

# Enhanced Huffman Coded OFDM with Index Modulation

Shuping Dang, *Member, IEEE*, Shuaishuai Guo, *Member, IEEE*, Justin P. Coon, *Senior Member, IEEE*, Basem Shihada, *Senior Member, IEEE*, Mohamed-Slim Alouini, *Fellow, IEEE*

**Abstract**—In this paper, we propose an enhanced Huffman coded orthogonal frequency-division multiplexing with index modulation (EHC-OFDM-IM) scheme. The proposed scheme is capable of utilizing all legitimate subcarrier activation patterns (SAPs) and adapting the bijective mapping relation between SAPs and leaves on a given Huffman tree according to channel state information (CSI). As a result, a dynamic codebook update mechanism is obtained, which can provide more reliable transmissions. We take the average block error rate (BLER) as the performance evaluation metric and approximate it in closed form when the transmit power allocated to each subcarrier is independent of channel states. Also, we propose two CSI-based power allocation schemes with different requirements for computational complexity to further improve the error performance. Subsequently, we carry out numerical simulations to corroborate the error performance analysis and the proposed dynamic power allocation schemes. By studying the numerical results, we find that the depth of the Huffman tree has a significant impact on the error performance when the SAP-to-leaf mapping relation is optimized based on CSI. Meanwhile, through numerical results, we also discuss the trade-off between error performance and data transmission rate and investigate the impacts of imperfect CSI on the error performance of EHC-OFDM-IM.

**Index Terms**—Orthogonal frequency-division multiplexing with index modulation (OFDM-IM), Huffman coding, dynamic codebook design, power allocation, channel state information (CSI).

## I. INTRODUCTION

**T**O cope with the emerging requirements of high reliability and high transmission rate in next generation networks, novel modulation schemes have been regarded as promising solutions and have been exhaustively studied in recent years [1]–[4]. One of the most noticeable directions is index modulation (IM), which is related to permutation modulation and parallel combinatory modulation but developed in an independent manner since 2001 [3]. In recent decades, with the proliferation of multi-antenna and multi-carrier systems, IM has found its suitable application scenarios, and has attracted

researchers' attention with proposals of spatial modulation (SM) and orthogonal frequency-division multiplexing with IM (OFDM-IM), which are applications of IM in the spatial and frequency domains [5], [6]. Because SM is mainly designed for multi-antenna systems, which demands relatively high system complexity and a large device size [7], OFDM-IM has become more popular by releasing these constraints since its proposal in 2013 [8]. Another advantage of OFDM-IM is that the mature standards, protocols, and existing communication infrastructure for OFDM in fourth generation (4G) wireless networks can be applied to OFDM-IM systems without a considerable amendment, because of the similarity between both paradigms in practice [9]. Rigorous studies have proved the performance superiority of OFDM-IM over plain OFDM without IM in terms of error performance and transmission efficiency [10], [11]. The solid theoretical basis and the easy-to-implement attribute pave the way for OFDM-IM to prosper in next generation networks, so as to meet the stringent requirements of high reliability and high transmission rate.

Apart from the aforementioned benefits of OFDM-IM, there is one question that has accompanied OFDM-IM since its proposal until now: what is an efficient way to design a codebook for OFDM-IM that maps the codewords to subcarrier activation patterns (SAPs) in a bijective manner? Even before the proposal of OFDM-IM, two cognate schemes called subcarrier-index modulation (SIM) OFDM<sup>1</sup> and enhanced SIM OFDM have paid special attention to this crucial question [12], [13]. In [12], a forward error control mechanism was introduced to realize a heuristic mapping procedure, which raises the system complexity unnecessarily, and in [13], the coding efficiency was significantly reduced, albeit circumventing the forward error control mechanism. The classic OFDM-IM proposed in [8] adopts a truncated subset of SAPs with a fixed number of active subcarriers to match fixed-length bit sequences, which still decreases a portion of the transmission rate. In [14], a combined mapping relation with variable numbers of active subcarriers was proposed, which increases the transmission rate at the cost of increased system complexity. To reduce the system complexity raised by the combined OFDM-IM, another mapping relation relying on a special always-on subcarrier was studied in [15]–[18], by which the number of active subcarriers also varies. Such a feature of variable numbers of active subcarriers would cause practical issues in currently existing infrastructure and

This work was supported by KAUST Office of Sponsored Research.

S. Dang, B. Shihada, and M.-S. Alouini are with Computer, Electrical and Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia (e-mail: {shuping.dang, basem.shihada, slim.alouini}@kaust.edu.sa).

S. Guo was with the Computer, Electrical and Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia. He is now with Shandong Provincial Key Laboratory of Wireless Communication Technologies and School of Control Science and Engineering, Shandong University, Jinan 250061, P.R. China (e-mail: shuaiguosdu@gmail.com).

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

<sup>1</sup>Because of the shared nature, SIM OFDM is also regarded as an interchangeable term with OFDM-IM in some papers.

is thereby undesirable.

To mitigate the dilemma between performance and complexity, Huffman coding is useful to construct a bijective mapping relation between SAPs consisting of a fixed number of active subcarriers to variable-length bit sequences [19]. Huffman coding was introduced to two spatial applications of IM in [20] and [21], which studied the effects of the architecture of the Huffman tree on the system performance. According to the feedback from the receiving side, the transmitter is capable of dynamically updating the codebook to maximize the transmission rate. When it comes to OFDM-IM, a mapping relation between variable-length bit sequences and SAPs was defined in [22], by which all legitimate SAPs can be employed in the codebook, and a low-complexity detection scheme for OFDM-IM systems is thereby enabled. Although Huffman coding has not been explicitly mentioned in this work, the proposed system applies the core principle of Huffman coding and selects the Huffman tree with the minimum depth to encode incoming bits. The optimization of Huffman coded OFDM-IM by adjusting the architecture of the Huffman tree was formulated in [23] and [24] with various optimization objectives. They both provide good thoughts on how the Huffman tree architecture could be adapted to provide a higher capacity and transmission rate when utilizing all legitimate SAPs. In [25], Huffman coding was applied to OFDM-IM with the help of channel state information (CSI). However, the Huffman coded OFDM-IM scheme proposed in [25] did not fully exploit all legitimate SAPs, and thereby only shows limited improvements. The latest research results with respect to the optimization of the achievable rate of binary-tree coded OFDM-IM (a generic version of Huffman coded OFDM-IM) were presented in [26]. In this work, the achievable rate is optimized by dynamically varying the SAP probability distribution and transmit power allocated to active subcarriers.

To attain a proper transmission rate by involving all legitimate SAPs in a codebook and dynamically update the codebook by CSI to enhance the error performance, we propose an enhanced Huffman coded OFDM-IM (EHC-OFDM-IM) scheme in this paper. In order to clarify the contributions of this paper and the novelty compared to prior works, we detail the difference between the proposed EHC-OFDM-IM scheme with other existing and related schemes as follows:

- Different from the classic Huffman coded OFDM-IM scheme proposed in [22] that does not rely on any channel feedback, we employ CSI at the transmitting side to provide an adaptive optimization mechanism.
- Different from the adaptive Huffman coded OFDM-IM schemes proposed in [23] and [24] that focus on the optimization of the architecture of the Huffman tree, we instead focus on the optimization of the bijective mapping relation between all legitimate SAPs and leaves on a *given* Huffman tree, i.e., the variable-length bit sequences. Therefore, EHC-OFDM-IM can be applied on top of the optimized systems constructed in [23] and [24] to further enhance their error performance.
- Different from the adaptive Huffman coded OFDM-IM scheme proposed in [25] that does not exploit all SAPs

but only involves a binary probability vector to pair up a portion of SAPs based on channel qualities, EHC-OFDM-IM includes all legitimate SAPs in the codebook and use them for index bit transmission.

- Different from the binary-tree coded OFDM-IM proposed in [26] that undertakes the task of optimizing the achievable rate subject to the availability of channel knowledge at the transmitter, we focus on the task of improving the error performance, which formulates a different optimization problem and employs a different optimization strategy.
- We borrow the SAP ordering principle given in [27] to our paper to construct the pairing relation between SAPs to the leaves on a given Huffman tree. Although both schemes employ the CSI at the transmitter to perform codebook optimization, the system model and optimization scenario in this paper are different from the OFDM-IM aided by lexicographic codebook design (LCD-OFDM-IM) proposed in [27]. In particular, the LCD-OFDM-IM scheme selects a set of SAPs by the priority and excludes the other SAPs in the codebook, which does not fully exploit the largest span of legitimate SAPs but could be able to harvest a diversity gain. The EHC-OFDM-IM scheme employs the priority of each SAP to pair it up to a leaf on the Huffman tree. In this way, the SAP with higher priority will be used more frequently, so as to provide better error performance with a coding gain. By sacrificing the diversity gain, the EHC-OFDM-IM scheme utilizes all legitimate SAPs and thereby has a higher rate.

In order to provide a comprehensive and qualitative comparison so as to emphasize the contributions of the proposed EHC-OFDM-IM in an explicit manner<sup>2</sup>, we list the key attributes of various OFDM-IM and sibling schemes in Table I.

Following the proposal of EHC-OFDM-IM, we analyze its error performance when maximum-likelihood (ML) detection is in use and the transmit power allocated to each active subcarrier does not depend on CSI, and we approximate the average block error rate (BLER) in closed form by the union bound and an exponential approximation of the Q-function.

Apart from the optimization of the mapping relation between SAPs and variable-length bit sequences, power allocation is another key technique that could be used to enhance error performance. Although there have existed a variety of power allocation schemes proposed for plain OFDM [34]–[38], for OFDM-IM, there are only a limited number of works dedicated to power allocation. The power allocation for OFDM-IM systems was first investigated in [39], in which two power allocation schemes aiming at optimizing error performance were proposed based on a simplistic approximation counting only two kinds of pairwise error events. Profile-based power allocation for OFDM-IM was studied in [40], which resorts to the maximization of the minimum Euclidean distance (MED). Better error performance is attainable with rea-

<sup>2</sup>Recently, an open-source IM toolkit has been released aiming at facilitating the quantitative comparisons among multiple IM schemes and accelerating open innovation in IM studies. Interested readers can refer to [28] for more details.

TABLE I: Comparisons and key attributes of various OFDM-IM and sibling schemes.

Scheme	Literature	Variable no. of active subcarriers?	Variable bit stream length?	CSI required at the Tx?	Adaptive codebook?	All legit. SAPs used?	Diversity gain harvested?	Coding gain harvested?
EHC-OFDM-IM (proposed)	–	NO	YES	YES	YES	YES	NO	YES
Classic Huffman coded OFDM-IM	[22]	NO	NO	NO	NO	YES	NO	NO
Adaptive Huffman coded OFDM-IM I	[23]	NO	YES	YES	YES	YES	NO	YES
Adaptive Huffman coded OFDM-IM II	[24]	NO	YES	NO	YES	YES	NO	NO
Adaptive Huffman coded OFDM-IM III	[25]	NO	YES	YES	YES	NO	NO	YES
Binary-tree coded OFDM-IM	[26]	NO	YES	YES	YES	YES	NO	YES
Classic OFDM-IM	[8]	NO	NO	NO	NO	NO	NO	NO
Adaptive OFDM-IM	[14], [29]	YES	YES	YES	YES	NO	YES	YES
OFDM-IM with an always-on subcarrier	[15]–[18]	YES	YES	NO	NO	NO	NO	NO
LCD-OFDM-IM	[27]	NO	NO	YES	YES	NO	YES	YES
SIM OFDM	[12]	YES	YES	NO	YES	NO	NO	NO
Enhanced SIM OFDM	[13]	NO	NO	NO	NO	YES	NO	NO
OFDM-SNM	[30]	YES	YES	NO	NO	NO	NO	NO
Enhanced OFDM-SNM	[31]	YES	YES	YES	YES	NO	NO	YES
FQAM/FPSK	[32], [33]	NO	NO	NO	NO	NO	NO	NO

sonable computational complexity. Power allocation problems for adaptive cooperative OFDM-IM and dual-mode OFDM-IM were formulated and discussed in [29] and [41], respectively. The power allocation problem for maximizing the rate of binary-tree coded OFDM-IM was also investigated in [26]. To further enhance the error performance of EHC-OFDM-IM, we study the corresponding power allocation problems and propose two suboptimal power allocation schemes based on the CSI at the transmitter with different requirements for computational complexity, which are capable of achieving a lower average BLER and are also easy to implement.

