Improved physical-layer security for OFDM using data-based subcarrier scrambling

Mohammad M. Banat*, *Senior Member, IEEE*, Joan Bas†, *Senior Member, IEEE*,
and Alexis A. Dowhuszko‡, *Senior Member, IEEE*
*Department of Electrical Engineering, Jordan University of Science and Technology, 22110 Irbid, Jordan
†Department of Array and Multi-Sensor Processing, CTTC/CERCA, 08860 Castelldefels, Spain
‡Department of Communications and Networking, Aalto University, 02150 Espoo, Finland
Email: banat@just.edu.jo; joan.bas@cttc.es; alexis.dowhuszko@aalto.fi

2021 IEEE Globecom Workshops (GC Wkshps) | 978-1-6654-2390-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/GCWkshps52748.2021.9682170

***Abstract*—This paper presents a novel physical-layer security approach to protect the information exchanged in a wireless communication system based on OFDM. In this method, QAM symbols that are fed into the IFFT block are split into two subsets. The first subset of symbols is placed on non-scrambled (indexing) subcarriers, whereas the remaining symbols are transmitted on scrambled (data) subcarriers. Based on the bits placed on the indexing subcarriers, a permutation matrix that defines the (data-based) scrambling sequence of the data subcarriers is determined using an algorithm that is known *a priori* between the transmitter and receiver. The mapping between indexing bits and scrambling sequences is designed to minimize error propagation when there are erroneous received indexing bits (*i.e. Gray-mapped* sequences). Closed form formulas that approximate the Bit Error Probability (BEP) of the baseline (non-scrambled) and proposed (scrambled) OFDM transmissions are determined for different link configurations. The impact of the proposed physical-layer security scheme on the BEP is minimal, while increasing notably the number of combinations that an eavesdropper must check in order to execute a brute-force search attack.**

***Index Terms*—Physical-layer security; wireless communications; OFDM; subcarrier scrambling; performance analysis.**

I. Introduction

The broadcast nature of wireless communication systems makes them inherently vulnerable to eavesdropping, as non-authorized receivers can intercept the data traffic if they lie within the coverage range. Conventional Orthogonal Frequency-Division Multiplexing (OFDM) is the dominating waveform in contemporary mobile (LTE/4G and NR/5G) [1], [2] and wireless (Wi-Fi) communication standards [3]. Nevertheless, OFDM is highly vulnerable to these attacks due to its distinctive stochastic characteristics (*i.e.*, time- and frequency-domain correlations and second-order cyclostationarity) [4], which can be exploited by an eavesdropper to estimate the transmission parameters and infer the transmitted information. Therefore, built-in security features that are backward compatible and easily incorporated in such systems are highly desired to tackle the security flaws of OFDM while taking advantage of its high spectral efficiency and robustness against multipath.

This work received funding from the Spanish ministry of science and innovation under project IRENE PID2020-115323RB-C31 (AEI/FEDER,UE) and from the Deanship of Scientific Research at Jordan University of Science and Technology, Irbid, Jordan.

Communication systems traditionally rely on upper layer security techniques to encrypt, at the bit level, the exchanged information between authorized users [5]. Upper layer security ensures protection based on the assumption that the computational power that a non-legitimate user needs to break into the communication is limited. However, exploiting the limited received signal power that an eavesdropper observes, physical-layer security (PLS) approaches have emerged to complement cryptographic security schemes under the umbrella of layered defenses [6]. PLS techniques exploit the fact that: *a*) The signal-to-interference-plus-noise ratio (SINR) of the eavesdropper is lower than that the one of the authorized receiver (*i.e.*, keyless SINR-based approach); *b*) The properties of the shared channel between authorized transmitter and receiver, which are unique and can be used to generate the secret key for encryption (*i.e.*, key-based approach) [7]. Although these PLS techniques can enable high levels of security, many of them are computationally complex and hard to implement. Moreover, they may require considerable changes of the hardware and/or protocols of the wireless system, which are usually incompatible with the OFDM-based communication standards. Therefore, this paper proposes a backward compatible PLS scheme with no spectral efficiency loss, where the loading order of OFDM subcarriers is scrambled based on the data bits placed on the predefined subcarrier subset. When compared to the proposed technique, most previous PLS works suffer loss on the spectral efficiency and/or significant added complexity.

Subcarrier ordering techniques have been recently proposed for enhancing the security of 5G and beyond wireless systems [8]. For example, subcarrier index selection (SIS), a frequency-domain technique, was proposed in [9]. In SIS, the subcarrier scrambling sequence is selected based on the magnitude of the channel between legitimate users, assuming that the eavesdropper has no access to this information. Chaotic discrete Hartley transform (DFT) was suggested in [10] to encrypt modulated data in the downstream of an OFDM-based passive optical network (PON) [11]. Subcarrier obfuscation and training symbol resequencing (SOTSR) was presented in [12], reserving some subcarriers for dummy symbols while introducing certain levels of randomization such that an eavesdropper cannot synchronize the OFDM signal, estimate the channel between legitimate users, and break into the link.

