

Set Partition Modulation

Ferhat Yarkin *Student Member, IEEE* and Justin P. Coon *Senior Member, IEEE*

Abstract

In this paper, a novel modulation scheme called set partition modulation (SPM) is proposed. In this scheme, set partitioning and ordered subsets in the set partitions are used to form codewords. We define different SPM variants and depict a practical model for using SPM with orthogonal frequency division multiplexing (OFDM). For the OFDM-SPM schemes, different constellations are used to distinguish between different subsets in a set partition. To achieve good distance properties as well as better error performance for the OFDM-SPM codewords, we define a codebook selection problem and formulate such a problem as a clique problem in graph theory. In this regard, we propose a fast and efficient codebook selection algorithm. We analyze error and achievable rate performance of the proposed schemes and provide asymptotic results for the performance. It is shown that the proposed SPM variants are general schemes, which encompass multi-mode OFDM with index modulation (MM-OFDM-IM) and dual-mode OFDM with index modulation (DM-OFDM-IM) as special cases. It is also shown that OFDM-SPM schemes are capable of exhibiting better error performance and improved achievable rate than conventional OFDM, OFDM-IM, DM-OFDM-IM, and MM-OFDM-IM.

Index Terms

Orthogonal frequency division multiplexing (OFDM), index modulation, set partitions.

I. INTRODUCTION

The idea of embedding information in the element permutations of a codeword was first proposed by Slepian in [2]. Slepian's idea, which he called *permutation modulation* (PM), was hinged upon constructing a codebook by permuting elements of a codeword. *Index modulation* (IM) techniques, which can be considered as a subclass of PM, have attracted remarkable interest

F. Yarkin and J. P. Coon are with the Department of Engineering Science, University of Oxford, Parks Road, Oxford, OX1 3PJ, U.K. E-mail: {ferhat.yarkin and justin.coon}@eng.ox.ac.uk

This paper was presented in part at the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), 8-11 September 2019 [1].

The authors acknowledge the support of the Bristol Innovation & Research Laboratory of Toshiba Research Europe Ltd.

due to their capabilities for achieving better error performance and improved energy/spectral efficiency compared to conventional systems [3]. IM encodes the information in the indices of active/inactive sources. For example, combinations of the (in)active transmit antenna indices form the IM codewords in a multi-antenna communication technique called *spatial modulation* (SM) [4]. In contrast, the activation patterns of the *subcarriers* are used to construct IM codewords in a multicarrier communication technique called orthogonal frequency-division multiplexing with index modulation (OFDM-IM) [5]–[7]. Other PM/IM-based applications exist with manifestations in space, time and frequency (see, e.g., [8]).

The application of IM to the well-known OFDM structure brings various advantages [5]–[7], [9]–[17]. For example, it was shown by [5]–[7] that OFDM-IM is capable of achieving substantially better error performance than conventional OFDM. To further improve the error performance, transmit diversity and trellis-coded modulation were applied to OFDM-IM in [9] and [10], respectively. These studies show how one can achieve a diversity gain for the OFDM-IM scheme by sacrificing the data rate. In [11], a coordinate interleaved OFDM-IM scheme was proposed, and it was shown that the proposed scheme can achieve an additional diversity gain without sacrificing the data rate. In [12], two different generalized OFDM-IM schemes were proposed to improve the spectral efficiency and error performance. Moreover, it was revealed by [13], [14] that OFDM-IM is capable of achieving a better achievable rate than conventional OFDM for small modulation orders. A useful guideline to design spectrally efficient OFDM-IM schemes is reported in [15]. In [16], a binary tree encoding method was applied to OFDM-IM to cover all of the subcarrier activation patterns and, therefore, improve the data rate beyond recent benchmarks. For the same purpose, a discrete cosine transform based solution for OFDM-IM was implemented in [17]; substantial data rate improvements compared to OFDM-IM and conventional OFDM were shown possible.

To encode information in the combinations of the subcarriers, a certain number of subcarriers are nulled in OFDM-IM. Although carrying information on the combinations of active/inactive subcarriers results in better error performance and improved spectral/energy efficiency for small modulation orders, it becomes difficult to achieve spectral efficiencies comparable to conventional OFDM for high modulation orders. To overcome this problem, the idea of employing distinguishable constellations on different subcarriers rather than nulling them has been considered [18]–[26]. In [18], an OFDM scheme called *dual mode OFDM with IM* (DM-OFDM-IM) was documented; this method uses two distinguishable constellations rather than active/inactive

subcarriers to encode information. In [19], a dual-mode index modulation aided OFDM scheme that employs two PSK constellations with different power levels is proposed. The authors of [20] further considered altering the number of subcarriers modulated by the same constellation and provided a more general IM scheme called *Generalized DM-OFDM-IM* (GDM-OFDM-IM). In [21], the authors proposed two different precoding-aided OFDM-IM schemes that use distinguishable constellations. The first scheme of [21] can be considered as a generalization of DM-OFDM-IM since such a scheme partitions the subcarriers into groups and uses the same amount of constellations as the number of groups to modulate the subcarriers. In [22], in addition to two distinguishable constellations, some subcarriers are allowed to remain unused, which yields a third mode that, as it turns out, enhances the bit-error-rate (BER) performance relative to DM-OFDM-IM and OFDM-IM. In [23], Wen et al. proposed a *multi-mode OFDM-IM* (MM-OFDM-IM) scheme, which uses distinguishable constellations on each subcarrier of OFDM sub-blocks to increase the spectral efficiency as well as to improve BER performance. In [24] and [25], the MM-OFDM-IM scheme is generalized. In [26], the MM-OFDM-IM scheme is extended to space and time domains by using Latin matrices.

Against this background, we develop a new codebook design method, which we call *set partition modulation* (SPM). Our novel contributions can be listed as follows:

- The proposed method uses a novel combinatorial tool, which is partitions of codeword elements rather than permutations or combinations of such elements. By using the set partitioning, we aim to design a more general codebook that can achieve a better error performance as well as a higher data rate than conventional OFDM and OFDM-IM schemes.
- We define different variants of SPM and give a practical model for applying SPM with OFDM transmissions.
- The proposed OFDM-SPM schemes employ distinguishable constellations, similar to MM-OFDM-IM; however, in OFDM-SPM, the different constellations are used to distinguish between subsets in a set partition, unlike MM-OFDM-IM where distinguishable constellations are used to construct permutations.
- The proposed OFDM-SPM variants are capable of encompassing DM-OFDM-IM and MM-OFDM-IM schemes as special cases. Moreover, they are also capable of achieving a better data rate than such schemes.
- We define a codebook selection problem for OFDM-SPM variants to design efficient codebooks, which are at least as good as the OFDM-IM schemes. We formulate such a problem

as a clique problem and develop an efficient solution for such a problem.

- The achievable data rate and BER of OFDM-SPM variants are investigated in this paper, and an upper-bound on the BER is obtained. We also provide asymptotic expressions for data rate and BER.
- Our analytical findings show that OFDM-SPM variants can outperform conventional OFDM, OFDM-IM, DM-OFDM-IM, and MM-OFDM-IM in terms of data rate and BER.

The rest of the paper is organized as follows. In Section II, we define SPM and its variants. OFDM-SPM and an efficient codebook selection algorithm are described in Sections III and IV, respectively. Performance analysis is undertaken in Section V. We present and compare analytical and numerical results in Section VI. Finally, we conclude the paper in Section VII.

II. SET PARTITION MODULATION

In this section, we describe the basic idea of SPM. We begin with some useful definitions.

Definition 1. Set Partition: A set partition is the grouping the elements of a set in a way that the groups are disjoint and the union of the groups gives the set. The set $\{x_1, x_2, x_3\}$, for example, has the following five partitions $\{\{x_1, x_2, x_3\}\}$, $\{\{x_1\}, \{x_2, x_3\}\}$, $\{\{x_2\}, \{x_1, x_3\}\}$, $\{\{x_3\}, \{x_1, x_2\}\}$ and $\{\{x_1\}, \{x_2\}, \{x_3\}\}$.

Definition 2. Stirling Number of the Second Kind: The Stirling number of the second kind, denoted by $\left\{ \begin{smallmatrix} N \\ K \end{smallmatrix} \right\}$, can be defined as the number of ways to partition an N -element set into K non-empty subsets. For example, the set $\{x_1, x_2, x_3\}$ has the following three set partitions with two non-empty subsets $\{\{x_1\}, \{x_2, x_3\}\}$, $\{\{x_2\}, \{x_1, x_3\}\}$ and $\{\{x_3\}, \{x_1, x_2\}\}$, thus $\left\{ \begin{smallmatrix} 3 \\ 2 \end{smallmatrix} \right\} = 3$.

Definition 3. Bell Number: The Bell number B_N enumerates the total number of partitions of a set of N elements. The Bell number is related to Stirling numbers of the second kind through the equation $B_N = \sum_{K=1}^N \left\{ \begin{smallmatrix} N \\ K \end{smallmatrix} \right\}$. Since a three-element set has five partitions, $B_3 = 5$.

Definition 4. Ordered Bell Number: The ordered Bell number \check{B}_N enumerates the total number of partitions of an N -element set considering all permutations of subsets for each partition. The ordered Bell number satisfies the equation $\check{B}_N = \sum_{K=1}^N K! \left\{ \begin{smallmatrix} N \\ K \end{smallmatrix} \right\}$. A three-element set has one set partition with one subset, three set partitions with two subsets and one set partitions with three subsets. Hence, the ordered Bell number, \check{B}_3 , can be calculated as $\check{B}_3 = 1!(1) + 2!(3) + 3!(1) = 13$.

A. SPM

In an SPM system, a codebook of L codewords $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_L$ is constructed such that each codeword is a sequence of N elements, which are drawn from a constellation diagram in the complex plane, i.e., $\mathbf{x}_l = \{x_{l1}, x_{l2}, \dots, x_{lN}\}$, $l = 1, \dots, L$, where $x_{ln} \in \mathbb{C}$, $n \in \{1, \dots, N\}$, is an M -ary symbol. Each codeword is mapped to a partition of an N -element set \mathcal{X} into K subsets. Since the number of ways one can form the partition is $\{N \atop K\}$, the SPM codebook size is given by $L_{SPM} = \{N \atop K\}$. We call \mathcal{X} the generator set of an SPM codeword. To produce unique codewords for SPM, an N -element codeword should have at least K distinguishable elements, and each element, $\mu_k, k = 1, \dots, K$, is used to specify which element in the codewords belongs to which subset. Therefore, distinguishable elements differentiate each subset from the other(s).

B. Ordered SPM

Ordered SPM (OSPM) is an extended version of SPM in which the codebook size is increased by considering permutations of the subsets, thus the codebook size is given by $L_{OSPM} = K! \{N \atop K\}$.

C. Full SPM

In Full SPM (FSPM), the codewords $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{L_{FSPM}}$ are generated by partitioning an N -element set \mathcal{X} into non-empty disjoint subsets in such a way that the number of these subsets takes any possible value, $K \in \{1, \dots, N\}$. In other words, all partitions of an N -element set \mathcal{X} into non-empty disjoint subsets are used to form the FSPM codebook. For FSPM, the codebook size is equal to the Bell number B_N , i.e., $L_{FSPM} = B_N = \sum_{K=1}^N \{N \atop K\}$.

D. Ordered Full SPM

We can further increase the number of codewords in an FSPM codebook by considering the permutations of the subsets in a partition. In this regard, we define ordered full SPM (OFSPM) as a modulation scheme that forms its codebook by using all partitions of an N -element set \mathcal{X} along with all permutations of the subsets in each partition. Hence, the OFSPM codebook size is given by the ordered Bell number \check{B}_N , i.e., $L_{OFSPM} = \check{B}_N = \sum_{K=1}^N K! \{N \atop K\}$.

Example 1. In Table I, we list unordered and ordered set partitions of a three-element set $\mathcal{X} := \{x_1, x_2, x_3\}$ to give an example of a mapping between unordered/ordered set partitions and the variants of SPM for $N = 3$. For SPM and OSPM, we set $K = 2$. As seen from the table, to

Table I: Set Partitioning and Codeword Generation Example For SPM Variants When $N = 3$.