Numerical results generated by Monte Carlo simulations are provided to verify the error performance superiority and the error performance analysis of EHC-OFDM-IM either with or without applying the CSI-based power allocation schemes. Furthermore, we also investigate the trade-off in the architecture of the Huffman tree between error performance and data transmission rate as well as the impacts of imperfect CSI on the error performance through the numerical results provided. To study the bitwise error event in the context of variable-length bit sequences by EHC-OFDM-IM, we design an appropriate way to count the number of bit errors and numerically investigate the bit error rate (BER) by numerical results. The proposed scheme and algorithms in this paper can be easily extended to more sophisticated configurations, e.g.,

space-time-frequency IM, dual-mode OFDM-IM, and multi-mode OFDM-IM [42]–[44].

The rest of the paper is organized as follows. In Section II, we present the system model. We then give the details on how Huffman coding can be enhanced with CSI at the transmitter in Section III. Then, we carry out error performance analysis in Section IV and investigate the power allocation problems in Section V. Following the analysis, numerical results are illustrated, compared, and discussed in Section VI. Finally, the paper is concluded in Section VII.

## II. SYSTEM MODEL

Although a full set of subcarriers will normally be separated into different groups by OFDM-IM, for simplicity, we now only focus on a single group of  $N$  subcarriers in this paper. For each transmission attempt,  $K$  out of these  $N$  subcarriers are chosen to be activated so as to represent an SAP for conveying extra index bits, which are assumed to be equiprobable. Therefore, the number of all legitimate SAPs is  $S = \binom{N}{K}$ , where  $\binom{N}{K}$  is the binomial coefficient. Further assume that  $M$ -ary phase-shift keying ( $M$ -PSK) is in use to modulate equiprobable data bits due to its *constant modulus nature* and *rotational symmetry* [10], [45]. We write the transmitted

OFDM block as

$$\mathbf{x}(s, \mathbf{m}) = [x(s, m_1), x(s, m_2), \dots, x(s, m_N)]^T \in \mathbb{C}^{N \times 1}, \quad (1)$$

where  $s \in \mathcal{S} = \{1, 2, \dots, S\}$  is the index of the SAP;  $\mathbf{m} = [m_1, m_2, \dots, m_N] \in \mathcal{M}^N = \{0, 1, 2, \dots, M\}^N$  is the joint index vector of  $K$   $M$ -ary data symbols conveyed on active subcarriers and  $N - K$  nulls on inactive subcarriers;  $x(s, m_n)$  is a normalized data symbol only if the  $n$ th subcarrier is active, or null otherwise<sup>3</sup>;  $(\cdot)^T$  denotes the transpose operation on the enclosed vector/matrix.

Assuming perfect synchronization, pulse shaping, equalization, sampling and knowledge of CSI, the received OFDM block after sampling and discarding the cyclic prefix can be expressed as [35]

$$\mathbf{y}(s, \mathbf{m}) = [y(s, m_1), y(s, m_2), \dots, y(s, m_N)]^T = \mathbf{\Omega}(s) \mathbf{H} \mathbf{x}(s, \mathbf{m}) + \mathbf{w} \in \mathbb{C}^{N \times 1}, \quad (2)$$

where  $\mathbf{\Omega}(s) = \text{diag}\{\sqrt{P_t(s, 1)}, \sqrt{P_t(s, 2)}, \dots, \sqrt{P_t(s, N)}\}$  and  $P_t(s, n)$  is the transmit power allocated to the  $n$ th subcarrier in the  $s$ th SAP<sup>4</sup>;  $\mathbf{H} = \text{diag}\{h(1), h(2), \dots, h(N)\}$  and  $\mathbf{w} = [w(1), w(2), \dots, w(N)]^T$  represent the diagonal channel state matrix and the additive white Gaussian noise (AWGN) vector, whose entries are the channel coefficients and AWGN samples complying with complex Gaussian distributions  $\mathcal{CN}(0, \mu)$  and  $\mathcal{CN}(0, N_0)$ , respectively;  $\mu$  is the average channel power gain, and  $N_0$  is the average noise power. Note that, all channel coefficients are assumed to be independent, which can be easily achieved by interleaved subcarrier grouping [46]. Here, we adopt the Rayleigh fading model, and all channel power gains  $G(1), G(2), \dots, G(N)$  ( $G(n) = |h(n)|^2, \forall n \in \mathcal{N} = \{1, 2, \dots, N\}$ ) are assumed to be independently and exponentially distributed with the same average channel power gain  $\mu$ , which have the probability density function (PDF) and cumulative distribution function (CDF) given by [47]

$$f_G(\epsilon) = \frac{1}{\mu} \exp\left(-\frac{\epsilon}{\mu}\right) \Leftrightarrow F_G(\epsilon) = 1 - \exp\left(-\frac{\epsilon}{\mu}\right). \quad (3)$$

Receiving the contaminated OFDM block, the receiver can utilize ML detection scheme to estimate the transmitted OFDM block by the criterion [46]

$$\hat{\mathbf{x}}(\hat{s}, \hat{\mathbf{m}}) = \arg \min_{\hat{\mathbf{x}}(\hat{s}, \hat{\mathbf{m}}) \in \mathcal{X}} \|\mathbf{y}(s, \mathbf{m}) - \mathbf{\Omega}(\hat{s}) \mathbf{H} \hat{\mathbf{x}}(\hat{s}, \hat{\mathbf{m}})\|_F, \quad (4)$$

where  $\|\cdot\|_F$  denotes the Frobenius norm of the enclosed vector/matrix, and  $\mathcal{X}$  is the set of all legitimate transmit OFDM blocks, which has a cardinality of  $SM^K$ .

<sup>3</sup>For convenience, we let  $m_n = 0$  indicate that the  $n$ th subcarrier is inactive when the  $s$ th SAP is adopted for transmission.

<sup>4</sup>For simplicity, we tentatively suppose that  $P_t(s, n)$  is independent of CSI and has been set when analyzing the error performance of EHC-OFDM-IM systems. To further enhance the error performance, CSI-based power allocation can be applied, which, however, can not be analyzed thoroughly to the best of authors' knowledge. Details of the CSI-based power allocation can be found in Section V.

When the OFDM block  $\mathbf{x}(s, \mathbf{m})$  is transmitted, the corresponding BLER can be written as<sup>5</sup>

$$P_e(\mathbf{x}(s, \mathbf{m})) = \mathbb{P}\{\hat{\mathbf{x}}(\hat{s}, \hat{\mathbf{m}}) \neq \mathbf{x}(s, \mathbf{m})\}, \quad (5)$$

where  $\mathbb{P}\{\cdot\}$  denotes the probability of the random event enclosed. Consequently, by considering all legitimate transmit OFDM blocks, we can write the average BLER as

$$\bar{P}_e = \mathbb{E}_{\mathbf{x}(s, \mathbf{m}) \in \mathcal{X}} \{P_e(\mathbf{x}(s, \mathbf{m}))\}, \quad (6)$$

where  $\mathbb{E}\{\cdot\}$  denotes the expected value of the enclosed random variable.

### III. ENHANCED HUFFMAN CODING WITH CSI

By introducing Huffman coding, a full binary tree with  $S$  leaves representing  $S$  SAPs is constructed, which is also termed the Huffman tree. Here, we are only interested in the full binary tree because of its *prefix-free property* and *instantaneous decodability* [48]. For a full binary tree with  $S$  leaves, its architecture can be uniquely characterized by a set of depths  $\{d(s)\}_{s=1}^S$  or equivalently a set of occurrence probabilities  $\{p(s)\}_{s=1}^S$ , where  $p(s) = \left(\frac{1}{2}\right)^{d(s)}$  and  $\sum_{s=1}^S p(s) = 1$  [23]. The number of full binary trees with  $S$  leaves is the Catalan number  $C(S) = \frac{1}{(S+1)} \binom{2S}{S}$ , among which the minimum and maximum depths are  $\lceil \log_2(S-1) \rceil + 1$  and  $S-1$ , respectively [26]. Once the architecture of the Huffman tree used for Huffman encoding is given, the average encoding rate is fixed, which can be calculated by [24]

$$B = \underbrace{\sum_{s=1}^S \left[ \left(\frac{1}{2}\right)^{d(s)} d(s) \right]}_{B_I} + \underbrace{K \log_2(M)}_{B_D}, \quad (7)$$

where  $B_I$  is the encoding rate of the index bits modulated by the SAP in units of bit per channel use (bpcu), and  $B_D$  is the encoding rate of the data bits modulated by the data constellation symbols conveyed on  $K$  active subcarriers in units of bpcu.

Meanwhile, with CSI at the transmitter, there is still a possibility to further enhance the error performance by designing a bijective mapping relation between  $S$  SAPs and  $S$  leaves on the generated Huffman tree. This is because the occurrence probabilities of SAPs represented by different leaves are different, as long as the Huffman tree is not *perfect*<sup>6</sup>.

<sup>5</sup>Although in most publications associated with OFDM-IM, bit error rate (BER) is adopted as an error performance evaluation metric, BER is intuitive and more suited for measuring the error performance when bit streams with a fixed and equal length are transmitted. However, by involving Huffman coding, the length of the transmitted bit stream varies when different SAPs are chosen. When variable-length bit streams are taken into consideration, we cannot 'intuitively' determine the number of bit errors without stipulating a unified quantifying rule. However, defining such a rule is not a trivial task. The quantifying rule must be general enough to cover all possible scenarios, and in the meantime aligned with common sense, e.g., the number of bit errors cannot be higher than the number of transmitted bits. Moreover, the rule must also lead to consistent evaluations for all scenarios. Therefore, to avoid unnecessary discussion regarding how to count bit errors when transmitted and erroneously decoded bit streams have different lengths, we adopt BLER as the error performance evaluation metric in this paper.

<sup>6</sup>A perfect Huffman tree is defined as the tree, on which all interior nodes have two children and all resultant leaves are with the same depth [49].

To enable the error performance optimization by CSI at the transmitter with channel power gains  $G(1), G(2), \dots, G(N)$ , we order the subchannels by their channel power gains and write the ranked channel power gains in ascending order as<sup>7</sup>

$$G_{(1)}(\xi_1) < G_{(2)}(\xi_2) < \dots < G_{(N)}(\xi_N), \quad (8)$$

which characterizes the utilization preference of all  $N$  subcarriers, where  $\xi_n$  is the index of the subcarrier with the  $n$ th smallest channel power gain.

However, the performance of OFDM-IM systems is associated with an SAP with  $K$  active subcarriers instead of a single active subcarrier; we have to derive the utilization preference of  $S$  SAPs. To obtain the utilization preference of  $S$  SAPs, we resort to the pattern ordering principle by reverse lexicographic order presented in [27]. Then, we can rank the  $S$  SAPs by priority according to the following steps:

1) *Generate the first and the last ASVs*: The first activation state vector (ASV) corresponding to the best SAP can be constructed by creating a  $N \times 1$  zero vector and then replacing the rightmost  $K$  zeros by ones. Mathematically, we have the initial ASV in the following form

$$\mathbf{v}_1 = \underbrace{[0, 0, \dots, 0]_{N-K}}_{N-K} \underbrace{[1, 1, \dots, 1]_K}_K. \quad (9)$$

Similarly, the last ASV corresponding to the worst SAP can be constructed by creating a  $N \times 1$  zero vector and then replacing the leftmost  $K$  zeros by ones. Mathematically, we have the initial ASV in the following form

$$\mathbf{v}_S = \underbrace{[1, 1, \dots, 1]_K}_K \underbrace{[0, 0, \dots, 0]_{N-K}}_{N-K}. \quad (10)$$

Both are generated as the reference points for the other ASVs.