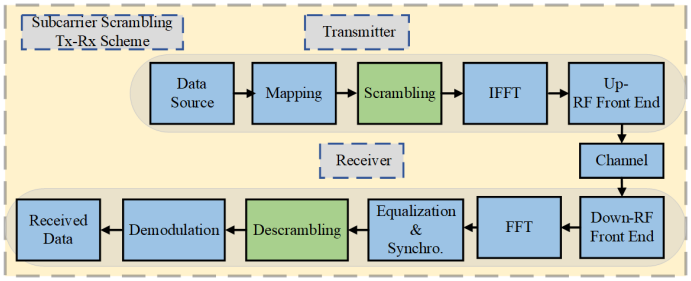

Fig. 1: Functional blocks of an simplified OFDM transmitter (upper blocks) and OFDM receiver (lower blocks). The green boxes show the position of the new proposed subcarrier scrambling and descrambling processing that is used to improve the physical-layer security.

In this paper, it is assumed that the permutation sequence of data subcarriers is based on the QAM symbols placed on indexing subcarriers. Indexing positions could be known *a priori*, or defined according to channel gains between authorized users. The mapping between indexing bits and scrambling sequences minimizes the error propagation on data subcarriers (*i.e.*, *Gray-mapped* sequences). Closed form Bit Error Probability (BEP) approximations for the proposed scheme are also derived. Thanks to this backward compatible PLS approach, the security of an OFDM system can be notably improved with negligible impact on the BEP performance.

The rest of the paper is organized as follows: Section II presents the system model and the key concept behind data-based scrambling. Section III explains the subcarrier scrambling process, whereas Section IV derives the closed form expressions that estimate the impact of the proposed PLS method in an OFDM system. Finally, Section V shows the obtained performance results and Section VI draws the conclusions.

II. System Model and Main Assumptions

The proposed data-based subcarrier scrambling scheme is suitable to improve the security of OFDM-based communications standards, such as 4G and 5G [2]. Fig. 1 illustrates the general block diagram of an OFDM system after adding the proposed subcarrier scrambling and descrambling stages (green boxes). In the OFDM transmitter, the scrambling block is inserted before the IFFT processing, whereas in the OFDM receiver, the descrambling process is placed after the equalizer.

Let N be the number of subcarriers that, when being a power of 2, enables the use of the the Cooley-Tuckey fast algorithm for OFDM encoding and decoding [13]. Let us assume that all the N subcarriers are loaded with M-QAM symbols, where $M = 2^b$ is the constellation size and b is the number of bits per symbol. Let $\mathbf{x} = [x_1 \cdots x_N]^{\mathrm{T}}$ be the vector of transmitted symbols, where $(\cdot)^{\mathrm{T}}$ denotes the transpose operation. The indexes of the first N_{s} symbols are scrambled by matrix $\mathbf{P} \in \mathbb{R}^{N_{\mathrm{s}} \times N_{\mathrm{s}}}$. The resulting symbols are referred to as the *scrambled symbols*, whereas the subcarriers that carry the scrambled symbols are known as *scrambled subcarriers*. Scrambling matrix $\mathbf{P}$ is determined by the data transported on the last N_{i} input symbols, known as *indexing symbols*. The subcarriers that carry the indexing symbols are known as the

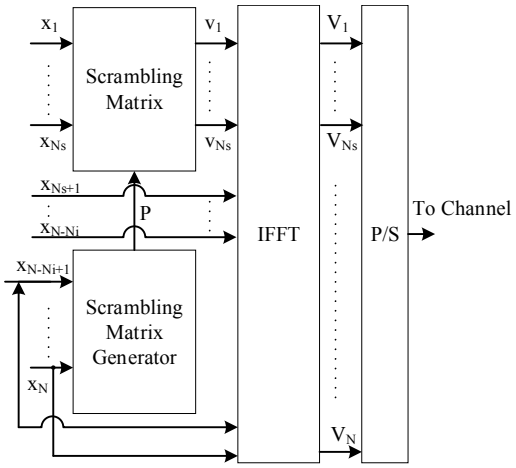

Fig. 2: Block diagram of the scrambled OFDM transmitter.

indexing subcarriers. The (non-scrambled) indexing symbols and the (scrambled) data symbols can be any two disjoint subsets of input symbols, such as the one shown in Fig. 2. The data bits carried by the indexing subcarriers are referred to as the *indexing bits, which are equal to* $b_{\mathrm{i}} = b\,N_{\mathrm{i}}$.

Since the number of sequences that can be identified with b_{i} indexing bits is $2^{b_{\mathrm{i}}}$, and the number of subcarrier scrambling sequences that can be obtained with N_{s} subcarriers is $N_{\mathrm{s}}!$, the number of subcarrier scrambling sequences, N_{scr}, must verify

$$N_{\mathrm{scr}} \leq \min(2^{b_{\mathrm{i}}}, N_{\mathrm{s}}!). \tag{1}$$

There are countless possible associations between indexing bits and subcarrier scrambling sequences that verify (1). The number of scrambled subcarriers is a function of b_{i}, *i.e.*, $N_{\mathrm{s}} = f(b_{\mathrm{i}})$. Ideally, it is convenient that $N_{\mathrm{s}} + N_{\mathrm{i}} = N$ is verified, such that each subcarrier is either an indexing or a data subcarrier. However, this cannot be the case for arbitrary b and N_{i}, since $f(b_i)$ can take many forms. Let us define

$$N_{\mathrm{a}} = N - (N_{\mathrm{i}} + N_{\mathrm{s}}) = N - N_{\mathrm{i}} - f(b_{\mathrm{i}}) \tag{2}$$

as the number of non-indexing OFDM subcarriers that are not scrambled, which can be different from 0 for given b and N_{i}. Thus, the total number of unscrambled subcarriers becomes $N_{\mathrm{u}} = N_{\mathrm{i}} + N_{\mathrm{a}}$, and the ratio between the number of unscrambled subcarriers and the total number of subcarriers is given by

$$\rho = \frac{N_{\mathrm{u}}}{N} = \frac{N_{\mathrm{i}} + N_{\mathrm{a}}}{N_{\mathrm{i}} + N_{\mathrm{a}} + N_{\mathrm{s}}}. \tag{3}$$

III. Subcarrier Scrambling Process

This section explains the concept of scrambling sequence generation, including a simplified algorithm that minimizes the error propagation due to indexing bit errors.