	SPM	OSPM	FSPM	OFSPM
$\mathcal{S}_1 = \{\{x_1, x_2, x_3\}\}$	-	-	$\mathbf{x}_1 = \{\mu_1, \mu_1, \mu_1\}$	$\mathbf{x}_1 = \{\mu_1, \mu_1, \mu_1\}$
$\mathcal{S}_2 = \{\{x_1\}, \{x_2, x_3\}\}$	$\mathbf{x}_1 = \{\mu_1, \mu_2, \mu_2\}$	$\mathbf{x}_1 = \{\mu_1, \mu_2, \mu_2\}$	$\mathbf{x}_2 = \{\mu_1, \mu_2, \mu_2\}$	$\mathbf{x}_2 = \{\mu_1, \mu_2, \mu_2\}$
$\mathcal{S}_3 = \{\{x_2\}, \{x_1, x_3\}\}$	$\mathbf{x}_2 = \{\mu_2, \mu_1, \mu_2\}$	$\mathbf{x}_2 = \{\mu_2, \mu_1, \mu_2\}$	$\mathbf{x}_3 = \{\mu_2, \mu_1, \mu_2\}$	$\mathbf{x}_3 = \{\mu_2, \mu_1, \mu_2\}$
$\mathcal{S}_4 = \{\{x_3\}, \{x_1, x_2\}\}$	$\mathbf{x}_3 = \{\mu_2, \mu_2, \mu_1\}$	$\mathbf{x}_3 = \{\mu_2, \mu_2, \mu_1\}$	$\mathbf{x}_4 = \{\mu_2, \mu_2, \mu_1\}$	$\mathbf{x}_4 = \{\mu_2, \mu_2, \mu_1\}$
$\mathcal{S}_5 = \{\{x_2, x_3\}, \{x_1\}\}$	-	$\mathbf{x}_4 = \{\mu_2, \mu_1, \mu_1\}$	-	$\mathbf{x}_5 = \{\mu_2, \mu_1, \mu_1\}$
$\mathcal{S}_6 = \{\{x_1, x_3\}, \{x_2\}\}$	-	$\mathbf{x}_5 = \{\mu_1, \mu_2, \mu_1\}$	-	$\mathbf{x}_6 = \{\mu_1, \mu_2, \mu_1\}$
$\mathcal{S}_7 = \{\{x_1, x_2\}, \{x_3\}\}$	-	$\mathbf{x}_6 = \{\mu_1, \mu_1, \mu_2\}$	-	$\mathbf{x}_7 = \{\mu_1, \mu_1, \mu_2\}$
$\mathcal{S}_8 = \{\{x_1\}, \{x_2\}, \{x_3\}\}$	-	-	$\mathbf{x}_5 = \{\mu_1, \mu_2, \mu_3\}$	$\mathbf{x}_8 = \{\mu_1, \mu_2, \mu_3\}$
$\mathcal{S}_9 = \{\{x_1\}, \{x_3\}, \{x_2\}\}$	-	-	-	$\mathbf{x}_9 = \{\mu_1, \mu_3, \mu_2\}$
$\mathcal{S}_{10} = \{\{x_2\}, \{x_1\}, \{x_3\}\}$	-	-	-	$\mathbf{x}_{10} = \{\mu_2, \mu_1, \mu_3\}$
$\mathcal{S}_{11} = \{\{x_2\}, \{x_3\}, \{x_1\}\}$	-	-	-	$\mathbf{x}_{11} = \{\mu_3, \mu_1, \mu_2\}$
$\mathcal{S}_{12} = \{\{x_3\}, \{x_1\}, \{x_2\}\}$	-	-	-	$\mathbf{x}_{12} = \{\mu_2, \mu_3, \mu_1\}$
$\mathcal{S}_{13} = \{\{x_3\}, \{x_2\}, \{x_1\}\}$	-	-	-	$\mathbf{x}_{13} = \{\mu_3, \mu_2, \mu_1\}$

obtain the codewords in an SPM codebook, we first partition the elements of the generator set \mathcal{X} into subsets, and then permute the orders of the subsets if we are constructing the codewords for one of the ordered counterparts of SPM, i.e., OSPM or OFSPM. The unordered and ordered set partitions, $\mathcal{S}_j, j = 1, \dots, 13$, are used to form SPM codewords, $\mathbf{x}_i, i = 1, \dots, L$ where $L = 3, 6, 5$ and 13 for SPM, OSPM, FSPM and OFSPM, respectively, since $\binom{3}{2} = 3, 2!\binom{3}{2} = 6, B_3 = 5$ and $\check{B}_3 = 13$. Then, we use the subset identifier $\mu_k, k \in \{1, \dots, N\}$, to represent the elements that belong to the k th subset. For example, for the first codeword of SPM in Table I, we have the partition $\mathcal{S}_2 := \{\{x_1\}, \{x_2, x_3\}\}$. Since the element x_1 and the elements x_2 and x_3 are in the first and second subsets, respectively, we assign μ_1 to first element and μ_2 to remaining elements. For OSPM, the first three codewords are SPM codewords and we further obtain an extended codebook by taking the permutations of the subsets in a set partition into account. In FSPM, we consider all set partitions without permuting the orders of the subsets. However, in OFSPM, we attain a considerably extended codebook by considering all set partitions as well as the permutations of the subsets in those of the partitions.

III. PRACTICAL MODEL FOR OFDM

We present a practical system model in which we apply SPM schemes to OFDM transmissions. The transmitter structure of the OFDM-SPM scheme is shown in Fig. 1. In this scheme, m input

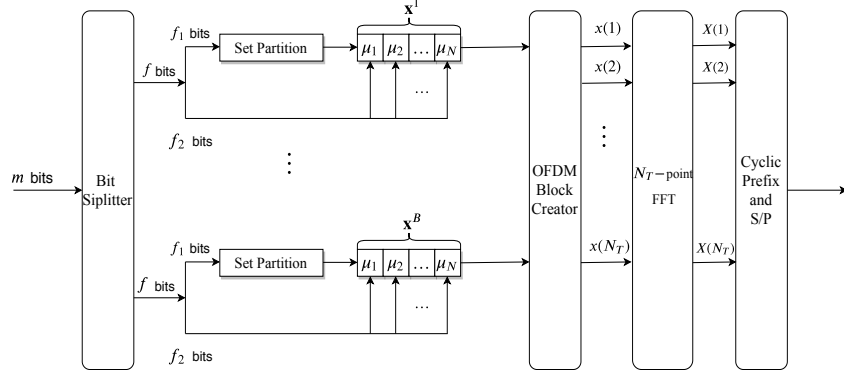


Figure 1: Transmitter structure of OFDM-SPM scheme.

bits enter the SPM transmitter, and these bits are divided into $B = m/f$ blocks, each having f input bits. Similarly, the total number of subcarriers N_T is also divided into $B = N_T/N$ blocks, each having N subcarriers. For each block of input bits, f information bits are modulated by an SPM encoder and the resulting modulated symbols are carried by N subcarriers.

Since each bit and each subcarrier block have the same mapping operation, we focus on a single block, the b th block (where $b \in \{1, 2, \dots, B\}$), in what follows. In the b th block, the SPM encoder further divides f information bits into two parts, one of them having f_1 bits and the other one having f_2 bits with $f_1 + f_2 = f$. The first f_1 bits are used to determine the specific set partition \mathcal{S}_i^b , $i = 1, \dots, L$ ($L \in \{L_{SPM}, L_{OSPM}, L_{FSPM}, L_{OFSPM}\}$), of the N -element generator set $\mathcal{X} := \{x_1, x_2, \dots, x_N\}$ belonging to one of the variants of SPM defined above. The chosen partition is mapped to the corresponding SPM codeword \mathbf{x}_i^b where the superscript b stands for the b th block. Here, each element in the SPM codeword corresponds to a subcarrier in the b th block. The remaining f_2 bits are used to modulate symbols on the N subcarriers, considering the corresponding mapping of the set partition determined by the first f_1 bits. As discussed in the previous subsection, we use different subset identifiers in order to produce unique SPM codewords. To preserve the uniqueness property of SPM codewords and modulate the symbols on each subcarrier, we further assume that each subset identifier μ_k is an element of a disjoint M -ary signal constellation \mathcal{M}_k , i.e., $\mu_k \in \mathcal{M}_k$ and $\mathcal{M}_k \cap \mathcal{M}_{\hat{k}} = \emptyset$, where $k, \hat{k} \in \{1, 2, \dots, N\}$

and $k \neq \hat{k}^1$. For convenience, we choose the size of each constellation as M and, therefore, $f_2 = N \log_2 M$. By following the useful design guidelines in [23], we obtain the distinguishable PSK constellations \mathcal{M}_k by rotating each constellation with the angle of $2(k-1)\pi/(MN)$, $k = 1, \dots, N$, to maximize the distance between constellation points. To obtain distinguishable QAM constellations, likewise [23], we employ the well-known set partitioning technique in [27].

The mapping of f_1 bits to the set partitions can be performed by using a look-up table in a similar manner as was proposed to map information bits to subcarrier activation patterns in [5], or to permutation indices as detailed in [23]. Once the mapping between f_1 bits and set partitions has been completed, $f_2 = N \log_2 M$ bits are used to determine the modulated symbols, or in other words subset identifiers, on each subcarrier. Hence, one of the SPM codewords $\mathbf{x}^b \in \{\mathbf{x}_1^b, \dots, \mathbf{x}_L^b\}$ along with the corresponding modulation symbols $\{\mu_k\}$ constitutes the symbol vector of the b th block. After obtaining symbol vectors for all blocks, an OFDM block creator forms the overall symbol vector $\mathbf{x} := [x(1), x(2), \dots, x(N_T)]^T = [\mathbf{x}^1, \dots, \mathbf{x}^b, \dots, \mathbf{x}^B]^T \in \mathcal{C}^{N_T \times 1}$. Here, we assume that each element of \mathbf{x} is distributed among equally spaced subcarriers to ensure diversity in frequency, and each modulated symbol carried by a subcarrier has unit energy, i.e., $E[|x(t)|^2] = 1$, $t = 1, \dots, N_T$. After this point, exactly the same operations as conventional OFDM are applied. The symbol vector is processed with an N_T -point IFFT, and a cyclic prefix of sufficient length, which is not lower than the memory of the discrete channel impulse response, is attached to the beginning of each time-domain symbol vector. After parallel-to-serial and up-conversion, transmission is operated over a frequency-selective Rayleigh fading channel.

At the receiver, the received signal is down converted and the cyclic prefix is then removed from each received baseband symbol vector before processing with an FFT. After employing an N_T -point FFT operation, the frequency domain received signal vector can be written as

$$\mathbf{y} := [y(1), y(2), \dots, y(N_T)]^T = \sqrt{E_S} \mathbf{X} \mathbf{h} + \mathbf{n} \quad (1)$$

where E_S is the energy of the transmitted symbol vector and $\mathbf{X} = \text{diag}(\mathbf{x})$. Moreover, \mathbf{h} and \mathbf{n} are $N_T \times 1$ channel and noise vectors, respectively. Elements of these vectors follow the

¹Note that the subset identifiers are not necessarily elements of disjoint M -ary signal constellations and each of them may be chosen as a single constellation point in the same M -ary constellation. For example, four constellation points in a single QPSK modulation can be used to specify the subsets of set partitions regarding a four-element set. Hence, assigning a unique constellation point to each subset identifier would be enough to constitute an SPM scheme. However, in this special case, the number of information bits transmitted by an OFDM block is decreased by f_2 bits since f_2 bits are not used to modulate the subset identifiers. Note also that this special case is not shown in Fig. 1.

