Energy Efficient OFDM with Intelligent PAPR-aware Adaptive Modulation

Article *in* IEEE Communications Letters · December 2023

DOI: 10.1109/LCOMM.2023.3324137

4 authors, including:

Nikos A. Mitsiou
Aristotle University of Thessaloniki
35 PUBLICATIONS **119** CITATIONS
SEE PROFILE

Panagiotis G. Sarigiannidis
University of Western Macedonia
474 PUBLICATIONS **7,916** CITATIONS
SEE PROFILE

George K. Karagiannidis
Aristotle University of Thessaloniki
853 PUBLICATIONS **35,236** CITATIONS
SEE PROFILE

Energy Efficient OFDM with Intelligent PAPR-aware Adaptive Modulation

Nikos A. Mitsiou, *Graduate Student Member, IEEE*, Panagiotis D. Diamantoulakis, *Senior Member, IEEE*, Panagiotis G. Sarigiannidis, *Member, IEEE*, and George K. Karagiannidis, *Fellow, IEEE*

***Abstract*—In this paper, we propose a novel peak-to-average power ratio (PAPR)-aware adaptive modulation scheme for Orthogonal Frequency Division Multiplexing (OFDM). The aim of the proposed PAPR-aware protocol is to maximize the energy efficiency of the OFDM transmission, where the energy efficiency is defined as the ratio of the OFDM throughput to the PA consumption, and is thus strongly affected by PAPR. For this purpose, we formulate the energy efficiency maximization problem with respect to power and adaptive modulation constraints per subcarrier. Due to the joint optimization of power allocation, adaptive modulation, and the non-convex terms introduced by PAPR, the problem cannot be solved by using convex optimization tools. Therefore, a solution based on online deep learning is proposed. Numerical results showcase that the energy efficiency of the proposed PAPR-aware protocol is greater up to 3 dB compared to the conventional PAPR-unaware protocol, while the PAPR is also reduced by 3 dB.**

***Index Terms*—orthogonal frequency-division multiplexing (OFDM), peak-to-average ratio (PAPR), online learning, 6G.**

I. Introduction

Recent studies have highlighted the urgent need to reduce the environmental impact of future networks [1], [2]. To that end, sixth generation (6G) wireless networks are expected to be 10-100 times more energy efficient than the fifth generation (5G) networks [1]. In addition, 6G networks are expected to offer increased capacity by utilizing higher frequency bands, including Millimeter (MM) and Terahertz (THz) waves. However, the increase in available spectrum may limit the energy efficiency of 6G networks due to high path loss and reduced power amplifier (PA) efficiency [1].

Orthogonal Frequency Division Multiplexing (OFDM), has been identified as a key enabler of beyond 5G and 6G networks [1]. However, the energy efficiency of OFDM is limited by the high peak-to-average power ratio (PAPR), which is a fundamental problem of OFDM and increases with the number of data subcarriers [1], [3]. Therefore, the design of energy-efficient medium access control (MAC) protocols that optimize the physical layer (PHY-L) of the network, taking into account hardware constraints, is imperative for 6G networks. To that end, 6G networks will aim to build intelligence into the PHY-L and MAC layers of networks and provide new functionalities such as intelligent power control, by relying on native artificial intelligence (AI) [2]. In fact, AI-based optimization is a viable approach to solve practical resource allocation problems where the impact of hardware cannot be ignored [1], [2].

PAPR reduction and adaptive modulation in OFDM transmission has been studied extensively in the past, but there is currently no work which combines both. In [4] and [5] PAPR minimization was formulated as a convex optimization problem, but none of the studies took into account the data rate of the OFDM transmission. Similarly, in [6], [7] different clipping methods have been proposed that can reduce PAPR, but adaptive modulation was not considered. Furthermore, in [8], [9], adaptive modulation for OFDM was studied, but the PAPR of the OFDM time-domain signal was not investigated. Moreover, deep learning-based adaptive modulation OFDM protocols were designed in [10]–[12], but the PAPR of the OFDM transmission was not investigated.

