basic element of the average BLER, i.e. the conditional probability of the event that the original transmitted block

[12]For two arbitrary functions $f(x)$ and $g(x)$, let $x \to \lambda$ where λ can be either a real number, infinity or infinitesimal. It is defined that $f(x) = \mathcal{O}(g(x))$, if $\lim_{x\to\lambda}\{f(x)/g(x)\}$ is bounded.

Algorithm 1 Proposed GA for solving optimization problems for OFDM-IM systems applying the lexicographic codebook design scheme.

BEGIN
-Initialization
Input: V_p, $\overline{\mathsf{G}}$, P_c, P_m, $\overline{N}$ and M as well as objective and constraints;
$\mathsf{G} \leftarrow 1$;
-Generate initial population
for $i = 1 : \mathsf{V}_p$ **do**
  $N_i \leftarrow \text{Random}\{2, 3, \ldots, \overline{N}\}$;
  $K_i \leftarrow \text{Random}\{1, 2, \ldots, N_i - 1\}$;
  $\mathsf{C}_i = \{N_i, K_i\}$;
  $\mathsf{m}_i = 1/\mathsf{V}_p$;
end for
-Check termination conditions
while $\mathsf{G} \leq \overline{\mathsf{G}}$ **do**
  -Survivor selection
  Perform the roulette wheel selection process with inputs $\{\mathsf{C}_i\}$ and $\{\mathsf{m}_i\}$ (for details refer to [46]) and obtain new V_p parental chromosomes;
  -Crossover
  if V_p is even **then**
    Perform crossover mode 1) with probability P_c;
  else
    Perform crossover mode 2) with probability P_c;
  end if
  -Mutation
  Generating one random variable R_M in the interval $(0, 1)$.
  for $i = 1 : \mathsf{V}_p$ **do**
    if $\mathsf{R}_M < \mathsf{P}_m$ **then**
      $N_i \leftarrow \hat{N} = \text{Random}\{2, 3, \ldots, \overline{N}\}\backslash\{N_i\}$;
      $K_i \leftarrow \hat{K} = \text{Random}\{1, 2, \ldots, N_i - 1\}\backslash\{K_i\}$;
    end if
  end for
  -Fitness evaluation
  for $i = 1 : \mathsf{V}_p$ **do**
    if constraints held **then**
      $\mathsf{M}_i \leftarrow \text{Objective}(N_i, K_i)$;
    else
      $\mathsf{M}_i \leftarrow 0$;
    end if
  end for
  if $\sum_{i=1}^{\mathsf{V}_p} \mathsf{M}_i == 0$ **then**
    for $i = 1 : \mathsf{V}_p$ **do**
      $\mathsf{M}_i \leftarrow 1/\mathsf{V}_p$;
    end for
  end if
  Obtain $\{\mathsf{m}_i\}$ for all chromosomes by (22);
  $\mathsf{G} \leftarrow \mathsf{G} + 1$;
end while
$i_{\text{opt}} = \arg\max_i\{\mathsf{m}_i\}$;
return $\mathsf{C}_{i_{\text{opt}}} = \{N_{i_{\text{opt}}}, K_{i_{\text{opt}}}\}$;
END

TABLE III: Rearranging process from transmitted OFDM block to permuted OFDM block when $N = 4$ and $K = 2$.

$\mathbf{x}(c, \mathbf{b})$	ASV	$\mathbf{z}(\mathbf{b})$
$[0, \chi_{m_2}, \chi_{m_3}, 0]$	$\mathbf{v}_1 = [0, 0, 1, 1]$	$[0, 0, \chi_{m_2}, \chi_{m_3}]$
$[0, \chi_{m_2}, 0, \chi_{m_4}]$	$\mathbf{v}_2 = [0, 1, 0, 1]$	$[0, \chi_{m_2}, 0, \chi_{m_4}]$
$[0, 0, \chi_{m_3}, \chi_{m_4}]$	$\mathbf{v}_3 = [0, 1, 1, 0]$	$[0, \chi_{m_3}, \chi_{m_4}, 0]$
$[\chi_{m_1}, \chi_{m_2}, 0, 0]$	$\mathbf{v}_4 = [1, 0, 0, 1]$	$[\chi_{m_1}, 0, 0, \chi_{m_2}]$

$\dot{\mathbf{x}}(c, \dot{\mathbf{b}})$ is erroneously estimated to be $\hat{\mathbf{x}}(c, \hat{\mathbf{b}})$, which can be expressed as [27]

$$\begin{aligned} &P_e(\dot{\mathbf{x}}(c, \dot{\mathbf{b}}) \to \hat{\mathbf{x}}(c, \hat{\mathbf{b}})|\mathbf{H}) \\ &= Q\left(\sqrt{\frac{P_t}{KN_0}\left\|\mathbf{H}\left(\dot{\mathbf{x}}(c, \dot{\mathbf{b}}) - \hat{\mathbf{x}}(c, \hat{\mathbf{b}})\right)\right\|_F^2}\right) \\ &= Q\left(\sqrt{\frac{P_t}{KN_0}\sum_{n=1}^{N} G(n)|\dot{x}(\dot{m}_n, n) - \hat{x}(\hat{m}_n, n)|^2}\right), \end{aligned} \tag{25}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty \exp\left(-u^2/2\right) \mathrm{d}u$ is the Gaussian tail function (a.k.a. the Q-function). Because of the difficulty of processing the Q-function, to provide an insightful expression of the final result, we adopt the approximation of the Q-function by the sum of two linear transformations of the exponential function [15]:

$$Q(x) \approx \frac{1}{12}\exp\left(-\frac{x^2}{2}\right) + \frac{1}{4}\exp\left(-\frac{2x^2}{3}\right), \tag{26}$$

which becomes increasingly accurate for a large x. By such an approximation, we can approximate $P_e(\dot{\mathbf{x}}(c, \dot{\mathbf{b}}) \to \hat{\mathbf{x}}(c, \hat{\mathbf{b}})|\mathbf{H})$ as follows.

$$\begin{aligned} &P_e(\dot{\mathbf{x}}(c, \dot{\mathbf{b}}) \to \hat{\mathbf{x}}(c, \hat{\mathbf{b}})|\mathbf{H}) \\ &\approx \sum_{i=1}^{2} \rho_i \prod_{n=1}^{N} \exp\left(-\frac{\eta_i P_t}{KN_0} G(n)|\dot{x}(\dot{m}_n, n) - \hat{x}(\hat{m}_n, n)|^2\right), \end{aligned} \tag{27}$$

where $\{\rho_1, \rho_2\} = \left\{\frac{1}{12}, \frac{1}{4}\right\}$ and $\{\eta_1, \eta_2\} = \left\{\frac{1}{2}, \frac{2}{3}\right\}$.