Algorithm 1: Low-complexity maximum likelihood (LC-ML) detection of OFDM-SPM

Step 1: Initialize $1 \times N$ decoding vector, \mathbf{d} , of all zeros and $L \times N$ index matrix \mathbf{Z} ;

Step 2: Sort the channel gains of N subcarriers in descending order, i.e.,

$$|h(\lambda_1)|^2 \geq \dots \geq |h(\lambda_N)|^2 \text{ where } \lambda_n \in \{1, \dots, N\};$$

Step 3: LC-ML detection;

for $n = 1$ **to** N **do**

$$\mathcal{I}_n := \text{unique}(\mathbf{Z}(:, \lambda_n));$$

$$\mathcal{M} := \cup_{i \in \mathcal{I}_n} \mathcal{M}_i;$$

$$\hat{x}_{\lambda_n} = \arg \min_{x \in \mathcal{M}} |y(\lambda_n) - \sqrt{E_S} x h(\lambda_n)|^2, \hat{I}_{\lambda_n} = \arg_{i \in \mathcal{I}_n} \hat{x}_{\lambda_n} \in \mathcal{M}_i;$$

$$\mathbf{r} = \text{find}(\mathbf{Z}(:, \lambda_n) == \hat{I}_{\lambda_n});$$

$$\mathbf{Z} = \mathbf{Z}(\mathbf{r}, :);$$

$$\mathbf{d}(\lambda_n) = \hat{x}_{\lambda_n};$$

$$n \leftarrow n + 1;$$

end

return \mathbf{d} ;

complex-valued Gaussian distributions $\mathcal{CN}(0, 1)$ and $\mathcal{CN}(0, N_0)$, respectively, where N_0 is the noise variance.

Since the encoding procedure for each block is independent of others, decoding can be performed independently at receiver. Hence, using maximum likelihood (ML) detection, the detected symbol vector for the b th block can be written as

$$\hat{\mathbf{x}}^b = \arg \min_{\mathbf{S}_i, \mu_k} \|\mathbf{y}^b - \sqrt{E_S} \mathbf{X}^b \mathbf{h}^b\|^2 \quad (2)$$

where $\mathbf{y}^b = [y((b-1)N+1), \dots, y(bN)]^T$, $\mathbf{X}^b = \text{diag}(\mathbf{x}^b)$ and $\mathbf{h}^b = [h((b-1)N+1), \dots, h(bN)]^T$.

The optimum ML detector in (2) performs $8LM^N$ real multiplications since we have L index as well as M^N modulation symbols on N subcarriers and we perform eight real multiplications for each squared Euclidean distance calculation. To overcome the high complexity of the optimum ML detector when M is high, we design a look-up table-based low-complexity ML (LC-ML) detector. The proposed LC-ML detector for the OFDM-SPM schemes is summarized in

Algorithm 1². The algorithm starts by initializing a $1 \times N$ decoding vector, \mathbf{d} , of all zeros and an $L \times N$ index matrix, \mathbf{Z} , which contains the indices of distinguishable constellations that define all admissible partitions³. Since there are N distinguishable constellations, the l th row and n th column element, z_{ln} , of \mathbf{Z} is drawn from the set $\{1, 2, \dots, N\}$, i.e., $z_{ln} \in \{1, 2, \dots, N\}$. Then, the channel gains of N subcarriers are sorted in descending order, i.e., $|h(\lambda_1)|^2 \geq \dots \geq |h(\lambda_N)|^2$ where $\lambda_n \in \{1, \dots, N\}$. Starting from the subcarrier with the largest channel gain, the algorithm determines the possible constellations indices that a subcarrier can adopt. To do that, the algorithm looks at the unique elements in the λ_n th column of the index matrix, $\mathbf{Z}(:, \lambda_n)$, by using the function “unique(.)”. Next, the signals, x , drawn from the union of the possible distinguishable constellations, \mathcal{M} , are used in an ML detector to determine the constellation symbol, \hat{x}_{λ_n} , and index, \hat{I}_{λ_n} on the λ_n th subcarrier. Then, the row indices, \mathbf{r} , corresponding to elements of $\mathbf{Z}(:, \lambda_n)$ that equal to \hat{I}_{λ_n} are found by using the function “find(.)”. To prevent illegal constellation indices, the index matrix is limited with only these rows, i.e., $\mathbf{Z} = \mathbf{Z}(\mathbf{r}, :)$. \hat{x}_{λ} is assigned to the λ_n th element of decoding vector \mathbf{d} , i.e., $\mathbf{d}(\lambda_n) = \hat{x}_{\lambda_n}$. The algorithm continues until all symbols on N subcarriers are detected. The number of operations performed by the LC-ML detectors of OFDM-SPM and OFDM-OSPM can be upper bounded by $(NK - K + N + 2L - 1) \log_2 L + 8NKM$ operations, whereas it can be upper bounded by $(N^2 + 2L - 1) \log_2 L + 8N^2M$ operations for the LC-ML detectors of OFDM-FSPM and OFDM-OFSPM. These complexities become relatively small compared to the complexity of optimum ML detection when especially M is high. For example, when $N = 4$, $K = 2$ and $M = 8$, the proposed LC-ML detector performs at least 99.75 %, 99.87 %, 99.76 %, and 99.92 % less operations, respectively, for the OFDM-SPM, OFDM-OSPM, OFDM-FSPM and OFDM-OFSPM schemes than the optimum ML detector.

IV. CODEBOOK SELECTION

As reported by several studies [23], [28], for a fading channel, the BER performance of a codebook at a high signal-to-noise ratio (SNR) is limited by the minimum Euclidean distance

² Here, we omit the subblock superscript for convenience since the LC-ML decoding can be performed independently for each subblock.

³For example, for the OFDM-OFSPM scheme having $N = 3$ subcarriers, $\mathbf{Z} \in \mathbb{Z}^{13 \times 3}$ matrix consists the subscripts of subset identifiers, or in other words, the indices of distinguishable constellations, shown in Table 1.

between codeword pairs whose difference matrix has the minimum rank, i.e., the performance at high SNR is determined by

$$d_{min} = \min_{i \neq j, i, j \in \{1, \dots, 2^f\}} ||\mathbf{X}^i - \mathbf{X}^j||^2 \quad (3)$$

where $\mathbf{X}^i = \text{diag}(\mathbf{x}^i)$ and $\mathbf{X}^j = \text{diag}(\mathbf{x}^j)$. Note that here $\text{rank}(\mathbf{X}^i - \mathbf{X}^j)$ corresponds to the minimum rank among all \mathbf{X}^i and \mathbf{X}^j pairs. For OFDM-IM schemes, the minimum rank codeword pairs are formed by modulation symbols instead of the symbols carried by the subcarrier indices [23]. It is fairly easy to see that such minimum rank between codeword pairs is observed when the matrices \mathbf{X}^i and \mathbf{X}^j regarding OFDM-IM codewords have the same subcarrier indices and different modulation symbols on one of their subcarriers. Moreover, the rank of the difference matrix of two matrices corresponding to different subcarrier indices is at least two [23]. Although the statement on the modulation symbols is valid for OFDM-SPM schemes, the statement on symbols carried by subcarrier indices is not necessarily valid for OFDM-SPM schemes. This can be shown by comparing the OSPM codewords \mathbf{x}_1 and \mathbf{x}_5 in Table I. Here, the only difference between the two partitions is observed in the second elements. Such a condition limits the diversity order of the index symbols to one if we include both codewords in the same codebook. However, one can also construct the codebook in a way that unit rank codeword pairs are not included. This ensures the minimum rank properties for OFDM-SPM codewords are the same as for OFDM-IM and MM-OFDM-IM. Hence, our aim in codebook selection for OFDM-SPM variants can be summarized as obtaining codewords that are at least as good as OFDM-IM and MM-OFDM-IM. We know that the optimum codebook selection criterion in a fading channel at high SNR is the well-known rank-determinant criterion [28]. However, due to the complexity of the codebook selection problem, our selection algorithms do not consider product-distances.

The codebook selection problem can be formulated as a *maximum clique problem*⁴ or *k-clique problem* in a graph. Before formulating such problem, we begin with some definitions and notations related to graph theory for clarity. Consider an undirected graph $G = (V, E)$, where $V = \{v_1, v_2, \dots, v_L\}$ is the vertex set and $E \subseteq V \times V$ is the edge set of G . A graph is complete if all its vertices are connected by an edge, i.e., we have $(v_l, v_{\hat{l}}) \in E$ for $\forall v_l, v_{\hat{l}} \in V, v_l \neq v_{\hat{l}}$. A clique can be defined as a complete subgraph of a graph [29], [30]. The clique number or

⁴This problem is also equivalent to some other important graph problems such as the *maximum independent set problem* and *minimum vertex cover problem*.

the size of the maximum clique is denoted by $w(G)$. Moreover, the symmetric $L \times L$ matrix $A_G = (a_{v_l v_{\hat{l}}})_{v_l, v_{\hat{l}} \in V}$, where $a_{v_l v_{\hat{l}}} = 1$ if $(v_l, v_{\hat{l}}) \in E$ and $a_{v_l v_{\hat{l}}} = 0$ if $(v_l, v_{\hat{l}}) \notin E$, stands for the adjacency matrix of the graph G .

Now, we formulate our codebook selection problem as maximum clique and k -clique problems as follows. Assume that a graph's vertices represent the OFDM-SPM codewords, which map to set partitions, and the graph's edges represent the Hamming distance between such codeword pairs⁵. To ensure similar Hamming distance properties for index symbols as OFDM-IM benchmarks, we set a condition on the vertex connections such that two vertices are connected to each other if the Hamming distance between vertices (codewords) is greater or equal to two, i.e. $E = \{(v_l, v_{\hat{l}}) | v_l, v_{\hat{l}} \in V, \text{HamDist}(v_l, v_{\hat{l}}) \geq 2\}$ where the "HamDist" function is used to measure the Hamming distance between vertices. Here, our aim is to finding the maximum clique that includes the largest possible number of vertices.

On the other hand, if the problem is formulated as a k -clique problem, one should consider the cliques that includes a specific number, k , of vertices. Considering the number of the set partitions, the clique size, k , can be determined. For example, we have $\check{B}_N = 75$ ordered set partitions for $N = 4$ and we are interested in choosing 64 partitions out of 75 to send six data bits. Hence, we can formulate such a problem in the same way as the maximum clique problem. However, in this case, we are interested in the cliques with 64 vertices. One drawback of the k -clique problem is that we do not know whether the current graph has a clique of size k . However, we can proceed to check the cliques by decreasing the clique size in the absence of a k -clique in the given graph.

A. Brute-force Search for k -clique Problem

One straightforward way of selecting good codewords or, in other words finding a k -clique, is exhaustively searching among all possible subgraphs. In this regard, the brute-force search algorithm for the k -clique problem is given in Algorithm 2. The algorithm starts by initializing the graph $G(V, E)$ where $V = \{v_1, v_2, \dots, v_L\}$ is the vertex set representing SPM codewords and, $E = \{(v_i, v_j) | v_i, v_j \in V, \text{HamDist}(v_i, v_j) \geq 2\}$ is the edge set of G . In the second step, we set $\kappa = 0$ where κ is used to adjust the clique size. Then, in the third step, we divide the

⁵Note that the Hamming distance between \mathbf{x}_i and \mathbf{x}_j is equivalent to the rank of the difference matrix $\mathbf{X}^i - \mathbf{X}^j$ where $\mathbf{X}^i = \text{diag}(\mathbf{x}_i)$ and $\mathbf{X}^j = \text{diag}(\mathbf{x}_j)$. Here, we prefer to use Hamming distance rather than rank for convenience.

graph constituted by the partitions into subgraphs, V_i , in a way that each subgraph includes $k = 2^{\lfloor \log_2 w_U(G) \rfloor - \kappa}$ vertices where $w_U(G) = N_{-1} + 1$ is an upper-bound on the size of the maximum clique $w(G)$, i.e., $w(G) \leq w_U(G) = N_{-1} + 1$ ⁶ [31]. Here, we choose the number of vertices, k , in a subgraph as a power of two since the complexity of the algorithm would be high if we take all possible values of k into consideration. The choice of k can also be justified by the use of a fixed-length bit mapping scheme, since we wish to send the same amount of data bits even if we consider all possible values of k . Moreover, N_{-1} denotes the number of eigenvalues of the adjacency matrix A_G that do not exceed -1 . Then, for each subgraph, we calculate adjacency matrices, A_{G_i} , and initialize a selected graph as $\hat{V} \in \emptyset$. In the fifth step, we check each adjacency matrix to see whether the subgraph is a k -clique or not. If the subgraph is a k -clique, the algorithm returns the current subgraph as the selected graph. If no k -clique is encountered after checking all subgraphs, we increase κ by one and go to Step 3 to check whether there is a clique of smaller size. In that way, the brute-force search algorithm guarantees reaching the maximum number of bits carried by set partitions for fixed-length binary coding schemes. Note that subgraph initialization in Step 3 may require excessive memory when the number of vertices is high. To avoid such memory requirements, one can use a combinatorial number system to obtain vertex combinations from natural numbers. The combinatorial number system provides a bijective mapping between the natural numbers and k -combinations [32]. Hence, Steps 3 and 4 can be integrated into the for-loop in Step 5 and each subgraph can be constructed in each loop by using a combinatorial number system. Note also that the for-loop in Step 5 can be parallelized to speed up the algorithm.