In this work, a deep neural network (DNN)-based PAPR-aware adaptive modulation and power allocation scheme is proposed. The objective of this work is to maximize the energy efficiency of OFDM transmission, which in our case is defined as the ratio of throughput to average power consumption of the PA. We note that the PA's consumption depends on the PAPR, while the PAPR depends on the power allocation and modulation per subcarrier. The problem is intractable by standard optimization tools, due to the adaptive modulation following a categorical distribution and the non-convex terms introduced by the PAPR. Thus, an online learning and an unsupervised learning solution are given. Numerical simulations validate the improved energy efficiency of the proposed protocol compared to the typical PAPR-unaware adaptive modulation protocol.

II. System Model

We consider a point-to-point communication link between an access point (AP) and a user. There are K available data transmission subcarriers. The set of usable constellations is denoted as $\mathcal{J}$, with cardinality $|\mathcal{J}|=J$, and the order of each constellation as $M(j), \forall j \in \mathcal{J}$. The AP employs OFDM to communicate with the user and we assume that the AP has a maximum power constraint, denoted as $P_{\max}$. The AP utilizes square M-QAM constellations, while B is the bandwidth of the received signal at each subcarrier. The following analysis can be generalized to any constellation type.

A. OFDM Transmission

Let $\mathcal{C}_j$, $j \in \mathcal{J}$, be the j-th frequency-domain OFDM constellation, with order $M(j)$. Then, let $c_{k,j} \in \mathbb{C}_j$ be a symbol taken from the j-th constellation, that will be transmitted in the k-th subcarrier. We assume that all $c_{k,j}$ are chosen

N. A. Mitsiou, P. D. Diamantoulakis, and G. K. Karagiannidis are with the Wireless Communications and Information Processing (WCIP) Group, Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, 54636, Thessaloniki, Greece (e-mail: nmitsiou@ece.auth.gr, padiaman@ieee.org, geokarag@auth.gr).

P. G. Sarigiannidis is with the Department of Informatics and Telecommunications Engineering, University of Western Macedonia, 501 00 Kozani, Greece (e-mail: psarigiannidis@uowm.gr).

G. K. Karagiannidis is also with the Cyber Security Systems and Applied AI Research Center, Lebanese American University (LAU), Lebanon.

This work has received funding from the European Union's Horizon Europe Framework Programme under grant agreement No 101096456.

uniformly random from the j-th constellation $\mathcal{C}_j$. We note that the coordinates of any symbol $c_{k,j}$ are given as follows [3]

$$\mathrm{Re}\{c_{k,j}\} = \frac{2\psi - 2D - 1}{\Delta}, \ \mathrm{Im}\{c_{k,j}\} = \frac{2\zeta - 2D - 1}{\Delta}, \quad (1)$$

where $D = 2^{0.5\log_2(M(j)/4)}$, $\psi, \zeta \in [1, ..., D]$ and Δ is a normalization factor, which is chosen such that $\mathbb{E}[|c_{k,j}|^2] = 1$. For a square M-QAM, Δ is given below

$$\Delta = \sqrt{\frac{2}{3}(M(j) - 1)}. \quad (2)$$

Moreover, we define $a_{k,j}$ to be an index which equals one if the k-th subcarrier is scheduled to the j-th constellation, and zero otherwise. Therefore, the frequency-domain OFDM symbol is given as follows

$$\boldsymbol{c}_s = \sum_{j=1}^{J} \boldsymbol{c}_j \odot \boldsymbol{a}_j, \quad (3)$$

where $\boldsymbol{a}_j = [a_{1,j}, a_{2,j}, ..., a_{K,j}]$ and $\odot$ denotes the Hadamard product. We define $\boldsymbol{c}_j = [c_{1,j}, c_{2,j}, ..., c_{K,j}]$, $\boldsymbol{c}_j \in \mathbb{C}^K$, to be the vector containing the symbols of the j-th constellation that will be used to the k-th subcarrier, when $a_{k,j}$ equals one.