2) *Enumerate  $S$  ASVs*: Apart from both ends, we also need to have the ASVs in the middle. When the number of subcarriers  $N$  is small, the easiest way to enumerate all  $S$  ASVs is to generate an expanded set of  $2^N$  ASVs by releasing the requirement on the number of active subcarriers. We can thereby have the expanded set as

$$\begin{aligned} & \{[0, 0, \dots, 0, 0]^T, [0, 0, \dots, 0, 1]^T \\ & [0, 0, \dots, 1, 0]^T, [0, 0, \dots, 1, 1]^T \\ & \vdots \\ & [1, 1, \dots, 1, 0]^T, [1, 1, \dots, 1, 1]^T\} \end{aligned} \quad (11)$$

Subsequently, we check the Hamming weight of each element in the expanded set and obtain a subset by removing the elements whose Hamming weights do not equal  $K$ . The obtained subset contains all  $S$  ASVs corresponding to the legitimate SAPs with  $K$  active subcarriers. Albeit simple, the

<sup>7</sup>To be general, the relation between the channel power gains and ranked channel power gains can be explicitly expressed in the following way. Assuming the full set of channel power gains is denoted as  $\mathbf{G} = \{G(1), G(2), \dots, G(N)\}$ , we have  $G_{(1)}(\xi_1) = \min_{\mathbf{G}}\{G(n)\}$ ,  $G_{(2)}(\xi_2) = \min_{\mathbf{G} \setminus \{G_{(1)}(\xi_1)\}}\{G(n)\}$ ,  $G_{(3)}(\xi_3) = \min_{\mathbf{G} \setminus \{G_{(1)}(\xi_1), G_{(2)}(\xi_2)\}}\{G(n)\}$ ,  $\dots$ ,  $G_{(N)}(\xi_N) = \min_{\mathbf{G} \setminus \{G_{(1)}(\xi_1), G_{(2)}(\xi_2), \dots, G_{(N-1)}(\xi_{N-1})\}}\{G(n)\} = \max_{\mathbf{G}}\{G(n)\}$ .

above approach might not be applicable when  $N$  is large, because it is demanding to have the register space for  $2^N$  vectors.

For large  $N$ , we propose a *state shifting* approach based on the first ASV  $\mathbf{v}_1$  to yield the set of  $S$  ASVs. Imagining that ‘1’s in an ASV are entities and there is nothing in the positions of ‘0’, we shift the leftmost ‘1’ to the left by one place and obtain the next ASV as  $[0, 0, \dots, 1, 0, 1, \dots, 1]^T$ . Similarly, we shift the second leftmost ‘1’ to the left by one place and obtain the next ASV as  $[0, 0, \dots, 1, 1, 0, \dots, 1]^T$ . We subsequently repeat this process  $K$  times until a *consecutive* string of  $K$  ‘1’s emerging again, and have the  $(K+1)$ th ASV  $\underbrace{[0, 0, \dots, 0]_{N-K-1}}_{N-K-1} \underbrace{[1, 1, \dots, 1]_K}_K$ . After this, we shift the leftmost ‘1’ in the consecutive string to the left by one place and reset all other  $K-1$  ‘1’s starting back to the rightmost position in the ASV, which results in the  $(K+2)$ th vector  $\underbrace{[0, 0, \dots, 0]_{N-K-2}}_{N-K-2} \underbrace{[1, 0, 0, 1, 1, \dots, 1]_{K-1}}_{K-1}$ . We repeat this process until the next emergence of a consecutive string of  $K$  ‘1’s and so on. Finally, we terminate the state shifting process when the last ASV  $\mathbf{v}_S$  is obtained. In this way, we can obtain the set of  $S$  ASVs corresponding to all legitimate SAPs with  $K$  active subcarriers.

3) *Rank  $S$  ASVs by reverse lexicographic order*: By the ASV generating approaches introduced above, we can easily enumerate all legitimate SAPs by their ASVs  $\mathbf{v}_s = [v_s(\xi_1), v_s(\xi_2), \dots, v_s(\xi_N)]^T \in \{0, 1\}^{N \times 1}$ , where  $v_s(\xi_n)$  is either ‘0’ or ‘1’ depending on whether the  $\xi_n$ th subcarrier (i.e., the subcarrier with the  $n$ th smallest channel power gain) is activated<sup>8</sup>. In particular, we rank all  $S$  SAPs in ascending order by their ASVs:

$$\mathbf{v}_1 <_{\mathbf{L}} \mathbf{v}_2 <_{\mathbf{L}} \dots <_{\mathbf{L}} \mathbf{v}_S, \quad (12)$$

where  $<_{\mathbf{L}}$  is the lexicographic less-than sign, and we define  $\mathbf{v}_s <_{\mathbf{L}} \mathbf{v}_t$  if  $\omega_s < \omega_t$ , where  $\omega_s = \text{b2d}(\mathbf{v}_s)$  and  $\text{b2d}(\cdot)$  converts a binary vector to a decimal number given the most significant digit on the left side. It has been explained in [27] that the smallest SAPs in the lexicographic sense are more preferable in most cases when considering the error and outage performance for OFDM-IM systems with independent and identically distributed (i.i.d.) channel power gains. As a result, we rank the  $S$  SAPs by priority according to the three steps given above.

In contrast to the classic OFDM-IM systems without Huffman coding [27], which simply select a subset of  $2^{\lfloor \log_2(S) \rfloor}$  SAPs from the first to the  $2^{\lfloor \log_2(S) \rfloor}$ th SAPs ranked by (12) and discarding the rest, Huffman coded OFDM-IM systems need to make full use of all  $S$  SAPs in order to improve the data transmission rate. Although it would be impossible to harvest a diversity gain in Huffman coded OFDM-IM systems in this way, we can still exploit CSI at the transmitter and the ranked ASVs to obtain a coding gain. This is simply because, in classic OFDM-IM systems, all selected SAPs have the

<sup>8</sup>For example,  $\mathbf{v}_s = [0, 1, 0, 1]^T$  does not mean that the second and fourth subcarriers shall be activated by their indices, but implies that the subcarriers with the second and fourth smallest channel power gains, i.e., the second and fourth *ordered* subcarriers, shall be activated.

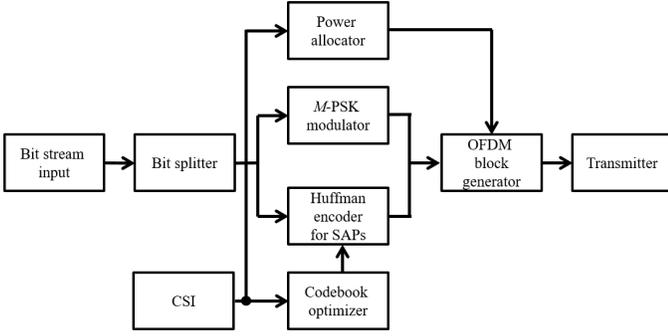


Fig. 1: System block diagram of the proposed EHC-OFDM-IM system.

same probability of being activated, while in Huffman coded OFDM-IM systems, the occurrence probabilities corresponding to  $S$  SAPs could be different. As we mentioned above, the occurrence probability of an SAP  $p(s)$  is associated with the depth  $d(s)$  of its representative leaf on the Huffman tree by the relation  $p(s) = (\frac{1}{2})^{d(s)}$ . That is, the SAPs paired to leaves with larger depths will be activated less, and vice versa. As a result, we can design and optimize the pairing relation between  $S$  SAPs to the  $S$  leaves on a given Huffman tree so as to reduce the activation probabilities of those undesirable SAPs (lexicographic larger ones). In this way, the average BLER is expected to be reduced, and hence a coding gain can be harvested. Note that the coding gain achieved by the proposed optimization strategy using CSI at the transmitter will disappear if the given Huffman tree is a perfect binary tree (i.e., all interior nodes on the tree have two children and all resultant leaves are with the same depth [49]).

In particular, given a Huffman tree, we rank all its leaves by their depths in ascending order as<sup>9</sup>

$$d_{(1)}(\sigma_1) \leq d_{(2)}(\sigma_2) \leq \dots \leq d_{(S)}(\sigma_S), \quad (13)$$

where  $d_{(s)}(\sigma_s)$  is the  $s$ th smallest depth and  $\sigma_s$  is the index of the leaf with the  $s$ th smallest depth.

With the ranked relations presented in (12) and (13), we define the optimized bijective mapping relation between SAPs represented by their ASVs and leaves as<sup>10</sup>

$$f_{\text{opt}} : \mathbf{v}_s \leftrightarrow \sigma_s. \quad (14)$$

By introducing the encoding procedure, we depict the system block diagram of the proposed EHC-OFDM-IM system in Fig. 1. We present an example here to show how to implement the optimization strategy elaborated above.

**Example** Assume  $\{N, K\} = \{4, 2\}$ , i.e., four subcarriers in total and two of them are activated to form an SAP. Obviously, we have  $S = \binom{4}{2} = 6$  SAPs, which can be expressed

<sup>9</sup>One should note that it is possible and even common that for  $s \neq t$ , we have  $d_{(s)}(\sigma_s) = d_{(t)}(\sigma_t)$ , which indicates that both leaves are equally preferable and can be ranked arbitrarily. When given a perfect binary tree, we will have  $d_{(1)}(\sigma_1) = d_{(2)}(\sigma_1) = \dots = d_{(s)}(\sigma_s)$ , and all leaves are equivalent. As a result, the ordering process is nullified and no performance gain can be attained in this special case.

<sup>10</sup>To enable the CSI-based mapping optimization, we assume that channels are slowly faded and the overhead caused by updating the optimized bijective mapping relation is negligible.

TABLE II: An example of the bijective mapping relation between six SAPs and six leaves.

ASV	Active subcarriers	Leaf's depth	Leaf's activation prob.	Leaf's index	Code words
$\mathbf{v}_1$	{2, 3}	2	1/4	5	10
$\mathbf{v}_2$	{1, 3}	2	1/4	6	11
$\mathbf{v}_3$	{1, 2}	3	1/8	1	000
$\mathbf{v}_4$	{3, 4}	3	1/8	2	001
$\mathbf{v}_5$	{2, 4}	3	1/8	3	010
$\mathbf{v}_6$	{1, 4}	3	1/8	4	011

by ASVs ranked by the lexicographic ordering principle as  $\mathbf{v}_1 = [0, 0, 1, 1]^T$ ,  $\mathbf{v}_2 = [0, 1, 0, 1]^T$ ,  $\mathbf{v}_3 = [0, 1, 1, 0]^T$ ,  $\mathbf{v}_4 = [1, 0, 0, 1]^T$ ,  $\mathbf{v}_5 = [1, 0, 1, 0]^T$ , and  $\mathbf{v}_6 = [1, 1, 0, 0]^T$ . Assuming the channel power gains of the four subcarriers are given by  $G(1) = 0.627$ ,  $G(2) = 0.884$ ,  $G(3) = 1.716$ , and  $G(4) = 0.337$ , we can then have the ranked channel power gains in ascending order as

$$G_{(1)}(4) < G_{(2)}(1) < G_{(3)}(2) < G_{(4)}(3). \quad (15)$$

As ASVs are the collections of activation indicators of *ordered* subcarriers, the six ASVs from  $\mathbf{v}_1$  to  $\mathbf{v}_6$  correspond to the subsets of active subcarriers {2, 3}, {1, 3}, {1, 2}, {3, 4}, {2, 4}, and {1, 4}, respectively. To provide an intuitive illustration of the full SAP ordering procedure, we pictorially demonstrate the given example in Fig. 2.

The Huffman tree with leaves numbered from left to right is shown in Fig. 3. Obviously, we have the ranked depths by

$$\underbrace{d_{(1)}(5) = d_{(2)}(6)}_{=2} < \underbrace{d_{(3)}(1) = d_{(4)}(2) = d_{(5)}(3) = d_{(6)}(4)}_{=3}. \quad (16)$$

Subsequently, we resort to (14) to yield the bijective mapping relation between the six SAPs and the six leaves in Table II. By observing this table, we find that the two most preferable SAPs represented by  $\mathbf{v}_1$  and  $\mathbf{v}_2$  are activated with higher probabilities than the other SAPs. By this optimized mapping relation, the resultant system is expected to provide a lower average BLER.