A. Generation of the scrambling matrix

Two sample algorithms are now presented to map the indexing bits into subcarrier scrambling sequences. Both algorithms require $b_{\mathrm{i}} = N_{\mathrm{s}} - 1$ indexing bits to generate the scrambling

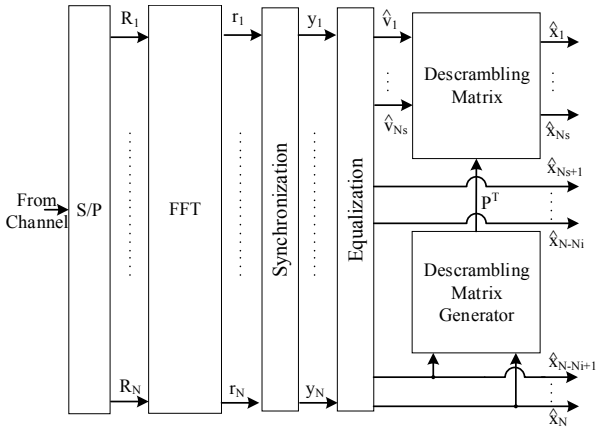

Fig. 3: Block diagram of the scrambled OFDM receiver.

Table I: Unscrambled symbols for different configurations.

b	N_i	b_i	N_scr	N_a	ρ (%)
6	9	54	$\approx 1.80 \times 10^{16}$	0	14.0625
6	6	36	$\approx 6.88 \times 10^{10}$	21	42.1875
6	3	18	262144	42	70.3125
4	12	48	$\approx 2.81 \times 10^{14}$	3	23.4375
4	9	36	$\approx 6.88 \times 10^{10}$	18	42.1875
4	6	24	$\approx 1.68 \times 10^{7}$	33	60.9375
2	21	42	$\approx 4.40 \times 10^{12}$	0	32.8125
2	15	30	$\approx 1.07 \times 10^{9}$	18	51.5625
2	9	18	262144	36	70.3125

matrix, which implies that $f(b_\mathrm{i}) = b_\mathrm{i} + 1$. Thus, the resulting number of subcarrier scrambling sequences is equal to

$$N_\mathrm{scr} = 2^{N_\mathrm{s}-1}. \tag{4}$$

Furthermore, based on (2), we have that

$$N_\mathrm{a} = N - N_\mathrm{i} - bN_\mathrm{i} - 1 = N - (b+1)N_\mathrm{i} - 1. \tag{5}$$

Table I shows the values that b_i, N_scr, N_a, and ρ can take for $N = 64$, when b and N_i are changed to obtain different configurations. From (2), we observe that the lower is the value of N_a, the larger are the values of N_i and N_s. From a PLS perspective, it is important to keep the number of unscrambled subcarriers low to augment N_scr as much as possible. Values of N_i that minimize ρ are shown in the shaded rows of Table I.

The b_i indexing bits are used by the scrambling matrix generator to produce matrix $\mathbf{P}$ of size $N_\mathrm{s} \times N_\mathrm{s}$. Algorithm 1 for scrambling matrix generation is based on Gray mapping. The algorithm starts with an identity matrix $\mathbf{P}$. If the first indexing bit is one, then rows 1 and 2 of $\mathbf{P}$ are exchanged. If not, $\mathbf{P}$ is left unchanged. Then, if the second indexing bit is one, rows 2 and 3 of $\mathbf{P}$ are exchanged . Otherwise, no row exchanges are made. The algorithm loops over the remaining indexing bits and acts on the corresponding rows of $\mathbf{P}$ in a similar manner. The above steps guarantee that if two distinct sequences of indexing bits differ in only one bit position, then the resulting scrambling matrices differ in only one pair of rows of $\mathbf{P}$ being exchanged. This is done on purpose, to minimize the effect that the reception of an erroneous indexing bit has on the overall BEP. Indexing bits are used by Algorithm 1 as a vector $\boldsymbol{\beta}_\mathrm{i}$ of size $(N_\mathrm{s} - 1) \times 1$.

Algorithm 1: Gray permutation matrix generation

Input : $N_\mathrm{s} - 1 \times 1$ vector of indexing bits $\boldsymbol{\beta}_\mathrm{i}$
Output: $N_\mathrm{s} \times N_\mathrm{s}$ permutation matrix $\mathbf{P}$

$\mathbf{P} \leftarrow \mathbf{I}_{N_\mathrm{s} \times N_\mathrm{s}}$ (identify matrix)
for $l \leftarrow 1$ **to** $N_\mathrm{s} - 1$ **do**
  $k \leftarrow l + 1$
  if $\boldsymbol{\beta}_\mathrm{i}(l) = 1$ **then**
    Exchange l-th and k-th rows of $\mathbf{P}$
  else
    Do nothing
  end
end