To ease the analysis with codebook design, we need to rearrange $\mathbf{x}(c, \mathbf{b})$ by the *orders* of subcarriers instead of the *indices*. Then, we can obtain the *permuted* OFDM block[13]

$$\mathbf{z}(\mathbf{b}) = [z(m_1, 1), z(m_2, 2), \ldots, z(m_N, N)]^T \in \mathbb{C}^{N\times 1}, \tag{28}$$

according to the incoming bit stream $\mathbf{b}$, which will not change with the channel states. Specifically, $\mathbf{z}(\mathbf{b})$ is formed by replacing the '1's of an ASV with the data symbols in the corresponding $\mathbf{x}(c, \mathbf{b})$ in sequence. We illustrate an example of this rearranging process in Table III, which follows the case shown in Table II. With the help of the concept of the permuted OFDM block, we can modify (27) to be (29) at the top of the next page.

We can remove the conditions on the channel state $\mathbf{H}$ in (29) and obtain $P_e(\dot{\mathbf{x}}(c, \dot{\mathbf{b}}) \to \hat{\mathbf{x}}(c, \hat{\mathbf{b}}))$ by the calculation given in (30) at the top of the next page, where $\Gamma(x) = \int_0^\infty u^{x-1}\exp(-u)\mathrm{d}u$ denotes the gamma function. After this, we can resort to the interchangeability between integral and summation operations to approximate the unconditional BLER by its union bound as follows:

$$P_e(\dot{\mathbf{x}}(c, \dot{\mathbf{b}})) \le \sum_{\hat{\mathbf{x}}(c,\hat{\mathbf{b}})\neq\dot{\mathbf{x}}(c,\dot{\mathbf{b}})} P_e(\dot{\mathbf{x}}(c, \dot{\mathbf{b}}) \to \hat{\mathbf{x}}(c, \hat{\mathbf{b}})). \tag{31}$$

Finally, as all bits are equiprobable, all subcarrier activation patterns will be used with a uniform probability. As a consequence of this uniformity, we can average $P_e(\dot{\mathbf{x}}(c, \dot{\mathbf{b}}))$ over all legitimate transmitted OFDM blocks and determine the average BLER to be

$$\bar{P}_e = \frac{1}{X}\sum_{\dot{\mathbf{x}}(c,\dot{\mathbf{b}})\in\mathcal{X}(c)} P_e(\dot{\mathbf{x}}(c, \dot{\mathbf{b}})), \tag{32}$$

which is the most simplified and general form that we can achieve for the proposed OFDM-IM system applying lexicographic codebook design, since the different components of the summation operation depend on the distribution of the ordered subchannels.

Furthermore, according to the above derivations, we can also derive the average BER, which is a fundamental metric to measure the error performance by considering different coding schemes. The average BER can be determined in (33) at the top of the next page, where $\delta(\dot{\mathbf{x}}(c, \dot{\mathbf{b}}) \to \hat{\mathbf{x}}(c, \hat{\mathbf{b}}))$ represents the number of bit errors for the pairwise error event, and is determined by the adopted coding scheme mapping B_M modulation bits to K data constellation points.

VI. Numerical Results

In this section, we first check the effectiveness of the solutions to the optimization problems provided by the GA compared to the optimal results provided by the brute-force method. Then, we verify the analysis of average BLER and BER given in Section V by numerical results provided by Monte Carlo simulations. Also, we provide a series of comparisons among the proposed OFDM-IM scheme assisted by the codebook design and relevant benchmarks, which illustrate the performance superiority of the proposed OFDM-IM scheme.

A. Verification of the Solutions to Optimization Problems by the GA

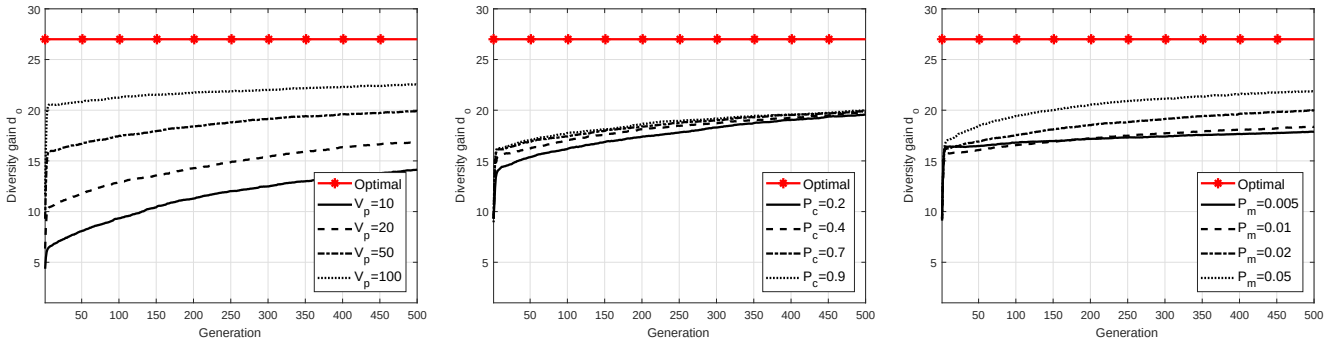
To verify the solutions to the formulated optimization problems by the GA, we need to check the results for all three optimization problems formulated in (17), (18) and (20) for different application scenarios. We set up the simulations by the parameters given as follows: $\overline{\mathsf{G}} = 500$, $\overline{N} = 100$, $M = 2$, i.e. the binary PSK (BPSK) is adopted as the APM scheme, $\underline{B} = 10$ and $\underline{d}_o = 2$ as well as $w_1 = w_2 = 0.5$. In addition, to reflect the average performance of the proposed optimization methodology based on the GA and obtain smoother curves, we average the optimization process over a thousand trials. We also vary the three key simulation parameters for the GA, i.e. the size of population V_p, crossover probability P_c and mutation probability P_m, to reflect the effects of these parameters. The results regarding the diversity gain, data rate and utility with the number of generations are plotted in Fig.