#### IV. ERROR PERFORMANCE ANALYSIS

For analytical simplicity, we temporarily assume that the transmit power allocated to each active subcarrier does not depend on CSI. Based on this regulation, we are able to derive a close-form and insightful expression of the average BLER. To do so, we first utilize the union bound to approximate (5) by [50]

$$P_e(\mathbf{x}(s, \mathbf{m})) \leq \sum_{\hat{\mathbf{x}}(\hat{s}, \hat{\mathbf{m}}) \neq \mathbf{x}(s, \mathbf{m})} P_e(\mathbf{x}(s, \mathbf{m}) \rightarrow \hat{\mathbf{x}}(\hat{s}, \hat{\mathbf{m}})), \quad (17)$$

where  $P_e(\mathbf{x}(s, \mathbf{m}) \rightarrow \hat{\mathbf{x}}(\hat{s}, \hat{\mathbf{m}}))$  denotes the pairwise error probability (PEP) that the actual transmitted OFDM block  $\mathbf{x}(s, \mathbf{m})$  is detected as  $\hat{\mathbf{x}}(\hat{s}, \hat{\mathbf{m}})$  by mistake. Further, PEP  $P_e(\mathbf{x}(s, \mathbf{m}) \rightarrow \hat{\mathbf{x}}(\hat{s}, \hat{\mathbf{m}}))$  can be derived by averaging the conditional PEP over channel coefficients as

$$P_e(\mathbf{x}(s, \mathbf{m}) \rightarrow \hat{\mathbf{x}}(\hat{s}, \hat{\mathbf{m}})) = \mathbb{E}_{\mathbf{H}} \{P_e(\mathbf{x}(s, \mathbf{m}) \rightarrow \hat{\mathbf{x}}(\hat{s}, \hat{\mathbf{m}})|\mathbf{H})\}. \quad (18)$$

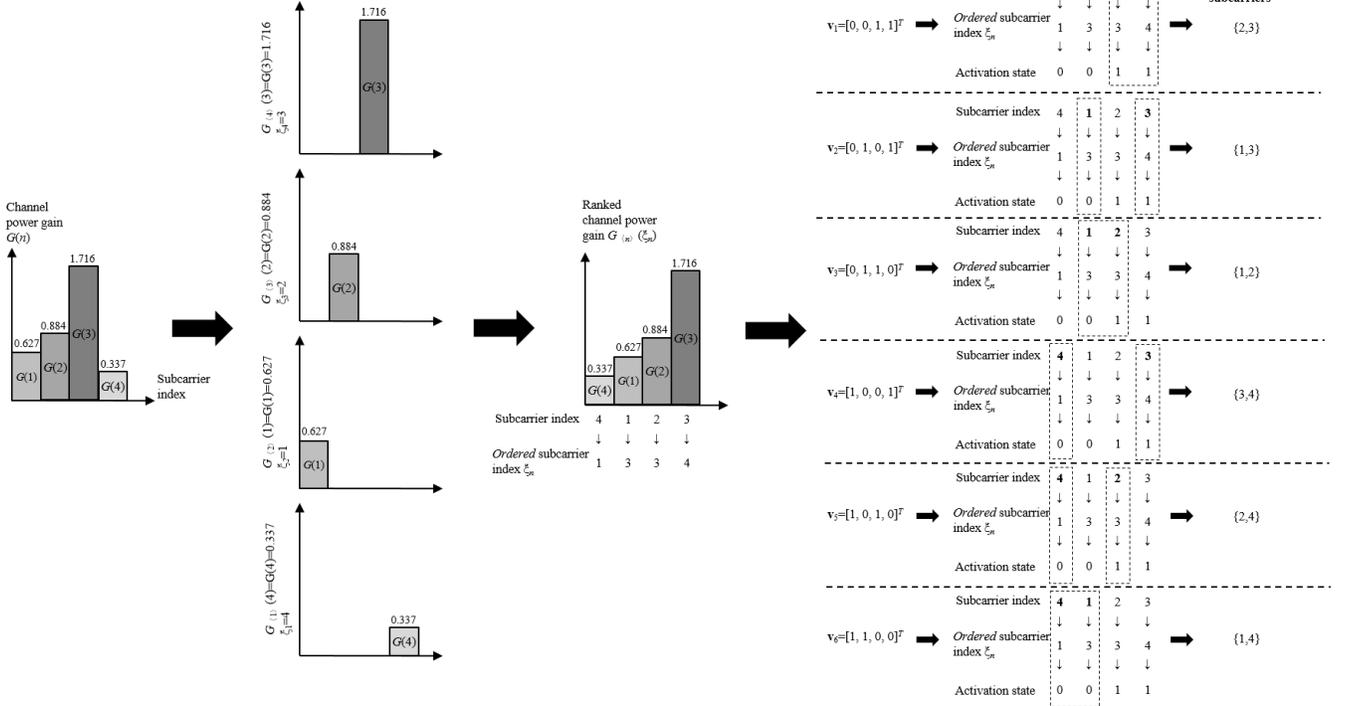


Fig. 2: SAP ordering procedure of the given example by the lexicographic ordering principle.

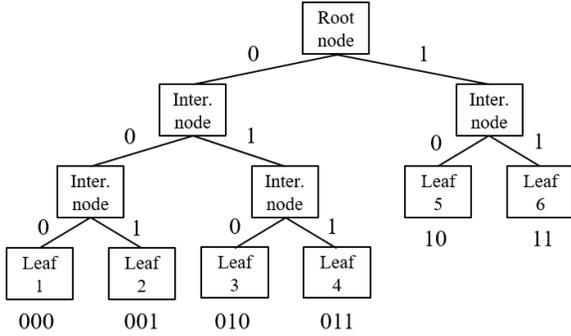


Fig. 3: An example of the Huffman tree with codewords attached.

According to [8], when the ML detection scheme is in use, the conditional PEP can be determined in (19), where  $Q(\epsilon) = \frac{1}{\sqrt{2\pi}} \int_{\epsilon}^{\infty} \exp\left(-\frac{u^2}{2}\right) du$  is the Q-function; (a) is approximated by an accurate exponential bound of the Q-function  $Q(\epsilon) \approx \prod_{q=1}^2 \rho_q \exp(-\eta_q \epsilon^2)$  proposed in [51];  $\{\rho_1, \rho_2\} = \{1/12, 1/4\}$  and  $\{\eta_1, \eta_2\} = \{1/2, 2/3\}$ .

To facilitate the following analysis and provide the simplest form of the final derivation, it is suggested to rearrange the form of the transmit OFDM block by the *orders* of subcarriers instead of their *indices*. Therefore, we resort to the concept of the *permuted* OFDM block proposed in [27] and have<sup>11</sup>

$$\mathbf{z}(s, \mathbf{m}) = [z(s, m_{\xi_1}), z(s, m_{\xi_2}), \dots, z(s, m_{\xi_N})]^T \in \mathbb{C}^{N \times 1}, \quad (20)$$

<sup>11</sup>One should note that the permuted OFDM block is simply utilized to facilitate the expression of analytical results and does not amend the actual transmission procedure at all.

which is built by replacing the ‘1’s of an ASV with the data symbols in the corresponding  $\mathbf{x}(s, \mathbf{m})$  in sequence. By establishing such a one-to-one mapping relation between  $\mathbf{x}(s, \mathbf{m})$  and  $\mathbf{z}(s, \mathbf{m})$ , (19) can be equivalently expressed as (21).

Meanwhile, the PDF of the  $n$ th smallest channel power gain  $G_{(n)}(\xi_n)$  is determined by the rudiments of order statistics as [52]

$$\phi_{(n)}(\epsilon) = \frac{N!(F_G(\epsilon))^{n-1}(1-F_G(\epsilon))^{N-n}f_G(\epsilon)}{(n-1)!(N-n)!}, \quad (22)$$

where  $(\cdot)!$  denotes the factorial of the enclosed argument. Subsequently, we can remove the condition on  $\mathbf{H}$  by averaging (21) over  $\{G_{(n)}(\xi_n)\}_{n=1}^N$  when the allocated power for active subcarriers is independent of  $\mathbf{H}$ , which yields (23), where (a) follows from by the independence among subcarriers and  $\Gamma(\epsilon) = \int_0^{\infty} u^{\epsilon-1} \exp(-u) du$  is the complete gamma function. Substituting (23) into (17) gives the approximate BLER when  $\mathbf{x}(s, \mathbf{m})$  is transmitted.

Finally, according to (6), (14) and the Huffman encoding rules, it is evident that there exist  $M^K$  legitimate OFDM blocks adopting the same SAP. Hence, it is straightforward to explicitly express the average BLER as

$$\bar{P}_e = \frac{1}{M^K} \sum_{\mathbf{x}(s, \mathbf{m}) \in \mathcal{X}} p(s) P_e(\mathbf{x}(s, \mathbf{m})). \quad (24)$$

## V. POWER ALLOCATION AND ERROR PERFORMANCE OPTIMIZATION

### A. Suboptimal Power Allocation Scheme I

When  $\mathbf{H}$  is fixed and known at the transmitter, we can resort to the conditional average BLER as an optimization

$$\begin{aligned}
P_e(\mathbf{x}(s, \mathbf{m}) \rightarrow \hat{\mathbf{x}}(\hat{s}, \hat{\mathbf{m}})|\mathbf{H}) &= Q \left( \sqrt{\frac{1}{N_0} \|\mathbf{H}(\boldsymbol{\Omega}(s)\mathbf{x}(s, \mathbf{m}) - \boldsymbol{\Omega}(\hat{s})\hat{\mathbf{x}}(\hat{s}, \hat{\mathbf{m}}))\|_F^2} \right) \\
&= Q \left( \sqrt{\frac{1}{N_0} \sum_{n=1}^N G(n) |\sqrt{P_t(s, n)}x(s, m_n) - \sqrt{P_t(\hat{s}, n)}\hat{x}(\hat{s}, \hat{m}_n)|^2} \right) \\
&\stackrel{(a)}{\approx} \sum_{q=1}^2 \rho_q \exp \left( -\frac{\eta_q}{N_0} \sum_{n=1}^N G(n) |\sqrt{P_t(s, n)}x(s, m_n) - \sqrt{P_t(\hat{s}, n)}\hat{x}(\hat{s}, \hat{m}_n)|^2 \right) \\
&= \sum_{q=1}^2 \rho_q \prod_{n=1}^N \exp \left( -\frac{\eta_q}{N_0} G(n) |\sqrt{P_t(s, n)}x(s, m_n) - \sqrt{P_t(\hat{s}, n)}\hat{x}(\hat{s}, \hat{m}_n)|^2 \right)
\end{aligned} \tag{19}$$

$$P_e(\mathbf{x}(s, \mathbf{m}) \rightarrow \hat{\mathbf{x}}(\hat{s}, \hat{\mathbf{m}})|\mathbf{H}) \approx \sum_{q=1}^2 \rho_q \prod_{n=1}^N \exp \left( -\frac{\eta_q}{N_0} G_{\langle n \rangle}(\xi_n) |\sqrt{P_t(s, \xi_n)}z(s, m_n) - \sqrt{P_t(\hat{s}, \xi_n)}\hat{z}(\hat{s}, \hat{m}_n)|^2 \right) \tag{21}$$

$$\begin{aligned}
P_e(\mathbf{x}(s, \mathbf{m}) \rightarrow \hat{\mathbf{x}}(\hat{s}, \hat{\mathbf{m}})) &= \underbrace{\int_0^\infty \int_0^\infty \cdots \int_0^\infty}_{N} P_e(\mathbf{x}(s, \mathbf{m}) \rightarrow \hat{\mathbf{x}}(\hat{s}, \hat{\mathbf{m}})|\mathbf{H}) \left( \prod_{n=1}^N \phi_{\langle n \rangle}(G_{\langle n \rangle}(\xi_n)) \right) dG_{\langle 1 \rangle}(\xi_1), \dots, dG_{\langle N \rangle}(\xi_N) \\
&\stackrel{(a)}{=} \sum_{q=1}^2 \rho_q \prod_{n=1}^N \frac{N! \Gamma(N - n + 1 + \frac{\eta_q}{N_0} |\sqrt{P_t(s, \xi_n)}z(s, m_{\xi_n}) - \sqrt{P_t(\hat{s}, \xi_n)}\hat{z}(\hat{s}, \hat{m}_{\xi_n})|^2)}{(N - n)! \Gamma(N + 1 + \frac{\eta_q}{N_0} |\sqrt{P_t(s, \xi_n)}z(s, m_{\xi_n}) - \sqrt{P_t(\hat{s}, \xi_n)}\hat{z}(\hat{s}, \hat{m}_{\xi_n})|^2)}
\end{aligned} \tag{23}$$