Algorithm 2: Non-Gray permutation matrix generation

Input : $N_\mathrm{s} - 1 \times 1$ vector of indexing bits $\boldsymbol{\beta}_\mathrm{i}$
Output: $N_\mathrm{s} \times N_\mathrm{s}$ permutation matrix $\mathbf{P}$

$\mathbf{P}_1 \leftarrow [\cdot]$ (empty matrix)
Initialize $\mathbf{P}_2$ to $\mathbf{u}_{L,L}$
for $l \leftarrow 1$ **to** $N_\mathrm{s} - 1$ **do**
  if $\boldsymbol{\beta}_\mathrm{i}(l) = 1$ **then**
    $\mathbf{P}_2 \leftarrow \begin{bmatrix}\mathbf{P}_2 & \mathbf{u}_{i,L}\end{bmatrix}$
  else
    $\mathbf{P}_1 \leftarrow \begin{bmatrix}\mathbf{P}_1 & \mathbf{u}_{i,L}\end{bmatrix}$
  end
end
$\mathbf{P} \leftarrow \begin{bmatrix}\mathbf{P}_1 & \mathbf{P}_2\end{bmatrix}^\mathrm{T}$

To assess the benefits of Gray mapping, Algorithm 2, which does not use Gray mapping, is considered. Vector $\mathbf{u}_{i,L}$ of size $L \times 1$ has a one in position i and zeros elsewhere. The algorithm starts with an empty matrix $\mathbf{P}_1$ and an $L \times 1$ matrix $\mathbf{P}_2 = \mathbf{u}_{L,L}$. Then, if the first indexing bit is one, it appends $\mathbf{u}_{1,L}$ at the end of $\mathbf{P}_1$. Otherwise, it appends $\mathbf{u}_{1,L}$ at the end of $\mathbf{P}_2$. If the second indexing bit is one, the algorithm appends $\mathbf{u}_{2,L}$ at the end of $\mathbf{P}_1$. Otherwise, it appends $\mathbf{u}_{2,L}$ at the end of $\mathbf{P}_2$. The algorithm loops over the remaining indexing bits and acts on $\mathbf{P}_1$ and $\mathbf{P}_2$ in a similar manner. Finally, the scrambling matrix is generated as $\mathbf{P} = \begin{bmatrix}\mathbf{P}_1 & \mathbf{P}_2\end{bmatrix}^\mathrm{T}$.

B. Signal model for the OFDM transmitter

Let $\mathbf{x} \in \mathbb{C}^{N \times 1}$ be the input vector of M-QAM symbols,

$$\mathbf{x} = \left[\mathbf{x}_\mathrm{s}^\mathrm{T} \mid \mathbf{x}_\mathrm{a}^\mathrm{T} \mid \mathbf{x}_\mathrm{i}^\mathrm{T}\right]^\mathrm{T}, \tag{6}$$

where $\mathbf{x}_\mathrm{s} \in \mathbb{C}^{N_\mathrm{s} \times 1}$ is the vector of M-QAM symbols to be scrambled, $\mathbf{x}_\mathrm{a} \in \mathbb{C}^{N_\mathrm{a} \times 1}$ is the vector of non-indexing symbols that are not scrambled, and $\mathbf{x}_\mathrm{i} \in \mathbb{C}^{N_\mathrm{i} \times 1}$ is the vector of indexing symbols. After the data-subcarrier permutation, the vector $\mathbf{v} \in \mathbb{C}^{N \times 1}$ that feeds the IFFT becomes

$$\mathbf{v} = \left[\mathbf{v}_\mathrm{s}^\mathrm{T} \mid \mathbf{v}_\mathrm{a}^\mathrm{T} \mid \mathbf{v}_\mathrm{i}^\mathrm{T}\right]^\mathrm{T}, \tag{7}$$

where $\mathbf{v}_\mathrm{s} \in \mathbb{C}^{N_\mathrm{s} \times 1}$ is the vector of scrambled symbols, *i.e.*,

$$\mathbf{v}_\mathrm{s} = \mathbf{P}\,\mathbf{x}_\mathrm{s}. \tag{8}$$

Then, vector $\mathbf{v}$ is fed into the IFFT block, and its output is transmitted over the channel towards the receiver (see Fig. 3).

C. Signal model for the OFDM receiver

The received signal vector after the FFT can be written as

$$\mathbf{r} = \mathbf{H}\,\mathbf{v} + \mathbf{z}, \tag{9}$$

where $\mathbf{H} \in \mathbb{C}^{N\times N}$ is a diagonal matrix that contains the channel gain of each subcarrier, *i.e.*, $\mathbf{H} = \text{diag}\{H_1, \ldots, H_N\}$, $\mathbf{v}$ is the vector with the input data of the IFFT block in the OFDM transmitter, and $\mathbf{z}$ denotes a vector of zero-mean Additive White Gaussian Noise (AWGN) samples with single-sided power spectral density N_0. After that, the received samples are synchronized and equalized. Let $\mathbf{y}$ be the vector with the data samples after the synchronization step, and let $\mathbf{G} \in \mathbb{C}^{N\times N}$ be a diagonal OFDM equalization matrix; then, the equalized received vector is computed as follows:

$$\widehat{\mathbf{v}} = \mathbf{G}\,\mathbf{y}, \qquad \mathbf{G} = \text{diag}\{G_1, \ldots, G_N\}. \tag{10}$$