Moreover, it is assumed that the AP has the ability to allocate different average power levels at each subcarrier. Originally, the average power of the OFDM symbol in (3) equals one for all subcarriers, since the power of all constellations is normalized. However, by allocating different average power per subcarrier, the constellation of the k-th subcarrier ends up with different average power than the constellation of the k'-th subcarrier. Let us denote with $\boldsymbol{p} = [p_1, p_2, ..., p_K]$ the vector containing the power allocated at each subcarrier. This implies that $\mathbb{E}[|c_{k,j}|^2] = \boldsymbol{p}[k]$. Then, the overall frequency-domain OFDM symbol is given as

$$\boldsymbol{c} = \boldsymbol{c}_s \odot \sqrt{\boldsymbol{p}}. \quad (4)$$

In order to generate the time-domain OFDM signal, $\boldsymbol{x} \in \mathbb{C}^K$, the frequency components are converted into time samples by performing an inverse FFT (IFFT) on these K frequency-domain symbols, which is given below

$$\boldsymbol{x}[i] = \frac{1}{\sqrt{K}} \sum_{k=1}^{K-1} \boldsymbol{c}[k] e^{\frac{j2\pi ki}{K}}, i \in \{0, ..., K-1\}. \quad (5)$$

Therefore, the received frequency-domain signal at the input of the QAM demodulator, after the FFT, is given as [3]

$$\boldsymbol{Y}[k] = \sqrt{\boldsymbol{p}[k]}\boldsymbol{H}[k]\boldsymbol{c}_s[k] + \boldsymbol{N}[k], \forall k \in K, \quad (6)$$

where $\boldsymbol{H} \in \mathbb{C}^{K\times 1}$ is the flat fading channel coefficient associated with the k-th subcarrier, and $\boldsymbol{N} \in \mathbb{C}^K$ is the FFT of the time-domain noise signal at the k-th subcarrier. We note that the power allocation per subcarrier is constant throughout the whole coherence time of the channel, and is considered known at the receiver. We also assume additive white Gaussian noise (AWGN) with power spectral density equal to N_0. Thus, the average signal-to-noise ratio (SNR) at the k-th subcarrier is given as

$$\gamma_k = \frac{\boldsymbol{p}[k]\,|\boldsymbol{H}[k]|^2}{N_0 B}. \quad (7)$$

Moreover, the average symbol error rate (SER) of a square M-QAM with Gray bit mapping, as a function of the received SNR, γ_k, is given as [3]

$$\mathrm{P}^n_{k,j} = 1 - \left(1 - \frac{2\left(\sqrt{M(j)} - 1\right)}{\sqrt{M(j)}} \mathcal{Q}\left(\sqrt{\frac{3\gamma_k}{M(j) - 1}}\right)\right)^2. \quad (8)$$

B. PAPR Estimation of the OFDM Signal

Next, we define the discrete-time PAPR of the OFDM signal, as follows [13]

$$\omega = \frac{\max_i |\boldsymbol{x}[i]|^2}{\frac{1}{lK}\sum_{i=0}^{K-1} |\boldsymbol{x}[i]|^2} = \frac{\|\boldsymbol{x}\|^2_\infty}{\frac{1}{K}\|\boldsymbol{x}\|^2_2}. \quad (9)$$

The PAPR of the discrete time sequences typically determines the complexity of the digital circuitry. However, in practice, we are often more concerned with reducing the PAPR of the continuous-time signals, since the cost and power dissipation of the analog components often dominate. Thus, to better approximate the PAPR of the continuous-time OFDM signal, the OFDM signal samples will be obtained by l-times oversampling [13]. Thus, we define $\boldsymbol{x}_l \in \mathbb{C}^{lK}$, where l is the oversampling factor, to be the corresponding oversampled time-domain OFDM signal, which can be calculated by the inverse FFT (IFFT) of $\tilde{\boldsymbol{c}}$ as follows

$$\boldsymbol{x}_l[i] = \frac{1}{\sqrt{K}} \sum_{k=1}^{lK-1} \tilde{\boldsymbol{c}}[k] e^{\frac{j2\pi ki}{lK}}, i \in \{0, ..., lK-1\}, \quad (10)$$

where $\tilde{\boldsymbol{c}}$ occurs from zero-padding $(l-1)K$ zeros to $\boldsymbol{c}$, as shown

$$\tilde{\boldsymbol{c}} = \left[\boldsymbol{c}[0], ..., \boldsymbol{c}\left[\tfrac{K}{2} - 1\right], 0, ..., 0, \boldsymbol{c}\left[\tfrac{K}{2} - 1\right], ..., \boldsymbol{c}[K-1]\right]. \quad (11)$$

Then, the discrete time-domain PAPR will be given by

$$\omega = \frac{\|\boldsymbol{x}_l\|^2_\infty}{\frac{1}{lK}\|\boldsymbol{x}_l\|^2_2}. \quad (12)$$

We note that $l \geq 4$ is sufficient to approximate the continuous-time PAPR via its discrete-time representation [13].

