A Novel Online Subcarrier-Wise Extreme Learning Machine Receiver for OFDM Systems

Michel Saideh[1], Eric Pierre Simon[1], Joumana Farah[2]

[1]Univ. Lille, IEMN lab, France, {firstname.lastname@univ-lille.fr}
[2]Faculty of Engineering, Lebanese Univ., Roumieh, Lebanon {joumanafarah@ul.edu.lb}

Abstract—Recently, Extreme Learning Machine (ELM) started gaining interest among researchers in wireless communications as an online training solution for machine learning based receivers. ELM has proven to provide high training speed and global optimization capabilities. However, the number of needed training pilots is still relatively high and increases rapidly with the number of subcarriers, thus rendering its deployment impractical. In this paper, we propose subcarrier-wise ELM receivers that are robust to the increase in the number of used subcarriers; we then extend them to exploit adjacent channel knowledge, hence providing superior performance in frequency selective channels. In addition, we propose a novel training architecture based on interpolated training that saves more than 50% of the computational and spectral resources of conventional ELM receivers. We show the robustness of the proposed technique in different channel scenarios and OFDM settings by means of both practical channel measurements and numerical simulations.

Index Terms—Extreme Learning Machine, Interpolated Training, OFDM, Equalization, Channel Estimation.

I. INTRODUCTION

Machine Learning (ML) techniques have recently gained enormous attention in the wireless communications research field. Unlike traditional model-based communication systems, the non-linear capabilities of ML-based communication systems have led to superior performance when dealing with non-linear conditions and interfering sources, *e.g.,* insufficient cyclic prefix. This superiority was shown in various studies where deep Neural Networks (NN) were mainly adopted such as in [1], [2]. However, the majority of studies considered an offline training approach that depends on launching huge amounts of numerical simulations to accumulate the needed training sets. These sets are used afterwards in iterative optimization processes, such as stochastic gradient descent, in order to tune the NN weights in a global optimum sense. Although this strategy has shown impressing learning capabilities, a performance degradation is observed when the channel statistics faced in the real-world prediction phase differ from those used in the numerical training phase.

One way to handle this problem is to adopt an online training approach by means of extreme learning machine [3]–[5]. ELM is a single hidden-layer feed-forward NN that allows extremely fast training capabilities and ensures globally optimum weights in the least square sense. This means that ELM could be quickly trained in an online manner using pilots received from any surrounding radio environment, and that can be exploited afterwards for data detection. Although some recent studies have started proposing online ELM-based receivers [4], [6]–[8], they still require a large amount of pilots for online training, thus causing a significant loss in spectral efficiency. In addition, the number of needed training pilots increases with the number of ELM inputs [8], *i.e.,* the number of subcarriers in Orthogonal Frequency Division Multiplexing (OFDM) systems. In fact, an ELM architecture of $(N_{in} \times \tilde{N} \times N_{out})$ is normally designed with $N_p > \tilde{N} > N_{in}$, where N_p is the number of OFDM pilot symbols and $N_{in}, \tilde{N}, N_{out}$ represent the number of input, hidden, and output nodes, respectively. One straightforward way to reduce N_p is to decrease the number of inputs N_{in}, hence $\tilde{N}$, which corresponds to a smaller number of weights to be tuned. This can be done by dedicating one specific ELM receiver for each subcarrier, as was initially proposed in [4] with a $(1 \times \tilde{N} \times 1)$ architecture. This approach is denoted by Subcarrier-Wise ELM (SW-ELM). Even though the scheme in [4] needs a smaller amount of training symbols, regardless of the number of OFDM subcarriers, the $(1 \times \tilde{N} \times 1)$ architecture is incapable of handling non-linear situations such as the cases of insufficient or lack of Cyclic Prefix (CP).