[13]Note that, the concept of permuted OFDM block introduced here is just used to facilitate the performance analysis and will do nothing to the actual transmission at OFDM-IM transmitters.

$$P_e(\dot{\mathbf{x}}(c,\dot{\mathbf{b}}) \to \hat{\mathbf{x}}(c,\hat{\mathbf{b}})|\mathbf{H}) \approx \sum_{i=1}^{2} \rho_i \prod_{\xi=1}^{N} \exp\left(-\frac{\eta_i P_t}{KN_0} G_{\langle\xi\rangle}(n_\xi)|\dot{z}(\dot{m}_{n_\xi}, n_\xi) - \hat{z}(\hat{m}_{n_\xi}, n_\xi)|^2\right). \tag{29}$$

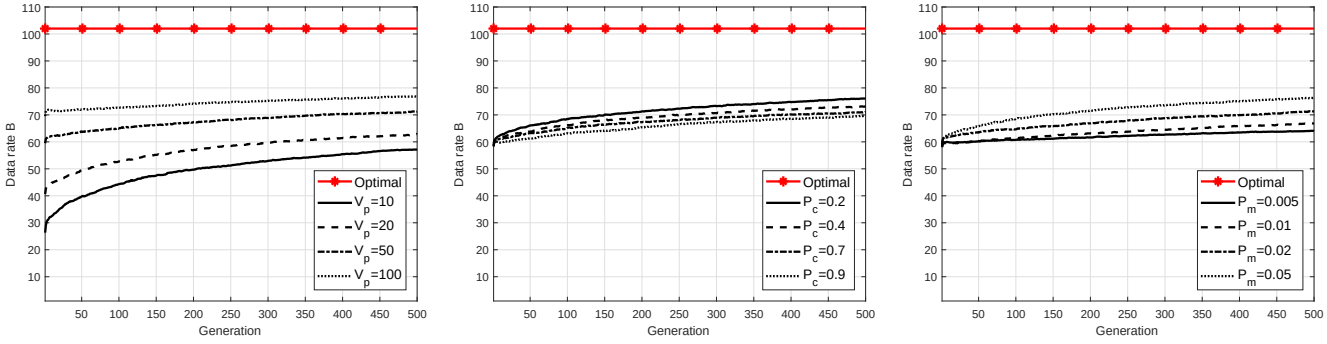
$$\begin{aligned} P_e(\dot{\mathbf{x}}(c,\dot{\mathbf{b}}) \to \hat{\mathbf{x}}(c,\hat{\mathbf{b}})) &= \sum_{i=1}^{2} \rho_i \prod_{\xi=1}^{N} \int_0^{\infty} \exp\left(-\frac{\eta_i P_t}{KN_0} G_{\langle\xi\rangle}(n_\xi)|\dot{z}(\dot{m}_{n_\xi}, n_\xi) - \hat{z}(\hat{m}_{n_\xi}, n_\xi)|^2\right) \phi_{\langle\xi\rangle}(G_{\langle\xi\rangle}(n_\xi)) \mathrm{d}G_{\langle\xi\rangle}(n_\xi) \\ &= \sum_{i=1}^{2} \rho_i \prod_{\xi=1}^{N} \frac{N!\Gamma\left(N-\xi+1+\frac{\eta_i P_t \mu |\dot{z}(\dot{m}_{n_\xi}, n_\xi) - \hat{z}(\hat{m}_{n_\xi}, n_\xi)|^2}{KN_0}\right)}{(N-\xi)!\Gamma\left(N+1+\frac{\eta_i P_t \mu |\dot{z}(\dot{m}_{n_\xi}, n_\xi) - \hat{z}(\hat{m}_{n_\xi}, n_\xi)|^2}{KN_0}\right)} \end{aligned} \tag{30}$$

$$\bar{P}_b = \frac{1}{XB} \sum_{\dot{\mathbf{x}}(c,\dot{\mathbf{b}}) \in \mathcal{X}(c)} \left(\sum_{\hat{\mathbf{x}}(c,\hat{\mathbf{b}}) \neq \dot{\mathbf{x}}(c,\dot{\mathbf{b}})} P_e(\dot{\mathbf{x}}(c,\dot{\mathbf{b}}) \to \hat{\mathbf{x}}(c,\hat{\mathbf{b}})) \delta(\dot{\mathbf{x}}(c,\dot{\mathbf{b}}) \to \hat{\mathbf{x}}(c,\hat{\mathbf{b}})) \right) \tag{33}$$



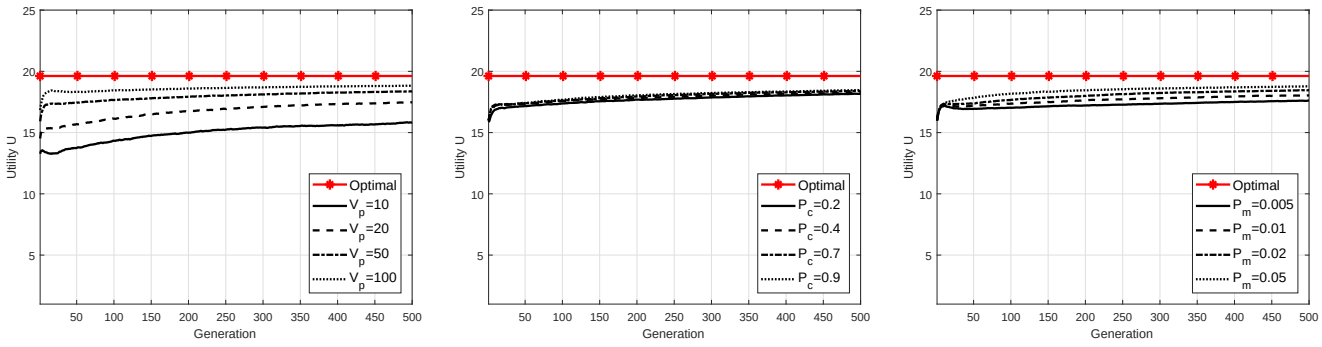
(a) (b) (c)

Fig. 6: Average diversity gain after performing the optimization based on the GA vs. the number of generations: (a) with different $\mathsf{V}_p \in \{10, 20, 50, 100\}$, given $\mathsf{P}_c = 0.7$ and $\mathsf{P}_m = 0.02$; (b) with different $\mathsf{P}_c \in \{0.2, 0.4, 0.7, 0.9\}$, given $\mathsf{V}_p = 50$ and $\mathsf{P}_m = 0.02$; (c) with different $\mathsf{P}_m \in \{0.005, 0.01, 0.02, 0.05\}$, given $\mathsf{V}_p = 50$ and $\mathsf{P}_c = 0.7$.