When using MMSE criterion, the equalization coefficient for the k-th subcarrier is given by $G_k = \gamma_k H_k^*/(\gamma_k |H_k|^2 + 1)$, where γ_k is the signal-to-noise ratio (SNR) that corresponds to the subcarrier with index k. In case of zero-forcing criterion, $G_k = H_k^{-1}$ [14]. Thus, vector $\widehat{\mathbf{v}}$ provides the estimation of the impairment introduced by the channel, *i.e.*,

$$\widehat{\mathbf{v}} = \left[\widehat{\mathbf{v}}_{\rm s}^{\rm T} \mid \widehat{\mathbf{x}}_{\rm a}^{\rm T} \mid \widehat{\mathbf{x}}_{\rm i}^{\rm T}\right]^{\rm T}, \tag{11}$$

where $\widehat{\mathbf{v}}_{\rm s} \in \mathbb{C}^{N_{\rm s}\times 1}$, $\widehat{\mathbf{x}}_{\rm a} \in \mathbb{C}^{N_{\rm a}\times 1}$, and $\widehat{\mathbf{x}}_{\rm i} \in \mathbb{C}^{N_{\rm i}\times 1}$ denote the estimated vector of scrambled, unscrambled, and indexing versions of the transmitted symbols, respectively. Finally, the descrambling of vector $\widehat{\mathbf{v}}_{\rm s}$ results in

$$\widehat{\mathbf{x}}_{\rm s} = \mathbf{P}^{-1}\,\widehat{\mathbf{v}}_{\rm s} = \mathbf{P}^{\rm T}\,\widehat{\mathbf{v}}_{\rm s}. \tag{12}$$

Note that in (12), the replacement of $\mathbf{P}^{-1}$ by $\mathbf{P}^{\rm T}$ is valid because the scrambling matrix is a permutation matrix (see Algorithm 1). It is also possible to observe that the presence of channel impairments may introduce errors in the descrambling procedure at the legitimate receiver, increasing the BEP when compared to the no-scrambling case. To minimize the propagation of errors in such situation, the algorithm that obtains the permutation matrix is designed based on the Gray mapping principle. We are now ready to derive the BEP formulas for the proposed scrambling strategy.

IV. Bit Error Probability for an OFDM Symbol

Let $\boldsymbol{\gamma}$ be the vector that stacks the SNRs for all subcarriers, *i.e.*, $\boldsymbol{\gamma} = \left[\gamma_1 \cdots \gamma_N\right]^{\rm T}$. Then, the BEP when using M-QAM symbols (with Gray mapping) can be expressed as

$$P_{\rm b}(M,\boldsymbol{\gamma})_{\rm t} = \rho\, P_{\rm b}(M,\boldsymbol{\gamma})_{\rm ns} + (1-\rho)\, P_{\rm b}(M,\boldsymbol{\gamma})_{\rm s}, \tag{13}$$

where $P_{\rm b}(M,\boldsymbol{\gamma})_{\rm t}$, $P_{\rm b}(M,\boldsymbol{\gamma})_{\rm ns}$, and $P_{\rm b}(M,\boldsymbol{\gamma})_{\rm s}$ represent the BEP of the total number of subcarriers, the non-scrambled subcarriers, and the scrambled subcarriers, respectively. Specifically, $P_{\rm b}(M,\boldsymbol{\gamma})_{\rm ns}$ corresponds to the BEP of the M-QAM subcarriers, which can be computed as

$$P_{\rm b}(M,\boldsymbol{\gamma})_{\rm ns} = \frac{1}{N_{\rm u}} \sum_{k=N_{\rm s}+1}^{N} P_{{\rm b},k}(M,\gamma_k), \tag{14}$$

being $P_{{\rm b},k}(M,\gamma_k)$ the BEP for an M-QAM transmission on the k-th non-scrambled subcarrier with SNR γ_k. The exact BEP theoretical formula in case of AWGN is given by [15]

$$P_{\rm b}(M,\gamma_k) = \sum_{m=1}^{b} \frac{1}{b\sqrt{M}} \sum_{l=0}^{(1-2^{-k})\sqrt{M}-1} P_{\rm b}(M,\gamma_k)_{m,l}, \tag{15}$$

where $P_{\rm b}(M,\gamma_k)_{m,l}$ attains the form

$$P_{\rm b}(M,\gamma_k)_{m,l} = (-1)^{\left\lfloor \frac{l\,2^{m-1}}{\sqrt{M}} \right\rfloor} \left(2^{m-1} - \left\lfloor \frac{l\,2^{m-1}}{\sqrt{M}} + \frac{1}{2} \right\rfloor\right) \times \text{erfc}\left((2l+1)\sqrt{\frac{3\log_2(M)\gamma_k}{2(M-1)}}\right), \tag{16}$$

where $\lfloor x \rfloor$ is the greatest integer less than or equal to x, and $\text{erfc}(x) = 1/\sqrt{2\pi}\int_x^\infty \exp(-t^2)\,dt$ is the complementary error function. The BEP of the scrambled subcarriers is equal to

$$P_{\rm b}(M,\boldsymbol{\gamma})_{\rm s} = \Pr\{q=0\} \sum_{k=1}^{N_{\rm s}} \frac{P_{{\rm b},k}(M,\boldsymbol{\gamma})}{N_{\rm s}} + \sum_{m=1}^{b_{\rm i}} \Pr\{q=m\} \sum_{k=1}^{N_{\rm s}} \frac{P_{{\rm b},k}(M,\boldsymbol{\gamma}|q=m)}{N_{\rm s}}, \tag{17}$$

where $\Pr\{q=m\}$ is the probability of having m bits in error in the $b_{\rm i}$ indexing bits (carried by the $N_{\rm i}$ indexing subcarriers), and $P_{{\rm b},k}(M,\boldsymbol{\gamma}|q=m)$ is the BEP of the M-QAM symbol that is carried by the subcarrier with index k, assuming that an q-bit-error event occurred in the indexing bits. Note that in the proposed data-based scrambling scheme, $q = 0, 1, \ldots, b_{\rm i}$.