In this paper, we first revisit the SW-ELM approach by proposing new architectures that exploit adjacent channel information in order to overcome the aforementioned limitations. Then, we use the proposed SW-ELM to introduce a novel *interpolated training* concept which permits more than 50% reduction of the needed computational and spectral resources. *Interpolated training* means that only a portion of subcarriers are dedicated for pilots in the training phase and thus used to train the corresponding SW-ELM receivers. The SW-ELM receivers on the non-pilot subcarriers are trained by directly calculating their output weights through interpolation. We consider different pilot spacing schemes and assess the receivers performance in various channel scenarios. In addition to numerical simulations, we conduct a real-world channel measurement campaign in an indoor environment. Our measurements support the SW-ELM based interpolated training superiority in terms of spectral and computational efficiencies.

The remainder of this paper is organized as follows. In Section II, we present the system model and the channel measurement campaign. Section III details our proposed interpolated training concept. Section IV shows the performance evaluation results, while Section V concludes the paper.

In this paper, $\boldsymbol{X}^T$, $\boldsymbol{X}^H$ and $\boldsymbol{X}^\dagger$ are the transposition, Hermi-

tian conjugate, and Pseudo-inverse of X, respectively. $\otimes$ is the Kronecker product. $diag(.)$ transforms a vector into a diagonal matrix and a matrix into a vector of its diagonal elements. $\mathbb{C}$ is the set of complex numbers.

II. SYSTEM MODEL

A. OFDM System

Consider a block of K OFDM symbols to be transmitted over M subcarriers. Therefore, the transmitted vector $x \in \mathbb{C}^{MK \times 1}$ of the whole transmission block is written as $x = [x_1^T\, x_2^T \,...\, x_K^T]^T$, where $x_k = [x_{1,k}\, x_{2,k}\, \ldots\, x_{M,k}]^T$. $x_{m,k}$ is the $\mathcal{M}$-QAM symbol transmitted on the mth subcarrier and the kth OFDM symbol. We denote by N the number of samples of the whole transmission block and N_{CP} the Cyclic Prefix (CP) length: $N = N_{CP\text{-OFDM}}K$, where $N_{CP\text{-OFDM}} = M + N_{CP}$ is the length of the CP-OFDM symbol. The demodulated vector $y \in \mathbb{C}^{MK \times 1}$ of the whole transmitted block is defined as $y = [y_1^T\, y_2^T \,...\, y_K^T]^T$, where $y_k = [y_{1,k}\, y_{2,k}\, \ldots\, y_{M,k}]^T$. It can be written as:

$$y = \underbrace{F_{rx}^H C F_{tx}}_{\mathcal{D}} x + F_{rx}^H \eta, \tag{1}$$

where $\eta \in \mathbb{C}^{N \times 1}$ is the additive white Gaussian noise vector. $F_{tx/rx} = diag(\mathbf{1}_K) \otimes W_{tx/rx}$, with $W_{tx/rx} \in \mathbb{C}^{N_{CP\text{-OFDM}} \times M}$ being the OFDM modulation and demodulation matrices. $\mathbf{1}_K$ is a vector of K ones. At the transmitter, $W_{tx}(n, m) = \frac{1}{\sqrt{M}} e^{j\frac{2\pi}{M} m(n - N_{CP})}$. At the receiver, $W_{rx}(n, m) = 0$ when $0 < n \leq N_{CP}$ and $W_{rx}(n, m) = \frac{1}{\sqrt{M}} e^{j\frac{2\pi}{M} mn}$ when $N_{CP} < n \leq N_{CP\text{-OFDM}}$. We denote by $C \in \mathbb{C}^{N \times N}$ the channel convolution matrix of the transmitted signal block, where $C(i, j) = c(i - j)$, with $c(l)$ being the lth tap of the channel impulse response. $\mathcal{D} = F_{rx}^H C F_{tx}$ represents the System Transmission Matrix (STM) of the CP-OFDM system. When the channel is time unvarying and the CP length is bigger than the maximum delay spread of the channel, $\mathcal{D}$ becomes diagonal with $\mathfrak{c} = diag(\mathcal{D})$ denoting its diagonal vector. Hence, $y_{m,k}$ can be equalized by dividing it with the corresponding STM value $\mathfrak{c}_{m,k} = \mathcal{D}(mk, mk)$:

$$\hat{x}_{m,k} = \frac{y_{m,k}}{\hat{\mathfrak{c}}_{m,k}}, \tag{2}$$

To obtain the channel estimate $\hat{\mathfrak{c}}_{m,k}$, different techniques could be considered such as the Least Square (LS) and Linear Minimum Mean Square Error (LMMSE) channel estimators.