The total consumed power at the PA can be modeled as follows [13], [14]

$$P_{\mathrm{cons}}(\eta) = \frac{1}{\eta}\left(\sum_{k=1}^{K} \boldsymbol{p}[k]\right)^{\epsilon} P^{1-\epsilon}_{\max}, \quad (13)$$

where $\epsilon \in [0, 0.64]$ [14], and η is the PA efficiency, which for an ideal class B PA is given as

$$\eta = \frac{\pi}{4\omega}. \quad (14)$$

We note that the presented PAPR-aware adaptive modulation is also valid for different PA consumption models. It is obvious that to increase the PA's energy efficiency, PAPR has to be as low as possible. Nevertheless, it is noticed that PAPR, and so the PA's efficiency, are random variables, since they depend on the transmitted symbols. Therefore, we denote the expected PA energy efficiency as follows

$$\bar{\eta} = \mathbb{E}\left[\eta^{-1} \middle| \boldsymbol{p}, \boldsymbol{a}\right] = \frac{1}{I}\sum_{i=1}^{I}\left(\frac{\pi}{4\omega_i}\right)^{-1}, \quad (15)$$

where ω_i is the discrete PAPR of the i-th realization, the expectation is taken with respect to all possible transmitted symbols $\boldsymbol{c}$, and I denotes the number of all possible symbol combinations. We note that the average PA's efficiency can be calculated offline by a Monte Carlo approach, while the average PA power consumption, $P_{\text{cons}}(\bar{\eta})$, is given by (13) by substituting η^{-1} with $\bar{\eta}$, since $\sum_{k=1}^{K} \boldsymbol{p}[k]$ is constant for all possible symbol combinations I.

III. Adaptive Modulation and Power Optimization

A. Problem Formulation

In this section we aim to maximize the energy efficiency of the OFDM transmission by proper adaptive modulation and power allocation per subcarrier. The SER-based throughput of the uncoded OFDM transmission is given as

$$R = \sum_{k=1}^{K} \sum_{j=1}^{J} a_{k,j} \left(1 - \mathrm{P}_{k,j}^{n}\right) \log_2 M(j). \tag{16}$$

We note that in the low SNR regime the BER approximates 0.5 for all constellation orders [3]. Thus, the BER-based throughput appears bigger for greater constellation orders. However, this is incorrect, since the SER metric approaches zero more rapidly for greater constellation orders [3]. As such, the BER-based throughput cannot capture the considered trade-off between the transmission rate and error rate in the low SNR regime, and the SER-based throughput is used instead. Then, the energy efficiency maximization problem can be formulated as follows

$$\begin{aligned} & \max_{\tilde{\boldsymbol{p}}, \boldsymbol{a}} \quad \frac{R}{P_{\text{cons}}(\bar{\eta})} \\ \text{s.t.} \quad & \mathrm{C}_1: \quad \sum_{k=1}^{K} \tilde{\boldsymbol{p}}[k] \leq 1 \\ & \mathrm{C}_2: \quad \sum_{j=1}^{J} \alpha_{k,j} = 1, \forall k \in \{1, .., K\} \\ & \mathrm{C}_3: \quad \alpha_{k,j} \in \{0,1\}, \forall k \in \{1, .., K\}, \forall j \in \mathcal{J}, \end{aligned} \tag{17}$$

where constraint C_1 limits the overall transmission power of the OFDM transmitter and constraint C_2 imposes that each subcarrier is scheduled with only one constellation., while a constellation can be allocated to many subcarriers. For convenience we have set that $\tilde{\boldsymbol{p}} = \frac{\boldsymbol{p}}{P_{\max}}$.

B. Problem Solution

A possible approach to find the optimal solution of (18) is exhaustive search. However, this is not a viable option even for the case of 16 subcarriers, since there are approximately a total of 10^6 constellation-subcarrier associations for $M \in \{4, 16, 64, 256, 1024\}$, while problem (18) is non-convex with respect to power allocation due to PAPR and is not tractable by standard convex optimization methods. Therefore, motivated by the encouraging results of DNN in addressing non-convex problems we propose an online learning-based and an unsupervised learning-based solution.