When the SNR grows large, the most likely error events when using M-QAM symbols with Gray mapping are single-bit-error events; in such situation, an q-bit-error event in the indexing bits can be approximated as having simultaneously q M-QAM symbols in error on the indexing subcarriers. At moderate to large SNR values, it can be assumed that the number of indexing subcarriers $N_{\rm i}$ is much larger than the simultaneous number of errors in the $b_{\rm i}$ indexing bits. Then, if the mapping between indexing bits and scrambling sequence is selected such that m bits in errors change the position (after descrambling) of $m+1$ QAM symbols in the data subcarriers, it is possible to approximate

$$P_{{\rm b},k}(M,\boldsymbol{\gamma}|q=m) \;\cong\; \left(\frac{N_{\rm s}-m-1}{N_{\rm s}}\right) P_{{\rm b},k}(M,\boldsymbol{\gamma}) + \frac{1}{2}\left(\frac{m+1}{N_{\rm s}}\right). \tag{18}$$

The probability of m erroneous indexing bits is given by

$$\Pr\{q=m\} = \binom{b_{\rm i}}{m} \left\{P_{{\rm b},k}(M,\boldsymbol{\gamma})\right\}^m \left\{1 - P_{{\rm b},k}(M,\boldsymbol{\gamma})\right\}^{b_{\rm i}-m}, \tag{19}$$

where $\binom{n}{k}$ is the binomial coefficient of n and k. Thus, the increment on the BEP that is reached by using the proposed subcarrier scrambling strategy can be written as

$$\Delta P_b(M,\boldsymbol{\gamma}) = P_b(M,\boldsymbol{\gamma})_{\rm t} - \frac{1}{N}\sum_{k=1}^{N} P_{{\rm b},k}(M,\gamma_k). \tag{20}$$

For the sake of simplicity, the channel is modelled as frequency flat (*i.e.*, all subcarriers have the same average SNR $\overline{\gamma}$). In this case, when $\overline{\gamma}$ grows large, it is highly probable that there is only a single erroneous subcarrier after the descrambling. Thus, it is possible to show that (13) reduces to

$$P_{\mathrm{b}}(M,\overline{\gamma})_{\mathrm{t}} \approx \frac{\Big(N_{\mathrm{u}} + N_{\mathrm{s}}\Pr\{q=0\}\Big)P_{\mathrm{b},k}(M,\overline{\gamma})}{N} + \frac{\Big((N_{\mathrm{s}}-2)P_{\mathrm{b},k}(M,\overline{\gamma})+1\Big)\Pr\{q=1\}}{N}. \tag{21}$$

Note that when $\overline{\gamma} \to \infty$ in (21), then $\Pr\{q=0\} \to 1$ and $\Pr\{q=1\} \to 0$; therefore, the total BEP, $P_{\mathrm{b}}(M,\overline{\gamma})_{\mathrm{t}}$, tends to the BEP of the M-QAM scheme, which is $P_{\mathrm{b},k}(M,\overline{\gamma})$ (note that $N = N_{\mathrm{u}} + N_{\mathrm{s}}$). As a result, at very high SNR values, the gap between the total BEP and the M-QAM BEP tends to zero (*i.e.*, $\Delta P_{\mathrm{b}}(M,\overline{\gamma}) \to 0$). In this situation, from the perspective of the legitimate user, the BEP of its M-QAM symbols (placed either on scrambled or non-scrambled subcarriers) is practically the same.

V. Simulation Results and Performance Analysis

This section presents the BEP curves for the legitimate user when using the proposed scrambling algorithms, and studies the SNR gap that is observed with respect to the baseline (non-scrambled) case at different target BEP values. In all cases, the number of subcarriers is set to $N = 64$, the constellation size of the M-QAM symbols is $M = 4$, 16, and 64, and the SNR is varied from 0 to 27 dB in steps of 1 dB. For the sake of simplicity, the use case could be a 4G (LTE) or 5G (NR) waveform, with 15 kHz sub-carrier spacing, 1 MHz bandwidth, and zero-padded subcarriers placed next to the passband edges.

A. Effect of the modulation order (M) and the number of indexing subcarriers (N_{i}) on the BEP performance

The larger the number of indexing subcarriers N_{i}, the higher the number of subcarrier scramblings sequences N_{scr} that an eavesdropper must check to break the subcarrier scrambling algorithm with a brute-force search attack. Unfortunately, the larger is the number of indexing subcarriers, the higher is the SNR loss of the legitimate user to achieve the same BEP when compared to the non-scrambled OFDM case.

To visualize this effect, Fig. 4 shows the overall BEP of the Gray-mapped scrambling strategy (Algorithm 1) for different number of indexing subcarriers. The cases under study are:

- $N_{\mathrm{i}} = 9$, 15, and 21 for 4-QAM,
- $N_{\mathrm{i}} = 6$, 9, and 12 for 16-QAM, and
- $N_{\mathrm{i}} = 3$, 6, and 9 for 64-QAM.