B. Subcarrier-wise ELM Receiver

Contrary to the previous work in [4], in our study, we give the SW-ELM the opportunity to have surrounding subcarriers as input along with the desired subcarrier. Fig. 1 represents the proposed SW-ELM system. We denote the input matrix of the ith SW-ELM receiver, i.e., the SW-ELM receiver dedicated to the ith subcarrier, by $Y_i \in \mathbb{C}^{N_x \times N_{in}}$. $N_x \in \{N_p, N_d\}$ represents the number of concatenated OFDM demodulated symbols and N_p is the number of OFDM symbols transmitted in the training phase, while N_d is the number of OFDM symbols transmitted in the prediction phase, with $K = N_p + N_d$.

Let $X_i = \mathcal{F}(Y_i) \in \mathbb{C}^{N_x \times 1}$ the output of the ith SW-ELM receiver that regroups the transmitted symbols on the ith subcarrier in the training phase or the corresponding detected symbols in the prediction phase. It can be written as follows:

$$\mathcal{F}(Y_i) = \underbrace{G_i\Big(Y_i \alpha_i + B_i\Big)}_{H_i} \beta_i, \tag{3}$$

where $\alpha_i \in \mathbb{C}^{N_{in} \times \tilde{N}}$ is the matrix of hidden layer weights with its jth column representing the weights connecting the N_{in} input nodes to the jth hidden node. In other words, in the ith receiver, the weight $\alpha_{i,nj}$ connects the nth input node to the jth hidden node, as can be seen in Fig. 1. $B_i = b_i^T \otimes \mathbf{1}_{N_x}$, with $b_i \in \mathbb{C}^{\tilde{N} \times 1}$ representing the bias vector added to the hidden nodes, the term $b_{i,j}$ being added to the jth hidden node. $\beta_i \in \mathbb{C}^{\tilde{N} \times 1}$ is the output weights vector that contains the weights connecting the $\tilde{N}$ hidden nodes to the single output node, such that the weight $\beta_{i,j}$ connects the jth hidden node to the output node. $G_i(.)$ denotes the element-wise activation function, taken in this work as an inverse hyperbolic sine function. Each SW-ELM receiver will go through three main phases: Initialization, training and prediction.

1) Initialization phase: In ELM, the hidden layer weights α and biases b are not trained, but randomly generated and fixed afterward. As stated by [3], any continuous random distribution could be assigned for both α and b, and we consider a uniform distribution in [-0.1, 0.1].

2) Training phase: It aims to calculate the output weights parameters, *i.e.*, the matrix β. To do so, the ith SW-ELM receiver is first provided with a training set $\left(Y_i^{(p)}, X_i^{(p)}\right)$. It is defined similarly to Y_i, while the superscript (p) denotes the nature of these symbols, pilots in this phase. $Y_i^{(p)}$ is fed to the SW-ELM receiver as its input $Y_i = Y_i^{(p)}$ in Eq. (3). Accordingly, the hidden layer output is written as:

$$H_i^{(p)} = G_i\Big(Y_i^{(p)} \alpha_i + B_i\Big). \tag{4}$$

In the end, $X_i^{(p)}$ is fed to the SW-ELM receiver as its corresponding desired output: $\mathcal{F}(Y_i^{(p)}) = X_i^{(p)}$. Hence, using Eqs. (3) and (4), we can write:

$$X_i^{(p)} = H_i^{(p)} \beta_i. \tag{5}$$

Unlike most of ML training techniques, Eq. (5) shows that ELM is trained in one shot by minimizing the corresponding LS cost function and using the pseudo-inverse method [9]:

$$\hat{\beta}_i = \left(H_i^{(p)}\right)^\dagger X_i^{(p)}, \tag{6}$$

where
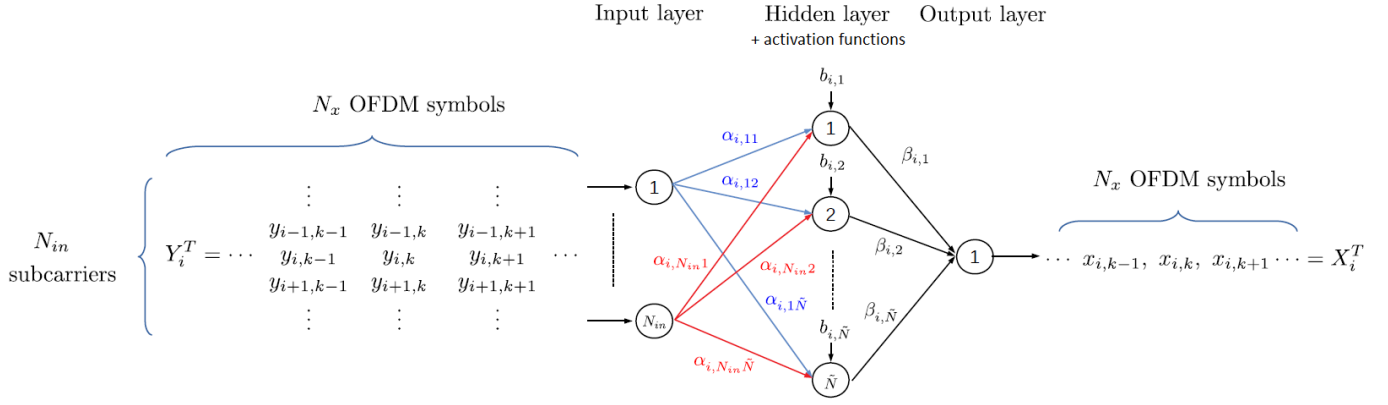
$$H_i^\dagger = \left(H_i^H H_i\right)^{-1} H_i^H. \tag{7}$$

Figure 1. Structure of the proposed subcarrier-wise ELM receiver

3) Prediction phase: Based on the calculated output weights matrix $\hat{\boldsymbol{\beta}}_i$, the ith SW-ELM receiver can estimate the N_d transmitted data symbols of the ith subcarrier $\boldsymbol{X}_i^{(d)} \in \mathbb{C}^{N_d \times 1}$ in one shot. To do so, the received data symbols at the ith subcarrier and its surrounding $\boldsymbol{Y}_i^{(d)} \in \mathbb{C}^{N_d \times N_{in}}$ are first fed to the corresponding SW-ELM to calculate its hidden layer output $\boldsymbol{H}_i^{(d)}$ using Eq. (4). Afterward, and based on Eq. (3), $\boldsymbol{H}_i^{(d)}$ is used together with the previously calculated $\hat{\boldsymbol{\beta}}_i$ to estimate the corresponding transmitted data symbols:

$$\hat{\boldsymbol{X}}_i^{(d)} = \boldsymbol{H}_i^{(d)} \hat{\boldsymbol{\beta}}_i. \tag{8}$$

C. Measurement Campaign

In this section, we present the used equipment and the considered environment for the channel measurement campaign that we conducted to validate the proposed SW-ELM receivers. A vector network analyzer (VNA) was used to sound the radio channel in the frequency domain. We considered a 3.5 GHz frequency with 1200 OFDM subcarriers of 15 KHz, thus a total of 18 MHz bandwidth. MegaPhase feeder cables were used for the transmit and receive antennas and were included in the VNA calibration to cancel their effects in the channel sounder. To reduce the measurement noise, the VNA acquires 10 successive realizations of the whole frequency range and takes their average for each measurement. The transmitter/receiver were positioned at 2m/1.50m above ground level and equipped with a patch / EM-6116 omnidirectional antennas, respectively. These Tx and Rx were placed in different rooms of a typical building in its first floor that consists of a 35m long corridor with offices on both sides. In Fig. 2, we present the different locations chosen for the transmitters and receivers to perform channel sounding, where only two concrete walls exist, while all other rooms are separated by plaster walls. The orientation of the transmitter patch antenna is shown by an arrow, while the black circle denotes the receiver location.