B. Vertex Exclusion for the Maximum Clique Problem

Since the maximum clique problem is NP-complete, the time of execution of exact algorithms will increase exponentially with the number of vertices in the graph [30]. In this regard, the brute-force search algorithm exhibits high complexity when the number of vertices is high. To overcome this problem, we propose a more practical algorithm based on excluding the vertices that have the minimum number of connections. Although such an algorithm provides a sub-optimal solution for the codeword selection problem, it is much more efficient than a brute-force

⁶One may find many bounds on the clique number [30]. The safest strategy would be calculating all bounds and using the tightest. We found the upper-bound in [31] sufficiently tight for our problem.

Algorithm 2: Brute-force search for the k -clique problem

Step 1: Initialize a graph $G(V, E)$ with $V = \{v_1, v_2, \dots, v_L\}$, the vertex set of SPM codewords, and, $E = \{(v_l, v_{\hat{l}}) \mid v_l, v_{\hat{l}} \in V, \text{HamDist}(v_l, v_{\hat{l}}) \geq 2\}$;

Step 2: $\kappa = 0$;

Step 3: Initialize subgraphs, $G_i(V_i, E)$, $i = 1, \dots, C$, with $V_i = \{v_1^i, v_2^i, \dots, v_k^i\}$ where $C = \binom{L}{k}$, $k = 2^{\lfloor \log_2 w_U(G) \rfloor - \kappa}$, $v_l^i, v_{\hat{l}}^i \in V$, $l, \hat{l} = \{1, 2, \dots, k\}$, $v_l^i \neq v_{\hat{l}}^i$;

Step 4: Initialize adjacency matrix, A_{G_i} , of G_i and $\hat{V} \in \emptyset$ where $A_{G_i} = (a_{l, \hat{l}}^i)_{(v_l^i, v_{\hat{l}}^i) \in V_i}$;

Step 5: Finding the k -clique;

for $i = 1$ **to** C **do**

if $a_{l, \hat{l}}^i = 1, \forall l, \hat{l} = \{1, 2, \dots, k\}$ **and** $l \neq \hat{l}$ **then**
 style="padding-left: 2em;"> $\hat{V} = V_i$;
 style="padding-left: 2em;">**break**;
end
 style="padding-left: 2em;"> $i \leftarrow i + 1$;
end

Step 6: Checking the presence of k -clique in G ;

if $\hat{V} \in \emptyset$ **then**

$\kappa \leftarrow \kappa + 1$;
 Go to Step 3;
end

return \hat{V} ;

search in terms of complexity. Our proposed algorithm is given in Algorithm 3. The algorithm starts by initializing the Hamming graph. Then, in the second step, we calculate the number of neighbors, $|N(v_l)|$, of each node $v_l \in V$ where $N(v_l) = \{v_{\hat{l}} \in V; a_{v_l v_{\hat{l}}} = 1\}$ is the neighborhood of v_l in G . In the third step, we remove the nodes that have the minimum number of neighbors until the remaining graph is a clique, i.e. $|N(v_l)| = |V| - 1, \forall v_l \in V$. In this step, $|N_c(v_l)|$ is used to update the cardinality of the neighborhood of v_l after removing the node that has the minimum number of neighbors. As a result, the algorithm removes the set partitions having the maximum number of unit Hamming distance pairs one-by-one until we have no set partition pairs with unit Hamming distances.

Algorithm 3: Vertex exclusion for maximum clique problem

Step 1: Initialize a graph $G(V, E)$ with $V = \{v_1, v_2, \dots, v_L\}$, the vertex set of SPM codewords, and, $E = \{(v_l, v_i) \mid v_l, v_i \in V, \text{HamDist}(v_l, v_i) \geq 2\}$;

Step 2: Calculate the cardinality of the neighborhood, $|N(v_l)|$, for each node where $v_l \in V$;

Step 3: Finding a clique by excluding vertices;

while $|N(v_l)| \neq |V| - 1, \forall v_l \in V$ **do**

$v_{min} = \arg \min_{v_l \in V} |N(v_l)|$;
 $G \leftarrow G - \{v_{min}\}$;
 $V \leftarrow V - \{v_{min}\}$;
 Calculate the neighborhood cardinality, $|N_c(v_l)|$, for each node of V where $\forall v_l \in V$;
 $|N(v_l)| \leftarrow |N_c(v_l)|$;

end

return V ;

C. Complexity Comparison

In this subsection, we compare the complexities of Algorithms 2 and 3 in terms of the algorithm run-times and the worst-case time-complexities.

The algorithm run-time results in units of seconds are shown in Tables II and III. In the tables, we also depict the numbers of vertices achieved by each algorithm in the brackets along with the upper-bound results on the achievable number of vertices according to [31]. Moreover, results using a built-in function called “FindClique[G]” in Wolfram Language are provided as a benchmark. Such a function searches for the maximal set of vertices where the corresponding subgraph is a clique [33]. Finally, we include the number of index symbols achieved by the DM-OFDM-IM and MM-OFDM-IM schemes in the rightmost columns of Tables II and III, respectively, to show the effectiveness of the proposed schemes against these schemes.⁷ Here, to reach the maximum number of index symbols, DM-OFDM-IM modulates half of the subcarriers by one of the distinguishable constellations and the other half by the other constellation.

⁷Here, we include the number of index symbols achieved by the DM-OFDM-IM and MM-OFDM-IM schemes to compare the number of index symbols achieved by these and the proposed schemes. Hence, we do not employ the codebook selection algorithms to get the same number of index symbols as those of these schemes.

Table II: The run-time complexities (in seconds) of the vertex selection algorithms along with the numbers of achieved vertices (in brackets) for OFDM-OSPM with $N \in \{4, 6, 8\}$ and $K = 2$.

	Algorithm 1	Algorithm 2	FindClique[G]	Upper-bound in [31]	DM-OFDM-IM
$N = 4$	0.4343 (8)	0.0033 (8)	0.05 (8)	(9)	(6)
$N = 6$	-	0.0176 (32)	0.06 (32)	(32)	(20)
$N = 8$	-	0.0761 (128)	0.20 (128)	(129)	(70)

In Table II, the complexity results are provided for OFDM-OSPM schemes having $N = 4, 6$ and, 8 subcarriers and $K = 2$ distinguishable constellations in each OFDM sub-block. In this case, OFDM-OSPM schemes can produce 14, 62 and 254 vertices (SPM codewords) for $N = 4, 6$ and, 8, respectively. On the other hand, in Table III, we provide complexity results for OFDM-OFSPM schemes having $N = 3, 4, 5$ and 6 subcarriers. In that case, OFDM-OFSPM schemes are capable of producing 13, 75, 541, and, 4683 vertices (SPM codewords) for $N = 3, 4, 5$ and 6, respectively. As can be seen from Tables II and III, Algorithm 2 is not able to provide a practical solution for most of the cases.⁸ However, Algorithm 3 provides efficient solutions for all of the cases by outperforming both Algorithm 2 and the “FindClique[G]” function in terms of the algorithm run-time. The effectiveness of Algorithm 3 becomes more evident when the number of subcarriers, and therefore the number of vertices, increases. Furthermore, the clique number achieved by Algorithm 3 is consistent with the clique numbers achieved by the “FindClique[G]” function and obtained by the upper-bound in [31]. Also, the OFDM-OSPM and OFDM-OFSPM schemes employing Algorithm 3 attain more index symbols, whose minimum Hamming distance is two, than those of the DM-OFDM-IM and MM-OFDM-IM schemes, respectively.

The worst-case time-complexities of Algorithms 2 and 3, on the other hand, can be calculated as follows. To perform Steps 1, 2, 4 and 6 of Algorithm 2, the required time is almost constant and small compared to the time required to perform Steps 3 and 5. The upper bound $\omega_U(G)$ in Step 3 can be executed in $O(L^3)$ operations [31]. In the worst case, it can be shown that the dominant term in the time-complexity becomes order $O((2e)^{L/2})$. Moreover, Algorithm 3 makes a total of $\frac{L(L+1)}{2} + \binom{L+1}{3}$ comparisons in the worst-case. Hence, Algorithm 3 can be executed in $O(L^3)$ operations which is much lower than the complexity of Algorithm 2.

⁸Here, “-” denotes that the algorithm is not able to return a solution within a reasonable amount of time.

Table III: The run-time complexities (in seconds) of the vertex selection algorithms along with the numbers of achieved vertices (in brackets) for OFDM-OFSPM with $N \in \{3, 4, 5, 6\}$.

	Algorithm 1	Algorithm 2	FindClique[G]	Upper-bound in [31]	MM-OFDM-IM
$N = 3$	0.024 (4)	0.0022 (7)	0.03 (7)	(7)	(6)
$N = 4$	-	0.0088 (32)	0.06 (32)	(33)	(24)
$N = 5$	-	0.4929 (181)	-	(225)	(120)
$N = 6$	-	477.0625 (1321)	-	(1876)	(720)

V. PERFORMANCE ANALYSIS

In this section, we analyze the data rate and the BER of the proposed SPM schemes.

A. Data Rate

We analyze the data rate of the proposed schemes in terms of the number of bits corresponding to the OFDM-SPM codewords normalized by the number of subcarriers used to convey each codeword. Here, we do not consider cyclic prefix length for convenience. We also provide some useful expressions for the number of partitions obtained by SPM schemes along with comparisons regarding OFDM-IM benchmarks. We assume $f_2 = N \log_2 M$ for each SPM variant.

1) *OFDM-SPM*: As discussed in Section II, the number of set partitions produced by SPM is given by the Stirling number of the second kind $\left\{ \begin{smallmatrix} N \\ K \end{smallmatrix} \right\}$. The achievable data rate per subcarrier for an OFDM-SPM scheme with N subcarriers and K subset identifiers in each sub-block is

$$R_{SPM} = \frac{f_1 + f_2}{N} = \frac{\lfloor \log_2 \left\{ \begin{smallmatrix} N \\ K \end{smallmatrix} \right\} \rfloor + N \log_2 M}{N} \quad (4)$$

where $\lfloor \cdot \rfloor$ is the floor operation.