(a) (b) (c)

Fig. 7: Average data rate after performing the optimization based on the GA vs. the number of generations: (a) with different $\mathsf{V}_p \in \{10, 20, 50, 100\}$, given $\mathsf{P}_c = 0.7$ and $\mathsf{P}_m = 0.02$; (b) with different $\mathsf{P}_c \in \{0.2, 0.4, 0.7, 0.9\}$, given $\mathsf{V}_p = 50$ and $\mathsf{P}_m = 0.02$; (c) with different $\mathsf{P}_m \in \{0.005, 0.01, 0.02, 0.05\}$, given $\mathsf{V}_p = 50$ and $\mathsf{P}_c = 0.7$.



(a) (b) (c)

Fig. 8: Average utility after performing the optimization based on the GA vs. the number of generations: (a) with different $\mathsf{V}_p \in \{10, 20, 50, 100\}$, given $\mathsf{P}_c = 0.7$ and $\mathsf{P}_m = 0.02$; (b) with different $\mathsf{P}_c \in \{0.2, 0.4, 0.7, 0.9\}$, given $\mathsf{V}_p = 50$ and $\mathsf{P}_m = 0.02$; (c) with different $\mathsf{P}_m \in \{0.005, 0.01, 0.02, 0.05\}$, given $\mathsf{V}_p = 50$ and $\mathsf{P}_c = 0.7$.

6, Fig. 7 and Fig. 8, respectively. Also, we take the optimal results produced by the brute-force method as comparison benchmarks.

From these three figures, we can observe that with any parameter set, the optimized results provided by the GA gradually approach the optimal solutions to all three formulated optimization problems with an increasing number of generations, which verifies the effectiveness of the optimization algorithm based on the GA proposed in this paper. In addition, the convergence rate is affected by the size of population V_p, crossover probability P_c and mutation probability P_m. First, a larger V_p will undoubtedly lead to a higher convergence rate and is thus able to find out the optimal solution within fewer generations. However, one should always keep in mind that a larger V_p will on the other hand yield a higher complexity. Normally, a practical problem would require a population of several thousands of chromosomes [48]. Therefore, a performance-complexity trade-off emerges and should be taken into consideration when configuring V_p. One should note that although it is ensured that the results produced by the GA will approach the optimal with a sufficiently long operation time, we will never know whether a specific result is optimal or not, and even have no idea how far this result is from the optimal without comparison benchmarks generated by other optimization techniques [48]. This is an intrinsic problem always arising with the GA-based optimization processes. Second, for such a simple chromosome structure only containing two genes (i.e. $\mathsf{C} = \{N, K\}$), if there exits an intrinsic restrictive relation between these two genes (i.e. $K \leq N - 1$), the conventional crossover processing might not be suitable anymore. Therefore, a higher P_c could result in a poorer convergence rate. In addition, for optimization problems without constraints, i.e. (19), P_c will have little impact on the convergence rate. Third, although mutation could efficiently explore the search space, one should always keep in mind that there is a trade-off between the *exploration* and *exploitation* [40]. Hence, the generated performance curves become more variable with a higher P_m, especially at the beginning.

Meanwhile, it is noteworthy that the results provided by the GA have been averaged over a thousand trials. As a result of this averaging process, optimized results are seemingly far from the optimal by 500 generations. However, for an individual trial, normally two extreme cases are expected. The optimizer either finds out the optimal solution quickly within several thousands of generations, or gets trapped to a poor solution (we can illustrate this phenomenon by the probability mass function (PMF)[14] of the required generation to achieve the optimal solution for a variety of cases in Fig. 9). The difference is caused by different initial populations. In other words, if the initial population contains several chromosomes close to the optimum, the optimum will be found quickly. Otherwise, it will take a large number of generations to reach the optimal solution by crossover and mutation. This means that the approach to generate the initial population is crucial for finding out the optimal solution in an efficient manner, and therefore more effort should be devoted to the design of the initial population with domain knowledge. However, as an intuitive optimization approach proposed in this paper, we have not involved too many details for how to improve the GA itself and more research regarding these points are necessary before implementing the GA-based optimization in practice.

B. Verification of Average BLER Analysis

Without loss of generality, we normalize $\mu = 1$, $N_0 = 1$ as well as the bandwidth of each subcarrier and simulate the average BLER as a function of the ratio of transmit power to noise power P_t/N_0 for the proposed OFDM-IM system with BPSK ($M = 2$) and quadrature PSK (QPSK) ($M = 4$) in Fig. 10. The numerical results presented in this figure as well as other simulation figures in the sequel are averaged over more than 10^6 trials. Channel gains regarding all N subcarriers are exponentially i.i.d. with mean $\mu = 1$, which are produced by setting a sufficiently long length of channel impulse response. Here, we select S subcarrier activation patterns out of L total patterns by the lexicographic codebook design proposed in Section III. We also adopt different N and K to observe

[14]Here, we apply the kernel density estimation (KDE) (a.k.a. the Parzen-Rosenblatt window method) to estimate the PMF [49].

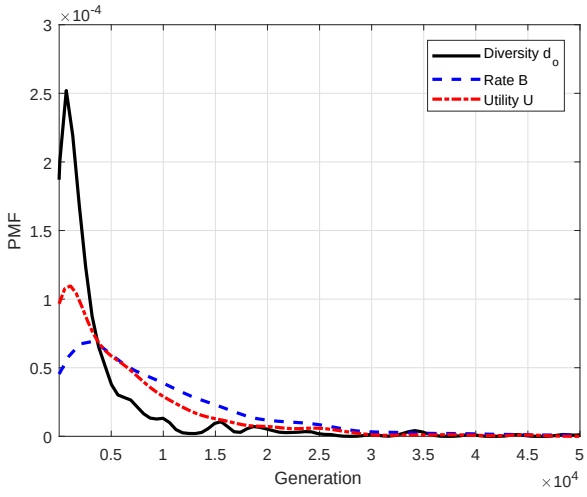


Fig. 9: Estimated PMF for the required number of generations to achieve the optimal for all three formulated optimization problems, given $\mathsf{V}_p = 50$, $\mathsf{P}_c = 0.7$ and $\mathsf{P}_m = 0.02$.