III. INTERPOLATED SW-ELM RECEIVERS

In this section, we use the SW-ELM receivers to propose an interpolated training concept that offers a spectral and computational efficient solution to the conventional block training
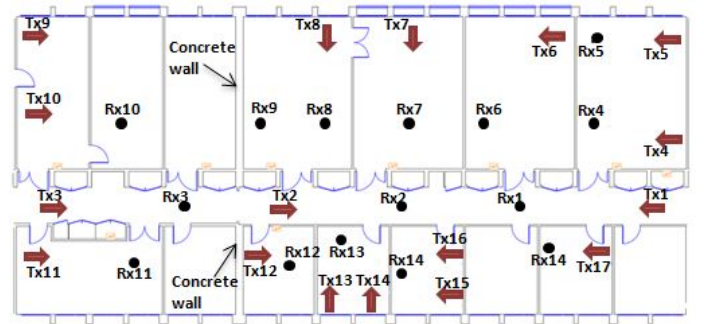

Figure 2. Measurement Environment

based (SB)-ELM receivers of Fig. 3a. For this purpose, we start by proposing a comb pilot pattern in the training phase where we send data among pilots, as is shown in Fig. 3b for a pilot spacing of two, *i.e.*, one pilot every two subcarriers. Using this concept, we propose to train SW-ELM receivers that
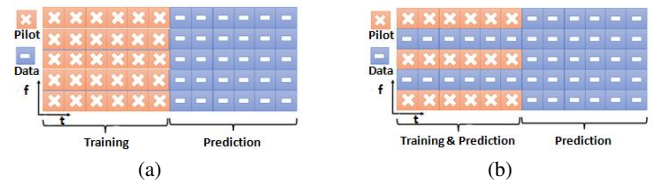

Figure 3. (a) Traditional training (Pilot spacing = 1) and prediction architecture (b) Proposed architecture with an example of a pilot spacing of 2

correspond only to pilot subcarriers by using the training in Eq. (6). On the other hand, SW-ELM receivers of data subcarriers are not trained but interpolated. In other words, the output weights matrix $\boldsymbol{\beta}_{data}$ of data-based ELM receivers are calculated by interpolating the corresponding output weights of the pilot-based ELM receivers. The interpolation is expressed as:

$$\boldsymbol{\beta}_{data} = \mathcal{W} \hat{\boldsymbol{\beta}}_{pilot}, \tag{9}$$

where $\hat{\boldsymbol{\beta}}_{pilot}$ contains the calculated output weights of pilot subcarriers and $\mathcal{W}$ is the considered interpolation matrix. Using Eq. (9) to interpolate the output weights of data-based

ELM receivers is only possible when the β variations along subcarriers are smooth. This means that the output weights of SW-ELM receivers should change smoothly from one subcarrier to the other. Nevertheless, this is not guaranteed since ELM hidden parameters are randomly chosen. To handle this, we propose to design all SW-ELM receivers while using the same randomly generated hidden layer parameters, *i.e.,* $\boldsymbol{\alpha}_i = \boldsymbol{\alpha}_j$, $\boldsymbol{B}_i = \boldsymbol{B}_j$, with $i, j \in \{1, 2, ..M\}$. By doing so, the output weights are now expected to vary smoothly from one subcarrier to the other. This is confirmed in Fig. 4 where we present the variations of a trained output weight over all subcarriers. We considered a CP-OFDM system with 256 subcarriers and a TDL-C power delay profile [10] with a Root Mean Square (RMS) delay spread of 200 ns. We note that the output weights variations are smooth and are highly proportional to the variations of the channel frequency response. The considered pilot spacing is taken such that pilots are separated by 4 data symbols which seems sufficient to conduct interpolation in this scenario. In what follows, we
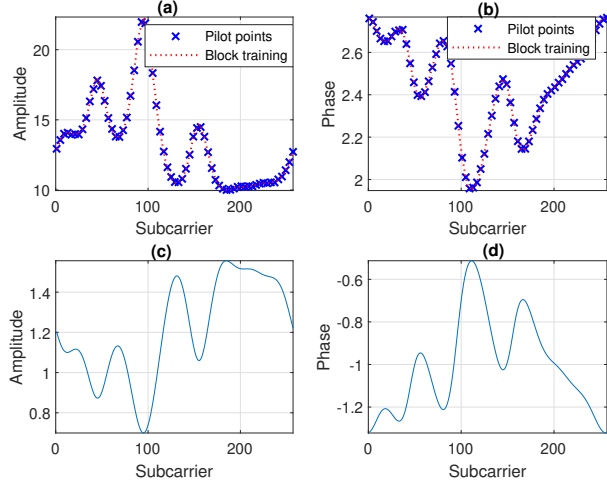