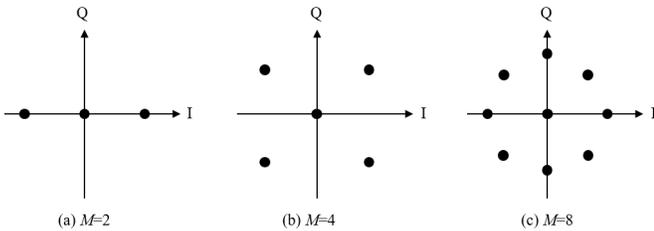


Fig. 4: Constellation diagrams for hypothetical quasi-OFDM-IM systems with  $M = 2$ ,  $M = 4$ , and  $M = 8$  when  $M$ -PSK is used for data constellation modulation.

objective to formulate a power allocation problem to further enhance the error performance of EHC-OFDM-IM. Different from plain OFDM, by which all subcarriers are activated and  $N$  symbols are transmitted, a subcarrier in EHC-OFDM-IM systems could be inactive. Also, because of the joint ML detection for symbols and SAP index (c.f. (4)), the power allocations for different SAPs are coupled. Both make the power allocation for EHC-OFDM-IM systems challenging. To enable the power allocation to optimize error performance, we make the following simplification. Specifically, from the perspective of transmitted symbols, deactivating a subcarrier is equivalent to transmitting a null on the subcarrier. In other words, we can reduce an EHC-OFDM-IM system utilizing  $M$ -PSK to a plain OFDM system utilizing a hypothetical  $M+1$  constellation diagram on  $K$  subcarriers, where the extra

constellation point is the origin in the two-dimensional signal plane. We term this paradigm a quasi-OFDM-IM system. For clarity, we illustrate constellation diagrams for quasi-OFDM-IM systems with  $M = 2$ ,  $M = 4$ , and  $M = 8$  in Fig. 4. This simplification allows us to decouple the power allocation procedures among different SAPs.

The BLER of the EHC-OFDM-IM system transmitting by the  $s$ th SAP can be approximated by the BLER of a hypothetical quasi-OFDM-IM system as<sup>12</sup>

$$\Upsilon_e(s|\mathbf{H}) = 1 - \prod_{i \in \mathcal{K}(s)} \left( 1 - Q \left( \sqrt{\frac{1}{N_0} P_t(s, i) G(i) d_{\min}^2} \right) \right), \tag{25}$$

where  $\mathcal{K}(s)$  is the set of  $K$  active subcarriers in the  $s$ th SAP;  $d_{\min}$  is the minimum Euclidean distance (MED) among  $M+1$  constellation points, which can be determined by

$$d_{\min} = \min \left\{ 1, \sqrt{2(1 - \cos(2\pi/M))} \right\}, \tag{26}$$

when  $M$ -PSK is used for data constellation modulation. In the high signal-to-noise (SNR) region, (25) can be simplified to [53]

$$\Upsilon_e(s|\mathbf{H}) \approx \tilde{\Upsilon}_e(s|\mathbf{H}) = \sum_{i \in \mathcal{K}(s)} Q \left( \sqrt{\frac{1}{N_0} P_t(s, i) G(i) d_{\min}^2} \right). \tag{27}$$

<sup>12</sup>Note that, this approximation is applicable for both  $M$ -PSK,  $M$ -QAM and other data constellation modulation schemes, although we adopt  $M$ -PSK as an illustrative example in the following part. Changing data constellation modulation scheme will only result in a different  $d_{\min}$ .

Now, we can formulate the power allocation problem for the  $s$ th SAP as (SP1):

$$\begin{aligned}
\text{(SP1)} \quad & \min_{\{P_t(s,n)\}} \left\{ \tilde{\Upsilon}_e(s|\mathbf{H}) \right\} \\
& \text{s.t.} \quad \sum_{i \in \mathcal{K}(s)} P_t(s,i) \leq \Psi \\
& \quad P_t(s,i) > 0, \quad \forall i \in \mathcal{K}(s) \\
& \quad P_t(s,i) = 0, \quad \forall i \notin \mathcal{K}(s)
\end{aligned}$$

where  $\Psi$  is the power budget allowed to be allocated to  $K$  active subcarriers. Because the Q-function  $Q(\cdot)$  is a monotone decreasing function, it is obvious that the formulated problem (SP1) is a quasi-convex optimization problem with respect to  $\{P_t(s,n)\}$  (it can be rigorously proved by the positive second-order partial derivative of the objective function with respect to  $\{P_t(s,n)\}$ ). As a consequence, we can construct a Lagrange function associated with the formulated problem (SP1) as follows:

$$\begin{aligned}
\mathcal{L}_1(P_t(s,1), P_t(s,2), \dots, P_t(s,N), \lambda) \\
= \tilde{\Upsilon}_e(s|\mathbf{H}) - \lambda \left( \Psi - \sum_{i \in \mathcal{K}(s)} P_t(s,i) \right), \quad (28)
\end{aligned}$$

where  $\lambda$  is the Lagrange multiplier. Subsequently, we resort to the Craig's expression of the Q-function and the Lebesgue's dominated convergence theorem to establish the Lagrange equation set as<sup>13</sup>

$$\begin{cases} \frac{\partial \mathcal{L}_1}{\partial \lambda} = - \left( \Psi - \sum_{i \in \mathcal{K}(s)} P_t(s,i) \right) = 0 \\ \frac{\partial \mathcal{L}_1}{\partial P_t(s,n)} = - \frac{1}{2\sqrt{\pi}P_t(s,n)} \sqrt{\frac{P_t(s,n)G(n)d_{\min}^2}{2N_0}} \\ \quad \times \exp\left(-\frac{P_t(s,n)G(n)d_{\min}^2}{2N_0}\right) + \lambda = 0, \quad \forall n \in \mathcal{K}(s) \end{cases} \quad (29)$$

Solving (29) yields the power allocation scheme to (SP1) when the  $s$ th SAP is in use.

However, as solving (29) involves transcendental equations, to the best of authors' knowledge, there is no closed-form expression of the allocated power to each active subcarrier. Iterative methods can be applied to obtain numerical solutions to (SP1). We take the bisection method as an example to numerically obtain the solution due to its simplicity and robustness [55]. To do so, we first carry out a simple transformation of the second equation in (29) and obtain the expression of  $P_t(s,t)$  as a function of  $\lambda$ :

$$P_t(s,n) = \frac{N_0}{G(n)d_{\min}^2} W \left( \frac{1}{8\pi} \left( \frac{G(n)d_{\min}^2}{\lambda N_0} \right)^2 \right), \quad (30)$$

where  $W(\epsilon)$  is the Lambert W function giving the principal solution for  $\varrho$  in  $\epsilon = \varrho \exp(\varrho)$ . We substitute (30) into the first equation in (29) and construct a function  $\Theta(\lambda)$  of  $\lambda$  as

$$\Theta(\lambda) = \Psi - \frac{N_0}{d_{\min}^2} \sum_{i \in \mathcal{K}(s)} \frac{1}{G(i)} W \left( \frac{1}{8\pi} \left( \frac{G(i)d_{\min}^2}{\lambda N_0} \right)^2 \right). \quad (31)$$

<sup>13</sup>The Lebesgue's dominated convergence theorem enables us to exchange the order of integration and differentiation operations in the constructed Lagrange function under certain conditions [54].

---

**Algorithm 1** Algorithm to find out the lower limit  $\lambda_-$  to start the bisection method based approaching procedure.

---

```

1: BEGIN
2: Input:  $N_0$ ,  $\mathbf{H}$ ,  $\mathcal{K}(s)$ , and  $M$ ;
3:  $j \leftarrow 0$ ;
4:  $\lambda_- \leftarrow \exp(j)$ ;
5: Calculate  $\Theta(\lambda_-)$  by (31);
6: while  $\Theta(\lambda_-) > 0$  do
7:    $j \leftarrow j + 1$ ;
8:    $\lambda_- \leftarrow \exp(j)$ ;
9: end while
10: return  $\lambda_-$ ;
11: END

```

---



---

**Algorithm 2** Algorithm to find out the upper limit  $\lambda_+$  to start the bisection method based approaching procedure.

---

```

1: BEGIN
2: Input:  $N_0$ ,  $\mathbf{H}$ ,  $\mathcal{K}(s)$ , and  $M$ ;
3:  $j \leftarrow 0$ ;
4:  $\lambda_+ \leftarrow \exp(j)$ ;
5: Calculate  $\Theta(\lambda_+)$  by (31);
6: while  $\Theta(\lambda_+) < 0$  do
7:    $j \leftarrow j + 1$ ;
8:    $\lambda_+ \leftarrow \exp(j)$ ;
9: end while
10: return  $\lambda_+$ ;
11: END

```

---

To perform the bisection method and start halving in an iterative manner, we first need to initialize the root interval  $[\lambda_-, \lambda_+]$ . However, because of the complication of the Lambert W function, the lower and upper limits  $\lambda_-$ ,  $\lambda_+$  that ensure  $\Theta(\lambda_-) < 0$  and  $\Theta(\lambda_+) > 0$  are not straightforward, which require elaborate design considering the trade-off between computational complexity and robustness. To determine these two limits, we apply heuristic algorithms as described in Algorithm 1 and Algorithm 2 to properly initialize  $\lambda_-$  and  $\lambda_+$ . Subsequently, with the initial root interval  $[\lambda_-, \lambda_+]$  and a stipulated precision  $\epsilon$ , the exact solution  $\lambda^*$  can be approached by the numerical solution  $\lambda^\&$  according to the bisection method as shown in Algorithm 3. Obviously, as long as we can find a  $\lambda^\&$  so as to achieve  $\Theta(\lambda^\&) \approx 0$ , we can substitute the value of  $\lambda^\&$  back into (30) to determine the amount of transmit power allocated to the  $n$ th active subcarrier when the  $s$ th SAP is in use. By repeating the procedure  $S$  times, the complete power allocation scheme for all  $S$  SAPs can be obtained.

### B. Suboptimal Power Allocation Scheme II

Because numerically solving (SP1) would render high computational complexity and a great amount of computational time, especially when the number of active subcarriers  $K$  is large, this could impair the practicality of power allocation for EHC-OFDM-IM in realistic scenarios where channel states vary rapidly. Therefore, to enhance the practicality, we are inspired by the exponential approximation of the Q-function  $Q(\epsilon) \approx \prod_{q=1}^2 \rho_q \exp(-\eta_q \epsilon^2)$  and reformulate another much

**Algorithm 3** Bisection method based procedure to approximate  $\lambda^*$  with  $\lambda^{\&}$  by a stipulated precision  $\varepsilon$ .