Results from Fig. 4 show that, regardless of the number of indexing subcarriers, the BEP for all configurations under study is very similar. Note that this observation becomes more evident when zooming-in the BEP curve for SNR values between 18 and 21 dB. Thus, when a high level of PLS is desirable (*i.e.*, largest possible N_{i}), the utilization of the proposed Gray-mapped scrambling algorithm provides an overall BEP that does not vary notably with respect to the minimum security level case (*i.e.*, lowest indexing subcarriers).

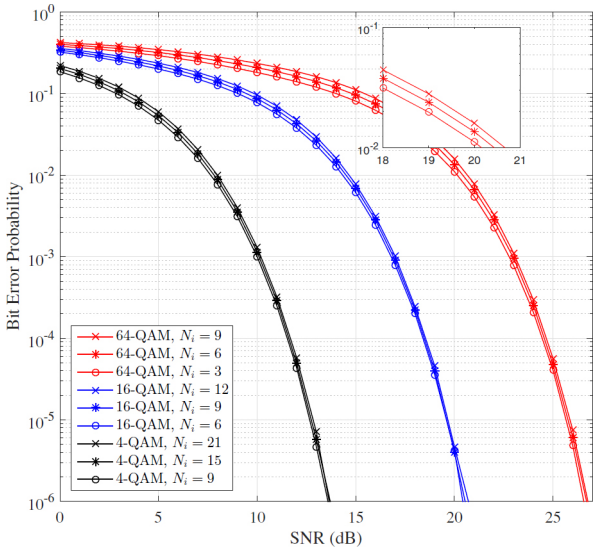

Fig. 4: Overall BEP versus mean SNR per subcarrier for different modulation orders and numbers of indexing subcarriers.

B. Comparison of the BEP for the different data-dependent subcarrier scrambling algorithms

Fig. 5 compares the BEP when using the two proposed subcarrier scrambling algorithms, namely Gray-mapped (Algorithm 1) and non-Gray-mapped (Algorithm 2). Fig. 5 also shows the closed form BEP curve obtained with (14) (no subcarrier scrambling), the closed form approximation in (13) for the overall BEP when using Gray-mapped scrambling in the data subcarriers, and the simulated BEP when using the data transmitted on the indexing bits to select the permutation matrix. The number of indexing subcarriers was kept constant to $N_{\mathrm{i}} = 9$, regardless the order M of the QAM scheme.

Based on the obtained results, it is possible to conclude that the indexing bits experience the same BEP of the theoretical M-QAM formula, as their order of transmission is not changed by the permutation matrix. Concerning the overall BEP when using subcarrier scrambling with(out) Gray mapping, it is observed that the Gray subcarrier scrambling provides a much lower SNR for a target BEP when compared to non-Gray case. Moreover, the simulated overall BEP for medium to high mean SNR values is almost identical to the case when approximation (13) is used, validating the mathematical derivation that was presented. However, when comparing the simulated overall BEP for Gray mapping (*i.e.*, *Overall BEP Alg.1* legend in Fig. 5) with respect to the cases when no subcarrier scrambling is used (*i.e.*, *Overall BEP No Scrambling* legend in Fig. 5), a minor SNR penalty is observed. For a more detailed observation of this effect, a zoom-in window that magnifies the BEP for 64-QAM when the mean SNR varies from 18 to 21 dB is also presented. Regarding the eavesdropper, it has to conduct at least a double force-brute attack to read the message: i) to the PLS strategy, and ii) to the network encryption process.

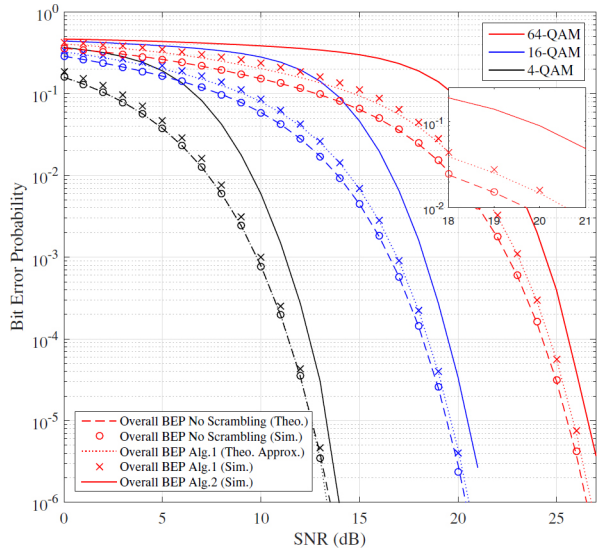

Fig. 5: Overall BEP vs. mean SNR for the different subcarrier scrambling algorithms and modulation orders. No scrambling (Theo.) (dashed lines); No scrambling (Sim.) (circles); Alg. 1 (Theo. Approx.) (dotted lines); Alg. 1 (Sim.) (crosses); Alg. 2 (Sim.) (solid lines).

C. SNR penalty originated by the subcarrier scrambling

Gray mapping subcarrier scrambling (Algorithm 1) provides a lower SNR penalty at any target BEP, when compared to the non-Gray mapped case (Algorithm 2). To quantify this effect, Tables II and III show the SNR penalty of the Gray and non-Gray subcarrier scrambling approaches, respectively, using the closed form BEP formula in (14) as reference (baseline M-QAM transmission). The values reported in these tables show that the higher is the modulation order M or the lower is the target BEP to be achieved, the larger is the SNR penalty that is observed. The same trend is observed when comparing the performance of Gray subcarrier scrambling (Algorithm 1) with respect to the non-Gray based scrambling (Algorithm 2).