Figure 4. Amplitude (a) and phase (b) variations of the output weight β along OFDM subcarriers compared with the amplitude (c) and phase (d) variations of a TDL-C channel frequency response with RMS delay spread = 200 ns

summarize the working principles of the interpolated SW-ELM receivers:

1) In the training phase, pilots are transmitted in a comb pattern with data in between. Different pilot spacing schemes could be considered based on the channel frequency selectivity.
2) The same randomly generated hidden layer parameters are used for all data and pilots SW-ELM receivers (*i.e.* the same among all subcarriers).
3) The pilot SW-ELM receivers are trained using Eq. (6).
4) The output weights of data SW-ELM receivers are calculated by applying an interpolation filter on the output weights of previously calculated pilot SW-ELMs.
5) The interpolated SW-ELM receivers are used to detect the data symbols sent in the training phase.

6) The overall trained and interpolated SW-ELM receivers are used to detect data symbols in the prediction phase.

IV. Results & Discussion

A. Practical Results

In this section, we start by conducting MonteCarlo simulations via Matlab to assess the numerical performance of the proposed architecture. Afterward, the system performance is analysed by incorporating the results of the measurement campaign. We consider the TDL-C power delay profile with an OFDM system of 256 subcarriers and a CP length of 4.7 μs, while the subcarriers spacing is set to 15 KHz. The following SW-ELM architectures are considered: $5 \times 6 \times 1$, $3 \times 4 \times 1$, $1 \times 3 \times 1$ and $1 \times 4 \times 1$, with a total of 200 training symbols for the training of all pilot-based SW-ELM receivers.
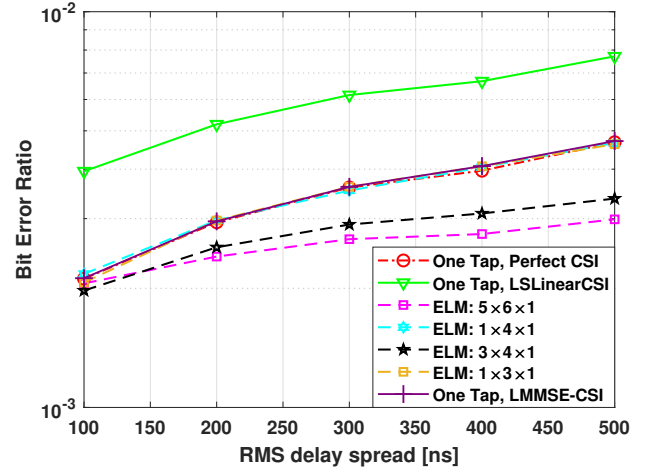