---

```

1: BEGIN
2: Input:  $\lambda_-$ ,  $\lambda_+$ ,  $\varepsilon$ ,  $N_0$ ,  $\mathbf{H}$ ,  $\mathcal{K}(s)$ , and  $M$ ;
3:  $\lambda_m \leftarrow (\lambda_- + \lambda_+)/2$ 
4: while  $\lambda_m - \lambda_- > \varepsilon$  do
5:   Calculate  $\Theta(\lambda_-)$  by (31);
6:   Calculate  $\Theta(\lambda_m)$  by (31);
7:   if  $\Theta(\lambda_-)\Theta(\lambda_m) < 0$  then
8:      $\lambda_+ \leftarrow \lambda_m$ ;
9:   else
10:     $\lambda_- \leftarrow \lambda_m$ ;
11:   end if
12:    $\lambda_m \leftarrow (\lambda_- + \lambda_+)/2$ 
13: end while
14:  $\lambda^{\&} \leftarrow \lambda_m$ ;
15: return  $\lambda^{\&}$ ;
16: END

```

---

simpler suboptimal power allocation problem (SP2) for a large power budget infra:

$$\begin{aligned}
(\text{SP2}) \quad & \min_{\{P_t(s,n)\}} \left\{ \sum_{i \in \mathcal{K}(s)} \exp\left(-\frac{1}{N_0} P_t(s,i) G(i) d_{\min}^2\right) \right\} \\
& \text{s.t.} \quad \sum_{i \in \mathcal{K}(s)} P_t(s,i) \leq \Psi \\
& \quad P_t(s,i) > 0, \quad \forall n \in \mathcal{K}(s) \\
& \quad P_t(s,i) = 0, \quad \forall n \notin \mathcal{K}(s)
\end{aligned}$$

which is also a convex optimization problem with respect to  $\{P_t(s,n)\}$  (this can also be easily proved by the positive second-order partial derivative of the objective function with respect to  $\{P_t(s,n)\}$ ). Analogously, the Lagrange function associated with the reformulated problem (SP2) is given by

$$\begin{aligned}
& \mathcal{L}_2(P_t(s,1), P_t(s,2), \dots, P_t(s,N), \lambda) \\
& = \sum_{i \in \mathcal{K}(s)} \exp\left(-\frac{1}{N_0} P_t(s,i) G(i) d_{\min}^2\right) \\
& \quad - \lambda \left( \Psi - \sum_{i \in \mathcal{K}(s)} P_t(s,i) \right).
\end{aligned} \tag{32}$$

We can construct the Lagrange equation set for (32) to be

$$\begin{cases} \frac{\partial \mathcal{L}_2}{\partial \lambda} = - \left( \Psi - \sum_{i \in \mathcal{K}(s)} P_t(s,i) \right) = 0 \\ \frac{\partial \mathcal{L}_2}{\partial P_t(s,n)} = -\frac{G(n) d_{\min}^2}{N_0} \exp\left(-\frac{P_t(s,n) G(n) d_{\min}^2}{N_0}\right) + \lambda = 0, \quad \forall n \in \mathcal{K}(s) \end{cases} \tag{33}$$

For an arbitrary active subcarrier  $n$ ,  $\forall n \in \mathcal{K}(s)$ , according to  $-\frac{G(n) d_{\min}^2}{N_0} \exp\left(-\frac{P_t(s,n) G(n) d_{\min}^2}{N_0}\right) + \lambda = 0$ , we can easily express  $P_t(s,n)$  as a function of  $\lambda$  by

$$P_t(s,n) = -\frac{N_0}{G(n) d_{\min}^2} \log\left(\frac{\lambda N_0}{G(n) d_{\min}^2}\right). \tag{34}$$

Substituting (34) for all active subcarriers,  $\forall n \in \mathcal{K}(s)$ , into  $\left(\Psi - \sum_{i \in \mathcal{K}(s)} P_t(s,i)\right) = 0$  and transforming the resulted expression yields

$$\Psi = -\frac{N_0}{d_{\min}^2} \sum_{i \in \mathcal{K}(s)} \frac{1}{G(i)} \log\left(\frac{\lambda N_0}{G(i) d_{\min}^2}\right). \tag{35}$$

By the fundamental properties of logarithm, we obtain the solution of  $\lambda$  to be

$$\lambda = \exp\left(-\frac{\Psi d_{\min}^2 + N_0 \sum_{i \in \mathcal{K}(s)} \frac{1}{G(i)} \log\left(\frac{N_0}{G(i) d_{\min}^2}\right)}{N_0 \sum_{i \in \mathcal{K}(s)} \frac{1}{G(i)}}\right). \tag{36}$$

Substituting (36) back into (34) gives the solutions of  $\{P_t(s,n)\}$ , the amount of power allocated to active subcarriers. By repeating the procedure  $S$  times, the complete power allocation scheme produced by the simplified optimization objective for all  $S$  SAPs can be obtained.

### C. Tailored Classic Power Allocation Schemes

In order to provide comparison benchmarks, we also borrow two classic power allocation schemes from plain OFDM and tailor them for EHC-OFDM-IM, which are the uniform power allocation and the equalization power allocation. The power allocation outcomes for the  $s$ th SAP produced by the two tailored classic power allocation schemes are given by [34], [56]

$$P_t(s,n) = \begin{cases} \Psi/K, & \forall n \in \mathcal{K}(s) \\ 0, & \forall n \notin \mathcal{K}(s) \end{cases} \tag{37}$$

and

$$P_t(s,n) = \begin{cases} \Psi / \left( G(n) \sum_{i \in \mathcal{K}(s)} \frac{1}{G(i)} \right), & \forall n \in \mathcal{K}(s) \\ 0, & \forall n \notin \mathcal{K}(s) \end{cases} \tag{38}$$

By repeating the above allocation procedure  $S$  times, the complete power allocation schemes for all  $S$  SAPs by the tailored uniform and equalization power allocation criteria can be obtained.

### D. Analysis of Computational Complexity

For comparison, we analyze the computational complexities of all depicted power allocation schemes in this subsection. The proposed sub-optimal power allocation scheme I needs to find  $\lambda^{\&}$  by using the bisection method in Algorithm 3. Provided that Algorithm 3 takes  $N_{\text{iter}}$  to converge, the function  $\Theta(\lambda)$  needs to be calculated by  $N_{\text{iter}}$  times. For each calculation of  $\Theta(\lambda)$ , we need to compute the Lambert  $W$ -function  $W(\epsilon)$   $K$  times. Since the Lambert  $W$ -function can be expanded by infinite series as  $W(\epsilon) = \sum_{n=1}^{\infty} \frac{(-n)^{n-1}}{n!} \epsilon^n$  [57], the computational complexity for calculating  $W(\epsilon)$  can be expressed by the Bachmann-Landau notation as  $\mathcal{O}(q)$  if the  $q$ th order approximation is adopted. Note that, because the complexities vis-à-vis Algorithm 1 and Algorithm 2 and the complexity caused by using  $\lambda^{\&}$  to obtain  $\{P_t(s,i)\}$  are

TABLE III: Comparisons among different power allocation schemes for EHC-OFDM-IM systems.

Schemes	Complexity	CSI requirement
SPA I	$\mathcal{O}(N_{\text{iter}}Kq)$	YES
SPA II	$\mathcal{O}(K)$	YES
UPA	$\mathcal{O}(1)$	NO
EPA	$\mathcal{O}(K)$	YES

much smaller than that of Algorithm 3, we thereby omit them for simplicity in the analysis. Summarizing all above, the complexity of the sub-optimal power allocation scheme I can be expressed as  $\mathcal{O}(N_{\text{iter}}Kq)$ .

On the contrary, the proposed sub-optimal power allocation II neither needs iterations nor the computation of  $W(\epsilon)$ . Therefore, its complexity is only of the order  $\mathcal{O}(K)$ , which proves its simplicity compared to the proposed sub-optimal power allocation scheme I.

For the two tailored classic power allocation schemes, the uniform power allocation requires the computational complexity of the order  $\mathcal{O}(1)$  and the equalization power allocation requires the computational complexity of the order  $\mathcal{O}(K)$ . For clarity, we compare these four power allocation schemes by listing their complexities and requirements of CSI in Table III.

## VI. NUMERICAL RESULTS

### A. Verification of Error Performance Analysis

To be general, we select two typical setups with  $\{N, K\} = \{4, 2\}$  ( $S = 6$ ) and  $\{N, K\} = \{4, 3\}$  ( $S = 4$ ) to simulate in this section. For the case of  $\{N, K\} = \{4, 2\}$ , there are five *unique* architectures of the Huffman trees as shown in Fig. 5(a). For the case of  $\{N, K\} = \{4, 3\}$ , there are two unique architectures of the Huffman trees as shown in<sup>14</sup> Fig. 5(b). We further normalize  $\mu = 1$ ,  $N_0 = 1$  and utilize binary PSK (BPSK,  $M = 2$ ) for data constellation modulation on  $K$  active subcarriers. To verify the error performance analysis when power allocation is independent of channel states, we tentatively adopt the uniform power allocation scheme that regulates  $P_t(s, n) = \Psi/K$  for active subcarriers and is zero otherwise. We also adopt LCD-OFDM-IM proposed in [27], classic OFDM-IM proposed in [8], classic Huffman coded OFDM-IM proposed in [22], and plain OFDM as error performance comparison benchmarks.

The analytical and numerical results regarding average BLER for the cases  $\{N, K\} = \{4, 2\}$  ( $S = 6$ ) and  $\{N, K\} = \{4, 3\}$  ( $S = 4$ ) are illustrated in Fig. 6. From this figure, the effectiveness of the analysis presented in Section IV is substantiated, as the analytical curves well approach the numerical curves at high SNR. The analytical errors existing in the low SNR region are mainly caused by the union bound (c.f. (17)), while the analytical errors existing in the high SNR region are mainly due to the exponential approximation of the

<sup>14</sup>The number of all possible architectures of full binary trees for the case with  $S$  leaves is the Catalan number  $C(S)$ , which increases exponentially when  $S$  increases and makes analysis difficult. However, as the tree architectures with symmetry can be regarded to be equivalent for error performance analysis, we can exclude them for simplicity, but without loss of generality. The approach to select unique tree architectures for an arbitrary number of SAPs has been detailed in [26].

Q-function (c.f. (19)). Meanwhile, by observing the numerical results presented in Fig. 6, one can also gain insight into how the architecture of the Huffman tree affects the error performance. It is clear that the depth of the Huffman tree has a significant impact on the error performance when the SAP-to-leaf mapping relation is optimized based on CSI. In particular, a Huffman tree with a larger depth produces a lower average BLER. Therefore, Huffman trees a-I and b-I (with the largest depths) provide the lowest average BLER for EHC-OFDM-IM systems with  $\{N, K\} = \{4, 2\}$  and  $\{N, K\} = \{4, 3\}$ , respectively.

On the other hand, the above observation does not imply that the Huffman tree with the largest depth is necessarily preferable. In fact, there is a trade-off in the architecture of the Huffman tree between error performance and data transmission rate. To clearly illustrate this trade-off, we plot the encoding rates of index bits associated with different tree architectures for EHC-OFDM-IM systems with  $\{N, K\} = \{4, 2\}$  and  $\{N, K\} = \{4, 3\}$  in Fig. 7. From this figure, we can see that for the case of  $\{N, K\} = \{4, 2\}$ , the Huffman tree a-I that produces the best error performance has the lowest encoding rate of index bits, while the Huffman tree a-V corresponding to the highest average BLER yields the highest encoding rate of index bits. The same trade-off can also be observed between Huffman trees b-I and b-II. To summarize, one should take both error performance and required transmission rate into consideration when choosing the architecture of the Huffman tree for EHC-OFDM-IM systems. As the optimization of the tree architecture is out of the scope of this paper, we omit in-depth discussion here.

In addition, compared to LCD-OFDM-IM and classic OFDM-IM, it is verified that one can adjust the tree architecture of EHC-OFDM-IM to achieve better error performance. However, if the tree architecture is improperly designed, EHC-OFDM-IM could even lead to a higher average BLER than LCD-OFDM-IM (refer to the cases of a-III and a-V). All cases of EHC-OFDM-IM outperform the cases of classic OFDM-IM and plain OFDM without IM.

Furthermore, by the comparison between EHC-OFDM-IM and classic Huffman coded OFDM-IM, we can observe that the proposed EHC-OFDM-IM scheme is capable of harvesting a coding gain and outperforms the classic Huffman coded OFDM-IM scheme when the tree architecture is not perfect. On the other hand, utilizing a perfect tree architecture, i.e., tree architecture b-II, the coding gain brought by EHC-OFDM-IM vanishes, which aligns with our expectation, because all leaves in this case have the same depth and can be regarded as equivalent.