The Stirling number of the second kind, $\left\{ \begin{smallmatrix} N \\ K \end{smallmatrix} \right\}$, can be written as [34]

$$\left\{ \begin{smallmatrix} N \\ K \end{smallmatrix} \right\} = \frac{1}{K!} \sum_{j=0}^K (-1)^j \binom{K}{j} (K-j)^N. \quad (5)$$

It is also straightforward to show that the following recurrence holds:

$$\left\{ \begin{smallmatrix} N \\ K \end{smallmatrix} \right\} = K \left\{ \begin{smallmatrix} N-1 \\ K \end{smallmatrix} \right\} + \left\{ \begin{smallmatrix} N-1 \\ K-1 \end{smallmatrix} \right\}. \quad (6)$$

Remark. Consider an OFDM-SPM block having two distinguishable constellations, i.e., $K = 2$, on N subcarriers and a DM-OFDM-IM block having two distinguishable constellations on N subcarriers in which $N - d$ of the subcarriers is modulated by one of two different constellations

and the remaining d of them is modulated by the other constellation.⁹ Except for the case where $N = 2$ and $d = 1$, the number of set partitions obtained by the OFDM-SPM encoder is equal to or greater than the number of subcarrier combinations obtained by the DM-OFDM-IM encoder, i.e., $\left\{ \begin{smallmatrix} N \\ 2 \end{smallmatrix} \right\} \geq \binom{N}{d}$ for $N \geq 2$. The equality holds for $N = 2$ and $d = 2$. This means that the achievable data rate for OFDM-SPM is equal to or greater than that of DM-OFDM-IM when the subcarriers of both schemes carry symbols that have the same modulation order. One may check that $\left\{ \begin{smallmatrix} N \\ 2 \end{smallmatrix} \right\} = 2^{N-1} - 1$. Moreover, from the recurrence relation of the binomial coefficient, we have $\binom{N}{d} = \binom{N-1}{d} + \binom{N-1}{d-1}$. If we compare $\binom{N}{d}$ with $\left\{ \begin{smallmatrix} N \\ 2 \end{smallmatrix} \right\}$, we see that $\left\{ \begin{smallmatrix} N \\ 2 \end{smallmatrix} \right\} \geq \binom{N}{d}$ for $N \geq 2$ except for the case where $N = 2$ and $d = 1$.

For a fixed and relatively small value of K , the asymptotic value of the Stirling number of the second kind as $N \rightarrow \infty$ can be written as $\left\{ \begin{smallmatrix} N \\ K \end{smallmatrix} \right\} \sim \frac{K^N}{K!}$. Hence, the asymptotic value of the achievable data rate per subcarrier for an OFDM-SPM scheme as $N \rightarrow \infty$ can be written as

$$R_{SPM} \sim \frac{N \log_2 K - \log_2 K! + N \log_2 M}{N} \sim \log_2(KM). \quad (7)$$

Remark. Consider a special MM-OFDM-IM scheme as in [21] having N subcarriers and K distinguishable constellations in each OFDM sub-block along with M -PSK symbols on each subcarrier. To compare the achievable rate of such an MM-OFDM-IM scheme, we assume that each N/K subcarriers employ the same constellation/mode. Hence, for this scenario, we have $\frac{N!}{(N/K)!^K}$ mode combinations. Applying the Stirling approximation¹⁰, the achievable data rate per subcarrier for this MM-OFDM-IM scheme as $N \rightarrow \infty$ can be written as $R_{MM-OFDM-IM} \sim \log_2(KM)$. Hence, the proposed OFDM-SPM scheme is capable of providing asymptotically the same achievable rate as this special MM-OFDM-IM scheme.

When N is large, the value of K that maximizes $\left\{ \begin{smallmatrix} N \\ K \end{smallmatrix} \right\}$ satisfies $K_N \sim \frac{N}{\ln N}$ [35]. More precisely, the following relation holds for sufficiently large N [36]

$$K_N \in \left\{ \lfloor e^{W(N)} - 1 \rfloor, \lceil e^{W(N)} - 1 \rceil \right\} \quad (8)$$

where $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ are floor and ceiling operations, respectively. $W(N)$ is the Lambert W function satisfying $W(N) \exp(W(N)) = N$ [37]. There is no exception to the relation in (8) for $1 \leq N \leq 1200$ [36]. More importantly, the maximum value of the Stirling number of the second

⁹It is fair to compare these two schemes since both of the schemes has two distinguishable constellations.

¹⁰ $\ln N! \sim N \ln N - N$ as $N \rightarrow \infty$.

kind adheres to the relation $\ln \left\{ \binom{N}{K_N} \right\} \sim N \ln N - N \ln \ln N - N$ [35]. Hence, the asymptotic maximum achievable rate of the OFDM-SPM scheme satisfies

$$R_{SPM}^{max} \sim \frac{N \log_2(N/e \ln N) + N \log_2 M}{N} \sim \log_2(N/\ln N) + \log_2(M) - \log_2 e. \quad (9)$$

This asymptotic result shows that a substantially improved data rate is attainable when we use K_N distinguishable constellations in the OFDM-SPM scheme.

2) *OFDM-OSPM*: The achievable data rate per subcarrier for an OFDM-OSPM scheme having N subcarriers and K subset identifiers in each sub-block can be written as

$$R_{OSPM} = \frac{f_1 + f_2}{N} = \frac{\lfloor \log_2 K! \left\{ \binom{N}{K} \right\} \rfloor + N \log_2 M}{N} \quad (10)$$

Remark. It is straightforward to show that $2! \left\{ \binom{N}{2} \right\} \geq \binom{N}{d}$. However, it is important to note that an OFDM-OSPM codebook, which incorporates $K = 2$ element partitions and their ordered counterparts, subsumes a DM-OFDM-IM codebook. This can easily be proved by considering set partitions along with permutations when $K = 2$. Hence, it can be concluded that OFDM-OSPM is a more general scheme, which encompasses the subcarrier combinations generated by a DM-OFDM-IM encoder. It is also important to note that an OFDM-OSPM encoder produces the same index patterns as an MM-OFDM-IM encoder for $K = N$. Hence, MM-OFDM-IM is a special case of OFDM-OSPM when $K = N$. Moreover, the partitions of an N -element set into two subsets would result in the same index symbols as GDM-OFDM-IM when the set, \mathcal{K} , containing possible numbers of subcarriers having one of a number of distinguishable constellations in each OFDM sub-block is given by $\mathcal{K} = \{1, 2, \dots, N-1\}$ for GDM-OFDM-IM. *Despite these similarities, it is important to recognize that OSPM is inherently different due to the use of set partitions to encode information instead of index patterns or permutations.*

Assuming a fixed K and using the asymptotic representation for the Stirling number of the second kind, we can write the following achievable data rate expression as $N \rightarrow \infty$:

$$R_{OSPM} \sim \frac{N \log_2 K + N \log_2 M}{N} \sim \log_2(KM). \quad (11)$$

Hence, OFDM-OSPM achieves the same asymptotic data rate as OFDM-SPM when $K \ll N$. On the other hand, it is known that, the value of K that maximizes $K! \left\{ \binom{N}{K} \right\}$ satisfies $\check{K}_N \sim \frac{N}{2 \ln 2}$ as $N \rightarrow \infty$ [38]. Moreover, we have the asymptotic relation $\check{K}_N! \left\{ \binom{N}{\check{K}_N} \right\} \sim N!/2(\ln 2)^{N+1}$ [38].

Hence, the asymptotic maximum achievable rate of the OFDM-OSPM scheme satisfies

$$R_{OFSPM}^{max} \sim \frac{\log_2(N!/2(\ln 2)^{N+1}) + N \log_2 M}{N} \sim \log_2(N) + \log_2(M) - \log_2(e \ln 2). \quad (12)$$

where the second asymptotic relation follows from the Stirling's approximation.

We can conclude that the OFDM-OSPM scheme is capable of achieving asymptotically better achievable rate than the MM-OFDM-IM scheme while the number of exploited distinguishable constellations is less than N .

3) *OFDM-FSPM*: The achievable data rate per subcarrier for an OFDM-FSPM scheme having N subcarriers can be written as

$$R_{FSPM} = \frac{f_1 + f_2}{N} = \frac{\lfloor \log_2 B_N \rfloor + N \log_2 M}{N}. \quad (13)$$

Using the asymptotic expression for the Bell number given in [39], the asymptotic achievable data rate of OFDM-FSPM can be written as

$$R_{FSPM} \sim \log_2(N/\ln N) + \log_2(M) - \log_2 e. \quad (14)$$

4) *OFDM-OFSPM*: The achievable data rate per subcarrier for an OFDM-OFSPM scheme having N subcarriers can be written as

$$R_{OFSPM} = \frac{f_1 + f_2}{N} = \frac{\lfloor \log_2 \check{B}_N \rfloor + N \log_2 M}{N}. \quad (15)$$

Remark. It is straightforward to show that $\check{B}_N > N!$ for $N \geq 2$, since the ordered set partitions include all permutations of N -element set partitions and the number of the permutations of N -element partitions is equal to $N!$. In other words, for an N -element set \mathcal{X} with $K = N$, $\mathcal{S} = \{\{x_1\}, \{x_2\}, \dots, \{x_N\}\}$ is a valid set partition, and ordering the subsets of this set would result in $N!$ different partitions. Hence, the resulting OFDM-OFSPM codebook contains the codeword $\mathbf{x}^b = \{\mu_1, \mu_2, \dots, \mu_N\}$ along with the codewords representing all permutations of the elements of \mathbf{x}^b . Note that these codewords form the MM-OFDM-IM codebook, and OFDM-OFSPM is a more general scheme compared to MM-OFDM-IM since it covers all codewords formed by an MM-OFDM-IM encoder.

Using the asymptotic relation for ordered Bell numbers given in [40], the asymptotic achievable data rate of OFDM-OFSPM scheme as $N \rightarrow \infty$ can be written as

$$R_{OFSPM} \sim \frac{\log_2(N!/2(\ln 2)^{N+1}) + N \log_2 M}{N} \sim \log_2(N) + \log_2(M) - \log_2(e \ln 2). \quad (16)$$

As can be observed from (9), (12), (14), and (16), the maximum achievable data rate of each SPM variant exhibits the same asymptotic behavior as its full SPM counterpart while exploiting fewer distinguishable constellations. Hence, OFDM-SPM and OFDM-OSPM are capable of

achieving the same asymptotic data rate as OFDM-FSPM and OFDM-OFSPM, respectively, by utilizing fewer constellation points.

B. Bit-Error Rate

Let $P(\mathbf{X}^i \rightarrow \mathbf{X}^j)$ denote the pairwise error probability (PEP) associated with the erroneous detection of \mathbf{X}^i as \mathbf{X}^j where $\mathbf{X}^i = \text{diag}(\mathbf{x}^i)$ and $\mathbf{X}^j = \text{diag}(\mathbf{x}^j)$. From (2), the PEP conditioned on the channel coefficients is given by

$$P(\mathbf{X}^i \rightarrow \mathbf{X}^j | \mathbf{h}) = Q\left(\sqrt{\frac{E_S ||(\mathbf{X}^i - \mathbf{X}^j)\mathbf{h}||^2}{2N_0}}\right). \quad (17)$$

By using the identity $Q(x) \approx \frac{1}{12}e^{-x^2/2} + \frac{1}{4}e^{-2x^2/3}$ and averaging over \mathbf{h} , an approximate unconditional PEP expression can be obtained [5]:

$$P(\mathbf{X}^i \rightarrow \mathbf{X}^j) = E_{\mathbf{h}} [P(\mathbf{X}^i \rightarrow \mathbf{X}^j | \mathbf{h})] \approx \frac{1/12}{\det(\mathbf{I}_N + \frac{E_S}{4N_0} \mathbf{C} \mathbf{Z}_{ij})} + \frac{1/4}{\det(\mathbf{I}_N + \frac{E_S}{3N_0} \mathbf{C} \mathbf{Z}_{ij})} \quad (18)$$

where \mathbf{I}_N denotes the $N \times N$ identity matrix, $\mathbf{C} = E_{\mathbf{h}}[\mathbf{h}\mathbf{h}^H]$ and $\mathbf{Z}_{ij} = (\mathbf{X}^i - \mathbf{X}^j)^H(\mathbf{X}^i - \mathbf{X}^j)$.

An upper-bound on the average BER is given by the well-known union bound

$$P_b \leq \frac{1}{f2^f} \sum_{i=1}^{2^f} \sum_{j=1}^{2^f} P(\mathbf{X}^i \rightarrow \mathbf{X}^j) D(\mathbf{X}^i \rightarrow \mathbf{X}^j) \quad (19)$$

where $D(\mathbf{X}^i \rightarrow \mathbf{X}^j)$ is the number of bits in error for the corresponding pairwise error event.