Figure 5. BER vs RMS delay spread for different SW-ELMs at SNR=25 dB

We start in Fig. 5 by presenting the performance of different SW-ELM architectures while using the conventional block training concept that does not use interpolation. We simulate the Bit Error Rate (BER) performance for different RMS delay spread values in a CP-free OFDM system. The aim here is to focus, in a first step, only on the impact of including adjacent subcarriers on the SW-ELM performance in frequency selective channels. The first thing to note is that SW-ELM with a single input cannot outperform the case of perfect channel state information (PCSI), even when increasing the number of hidden nodes. On the other hand, allowing the SW-ELM receiver to exploit adjacent subcarriers information succeeds in outperforming the PCSI case. This happens while using relatively simple architectures with a small number of hidden nodes, where we considered the cases of two surrounding subcarriers ($3\times4\times1$) and 4 surrounding subcarriers ($5\times6\times1$). Using a small number of inputs reflects into the need for a low number of hidden nodes, and therefore a smaller number of training symbols compared to more complex architectures. Indeed, we used only 200 training symbols compared to [8] where a number of 10^4 training symbols was necessary.
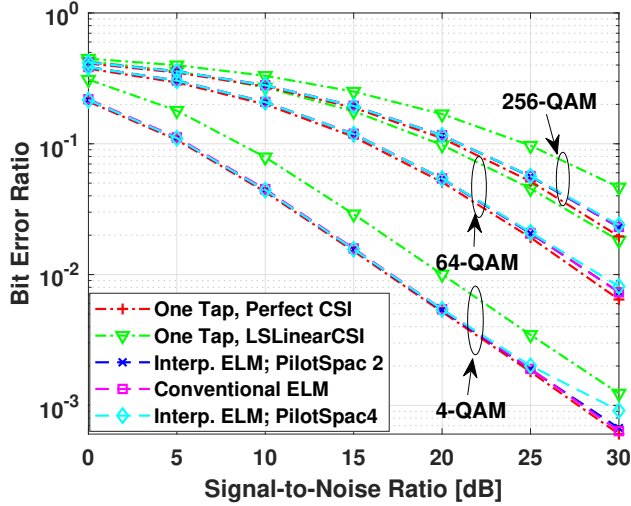
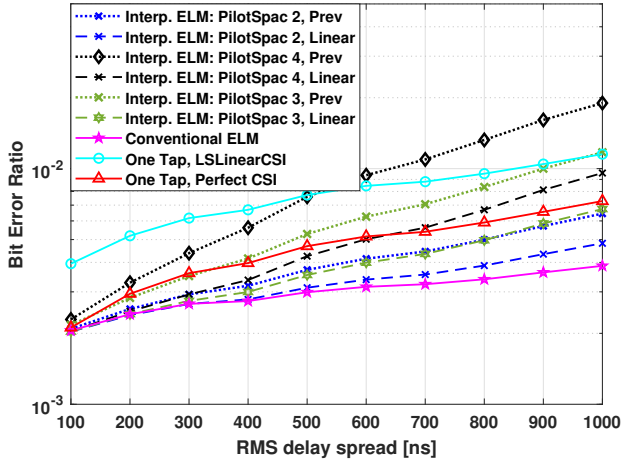Figure 6. BER vs SNR at channel RMS delay spread 200 ns

Figure 7. BER for different TDL-C channel RMS delay spread values and SNR= 25 dB

In Figs. 6 and 7, we adopt the $(5 \times 6 \times 1)$ architecture and assess the performance of the interpolated training in comparison with that of the traditional block training concept. The comparative analysis is provided for different pilot spacing schemes and interpolation techniques. Fig. 6 presents the BER performance obtained for different values of the Signal-to-Noise Ratio (SNR) and a channel RMS delay spread of 200 ns, while considering three constellation maps, namely the 4-QAM, 64-QAM, and 256-QAM. The performance of interpolated SW-ELM is simulated with linear interpolation and two pilot spacing schemes of 2 and 4, corresponding to pilots being separated by 1 and 3 data symbols respectively. The major observations to make are that our proposed interpolated architecture with linear interpolation performs close to the fully conventional architecture and the PCSI case, with the pilot spacing schemes of 2 and 4. Also, it should be noted that, similarly to the conventional case, the

proposed architecture is robust to the increase of the QAM constellations sizes.

The advantages of using ML-based receivers lie within their non-linear capabilities. In Fig. 7, we consider a CP-free OFDM system with a 4-QAM constellation map and compare the BER obtained for different RMS delay spread values between 50 and 1000 ns. Also, different pilot spacing schemes (2, 3 and 4) are considered, corresponding to 1, 2 and 3 data symbols between pilots, respectively. We consider both linear and previous interpolation techniques for data ELM receivers. Previous interpolation means that the trained Beta matrix of a specific pilot symbol will be used for the detection of subsequent data symbols. In other words, it is equivalent to applying the same pilot-based ELM receiver to subsequent data subcarriers. We first note the expected superiority of the conventional SW-ELM receiver over the PCSI and LS cases. This superiority gap increases with the increase of the RMS delay spread. On the other hand, the interpolated SW-ELM performs as good as the conventional SW-ELM for low values of the RMS delay spread. It can be noted that pilot spacing values of 3 and 4 maintain almost the same performance as that of the conventional case until an RMS value of about 300 ns and 200 ns, respectively. Besides, the pilot spacing 2 can still guarantee a similar performance to the conventional case even at an RMS delay spread of 600 ns, which is already higher than the measured RMS of many transmission scenarios such as the Pedestrian A and Indoor channel models. Also, one can clearly note the advantages of using the linear interpolation scheme with respect to a direct application of pilot-based SW-ELM receivers to data subcarriers.