### B. Verification of the Proposed Power Allocation Schemes

To verify and compare the efficiency of the power allocation schemes introduced in Section V, we select EHC-OFDM-IM systems with tree architectures a-I, a-V, b-I, and b-II as examples. Numerical results generated by Monte Carlo simulations are demonstrated in Fig. 8. From this figure, it is clear that with perfect CSI, the CSI-based power allocation schemes outperform the CSI-free uniform power allocation scheme

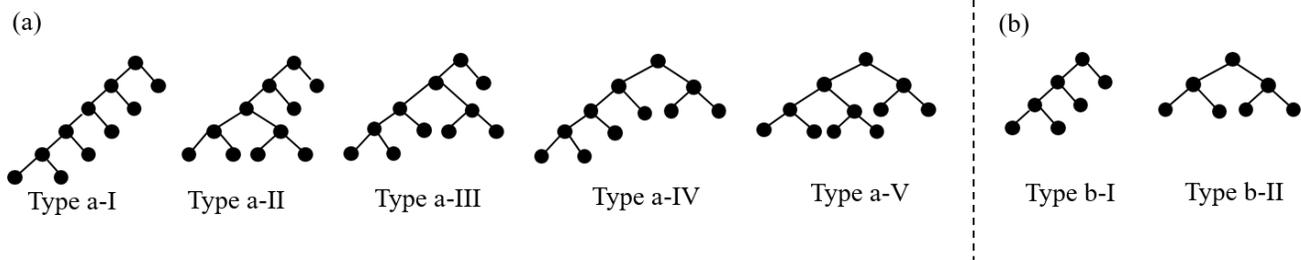


Fig. 5: Unique architectures of the Huffman trees for simulations: (a) Five types for  $\{N, K\} = \{4, 2\}$  ( $S = 6$ ); (b) Two types for  $\{N, K\} = \{4, 3\}$  ( $S = 4$ ).

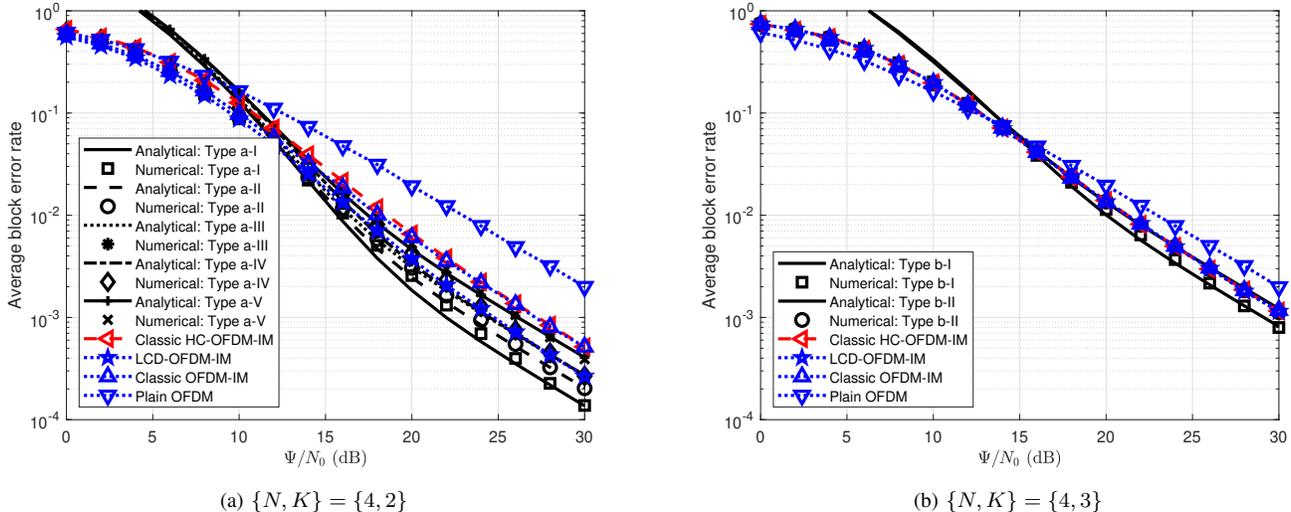


Fig. 6: Average BLER vs.  $\Psi/N_0$  with different tree architectures and different numbers of active subcarriers  $K$ , given  $N = 4$  and  $M = 2$ .

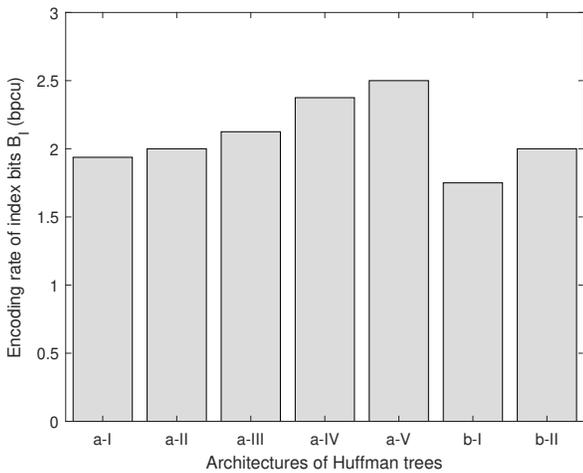


Fig. 7: Encoding rates of index bits associated with different tree architectures.

when  $\Psi$  is large. In addition, the two proposed suboptimal power allocation schemes provide almost the same outage performance, and the suboptimal power allocation scheme I is slightly worse than the suboptimal power allocation scheme II, because the optimization accuracy is lost when performing iterative algorithms. Hence, through these numerical results,

the effectiveness and efficiency of the proposed power allocation schemes with a large power budget are substantiated.

### C. Impact of Imperfect CSI on Error Performance

CSI is the key for implementing EHC-OFDM-IM systems, which is required at both transmitter and receiver for transmission adaptation and received OFDM block estimation. However, in practical situations, CSI might not be perfectly estimated, especially when parallel frequency channels are dynamic. In this subsection, we numerically evaluate the impact of imperfect CSI on error performance. We model the estimated CSI at the transmitter and receiver by the additive errors as follows:  $\mathbf{H}_T = \mathbf{H} + \mathbf{E}_T$  and  $\mathbf{H}_R = \mathbf{H} + \mathbf{E}_R$ , where  $\mathbf{E}_T$  and  $\mathbf{E}_R$  are the diagonal matrices of channel estimation errors at the transmitter and receiver. The diagonal entries of  $\mathbf{E}_T$  and  $\mathbf{E}_R$  follow the complex Gaussian distribution  $\mathcal{CN}(0, \delta_T)$  and  $\mathcal{CN}(0, \delta_R)$ , respectively [58].

Without loss of generality, we investigate the impact of imperfect CSI at the transmitter and receiver on EHC-OFDM-IM systems employing the tree architecture a-I, given  $N = 4$ ,  $K = 2$  and  $M = 2$  for the succinctness of illustration. Meanwhile, the suboptimal power allocation scheme II proposed in Section V-B and the tailored uniform power allocation scheme introduced in Section V-C are selected as examples for illustration. The simulation results are presented in Fig. 9,

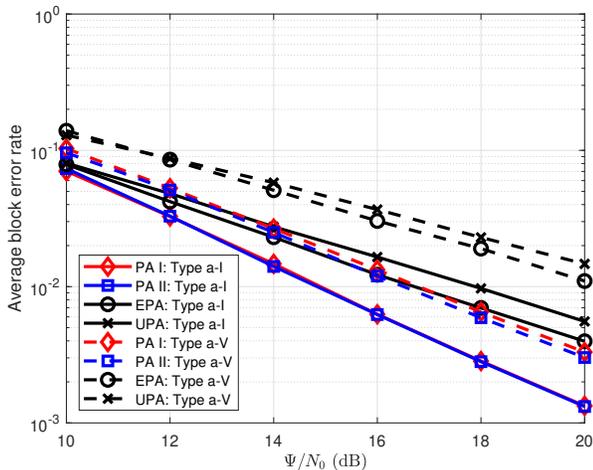
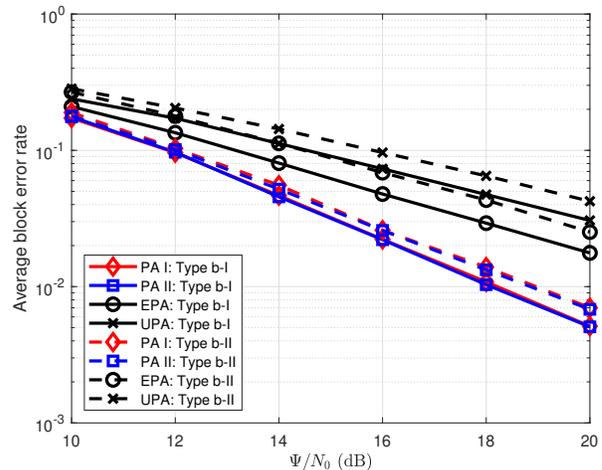
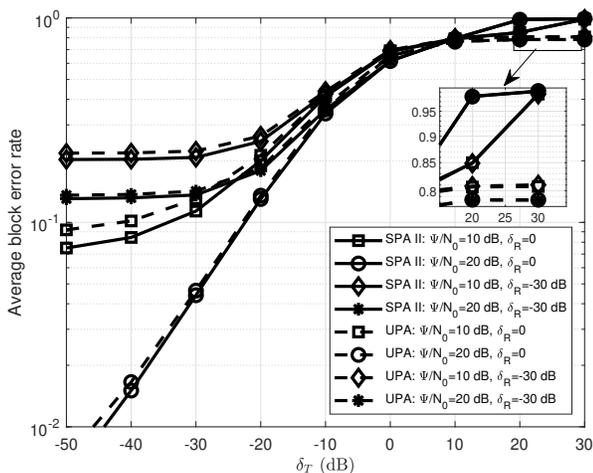
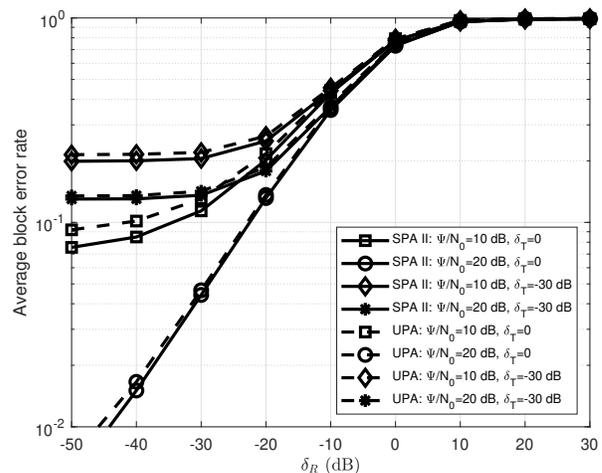
(a)  $\{N, K\} = \{4, 2\}$ (b)  $\{N, K\} = \{4, 3\}$ 

Fig. 8: Average BLER comparisons among different power allocation schemes with different tree architectures and different numbers of active subcarriers  $K$ , given  $N = 4$  and  $M = 2$ .



(a) Imperfect CSI at the transmitter



(b) Imperfect CSI at receiver

Fig. 9: Average BLER vs. the variance of channel estimation errors for imperfect CSI at the transmitter and receiver.

from which we can discover several key impacts of CSI on the error performance of EHC-OFDM-IM systems. First, when the estimation error becomes severe, whether at the transmitter or the receiver, the coding gain brought by the CSI-based power allocation scheme diminishes rapidly, which is caused by the mismatch of codebooks for encoding and decoding at the transmitter and the receiver. As the codebook mismatch is reciprocal, we can see that there exists symmetry between Fig. 9(a) and Fig. 9(b) when  $\delta_T$  and  $\delta_R$  are small.

When  $\delta_T$  and  $\delta_R$  become large, the CSI-free uniform power allocation scheme outperforms the CSI-based optimized power allocation schemes, since the erroneously estimated CSI only provides a misleading decision-making reference and yields a destructive impact on coded systems. Further, as used for detection purposes, the quality of CSI at the receiver dominates the error performance. Therefore, when  $\delta_R$  becomes large, the average BLERs corresponding to all cases converge to

a value close to unity. In other words, when  $\delta_R$  is large, the estimation process at the receiver is equivalent to a random guess, which thereby can hardly provide a correct estimation of the received OFDM block. On the contrary, when  $\delta_T$  is large and the uniform power allocation is adopted, there is still a possibility that the received OFDM block can be correctly decoded as long as the CSI at the receiver is accurate, even though a codebook mismatch exists.