1) Online Learning: Given the instant channel state information (CSI) $\boldsymbol{H}$, the DNN is trained in an online fashion to estimate the optimal variables, i.e., the transmit power allocation $\tilde{\boldsymbol{p}}$ and the modulation per subcarrier $\boldsymbol{a}$. It is clarified that unlike to offline training approaches where multiple samples are used for training and afterwards the testing stage follows, online learning has no testing stage, since for each new generated sample a new dedicated network is trained, and thus, the generalization issue does not exist. However, online learning cannot be used to first-seen data samples. The input of the online DNN, $\boldsymbol{F} \in \mathbb{R}_{+}^{K \times 1}$, will be given as $\boldsymbol{F} = \boldsymbol{H} \odot \boldsymbol{H}^*$, while the DNN's loss function is defined as the objective function of problem (18), and can be written as follows

$$\text{Loss} = -\frac{R}{P_{\text{cons}}(\bar{\eta})}, \tag{18}$$

where the minus sign was introduced for the DNN to maximize the energy efficiency of the OFDM transission. Moreover, we observe that the output of the DNN has both continuous and discrete variables. We denote the output estimation vector of the DNN as

$$\boldsymbol{o} = \Big[\hat{\boldsymbol{p}}, \underbrace{\boldsymbol{a}_1, \boldsymbol{a}_2, .., \boldsymbol{a}_J}_{J}\Big]^{\top} = m(\boldsymbol{F}, \boldsymbol{\theta}). \tag{19}$$

where $m : R_{+}^{K \times 1} \rightarrow R^{(K+1) \times 1} \times \{0,1\}^{KJ \times 1}$ is the mapping function of the DNN parametrized by its trainable parameters $\boldsymbol{\theta}$, while $\hat{\boldsymbol{p}} = [\tilde{\boldsymbol{p}}, p_s]$, and $p_s \in [0,1]$ is a slack variable introduced so that constraint C_1 is equivalently written as

$$\mathrm{C}_1: \sum_{k=1}^{K} \tilde{\boldsymbol{p}}[k] + p_s = 1. \tag{20}$$

Constraint C_1 can now be handled by using the *Softmax* function as one of the activation functions of the output layer, as shown in Fig. 1. Therefore, $\hat{\boldsymbol{p}}$ is obtained as follows

$$\hat{\boldsymbol{p}} = \text{Softmax}(\boldsymbol{z}_0) = \frac{e^{\boldsymbol{z}_0}}{\sum_i e^{\boldsymbol{z}_0[i]}}, \tag{21}$$

where $\boldsymbol{z}_0$ is the input from the previous layer.

However, the Softmax function provides a mapping to the continuous set $(0,1)$, and thus it cannot be utilized to estimate the values of $\boldsymbol{a}$. In fact, $\boldsymbol{a}_j, \forall j \in \mathcal{J}$ follows a categorical distribution, since for fixed j, $a_{k,j} = 1$ for only one $k \in \{1, ..., K\}$. As such, the Gumbel-Softmax activation function will be used. Therefore, $\boldsymbol{a}_j$ is estimated as follows

$$\boldsymbol{a}_j = \text{Gumbel-Softmax}(\boldsymbol{z}_j, \tau) = \frac{e^{(\log \boldsymbol{z}_j + \boldsymbol{g}_j)/\tau}}{\sum_i e^{(\log \boldsymbol{z}_j[i] + \boldsymbol{g}_j[i])/\tau}}, \tag{22}$$

where τ is the softmax temperature, $\boldsymbol{g}_j$ are i.i.d samples drawn from Gumbel(0,1) distribution, and $\boldsymbol{z}_j$ is the input to the j-th Gumbel-Softmax node from the previous layer. We note that when $\tau \rightarrow 0$, samples from the Gumbel-Softmax distribution become one-hot and the continuous Gumbel-Softmax distribution becomes identical to a categorical distribution. The total architecture of the proposed online DNN is given in Fig. 1.