their effects on the error performance and verify the diversity mechanism detailed in Section IV-A. From this figure, first of all, our analysis presented in Section V has been verified, as analytical curves approach the numerical curves when P_t/N_0 becomes large. The difference between the analytical and numerical results is mainly caused by the approximation given in (27). In addition, by comparing the results presented in Fig. 10a and Fig. 10b, it is obvious that a larger M will present poorer error performance, because it is more challenging to distinguish between two adjacent blocks. Furthermore, from this figure, it can be observed that the increase in N will lead to an uncertain effect (could be either constructive or destructive) depending on the value of K. On one hand, when increasing N, the number of subcarrier activation patterns $L = \binom{N}{K}$ increases and more suitable patterns could be found accordingly, which will yield a constructive effect on the error performance. On the other hand, a larger N will enlarge the search space for estimation at the receiver, which makes the system more error-prone. A similar dilemma can also be deduced for the change of K. Therefore, it would be unwise to assert whether it is good or bad to increase or decrease N and K. Instead, both should be considered as a whole and chosen holistically as a combination to attain a proper performance level. Most importantly, the diversity mechanism detailed in Section IV-A can also be substantiated by the case $\{N, K\} = \{6, 2\}$ and a diversity gain of two is achieved, which aligns with the expectation.

C. Performance Comparisons

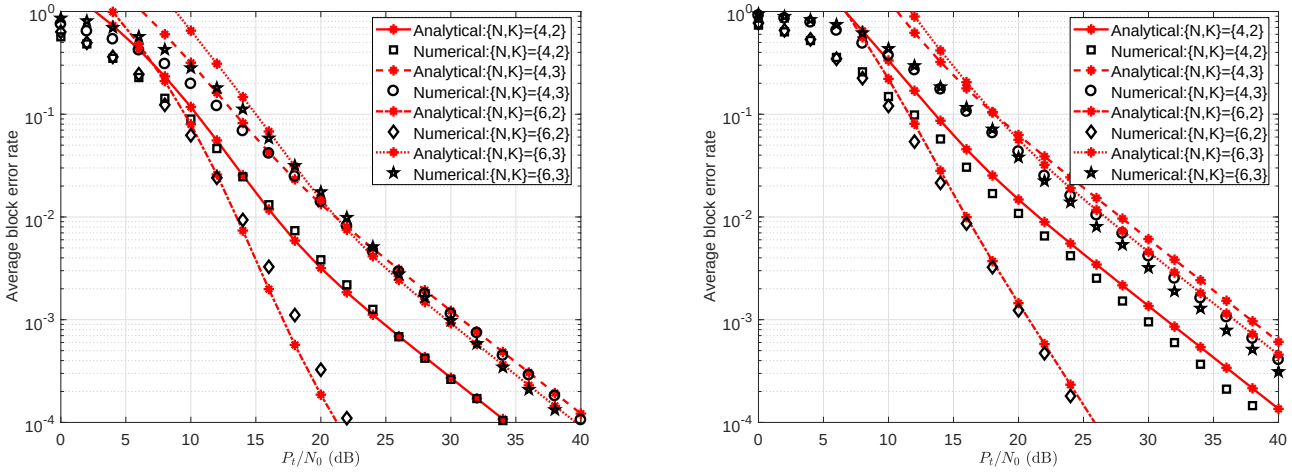
To verify the error performance superiority of the proposed OFDM-IM scheme with codebook design, we compare the average BLER of different OFDM schemes in this subsection by numerical simulations with the same simulation settings as in Section VI-B. We adopt the conventional OFDM-IM scheme without codebook design[15] and the traditional OFDM scheme without IM as benchmarks and the numerical results are presented in Fig. 11. From this figure, first, we can observe that with the same number of subcarriers and APM order, both proposed and conventional OFDM-IM outperform the traditional OFDM scheme without IM in terms of average BLER. For the case $\{N, K\} = \{6, 3\}$, a coding gain can be observed for the proposed OFDM-IM by the codebook design compared to the conventional OFDM-IM, while an obvious diversity gain is obtainable for the case $\{N, K\} = \{6, 2\}$. Meanwhile, as the same lexicographic coding is used for all proposed cases, it is expected that such a coding gain also holds for the case $\{N, K\} = \{6, 2\}$ when diversity gain is produced. All these performance gains come from the acquisition and utilization of instantaneous CSI. These observations confirm the superiority of the proposed OFDM-IM over the two benchmarks in terms of error performance. Furthermore, one should note that the transmission rate of the proposed OFDM-IM scheme with codebook design is exactly the same as that of the conventional OFDM-IM scheme proposed in [15], which means that the better error performance of our proposed scheme is not attained at the expense of transmission rate.

In this subsection, we also take different subcarrier activation methods into consideration. In existing literature, there are mainly three subcarrier activation methods. Except for the OFDM-IM adopting fixed-number subcarrier activation method in this paper, we also employ the OFDM-IM schemes with the OOK subcarrier activation method and the dual-mode subcarrier activation method as comparison benchmarks [17], [23]. All other simulation configurations are the same as specified in Section VI-B. We present the comparisons of transmission rate and average BLER among the aforementioned schemes in Fig. 12 and Fig. 13, respectively. From Fig. 12, it is shown that when the APM order M is relatively large, our proposed scheme can easily achieve the highest transmission rate by adjusting the number of active subcarriers K. In the meantime, as shown in Fig. 13, the adjustment of the number of active subcarriers K will also have an impact on the superiority of average BLER. To maintain the the superiority of average BLER at high SNR by harvesting an extra diversity gain, we have to abide by the rule summarized in (16).

D. Verification of Average BER with Different Coding Schemes

In this subsection, we mainly verify the effectiveness of (33) and compare the average BER when different coding schemes are in use. We adopt the same simulation configurations as for the verification of BLER specified in Section VI-B. Also, to facilitate the simulation in the sequel, we adopt two fundamental coding schemes for the proposed OFDM-IM system with QPSK, which are Gray code and binary code. The simulation results are presented in Fig. 14 for both coding schemes. By the results presented in this figure, we can verify the

[15]In the conventional OFDM-IM scheme, the mapping relation between incoming bit streams and subcarrier activation patterns are given in an arbitrary manner without considering the CSI. We follow this configuration for the benchmark and more relevant details can be found in [15].



(a) (b)

Fig. 10: Average BLER vs. the ratio of transmit power to noise power P_t/N_0 with different N and K: (a) BPSK; (b) QPSK.