Note that the upper-bound expression given in (19) is valid for all OFDM-SPM schemes.

Assuming $\mathbf{C} = E_{\mathbf{h}}[\mathbf{h}\mathbf{h}^H] \approx \mathbf{I}_N$ and considering the fact that \mathbf{Z}_{ij} is a diagonal matrix, one can rewrite (18) as

$$P(\mathbf{X}^i \rightarrow \mathbf{X}^j) \approx \frac{1/12}{\prod_{n=1}^N (1 + \frac{E_S}{4N_0} \lambda_n)} + \frac{1/4}{\prod_{n=1}^N (1 + \frac{E_S}{3N_0} \lambda_n)} \quad (20)$$

where λ_n is the n th diagonal element of \mathbf{Z}_{ij} . At high SNR, one can neglect the one in the denominator of (20) and write the following approximation

$$P(\mathbf{X}^i \rightarrow \mathbf{X}^j) \approx \frac{1/12}{\prod_{n \in \Gamma} \frac{E_S}{4N_0} \lambda_n} + \frac{1/4}{\prod_{n \in \Gamma} \frac{E_S}{3N_0} \lambda_n} \quad (21)$$

where Γ is the set of nonzero diagonal elements of \mathbf{Z}_{ij} . Finally, an asymptotic expression for the BER of the OFDM-SPM variants can be obtained by substituting (21) into (19).

As explained in the previous sections, OFDM-SPM codebooks can be designed in a way that the minimum Hamming distance between index symbols for the set partitions is equal to two. Hence, assuming the OFDM-SPM index symbols achieve this minimum Hamming distance property, the average BER expression will be dominated by the modulation symbols at high SNRs and, therefore, the diversity order of the BER curves is limited to one.

VI. NUMERICAL RESULTS

In this section, we provide numerical data rate and BER results for the proposed schemes. In figures, OFDM-SPM (N, K, M) and OFDM-OSPM (N, K, M) stand for OFDM-SPM schemes having N subcarriers and K distinguishable constellations in each OFDM sub-block along with M -PSK symbols on each subcarrier, whereas OFDM-FSPM (N, M) and OFDM-OFSPM (N, M) stand for variants of OFDM-SPM employing all set partitions and having N subcarriers with M -PSK symbols in each OFDM sub-block. OFDM-OSPM, (N, K, M) , Alg. 3 and OFDM-OFSPM, (N, M) , Alg. 3 form their codewords in the same way as OFDM-OSPM, (N, K, M) and OFDM-OFSPM, (N, M) , respectively, and then, they employ Algorithm 3 to cut some of their codewords.¹¹ Moreover, OFDM-IM (N, K_a, M) stands for the conventional OFDM-IM scheme having K_a activated subcarriers out of N in each sub-block and employing M -PSK modulation on the activated subcarriers. Finally, DM-OFDM-IM (N, M) signifies a dual-mode scheme having N subcarriers along with two distinguishable M -PSK constellations, and MM-OFDM-IM (N, M) represents a multi-mode scheme having N subcarriers along with N distinguishable M -PSK constellations in each sub-block.

In our simulations, we assume all schemes operate over a Rayleigh fading channel, whose elements are independent and identically distributed. ML detection is applied under the assumption that channel estimation is perfect. The use of the Rayleigh model in the simulations is realistic when considering environments with a large number of scatterers [41].

A. Data Rate

In Fig. 2, we compare the data rates of the proposed OFDM-SPM schemes with DM-OFDM-IM, GDM-OFDM-IM, and MM-OFDM-IM. The data rate results in terms of the number of index bits per subcarrier are given as a function of N in Fig. 2a and a function of detection complexity in Fig. 2b. Since all the schemes considered in this figure activate all subcarriers and the modulation order of the carried symbols on each subcarrier can be chosen to be the same, we ignore the modulation bits transmitted per subcarrier in Fig. 2a. Also, we do not restrict the codebook sizes to a power of two, since all the codewords in a codebook can be utilized by a binary

¹¹Here, we employ a slightly different version of Algorithm 3 to the overall codebook of the OFDM-OSPM and OFDM-OFSPM schemes. By defining the Hamming distance as the number of positions at which the corresponding symbols of the OFDM-SPM schemes are different, we use Algorithm 3 to cut the OFDM-OSPM and OFDM-OFSPM codewords having the maximum number of unit Hamming distance pairs one-by-one until we have the same data rate as those of the benchmarks.

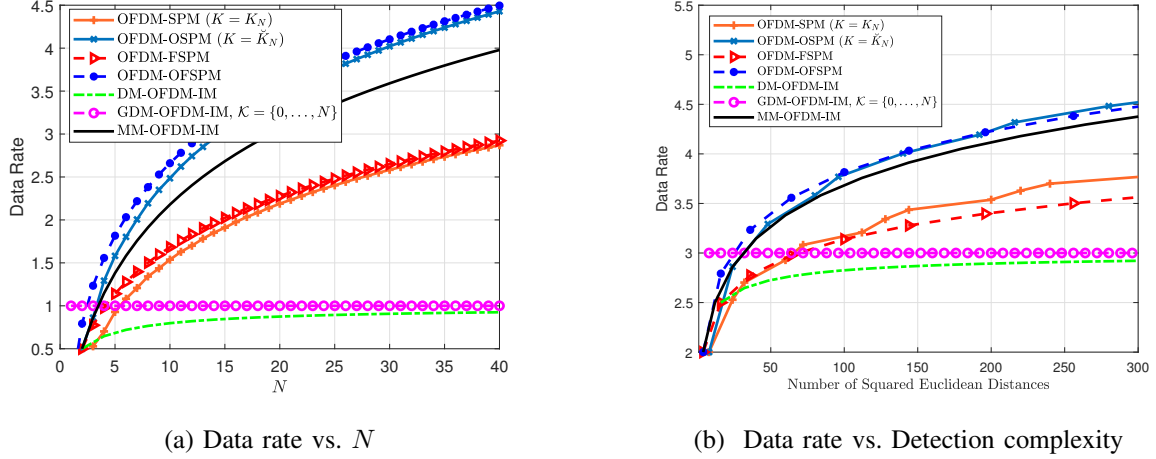


Figure 2: Data rate comparison of OFDM-SPM variants with DM-OFDM-IM, GDM-OFDM-IM and MM-OFDM-IM.

coding technique such as Huffman coding regardless of the number of codewords [16], [42]. To reach the maximum number of index bits, DM-OFDM-IM modulates half of the subcarriers by one of the distinguishable constellations and the other half by the other constellation. Moreover, we assume $\mathcal{K} = \{0, 1, \dots, N\}$ for GDM-OFDM-IM. To reach the maximum numbers of set partitions for OFDM-SPM and OFDM-OSPM, we substitute $K_N \in \{\lfloor e^{W(N)} - 1 \rfloor, \lceil e^{W(N)} - 1 \rceil\}$ and $\tilde{K}_N = \frac{N}{2 \ln 2}$, respectively, into K . The data rate results in Fig. 2a verify the remarks in the previous section and indicate that all SPM variants outperform DM-OFDM-IM and GDM-OFDM-IM for most of the values of N . Moreover, although MM-OFDM-IM considerably outperforms OFDM-SPM and OFDM-FSPM, it is outperformed by OFDM-OSPM and OFDM-OFSPM for almost all values of N . On the other hand, in Fig. 2b, we calculate the detection complexity as the number of squared Euclidean distance calculations performed by the low-complexity subcarrier-wise (LC-SW) or LC-ML detectors of the considered schemes. In this regard, the LC-SW detector of MM-OFDM-IM performs $(N^2 + N)M/2$ calculations, whereas the LC-SW detectors of DM-OFDM-IM and GDM-OFDM-IM perform $2NM$ calculations for an OFDM subblock of size N . Moreover, in the worst-case, the LC-ML detectors¹² of OFDM-SPM

¹²It is important to note that the proposed LC-ML detector does not require a look-up table when the codebook selection algorithms are not employed. In that case, it is enough to check that the detected distinguishable constellations form an admissible index symbol that can be mapped to a set partition. For this case, the detection complexity of the LC-ML detector is dominated by the number of squared Euclidean distance calculations.

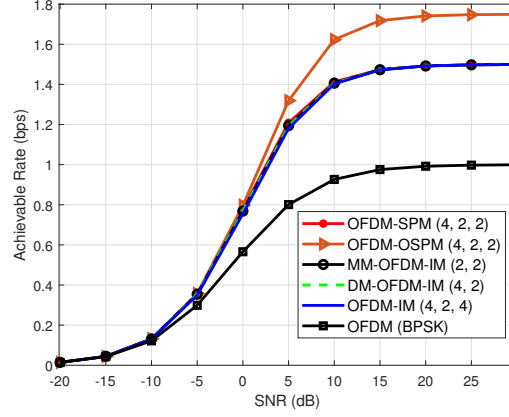


Figure 3: Achievable rate comparison of OFDM-SPM (4, 2, 2) and OFDM-OSPM (4, 2, 2) with MM-OFDM-IM (2, 2), DM-OFDM-IM (4, 2), OFDM-IM (4, 2, 4) and OFDM (BPSK).

and OFDM-OSPM, respectively, perform $NK_N M$ and $N\check{K}_N M$ calculations, whereas the LC-ML detectors of OFDM-FSPM and OFDM-OFSPM perform $N^2 M$ calculations for an OFDM subblock of size N . Here, we assume that $M = 4$ for all schemes. As seen from Fig. 2b, OFDM-OSPM and OFDM-OFSPM exhibit superior performance compared to other schemes. Moreover, the OFDM-SPM scheme outperforms the OFDM-FSPM, DM-OFDM-IM, and GDM-OFDM-IM schemes when the complexity increases.

To gain more insight into the achievable rate of the proposed schemes, one can substitute a finite input symbol set for such schemes into the standard mutual information expression. Assuming equally likely codewords and applying the same approach as discussed in [14], [43], the achievable rate of the proposed schemes can be computed by evaluating

$$R = \frac{1}{N} \left(f - \frac{1}{2^f} \sum_{i=1}^{2^f} E_{\mathbf{h}, \mathbf{n}} \left[\log_2 \sum_{j=1}^{2^f} e^{\delta(i,j)} \right] \right) \quad (22)$$

where $\delta(i, j) = \frac{-||\text{diag}(\mathbf{h})(\mathbf{x}^i - \mathbf{x}^j) + \mathbf{n}||^2 + ||\mathbf{n}||^2}{N_0}$. Although this approach does not yield a closed-form solution, we can obtain numerical data for the theoretical achievable rate of the proposed schemes.

We compare the achievable rates of OFDM-SPM variants with those of the MM-OFDM-IM, DM-OFDM, OFDM-IM and OFDM schemes in Figs. 3 and 4. The achievable rate curves provided in these figures are obtained by using (22). The results in Fig. 3 are given for OFDM-SPM (4, 2, 2), OFDM-OSPM (4, 2, 2), MM-OFDM-IM (2, 2), DM-OFDM-IM (4, 2), OFDM-IM (4, 2, 4) and OFDM (BPSK) schemes. Here, OFDM-SPM (4, 2, 2) and OFDM-OSPM (4, 2, 2)

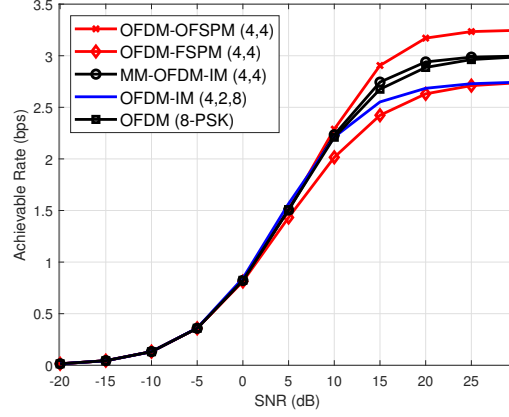


Figure 4: Achievable rate comparison of OFDM-OFSPM (4, 4) and OFDM-FSPM (4, 4) with MM-OFDM-IM (4, 4), OFDM-IM (4, 2, 8) and OFDM (8-PSK).

encoders produce $\left\{ \begin{smallmatrix} 4 \\ 2 \end{smallmatrix} \right\} = 7$ and $2! \left\{ \begin{smallmatrix} 4 \\ 2 \end{smallmatrix} \right\} = 14$ codewords, and then $2^{\lfloor \log_2 7 \rfloor} = 4$ and $2^{\lfloor \log_2 14 \rfloor} = 8$ set partitions are selected by Algorithm 2. As can be observed from the figure, OFDM-SPM (4, 2, 2), MM-OFDM-IM (2, 2), DM-OFDM-IM (4, 2) and OFDM-IM (4, 2, 4) exhibit almost the same achievable rate performance for all SNR values. On the other hand, OFDM-OSPM (4, 2, 2) outperforms all other IM schemes. Moreover, to achieve 1.5 bps, the SNR requirement for OFDM-OSPM (4, 2, 2) is approximately 12 dB lower than for other IM schemes.