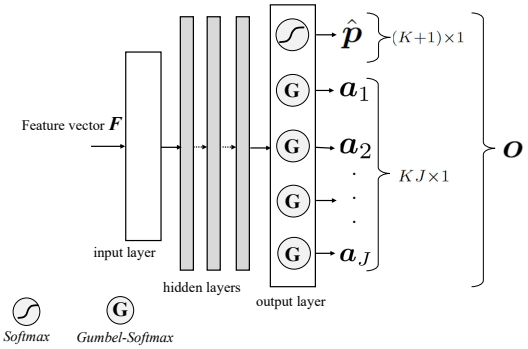


Fig. 1: DNN architecture.

2) Unsupervised Learning: The main difference between the online learning and the unsupervised learning, it is that in unsupervised learning, a batch of different channel realizations, with size B_{tch}, is given as input to the DNN during each epoch. The channel realizations are randomly drawn from a training dataset. The loss function of the unsupervised learning, per epoch, is the following

$$\text{Loss} = -\frac{1}{B_{\text{tch}}} \sum_{i=1}^{B_{\text{tch}}} E_i, \tag{23}$$

where $E = \frac{R}{P_{\text{cons}}(\bar{\eta})}$, and E_i is the energy efficiency of the i-th batch sample. From (18) it is concluded that the online learning provides a near-optimal solution with respect to instant CSI, while from (23) it is entailed that the unsupervised learning provides a solution with respect to the CSI distribution, thus it can generalize its output to first-seen channel realizations. The architecture of the unsupervised learning scheme also follows that of Fig. 1.

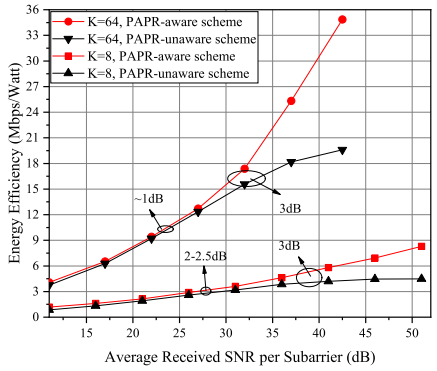
3) Complexity analysis: For the DNN architecture, without loss of generality, we consider a single hidden layer with L neurons. The forward-pass complexity, which is dominated by the weight matrix multiplication cost, is given as $\mathcal{O}(L + 2KL + KJL)$ and is equivalent to $\mathcal{O}(KJL)$. Given that the backward-pass has the same computational cost with the forward-pass, the overall back propagation algorithm for training the DNN, in an online fashion, is of the order of $\mathcal{O}(I_E KJL)$, where I_E is the number of training epochs. As such, the back-propagation complexity of the unsupervised learning is given as $\mathcal{O}(I_E B_{\text{tch}} KJL)$. Nonetheless, once trained, the unsupervised learning scheme requires only one forward-pass. In contrary, in online learning, a new DNN has to be trained for any new channel realization.

IV. Numerical Results

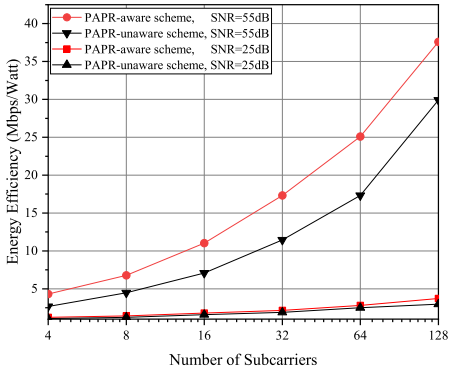
In this section, we validate the performance of the proposed PAPR-aware scheme. Unless otherwise stated, we consider Rayleigh fading, a path loss of -70 dB, $P_{\max} = 10$ W, $N_0 = -174$ dBm/Hz, $l = 4$, $M \in \{4, 16, 64, 256, 1024\}$, and $B = 1$ MHz. The benchmark is a PAPR-unaware scheme, similar to [8], which maximizes the throughput of the network, without considering the relation between the PAPR and the energy consumption of the PA. Furthermore, it was numerically observed that $P_{\text{cons}}(\bar{\eta})$ does not significantly change when considering more than $I = 500$ symbol combinations. Thus, $I = 500$ was used to reduce the computational burden.