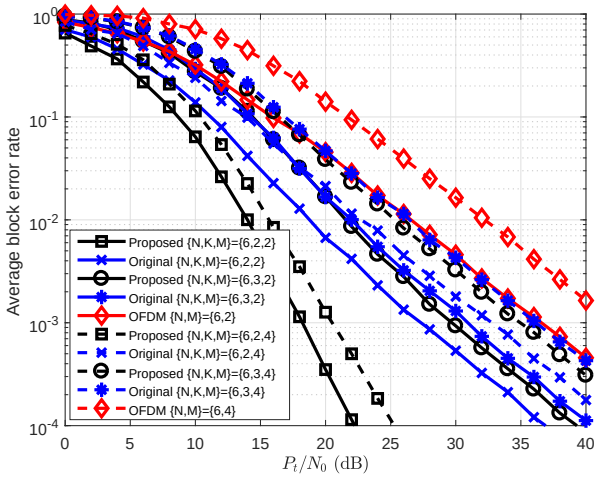


Fig. 11: Comparison of average BLER among the proposed OFDM-IM scheme enhanced by codebook design, conventional OFDM-IM scheme and traditional OFDM scheme without IM.

effectiveness of (33), as both numerical and analytical results converge at high SNR. The effects of the parameters N and K on the average BER share similar trends as on the average BLER, which are aligned with our expectation. Meanwhile, Gray code outperforms binary code, which is caused by its unique code design that only allows one bit (binary digit) to be different between two successive constellation points. As a result of this unique design, once a block error takes place, a smaller number of bit error events will concomitantly occur, which leads to a lower average BER under the same simulation setting. It can also be easily anticipated that such an error performance superiority brought by Gray code becomes significant with an increasing APM order M.

VII. Conclusion

In this paper, to enhance system performance, we proposed an OFDM-IM scheme with codebook optimization. The design process can be easily implemented by the lexicographic ordering principle. Also, it has been noticed that a diversity gain can be attained by such an optimization process. Subsequently, we investigated the diversity mechanism and formulated three diversity-rate optimization problems for the proposed OFDM-IM system in terms of the numbers of total subcarriers and active subcarriers N and K. Also, we provided details of a GA-based approach to carry out the optimization. With optimal N and K, we analyzed the average BLER and BER of OFDM-IM systems applying the codebook design and ML detection. Finally, all analyses were numerically verified by Monte Carlo simulations. In addition, we also provided a series of comparisons of the average BLER among the proposed OFDM-IM scheme with codebook design and a variety of benchmarks, which verifies the superiority of the proposed OFDM-IM in terms of error performance. Tailoring the conventional GA to fit the optimization scenarios of OFDM-IM systems would be worth investigating as future work, including the methods to initialize the first generation chromosomes and perform crossover as well as set up appropriate crossover and mutation probabilities. In addition, an in-depth investigation into coding gain yielded by this codebook optimization would be a worthwhile future research direction.

Acknowledgment

The authors would like to thank the editor and the anonymous reviewers for their constructive comments that help improve the quality of this manuscript.

References

[1] G. Femenias, F. Riera-Palou, X. Mestre, and J. J. Olmos, "Downlink scheduling and resource allocation for 5G MIMO-multicarrier: OFDM vs FBMC/OQAM," *IEEE Access*, vol. 5, pp. 13 770–13 786, July 2017.

[2] K. Liu, A. Sadek, W. Su, and A.Kwasinski, *Cooperative Communications and Networking*. Cambridge University Press, 2009.

[3] S. Dang, J. P. Coon, and G. Chen, "An equivalence principle for OFDM-based combined bulk/per-subcarrier relay selection over equally spatially correlated channels," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 1, pp. 122–133, Jan. 2017.

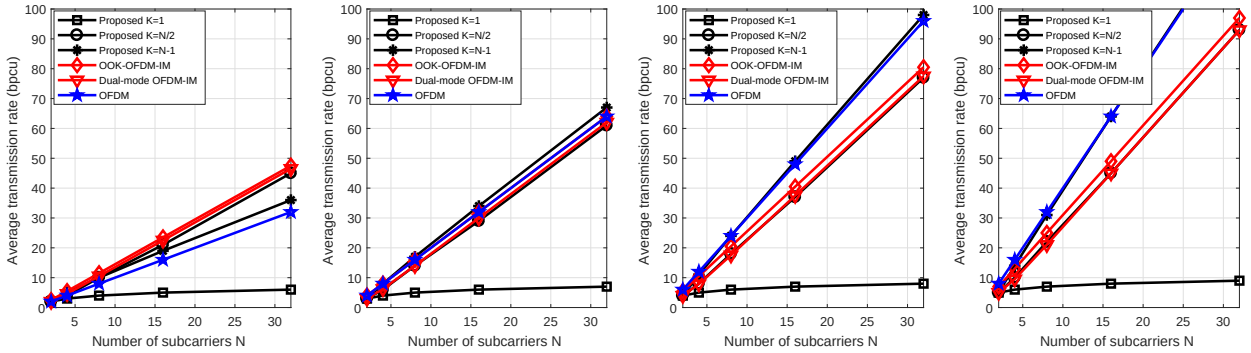


Fig. 12: Transmission rate comparisons among OFDM-IM schemes with different subcarrier activation methods with $M = 2, 4, 8, 16$ (from left to right).

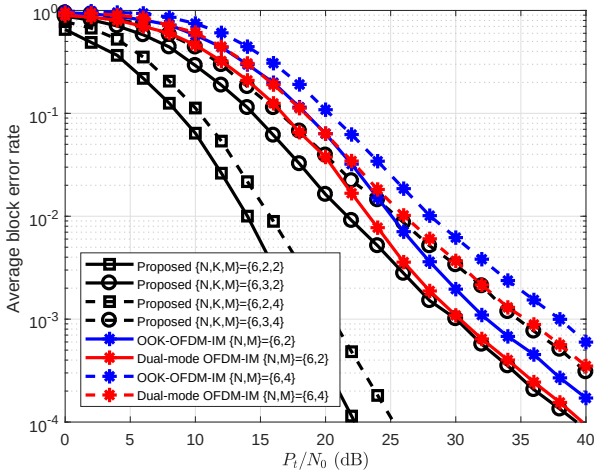


Fig. 13: Comparison of average BLER among different OFDM-IM schemes with different subcarrier activation methods.