For 64-QAM, the use of Gray subcarrier scrambling requires a mean SNR that is 2.2 dB and 0.8 dB lower than the non-Gray based case to obtain a BEP of 10^{-2} and 10^{-5}, respectively. Similar values are obtained for other modulation orders. These results confirm that Gray subcarrier scrambling results in a negligible SNR penalty. Note that this is a minor cost to pay in favor of having a larger PLS level. On the contrary, the use of non-Gray subcarrier scrambling provides an excessive penalization in terms of SNR.

Table II: SNR penalty of Gray subcarrier scrambling.

	BEP $= 10^{-2}$	BEP $= 10^{-3}$	BEP $= 10^{-4}$	BEP $= 10^{-5}$
$M = 64$	0.9 dB	0.6 dB	0.4 dB	0.3 dB
$M = 16$	0.6 dB	0.4 dB	0.25 dB	0.2 dB
$M = 4$	0.25 dB	0.2 dB	0.15 dB	0.125 dB

VI. Conclusion

This paper proposed, designed, and analyzed the performance of a data-based subcarrier scrambling algorithm for secured OFDM systems. The key idea was to make more difficult for an eavesdropper to detect correctly the data transmitted on the OFDM subcarriers. The proposed PLS approach does not require secret keys to be exchanged between authorized users, as scrambling sequence is determined based on actual data. Closed form approximations were derived for estimating the overall BEP at the legitimate receiver. Furthermore, by conducting extensive simulations, it was shown that similar BEP performance was observed for different numbers of indexing subcarriers when using a Gray subcarrier scrambling sequences. Therefore, the number of indexing subcarriers can be made as large as desired, since the penalty loss in terms of SNR is minimal. Finally, it was concluded that there is a substantial BEP performance benefit for using the principle of Gray mapping when creating the permutation matrix.

Table III: SNR penalty of non-Gray subcarrier scrambling.

	BEP $= 10^{-2}$	BEP $= 10^{-3}$	BEP $= 10^{-4}$	BEP $= 10^{-5}$
$M = 64$	3.1 dB	1.9 dB	1.4 dB	1.1 dB
$M = 16$	2.7 dB	1.7 dB	1.3 dB	0.95 dB
$M = 4$	2.15 dB	1.4 dB	1.0 dB	0.75 dB

References

[1] 3GPP, "TS 38.211 LTE. Evolved Universal Terrestrial Radio Access (E-UTRA). Physical channels and modulation," Tech. Rep. v.11.5, 2014.

[2] ——, "TS 38.211 5G; NR; Physical Channels and modulation," 3rd Generation Partnership Project (3GPP), Tech. Rep. v.15.3, 2018.

[3] IEEE, "Standard for Infor. Tech,–Telecom. and Infor. Exchange between Sys. - LAN and MAN–Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Spec." 2020.

[4] H. Li, X. Wang, and J.-Y. Chouinard, "Eavesdropping-resilient OFDM system using sorted subcarrier interleaving," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 1155–1165, Feb. 2015.

[5] K. Ren, H. Su, and Q. Wang, "Secret key generation exploiting channel characteristics in wireless communications," *IEEE Wireless Commun.*, vol. 18, no. 4, pp. 6–12, Aug. 2011.

[6] NIST, "Security and privacy controls for information systems and organizations," no. 5, pp. 1–492, Sept. 2020.

[7] J. Hamamreh, H. Furqan, and H. Arslan, "Classifications and applications of physical layer security techniques for confidentiality: A comprehensive survey," *IEEE Commun. Surv. & Tut.*, vol. 21, no. 2, pp. 1773–1828, 2Q 2019.

[8] T. Sedan, A. Tusha, E. Basar, and H. Arslan, "Multidimensional index modulation for 5G and beyond wireless networks," *IEEE Proceedings*, vol. 109, no. 2, pp. 170–199, Feb. 2020.

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

[10] A. Hajomer, X. Yang, and W. Hou, "Secure OFDM transmission precoded by chaotic discrete Hartley transform," *IEEE Photon. J.*, vol. 10, no. 2, pp. 1–9, Apr. 2018.

[11] M. Cheng, L. Deng, X. Gao, and H. Li, "Security-enhanced OFDM-PON using hybrid chaotic system," *IEEE Photon. Tech. Lett.*, vol. 27, no. 3, pp. 326–329, Feb. 2015.

[12] J. Zhang, A. Marshall, R. Woods, and T. Dong, "Design of an OFDM physical layer encryption scheme," *IEEE Trans. Veh. Tech.*, vol. 66, no. 3, pp. 2114–2126, Mar. 2017.

[13] J. Stoer and R. Burlisch, *Introduction to Numerical Analysis*. Springer-Verlag, 2017.

[14] M. Ahmed, S. Boussakta, B. Sharif, and C. Tsimenidis, "OFDM based on low complexity transform to increase multipath resilience and reduce PAPR," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5994–6007, Dec. 2011.

[15] K. Cho and D. Yoon, "On the general BER expression of one- and two-dimensional amplitude modulations," *IEEE Trans. Commun.*, vol. 50, no. 7, pp. 1074–1080, July 2002.