In Fig. 4, we compare the achievable rates of OFDM-OFSPM (4, 4) and OFDM-FSPM (4, 4) with MM-OFDM-IM (4, 4), OFDM-IM (4, 2, 8) and OFDM (8-PSK). Here, OFDM-FSPM (4, 4) and OFDM-OFSPM (4, 4) encoders produce $B_N = 15$ and $\check{B}_N = 75$ codewords, respectively. 32 codewords are then selected according to Algorithm 3 for OFDM-OFSPM (4, 4) in a way that the rank of the difference matrices for the codeword pairs in the selected codebook is at least two [23]. For OFDM-FSPM (4, 4), Algorithms 2 and 3 cannot return $2^{\lfloor \log_2 15 \rfloor} = 8$ codewords; they return only four and five codewords, respectively. To fully exploit the codewords generated by this scheme and to achieve a high data rate at high SNR, we use three other codewords in the OFDM-FSPM (4, 4) scheme in addition to the codewords generated by Algorithm 3. This limits the minimum Hamming distance of the codeword pairs to one. This explains why the OFDM-FSPM (4, 4) scheme is outperformed by all other schemes. However, Fig. 4 shows that OFDM-OFSPM (4, 4) is capable of providing a better achievable rate than all other schemes at high SNR. Moreover, OFDM-OFSPM (4, 4) is, theoretically, able to achieve 2 bps with an SNR

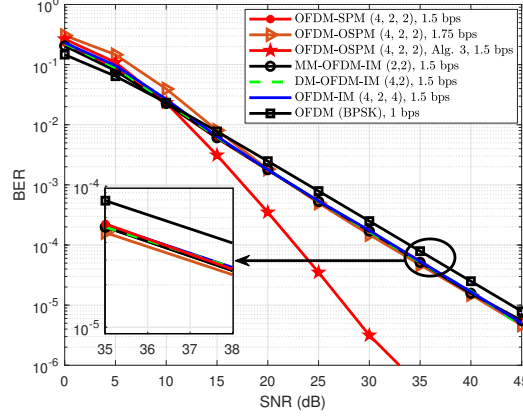


Figure 5: BER comparison of OFDM-SPM (4, 2, 2), OFDM-OSPM (4, 2, 2), and OFDM-OSPM (4, 2, 2), Alg. 3 with MM-OFDM-IM (2, 2), DM-OFDM-IM (4, 2), OFDM-IM (4, 2, 4) and OFDM (BPSK).

requirement that is almost 13 dB lower than that of MM-OFDM-IM (4, 4).

B. Bit-error Rate

In this subsection, we compare the proposed OFDM-SPM schemes with conventional OFDM and IM benchmarks in terms of BER performance. We also show the effectiveness of the proposed LC-ML detector compared to the optimum ML detector.

In Fig. 5, we compare the BER performance of OFDM-SPM (4, 2, 2), OFDM-OSPM (4, 2, 2), and OFDM-OSPM (4, 2, 2), Alg. 3 with MM-OFDM-IM (2, 2), DM-OFDM-IM (4, 2), OFDM-IM (4, 2, 4), and conventional OFDM (BPSK). Except for OFDM-OSPM (4, 2, 2) and conventional OFDM (BPSK), all schemes exhibit the same spectral efficiency, which is 1.5 bps. Spectral efficiencies for OFDM-OSPM (4, 2, 2) and conventional OFDM (BPSK) are 1.75 bps and 1 bps, respectively. Here, OFDM-SPM (4, 2, 2) and OFDM-OSPM (4, 2, 2) encoders produce $\binom{4}{2} = 7$ and $2!\binom{4}{2} = 14$ codewords, and then $2^{\lceil \log_2 7 \rceil} = 4$ and $2^{\lceil \log_2 14 \rceil} = 8$ set partitions are selected by Algorithm 2. To make a fair comparison in terms of data rate, OFDM-OSPM (4, 2, 2), Alg. 3 employs Algorithm 3 to cut the codewords of OFDM-OSPM (4, 2, 2) until it has the same spectral efficiency as those of OFDM-SPM (4, 2, 2), MM-OFDM-IM (2, 2), DM-OFDM-IM (4, 2) and OFDM-IM (4, 2, 4). As seen from the figure, OFDM-SPM (4, 2, 2) exhibits almost the same BER performance as OFDM-IM (4, 2, 4), DM-OFDM-IM (4, 2), and MM-OFDM-IM (2, 2), and it outperforms conventional OFDM (BPSK) at medium-to-high SNR. More importantly,

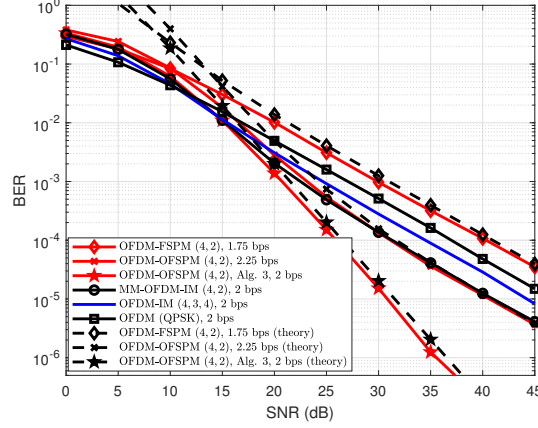


Figure 6: BER comparison of OFDM-FSPM (4, 2), OFDM-OFSPM (4, 2) and OFDM-OFSPM (4, 2), Alg. 3 with MM-OFDM-IM (4, 2), OFDM-IM (4, 3, 4) and OFDM (QPSK).

although OFDM-OSPM (4, 2, 2) has higher spectral efficiency than other schemes, it provides marginally better BER performance compared to other schemes at high SNR. The reason behind this is the codewords related to set partitions in the OFDM-OFSPM scheme have the capability of preserving the same Hamming distance properties as the codewords related to the permutations in the MM-OFDM-IM scheme or the combinations in the DM-OFDM-IM scheme, and the number of such codewords in the OFDM-OFSPM scheme is larger than those in the DM-OFDM-IM and MM-OFDM-IM schemes. Having a higher number of these codewords is desirable in the high SNR since they are capable of introducing diversity order of two unlike the codewords related to conventional modulation, which limit the diversity order to one. Therefore, they mitigate the effect of the codewords related to conventional modulation on the BER performance. OFDM-OSPM (4, 2, 2), Alg. 3, on the other hand, considerably outperforms the other schemes by introducing an additional diversity gain, as the codebook of OFDM-OSPM (4, 2, 2) is culled enough by Algorithm 3 that the minimum Hamming distance for the remaining codebook becomes two. This is only possible because the combinatorial structure that partitioning offers produces more index symbols than the benchmark schemes.

In Fig. 6, OFDM-FSPM (4, 2), OFDM-OFSPM (4, 2) and OFDM-OFSPM (4, 2), Alg. 3 are compared with MM-OFDM-IM (4, 2), OFDM-IM (4, 3, 4), and conventional OFDM (QPSK). Except for OFDM-FSPM (4, 2) and OFDM-OFSPM (4, 2), all schemes have the same spectral efficiency of 2 bps. Spectral efficiencies for OFDM-FSPM (4, 2) and OFDM-OFSPM (4, 2)

are 1.75 bps and 2.25 bps, respectively. Here, OFDM-FSPM (4, 2) and OFDM-OFSPM (4, 2) encoders produce $B_N = 15$ and $\check{B}_N = 75$ codewords. Then, 32 codewords are selected according to Algorithm 3 for OFDM-OFSPM (4, 2). As discussed in the previous section, Algorithms 2 and 3 are not capable of providing $2^{\lfloor \log_2 15 \rfloor} = 8$ codewords for OFDM-FSPM (4, 2). However, we use eight codewords of OFDM-FSPM (4, 2), including the five FSPM codewords provided by Algorithm 3, to achieve a higher data rate. OFDM-OFSPM (4, 2), Alg. 3 employs Algorithm 3 to remove the codewords of OFDM-OFSPM (4, 2) until it has the same spectral efficiency as those of MM-OFDM-IM (4, 2), OFDM-IM (4, 3, 4) and OFDM (QPSK). We also provide results on the theoretical upper-bound for the OFDM-SPM schemes. As observed from the figure, upper-bound curves are consistent with computer simulations, especially at high SNR. Although OFDM-FSPM (4, 2) cannot provide a BER advantage relative to OFDM-IM (4, 3, 4) and MM-OFDM-IM (4, 2), OFDM-OFSPM (4, 2) exhibits superior BER performance relative to all benchmarks at high SNR while achieving enhanced spectral efficiency. These results arise from the fact that the set partitions in the selected OFDM-FSPM codebook exhibit lower rank than the set partitions in the selected OFDM-OFSPM codebook. Moreover, the codewords related to set partitions in OFDM-OFSPM are capable of preserving the same minimum rank property as the codewords related to the permutations in the MM-OFDM-IM scheme, although the number of index bits for the former increases. As also seen from the figure, when we remove the codewords of OFDM-OFSPM (4, 2) by using Algorithm 3, the remaining codebook, in other words, OFDM-OFSPM (4, 2), Alg. 3, achieves a higher diversity order as well as considerably outperforms other schemes.

In Fig. 7, we compare the BER performance of OFDM-FSPM (4, 4), OFDM-OFSPM (4, 4) and OFDM-OFSPM (4, 4), Alg. 3 with MM-OFDM-IM (4, 4), OFDM-IM (4, 3, 8), and conventional OFDM (8-PSK) schemes. We also provide a BER curve (“OFDM-OFSPM (4, 4), QAM”) for OFDM-OFSPM employing four different 4-QAM constellations to distinguish the OFDM-OFSPM codewords. These constellations are obtained by employing the set partitioning technique described in [27] to a 16-QAM constellation. We use the same SPM codewords as Fig. 6 for OFDM-SPM variants. In this figure, OFDM-FSPM (4, 4) and OFDM-IM (4, 3, 8); OFDM-OFSPM (4, 4), Alg. 3, MM-OFDM-IM (4, 4) and OFDM (8-PSK); OFDM-OFSPM (4, 4) and OFDM-OFSPM (4, 4), QAM schemes have the same spectral efficiencies which are 2.75 bps, 3 bps, and 3.25 bps, respectively. Similar results to Fig. 6 can be observed in Fig. 7 for a higher modulation order. On the other hand, OFDM-OFSPM (4, 4), QAM outperforms both OFDM-

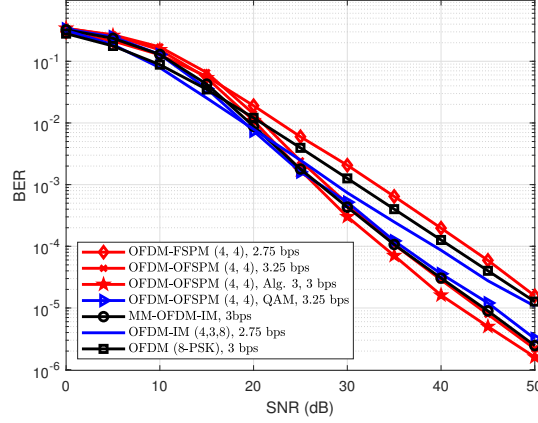


Figure 7: BER comparison of OFDM-FSPM (4, 4), OFDM-OFSPM (4, 4) and OFDM-OFSPM (4, 4), Alg. 3 with MM-OFDM-IM (4, 4), OFDM-IM (4, 3, 8) and OFDM (8-PSK).