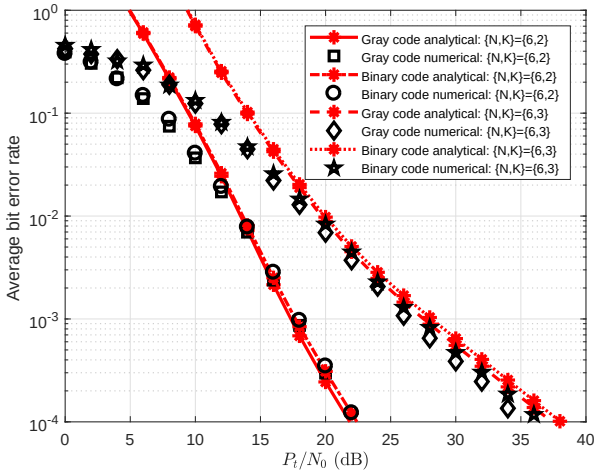


Fig. 14: Average BER vs. the ratio of transmit power to noise power P_t/N_0 for Gray coded and binary coded QPSK OFDM-IM systems applying the lexicographic ordering principle.

[4] M. Pejanovic-Djurisic, E. Kocan, and R. Prasad, *OFDM based relay systems for future wireless communications*. River Publishers, 2012.

[5] E. Basar, "Index modulation techniques for 5G wireless networks," *IEEE Communications Magazine*, vol. 54, no. 7, pp. 168–175, Jul. 2016.

[6] E. Basar, M. Wen, R. Mesleh, M. D. Renzo, Y. Xiao, and H. Haas, "Index modulation techniques for next-generation wireless networks," *IEEE Access*, vol. 5, pp. 16 693–16 746, 2017.

[7] 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. 21 774–21 790, 2017.

[8] N. Ishikawa, S. Sugiura, and L. Hanzo, "50 years of permutation, spatial and index modulation: from classic RF to visible light communications and data storage," *IEEE Communications Surveys Tutorials*, pp. 1–1, 2018.

[9] R. Fan, Y. J. Yu, and Y. L. Guan, "Generalization of orthogonal frequency division multiplexing with index modulation," *IEEE Transactions on Wireless Communications*, vol. 14, no. 10, pp. 5350–5359, Oct. 2015.

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

[11] M. Wen, E. Basar, Q. Li, B. Zheng, and M. Zhang, "Multiple-mode orthogonal frequency division multiplexing with index modulation," *IEEE Transactions on Communications*, vol. 65, no. 9, pp. 3892–3906, Sept. 2017.

[12] X. Cheng, M. Zhang, M. Wen, and L. Yang, "Index modulation for 5G: striving to do more with less," *IEEE Wireless Communications*, vol. 25, no. 2, pp. 126–132, Apr. 2018.

[13] R. Abu-alhiga and H. Haas, "Subcarrier-index modulation OFDM," in *Proc. IEEE PIMRC*, Tokyo, Japan, Sept. 2009.

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

[15] E. Basar, U. Aygolu, E. Panayırcı, and H. V. Poor, "Orthogonal frequency division multiplexing with index modulation," *IEEE Transactions on Signal Processing*, vol. 61, no. 22, pp. 5536–5549, Nov. 2013.

[16] M. Wen, Y. Zhang, J. Li, E. Basar, and F. Chen, "Equiprobable subcarrier activation method for OFDM with index modulation," *IEEE Communications Letters*, vol. 20, no. 12, pp. 2386–2389, Dec. 2016.

[17] S. Dang, J. P. Coon, and G. Chen, "Adaptive OFDM with index modulation for two-hop relay-assisted networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 1923–1936, Mar. 2018.

[18] S. Dang, G. Chen, and J. P. Coon, "Power allocation for adaptive OFDM index modulation in cooperative networks," in *Proc. IEEE GLOBECOM*, Singapore, Dec. 2017.

[19] J. Choi, "Coded OFDM-IM with transmit diversity," *IEEE Transactions on Communications*, vol. 65, no. 7, pp. 3164–3171, Jul. 2017.

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

[21] M. Wen, B. Ye, E. Basar, Q. Li, and F. Ji, "Enhanced orthogonal frequency division multiplexing with index modulation," *IEEE Transactions on Wireless Communications*, vol. 16, no. 7, pp. 4786–4801, July 2017.

[22] Q. Li, M. Wen, E. Basar, H. V. Poor, B. Zheng, and F. Chen, "Diversity enhancing multiple-mode OFDM with index modulation," *IEEE Transactions on Communications*, pp. 1–1, 2018.

[23] S. Dang, G. Chen, and J. P. Coon, "Outage performance of two-hop OFDM with index modulation and multi-carrier relay selections," *IEEE Wireless Communications Letters*, 2018.

[24] M. Wen, X. Cheng, M. Ma, B. Jiao, and H. V. Poor, "On the achievable rate of OFDM with index modulation," *IEEE Transactions on Signal Processing*, vol. 64, no. 8, pp. 1919–1932, Apr. 2016.

[25] Z. Wang, S. Dang, and D. T. Kennedy, "Multi-hop index modulation-aided OFDM with decode-and-forward relaying," *IEEE Access*, vol. 6, pp. 26 457–26 468, 2018.

[26] S. Dang, J. P. Coon, and G. Chen, "Outage performance of two-hop OFDM systems with spatially random decode-and-forward relays," *IEEE Access*, vol. 5, pp. 27 514–27 524, Nov. 2017.

[27] J. Proakis and M. Salehi, *Digital Communications*, ser. McGraw-Hill International Edition. McGraw-Hill, 2008.

[28] M. Agoston, *Computer Graphics and Geometric Modelling: Mathematics*, ser. Computer Graphics and Geometric Modeling. Springer London, 2005.

[29] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, ser. Wiley series in telecommunications. Cambridge University Press, 2005.

[30] A. Itai, "Generating permutations and combinations in lexicographical order," *Journal of the Brazilian Computer Society*, vol. 7, no. 3, pp. 65–68, 2001.

[31] E. Harzheim, *Ordered Sets*, ser. Advances in Mathematics. Springer US, 2006.

[32] Z. Wang and G. B. Giannakis, "A simple and general parameterization quantifying performance in fading channels," *IEEE Transactions on Communications*, vol. 51, no. 8, pp. 1389–1398, Aug. 2003.

[33] J. P. Coon and M. Sandell, "Combined bulk and per-tone transmit antenna selection in OFDM systems," *IEEE Communications Letters*, vol. 14, no. 5, pp. 426–428, May 2010.

[34] E. Kocan, M. Pejanovic-Djurisic, D. S. Michalopoulos, and G. K. Karagiannidis, "Performance evaluation of OFDM amplify-and-forward relay system with subcarrier permutation," *IEICE Transactions on Communications*, vol. 93, no. 5, pp. 1216–1223, May 2010.