## VII. CONCLUSION

In this paper, we proposed the EHC-OFDM-IM scheme for multi-carrier systems, which is jointly enhanced by Huffman coding, lexicographic codebook design, and a dynamic power allocation strategy. The proposed EHC-OFDM-IM scheme is capable of utilizing all legitimate SAPs, adapting the bijective mapping relation between SAPs and leaves on a given Huffman tree and allocating transmit power to each

active subcarrier by the CSI at the transmitter. In this way, the error performance is improved. To provide insights into the proposed system, we approximated the average BLER in closed form when the power allocation is independent of CSI. Then, we proposed two CSI-based dynamic power allocation schemes to minimize BLER. The first one utilizing the exact expression of the Q-function resulted in a transcendental equation set that requires high computational complexity for obtaining numerical solutions. The second scheme inspired by the exponential approximation of the Q-function is a reduced version of the first one, by which closed-form analytical solutions were provided. Numerical simulations verified our error performance analysis and the effectiveness of the two proposed power allocation schemes and revealed that the depth of the Huffman tree has a significant impact on the error performance when the SAP-to-leaf mapping relation is optimized based on CSI. In addition, because CSI is the key for both transmission adaptation and signal estimation, we numerically investigated the impacts of imperfect CSI at the transmitter and receiver on the error performance of EHC-OFDM-IM systems. Overall, this work provides an important step in OFDM-IM research.

#### ACKNOWLEDGMENT

We thank the editor and the anonymous reviewers for the constructive comments, which have helped us improve the quality of the paper.

#### REFERENCES

- [1] Q. Li, M. Wen, H. V. Poor, and F. Chen, "Information guided precoding for OFDM," *IEEE Access*, vol. 5, pp. 19 644–19 656, 2017.
- [2] 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.
- [3] 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*, vol. 20, no. 3, pp. 1905–1938, third quarter 2018.
- [4] T. Mao, Q. Wang, Z. Wang, and S. Chen, "Novel index modulation techniques: A survey," *IEEE Communications Surveys Tutorials*, vol. 21, no. 1, pp. 315–348, Firstquarter 2019.
- [5] 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.
- [6] Z. Wang, S. Dang, S. Shaham, Z. Zhang, and Z. Lv, "Basic research methodology in wireless communications: The first course for research-based graduate students," *IEEE Access*, vol. 7, pp. 86 678–86 696, 2019.
- [7] 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.
- [8] E. Basar, U. Aygolu, E. Panayirci, 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.
- [9] S. Gokceli, E. Basar, M. Wen, and G. K. Kurt, "Practical implementation of index modulation-based waveforms," *IEEE Access*, vol. 5, pp. 25 463–25 473, 2017.
- [10] 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.
- [11] N. Ishikawa, S. Sugiura, and L. Hanzo, "Subcarrier-index modulation aided OFDM - will it work?" *IEEE Access*, vol. 4, pp. 2580–2593, 2016.
- [12] R. Abu-alhiga and H. Haas, "Subcarrier-index modulation OFDM," in *Proc. IEEE PIMRC*, Tokyo, Japan, Sept. 2009, pp. 177–181.
- [13] D. Tsonev, S. Sinanovic, and H. Haas, "Enhanced subcarrier index modulation (SIM) OFDM," in *Proc. IEEE Globecom*, Houston, TX, USA, Dec. 2011, pp. 728–732.
- [14] 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.
- [15] 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*, pp. 1–1, 2018.
- [16] 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.
- [17] J. Zhao, S. Dang, and Z. Wang, "Full-duplex relay-assisted orthogonal frequency-division multiplexing with index modulation," *IEEE Systems Journal*, pp. 1–12, 2018.
- [18] S. Dang, J. Li, M. Wen, and S. Mumtaz, "Distributed processing for multi-relay assisted OFDM with index modulation," *IEEE Transactions on Wireless Communications*, 2019.
- [19] D. A. Huffman, "A method for the construction of minimum-redundancy codes," *Proceedings of the IRE*, vol. 40, no. 9, pp. 1098–1101, Sept. 1952.
- [20] R. Y. Chang, S. Lin, and W. Chung, "Energy efficient transmission over space shift keying modulated MIMO channels," *IEEE Transactions on Communications*, vol. 60, no. 10, pp. 2950–2959, Oct. 2012.
- [21] W. Wang and W. Zhang, "Huffman coding-based adaptive spatial modulation," *IEEE Transactions on Wireless Communications*, vol. 16, no. 8, pp. 5090–5101, Aug. 2017.
- [22] A. I. Siddiq, "Low complexity OFDM-IM detector by encoding all possible subcarrier activation patterns," *IEEE Communications Letters*, vol. 20, no. 3, pp. 446–449, Mar. 2016.
- [23] M. I. Kadir, H. Zhang, S. Chen, and L. Hanzo, "Entropy coding aided adaptive subcarrier-index modulated OFDM," *IEEE Access*, vol. 6, pp. 7739–7752, 2018.
- [24] Y. Liu and J. P. Coon, "Mitigating bit-synchronization errors in Huffman-coding-aided index modulation," *IEEE Communications Letters*, vol. 23, no. 3, pp. 426–429, Mar. 2019.
- [25] Z. Hu, J. Liu, and F. Chen, "On the mutual information and Huffman coding for OFDM-IM," in *Proc. IEEE IMCEC*, May 2018, pp. 298–302.
- [26] J. Coon, M. Badiu, Y. Liu, F. Yarkin, and S. Dang, "Binary-tree encoding for uniform binary sources in index modulation systems," *IEEE Journal of Selected Topics in Signal Processing*, pp. 1–1, 2019.
- [27] S. Dang, G. Chen, and J. P. Coon, "Lexicographic codebook design for OFDM with index modulation," *IEEE Transactions on Wireless Communications*, 2018.
- [28] N. Ishikawa, "IMToolkit: An open-source index modulation toolkit for reproducible research based on massively parallel algorithms," *IEEE Access*, vol. 7, pp. 93 830–93 846, 2019.
- [29] 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, pp. 1–6.
- [30] A. M. Jaradat, J. M. Hamamreh, and H. Arslan, "OFDM with subcarrier number modulation," *IEEE Wireless Communications Letters*, vol. 7, no. 6, pp. 914–917, Dec. 2018.
- [31] S. Dang, G. Ma, B. Shihada, and M.-S. Alouini, "Enhanced orthogonal frequency-division multiplexing with subcarrier number modulation," *IEEE Internet of Things Journal*, pp. 1–1, 2019.
- [32] S. Hong, M. Sagong, C. Lim, K. Cheun, and S. Cho, "FQAM : A modulation scheme for beyond 4G cellular wireless communication systems," in *Proc. IEEE Globecom*, Atlanta, GA, USA, Dec. 2013, pp. 25–30.
- [33] Y. Wang, W. Liu, M. Jin, S. Jang, and J. M. Kim, "FQAM/FPSK modulation for spatial modulation systems," in *Proc. IEEE ICTC*, Jeju, South Korea, Oct. 2016, pp. 511–515.
- [34] L. Goldfeld, V. Lyandres, and D. Wulich, "Minimum BER power loading for OFDM in fading channel," *IEEE Transactions on Communications*, vol. 50, no. 11, pp. 1729–1733, Nov. 2002.
- [35] D. J. Love and R. W. Heath, "OFDM power loading using limited feedback," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 5, pp. 1773–1780, Sept. 2005.
- [36] P. Zhang and H. Yang, "Minimum-BER power allocation for multicarrier systems with outdated channel state information," in *Proc. IEEE CCECE*, Vancouver, BC, Canada, Apr. 2007, pp. 180–183.
- [37] S. Sadr, A. Anpalagan, and K. Raahemifar, "Radio resource allocation algorithms for the downlink of multiuser OFDM communication systems," *IEEE Communications Surveys Tutorials*, vol. 11, no. 3, pp. 92–106, third quarter 2009.
- [38] R. O. Afolabi, A. Dadlani, and K. Kim, "Multicast scheduling and resource allocation algorithms for OFDMA-based systems: A survey," *IEEE Communications Surveys Tutorials*, vol. 15, no. 1, pp. 240–254, First quarter 2013.

- [39] Q. Ma, P. Yang, P. Wang, L. Peng, X. He, B. Fu, and Y. Xiao, "Power allocation for OFDM with index modulation," in *Proc. IEEE VTC Spring*, Sydney, NSW, Australia, June 2017, pp. 1–5.
- [40] X. Liu, L. Dan, Q. Ma, Y. Xiao, and X. He, "Profile-based power allocation in OFDM with index modulation," in *Proc. IEEE VTC Spring*, Sydney, NSW, Australia, June 2017, pp. 1–5.
- [41] X. Zhang, H. Bie, Q. Ye, C. Lei, and X. Tang, "Dual-mode index modulation aided OFDM with constellation power allocation and low-complexity detector design," *IEEE Access*, vol. 5, pp. 23 871–23 880, 2017.
- [42] S. Lu, I. A. Hemadeh, M. El-Hajjar, and L. Hanzo, "Compressed-sensing-aided space-time frequency index modulation," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 7, pp. 6259–6271, July 2018.
- [43] T. Mao, Z. Wang, Q. Wang, S. Chen, and L. Hanzo, "Dual-mode index modulation aided OFDM," *IEEE Access*, vol. 5, pp. 50–60, 2017.
- [44] 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.
- [45] P. Yang, Y. Xiao, S. Li, and L. Hanzo, "A low-complexity power allocation algorithm for multiple-input–multiple-output spatial modulation systems," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 3, pp. 1819–1825, Mar. 2016.
- [46] Y. Xiao, S. Wang, L. Dan, X. Lei, P. Yang, and W. Xiang, "OFDM with interleaved subcarrier-index modulation," *IEEE Communications Letters*, vol. 18, no. 8, pp. 1447–1450, Aug. 2014.
- [47] S. Dang, G. Chen, and J. P. Coon, "Outage performance analysis of full-duplex relay-assisted device-to-device systems in uplink cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 5, pp. 4506–4510, May 2017.
- [48] E.-H. Yang and D.-K. He, "Huffman coding," *Wiley Encyclopedia of Telecommunications*, 2003.
- [49] G. A. V. Pai, *Data Structures*. McGraw-Hill Education (India) PVT Limited, 2008.
- [50] S. Dang, G. Ma, B. Shihada, and M.-S. Alouini, "A novel error performance analysis methodology for OFDM-IM," *IEEE Wireless Communications Letters*, 2019.
- [51] M. Chiani and D. Dardari, "Improved exponential bounds and approximation for the Q-function with application to average error probability computation," in *Proc. IEEE Globecom*, vol. 2, Taipei, Taiwan, ROC, Nov. 2002, pp. 1399–1402 vol.2.
- [52] H. David and H. Nagaraja, *Order Statistics*, ser. Wiley Series in Probability and Statistics. Wiley, 2004.
- [53] W. Yang and Y. Cai, "On the performance of the block-based selective OFDM decode-and-forward relaying scheme for 4G mobile communication systems," *Journal of Communications and Networks*, vol. 13, no. 1, pp. 56–62, Feb. 2011.
- [54] G. Folland, *Real Analysis: Modern Techniques and Their Applications*, ser. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 2013.
- [55] A. Gomes, I. Voiculescu, J. Jorge, B. Wyvill, and C. Galbraith, *Implicit Curves and Surfaces: Mathematics, Data Structures and Algorithms*. Springer London, 2009.
- [56] J. A. C. Bingham, "Multicarrier modulation for data transmission: an idea whose time has come," *IEEE Communications Magazine*, vol. 28, no. 5, pp. 5–14, May 1990.
- [57] R. M. Corless, D. J. Jeffrey, and D. E. Knuth, "A sequence of series for the Lambert W function," in *Proc. ISSAC*. ACM, 1997, pp. 197–204.
- [58] S. Guo, H. Zhang, P. Zhang, S. Dang, C. Liang, and M.-S. Alouini, "Signal shaping for generalized spatial modulation and generalized quadrature spatial modulation," *IEEE Transactions on Wireless Communications*, vol. 18, no. 8, pp. 4047–4059, Aug. 2019.