OFSPM (4, 4) and MM-OFDM-IM (4, 4) at low SNR values; however, OFDM-OFSPM (4, 4) exhibits slightly better BER performance than OFDM-OFSPM (4, 4), QAM and MM-OFDM-IM (4, 4) at high SNR values. The behavior of the BER curves at low SNR values can be explained by the minimum Euclidean distances between the codewords. The minimum Euclidean distance between “OFDM-OFSPM (4, 4), QAM” codewords is 0.8944; however, the minimum Euclidean distance between OFDM-OFSPM (4, 4) and MM-OFDM-IM (4, 4) codewords is just 0.5518. On the other hand, the BER performance at high SNR will be dominated by the Euclidean distances between the codewords whose difference matrix has the minimum rank. In that case, the Euclidean distance between OFDM-OFSPM (4, 4) and MM-OFDM-IM (4, 4) codewords is given by 1.4142; however, the Euclidean distance between “OFDM-OFSPM (4, 4), QAM” codewords is just 1.2649. The comparison between OFDM-OFSPM (4, 4) and MM-OFDM-IM (4, 4) results from the number of index bits. OFDM-OFSPM (4, 4) produces more index bits than MM-OFDM-IM (4, 4) by using set partitions rather than permutations. Such index bits become undesirable at low SNR due to the minimum Euclidean distance between the associated codewords. However, at high SNR, they become desirable since these codewords provide better Hamming distance properties than the codewords associated with conventional modulation bits. It is also important to note that OFDM-OFSPM (4, 4), Alg. 3 achieves the best performance at high SNR by providing almost 1.4 dB SNR gain compared to OFDM-OFSPM (4, 4) and MM-OFDM-IM (4, 4) at a BER value of 10^{-4} .

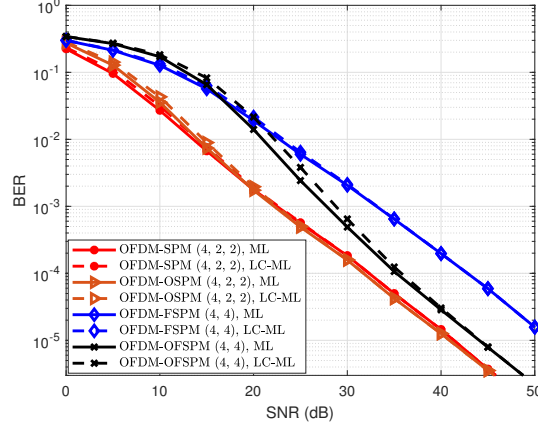


Figure 8: BER comparison of the proposed LC-ML detector with optimum ML detector for the OFDM-SPM schemes.

Fig. 8 demonstrates the BER performance of the proposed LC-ML detector compared to the optimum ML detector for OFDM-SPM (4, 2, 2), OFDM-OSPM (4, 2, 2), OFDM-FSPM (4, 4) and OFDM-OFSPM (4, 4). In this figure, curves with extensions “LC-ML” and “ML” in the legend relate to the OFDM-SPM schemes applying the proposed LC-ML and optimum ML detectors, respectively. As seen from the figure, the performance of the proposed LC-ML detector is very close to that of the optimum ML detector, especially at low and high SNRs.

VII. CONCLUSION

In this paper, we proposed a novel modulation concept, which we call set partition modulation. We represented several variants of the concept and showed a practical implementation in the context of OFDM. Moreover, we defined a codebook selection problem for the proposed techniques and expressed such a problem as a clique problem in graph theory. We further provided an efficient solution for the codebook selection. Then, we investigated the performance of the new techniques in terms of their data rates and BER characteristics, and we presented asymptotic results regarding the performance. We compared the proposed OFDM-SPM variants with the appropriate benchmarks. Through computer simulations and theoretical calculations, it is shown that the proposed SPM schemes can provide noteworthy improvements compared to conventional OFDM, OFDM-IM, DM-OFDM-IM, and MM-OFDM-IM in terms of data rate and BER.

REFERENCES

- [1] F. Yarkin and J. P. Coon, "Set partition modulation," in *IEEE 30th Int. Sym. Pers., Indoor Mobile Radio Commun.*, Sep. 2019, pp. 1–7.
- [2] D. Slepian, "Permutation modulation," *Proc. IEEE*, vol. 53, no. 3, pp. 228–236, Mar. 1965.
- [3] E. Basar, "Index modulation techniques for 5G wireless networks," *IEEE Commun. Mag.*, vol. 54, no. 7, pp. 168–175, Jul. 2016.
- [4] R. Y. Mesleh, H. Haas, S. Sinanovic, C. W. Ahn, and S. Yun, "Spatial modulation," *IEEE Trans. Veh. Tech.*, vol. 57, no. 4, pp. 2228–2241, Jul. 2008.
- [5] E. Basar, U. Aygolu, E. Panayirci, and H. V. Poor, "Orthogonal frequency division multiplexing with index modulation," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5536–5549, Nov. 2013.
- [6] R. Abu-alhiga and H. Haas, "Subcarrier-index modulation OFDM," in *Proc. IEEE Int. Sym. Pers., Indoor Mobile Radio Commun.*, Sep. 2009, pp. 177–181.
- [7] D. Tsonev, S. Sinanovic, and H. Haas, "Enhanced subcarrier index modulation (SIM) OFDM," in *Proc. IEEE GLOBECOM Workshops*, Dec 2011, pp. 728–732.
- [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 Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 1905–1938, 3rd Quart. 2018.
- [9] J. Choi, "Coded OFDM-IM with transmit diversity," *IEEE Trans. Commun.*, vol. 65, no. 7, pp. 3164–3171, Jul. 2017.
- [10] J. Choi and Y. Ko, "TCM for OFDM-IM," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 50–53, Feb. 2018.
- [11] E. Basar, "OFDM with index modulation using coordinate interleaving," *IEEE Wireless Commun. Lett.*, vol. 4, no. 4, pp. 381–384, Aug. 2015.
- [12] R. Fan, Y. J. Yu, and Y. L. Guan, "Generalization of orthogonal frequency division multiplexing with index modulation," *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, pp. 5350–5359, Oct. 2015.
- [13] M. Wen, X. Cheng, M. Ma, B. Jiao, and H. V. Poor, "On the achievable rate of OFDM with index modulation," *IEEE Trans. Signal Process.*, vol. 64, no. 8, pp. 1919–1932, Apr. 2016.
- [14] N. Ishikawa, S. Sugiura, and L. Hanzo, "Subcarrier-index modulation aided OFDM - will it work?" *IEEE Access*, vol. 4, pp. 2580–2593, 2016.
- [15] A. I. Siddiq, "Effect of subcarrier activation ratio on the performance of OFDM-IM over Rayleigh fading channel," *IEEE Commun. Lett.*, vol. 21, no. 6, pp. 1293–1296, Jun. 2017.
- [16] J. Coon, M. Badiu, Y. Liu, F. Yarkin, and S. Dang, "Binary-tree encoding for uniform binary sources in index modulation systems," *IEEE J. Sel. Topics Signal Process.*, pp. 1–1, 2019.
- [17] M. Chafii, J. P. Coon, and D. A. Hedges, "DCT-OFDM with index modulation," *IEEE Commun. Lett.*, vol. 21, no. 7, pp. 1489–1492, Jul. 2017.
- [18] T. Mao, Z. Wang, Q. Wang, S. Chen, and L. Hanzo, "Dual-mode index modulation aided OFDM," *IEEE Access*, vol. 5, pp. 50–60, Feb. 2017.
- [19] 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.
- [20] T. Mao, Q. Wang, and Z. Wang, "Generalized dual-mode index modulation aided OFDM," *IEEE Commun. Lett.*, vol. 21, no. 4, pp. 761–764, Apr. 2017.
- [21] 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.

- [22] T. Mao, Q. Wang, J. Quan, and Z. Wang, "Zero-padded tri-mode index modulation aided OFDM," in *Proc. GLOBECOM*, Dec. 2017, pp. 1–5.
- [23] M. Wen, E. Basar, Q. Li, B. Zheng, and M. Zhang, "Multiple-mode orthogonal frequency division multiplexing with index modulation," *IEEE Trans. Commun.*, vol. 65, no. 9, pp. 3892–3906, Sep. 2017.
- [24] M. Wen, Q. Li, E. Basar, and W. Zhang, "Generalized multiple-mode OFDM with index modulation," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6531–6543, Oct. 2018.
- [25] T. Mao, Q. Wang, Z. Wang, and S. Chen, "Novel index modulation techniques: A survey," *IEEE Commun. Surveys Tuts*, vol. 21, no. 1, pp. 315–348, 1st Quart. 2019.
- [26] J. Li, M. Wen, X. Jiang, and W. Duan, "Space-time multiple-mode orthogonal frequency division multiplexing with index modulation," *IEEE Access*, vol. 5, pp. 23 212–23 222, 2017.
- [27] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inf. Theory*, vol. 28, no. 1, pp. 55–67, Jan. 1982.
- [28] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.
- [29] R. Carraghan and P. M. Pardalos, "An exact algorithm for the maximum clique problem," *Oper. Res. Lett.*, vol. 9, no. 6, pp. 375 – 382, 1990.
- [30] I. M. Bomze, M. Budinich, P. M. Pardalos, and M. Pelillo, *The Maximum Clique Problem*. Boston, MA: Springer US, 1999, pp. 1–74.
- [31] A. Amin and S. Hakimi, "Upper bounds on the order of a clique of a graph," *SIAM J. of Appl. Math.*, vol. 22, no. 4, pp. 569–573, 1972.
- [32] D. E. Knuth, *The Art of Computer Programming, Volume 1 (3rd Ed.): Fundamental Algorithms*. Redwood City, CA, USA: Addison Wesley Longman Publishing Co., Inc., 1997.
- [33] E. W. Weisstein, "Clique." [Online]. Available: <http://mathworld.wolfram.com/Clique.html>
- [34] N. M. Temme, "Asymptotic estimates of Stirling numbers," *Studies in Applied Mathematics*, vol. 89, no. 3, pp. 233–243, 1993.
- [35] B. Rennie and A. Dobson, "On Stirling numbers of the second kind," *Journal of Combinatorial Theory*, vol. 7, no. 2, pp. 116 – 121, 1969.
- [36] E. R. Canfield and C. Pomerance, "On the problem of uniqueness for the maximum Stirling number(s) of the second kind," *INTEGERS*, vol. 2, 2002.
- [37] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the LambertW function," *Advances in Computational Mathematics*, vol. 5, no. 1, pp. 329–359, Dec. 1996.
- [38] I. Mezö, "Asymptotics of the modes of the ordered Stirling numbers," 2015. [Online]. Available: <https://arxiv.org/pdf/1504.06970.pdf>
- [39] N. G. de Bruijn, *Asymptotic Methods in Analysis*, 2nd ed. New York: Dover, 1981.
- [40] J. Barthelemy, "An asymptotic equivalent for the number of total preorders on a finite set," *Discrete Mathematics*, vol. 29, no. 3, pp. 311 – 313, 1980.
- [41] B. Vucetic and J. Yuan, *Space-Time Coding*. New York, NY, USA: John Wiley & Sons, Inc., 2003.
- [42] W. Wang and W. Zhang, "Huffman coding-based adaptive spatial modulation," *IEEE Trans. Wireless Commun.*, vol. 16, no. 8, pp. 5090–5101, Aug. 2017.
- [43] Soon Xin Ng and L. Hanzo, "On the MIMO channel capacity of multidimensional signal sets," *IEEE Trans. Veh. Tech.*, vol. 55, no. 2, pp. 528–536, Mar. 2006.