In Fig. 2(a), the energy efficiency against the average received SNR per subcarrier is plotted. We observe that for small SNR values the proposed PAPR-aware scheme is about 1-2.5 dB more energy efficient than the PAPR-unaware scheme. This small improvement is attributed to the PAPR-aware scheme considering the effect of PAPR to the energy efficiency. Nonetheless, we observe that as the average received SNR increases, the energy efficiency improves as well. This is due to the fact that in the low SNR regime the energy efficiency metric of (18) is limited by the achievable throughput of OFDM. However, for bigger SNR values, energy consumption is the limiting factor of energy efficiency, thus the proposed scheme greatly outperforms the benchmark by more than 3 dB. In Fig. 2(b), for 4-128 subcarriers, we observe that for smaller SNR the proposed scheme slightly outperforms the PAPR-unaware scheme, while for higher SNR values the improvement is significantly greater.

In Fig. 3, the power of the time-domain OFDM symbol, as given by (10), is plotted. The PAPR of each OFDM symbol is provided. From Fig. 3(a) and Fig. 3(b) it can be verified that the PAPR-aware scheme has lower PAPR than the PAPR-unaware scheme. Specifically, in the case of 5 dB average received SNR per subcarrier, the PAPR-aware scheme achieves 0.25 dB lower average PAPR compared to the PAPR-unaware scheme. This is due to fact that in the low SNR regime the energy efficiency is limited by throughput, thus, both schemes ignore the effect of the PAPR. In the case of 35 dB average received SNR per subcarrier, the PAPR-aware scheme has

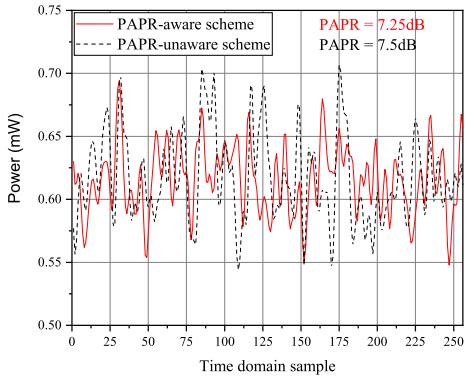


(a) Impact of the average received SNR per subcarrier.

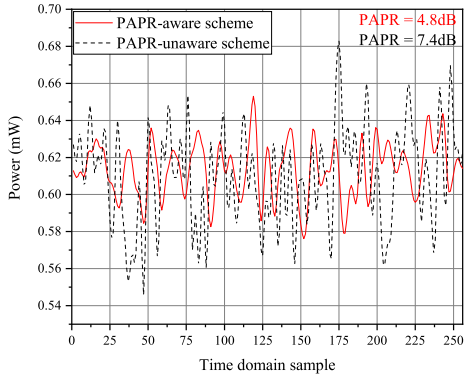


(b) Impact of the number of subcarriers.

Fig. 2: Energy efficiency performance.



(a) $K = 64, l = 4$, Average SNR per Subcarrier = 5 dB.



(b) $K = 64, l = 4$, Average SNR per Subcarrier = 35 dB.

Fig. 3: Average power of the time-domain symbol.

almost 3 dB lower PAPR compared to the PAPR-unaware scheme. For higher SNRs, the energy efficiency is mainly limited by the PAPR. As such, the PAPR-aware scheme attains a lower PAPR than the PAPR-unaware scheme, and achieves greater energy efficiency, as it was verified in Fig. 2(a).

In Fig. 4, the convergence of the proposed methods is shown. The training dataset of the unsupervised learning scheme consists of 20000 channel realizations, with batch size of 32, while testing has been averaged to 5000 different channel realizations. The hidden layers consist of three pairs of linear layers, both followed by the Hyperbolic tangent activation function. The hidden layers of the online learning scheme and the unsupervised learning scheme consist of 100 nodes and 150 nodes in respect. The Adam optimizer was used, with learning rate equal to 0.05 for the online learning and 0.1 for the unsupervised learning. We observe that both schemes roughly converge within 100 iterations to their local optimum. The unsupervised learning scheme is shown to be suboptimal, since the online learning scheme extracts its solution based on instant CSI, while the unsupervised learning scheme provides a solution according to the CSI distribution.