[35] Q. Ma, Y. Xiao, L. Dan, P. Yang, L. Peng, and S. Li, "Subcarrier allocation for OFDM with index modulation," *IEEE Communications Letters*, vol. 20, no. 7, pp. 1469–1472, July 2016.

[36] J. M. Hamamreh, E. Basar, and H. Arslan, "OFDM-subcarrier index selection for enhancing security and reliability of 5G URLLC services," *IEEE Access*, vol. 5, pp. 25 863–25 875, 2017.

[37] S. H. Lee, S. Lee, H. Song, and H. S. Lee, "Wireless sensor network design for tactical military applications: Remote large-scale environments," in *IEEE MILCOM*, Boston, MA, USA, Oct. 2009.

[38] Y. Toor, P. Muhlethaler, A. Laouiti, and A. D. L. Fortelle, "Vehicle ad hoc networks: applications and related technical issues," *IEEE Communications Surveys Tutorials*, vol. 10, no. 3, pp. 74–88, Third 2008.

[39] L. Atzori, A. Iera, and G. Morabito, "The internet of things: a survey," *Computer networks*, vol. 54, no. 15, pp. 2787–2805, 2010.

[40] T. Rondeau and C. Bostian, *Artificial Intelligence in Wireless Communications*, ser. Artech House mobile communications series. Artech House, 2009.

[41] Z. Lv, T. Yin, Y. Han, Y. Chen, G. Chen *et al.*, "WebVR-web virtual reality engine based on P2P network," *Journal of Networks*, vol. 6, no. 7, pp. 990–998, 2011.

[42] S. Sun, L. Gong, B. Rong, and K. Lu, "An intelligent SDN framework for 5G heterogeneous networks," *IEEE Communications Magazine*, vol. 53, no. 11, pp. 142–147, Nov. 2015.

[43] R. Hemmecke, M. Köppe, J. Lee, and R. Weismantel, "Nonlinear integer programming," in *50 Years of Integer Programming 1958-2008*. Springer, 2010, pp. 561–618.

[44] K. Murota, "Discrete convex analysis," *Mathematical Programming*, vol. 83, no. 1-3, pp. 313–371, 1998.

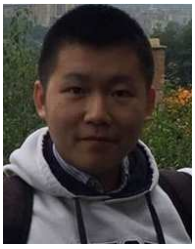
[45] J. McCall, "Genetic algorithms for modelling and optimisation," *Journal of Computational and Applied Mathematics*, vol. 184, no. 1, pp. 205–222, 2005.

[46] D. Goldberg, *Genetic Algorithms*. Pearson Education, 2006.

[47] H. David and H. Nagaraja, *Order Statistics*, ser. Wiley Series in Probability and Statistics. Wiley, 2004.

[48] M. Negnevitsky, *Artificial Intelligence: A Guide to Intelligent Systems*. Addison-Wesley, 2005.

[49] B. W. Silverman, *Density estimation for statistics and data analysis*. Routledge, 2018.



Shuping Dang (S'13, M'18) received B.Eng (Hons) in Electrical and Electronic Engineering from the University of Manchester (with first class honors) and B.Eng in Electrical Engineering and Automation from Beijing Jiaotong University in 2014 via a joint '2+2' dual-degree program. He also received D.Phil in Engineering Science from University of Oxford in 2018. Dr. Dang joined in the R&D Center, Huanan Communication Co., Ltd. after graduating from University of Oxford and is currently working as a Post-doctoral Fellow with the Computer, Electrical and Mathematical Science and Engineering Division, King Abdullah University of Science and Technology (KAUST). He serves as a reviewer for a number of key journals in communications and information science, including IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON COMMUNICATIONS and IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. His current research interests include artificial intelligence assisted communications, novel modulation schemes and cooperative communications.

Gaojie Chen (S'09-M'12) received the B.Eng. and B.Ec. degrees in electrical information engineering and international economics and trade from Northwest University, China, in 2006, and the M.Sc. (Hons.) and Ph.D. degrees in electrical and electronic engineering from Loughborough University, Loughborough, U.K., in 2008 and 2012, respectively. From 2008 to 2009, he was a Software Engineering with DTmobile, Beijing, China, and from 2012 to 2013, he was a Research Associate with the School of Electronic, Electrical and Systems Engineering, Loughborough University. He was a Research Fellow with 5GIC, Faculty of Engineering and Physical Sciences, University of Surrey, U.K., from 2014 to 2015. Then he was a Research Associate with the Department of Engineering Science, University of Oxford, U.K., from 2015 to 2018. He is currently a Lecturer with the Department of Engineering, University of Leicester, U.K. He has served as an Editor for IET ELECTRONICS LETTERS (2018-present). His current research interests include information theory, wireless communications, cooperative communications, cognitive radio, secrecy communication, and random geometric networks.

Justin P. Coon (S'02-M'05-SM'10) received a BSc. degree (with distinction) in electrical engineering from the Calhoun Honours College, Clemson University, USA and a Ph.D in communications from the University of Bristol, U.K. in 2000 and 2005, respectively. In 2004, he joined Toshiba Research Europe Ltd. (TREL) as a Research Engineer working in its Bristol based Telecommunications Research Laboratory (TRL), where he conducted research on a broad range of communication technologies and theories, including single and multi-carrier modulation techniques, estimation and detection, diversity methods, system performance analysis and networks. He held the position of Research Manager from 2010-2013, during which time he led all theoretical and applied research on the physical layer at TRL. Dr Coon was a Visiting Fellow with the School of Mathematics at the University of Bristol from 2010-2012, and held a position as Reader in the Department of Electrical and Electronic Engineering at the same university from 2012-2013. He joined the University of Oxford in 2013 where he is currently an Associate Professor with the Department of Engineering Science and a Tutorial Fellow of Oriel College.

Dr Coon is the recipient of TRL's Distinguished Research Award for his work on block-spread CDMA, aspects of which have been adopted as mandatory features in the 3GPP LTE Rel-8 standard. He is also a co-recipient of two 'best paper' awards for work presented at ISWCS 13 and EuCNC 14. Dr Coon has published in excess of 150 papers in leading international journals and conferences, and is a named inventor on more than 30 patents. He served as an Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (2007 - 2013), IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY (2013 - 2016), IEEE WIRELESS COMMUNICATIONS LETTERS (2016 - present) and IEEE COMMUNICATIONS LETTERS (2017 - present). Dr Coon's research interests include communication theory, information theory and network theory.