V. Conclusion

A PAPR-aware protocol was proposed, which aims to maximize the energy efficiency of the OFDM transmission. The energy efficiency maximization problem subject to power

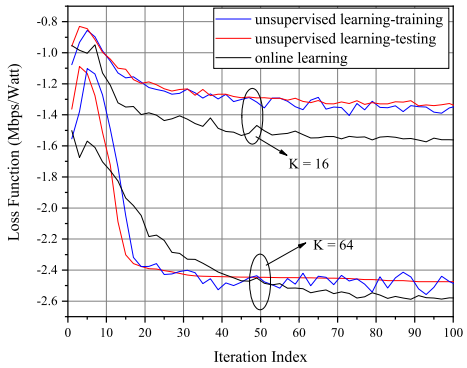


Fig. 4: Convergence of the proposed schemes.

and adaptive modulation constraints was formulated and a deep learning-based solution was designed. Numerical results showcase the increased energy efficiency of the proposed scheme, which improves the energy efficiency of the OFDM transmission by 3 dB. Thus, the proposed protocol can aid to satisfy the energy efficiency goals of beyond 5G networks.

References

[1] Z. Zhang, Y. Xiao, Z. Ma, M. Xiao, Z. Ding, X. Lei, G. K. Karagiannidis, and P. Fan, "6G Wireless Networks: Vision, Requirements, Architecture, and Key Technologies," *IEEE Veh. Technol. Mag.*, vol. 14, no. 3, pp. 28–41, 2019.

[2] B. Rong, "6g: The next horizon: From connected people and things to connected intelligence," *IEEE Wireless Communications*, vol. 28, no. 5, pp. 8–8, 2021.

[3] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[4] C. Nader, P. Händel, and N. Björsell, "Peak-to-Average Power Reduction of OFDM Signals by Convex Optimization: Experimental Validation and Performance Optimization," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 2, pp. 473–479, 2011.

[5] Q. Liu, R. J. Baxley, X. Ma, and G. T. Zhou, "Error Vector Magnitude Optimization for OFDM Systems With a Deterministic Peak-to-Average Power Ratio Constraint," *IEEE J. Sel. Top. Signal. Process.*, vol. 3, no. 3, pp. 418–429, 2009.

[6] Y.-C. Wang and Z.-Q. Luo, "Optimized Iterative Clipping and Filtering for PAPR Reduction of OFDM Signals," *IEEE Trans. on Commun.*, vol. 59, no. 1, pp. 33–37, 2011.

[7] X. Zhu, W. Pan, H. Li, and Y. Tang, "Simplified Approach to Optimized Iterative Clipping and Filtering for PAPR Reduction of OFDM Signals," *IEEE Trans. on Commun.*, vol. 61, no. 5, pp. 1891–1901, 2013.

[8] S. T. Chung and A. Goldsmith, "Degrees of freedom in adaptive modulation: a unified view," *IEEE Trans. on Commun.*, vol. 49, no. 9, pp. 1561–1571, 2001.

[9] ——, "Adaptive multicarrier modulation for wireless systems," in *Conference Record of the Thirty-Fourth Asilomar Conference on Signals, Systems and Computers (Cat. No.00CH37154)*, vol. 2, 2000, pp. 1603–1607 vol.2.

[10] D. Lee, Y. G. Sun, S. H. Kim, I. Sim, Y. M. Hwang, Y. Shin, D. I. Kim, and J. Y. Kim, "DQN-Based Adaptive Modulation Scheme Over Wireless Communication Channels," *IEEE Commun. Lett.*, vol. 24, no. 6, pp. 1289–1293, 2020.

[11] M. Li, X. Zhao, H. Liang, and F. Hu, "Deep Reinforcement Learning Optimal Transmission Policy for Communication Systems With Energy Harvesting and Adaptive MQAM," *IEEE Trans. Veh. Techol.*, vol. 68, no. 6, pp. 5782–5793, 2019.

[12] S. Mashhadi, N. Ghiasi, S. Farahmand, and S. M. Razavizadeh, "Deep Reinforcement Learning Based Adaptive Modulation With Outdated CSI," *IEEE Commun. Lett.*, vol. 25, no. 10, pp. 3291–3295, 2021.

[13] T. Jiang and Y. Wu, "An overview: Peak-to-average power ratio reduction techniques for ofdm signals," *IEEE Trans. Broadcast.*, vol. 54, no. 2, pp. 257–268, 2008.

[14] D. Persson, T. Eriksson, and E. G. Larsson, "Amplifier-aware multiple-input multiple-output power allocation," *IEEE Commun. Lett.*, vol. 17, no. 6, pp. 1112–1115, 2013.