Fig. 8 shows the BER performance obtained with the measurement campaign for different locations of the transmitter and receiver. The aim is to outline how the interpolated SW-ELM technique performs with respect to the conventional SW-ELM in different indoor channel scenarios. An SNR value of 30 dB is considered for all Tx-Rx combinations, with a 4-QAM constellation map in a CP-free OFDM system. Results confirm that the interpolated SW-ELM still performs similarly or closely to the conventional SW-ELM in the 2, 4 and 6 pilot spacing cases, depending on the channel scenario. Table I presents the BER for different SNR values of the conventional and interpolated SW-ELM receiver with the 6 pilot spacing case. The results are averaged over all conducted measurements of Fig. 2. They confirm that the interpolated SW-ELM performs as good as the block-trained SW-ELM. Also, it is clear that the 6 pilot spacing scheme with low and moderate SNR does not induce any significant loss in the average performance obtained in the practical indoor scenario.

Table I
AVERAGE BER OF THE MEASUREMENT CAMPAIGN FOR THE
INTERPOLATED AND CONVENTIONAL SW-ELM RECEIVERS

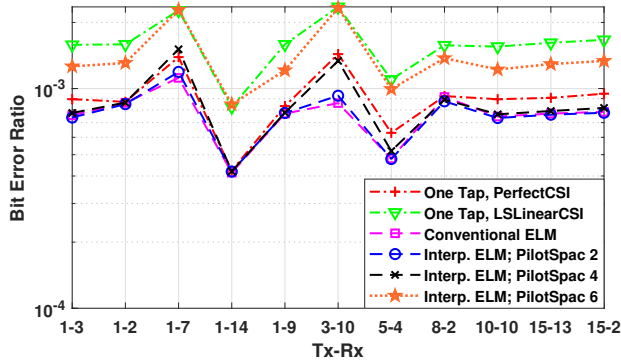| SNR [dB]	0	10	20	25	30
BER Interpolated	0.21	0.046	0.0057	0.0020	0.00082
BER Conventional	0.21	0.046	0.0057	0.0019	0.00065

Figure 8. BER for different Tx and Rx locations in the measurement campaign

Table II
COMPUTATIONAL AND SPECTRAL EFFICIENCY AT SNR= 25 DB

Pilot Spacing	RMS [ns]	$\frac{M_p}{M}100\%$	BER
1 (Block training)	200	100%	0.00241
2	200	50%	0.00241
3	200	33.33%	0.00241
4	200	25%	0.00248
1 (Block training)	500	100%	0.00299
2	500	50%	0.00312
3	500	33.33%	0.00354
4	500	25%	0.00425

B. Discussion on the Computational Load and Spectral Efficiency

In this part of the study, we compare the complexity and spectral efficiency of the conventional and proposed receivers. In the training phase of the SW-ELM receiver, the matrix inversion process of Eq. (7) has a complexity order of $O(\tilde{N}^3)$, while the matrix multiplication $\boldsymbol{H}^H\boldsymbol{H}$ results in $O(\tilde{N}^2N_p)$. Knowing that $N_p > \tilde{N} > M$, we can conclude that the asymptotic computational load of each SW-ELM receiver is $O(\tilde{N}^2N_p)$. Therefore, in block-trained based receivers, the overall complexity is given by $O(M\tilde{N}^2N_p)$. Let M_p and M_d the number of subcarriers used for pilots and data transmissions, respectively. In the interpolated SW-ELM, only M_p SW-ELM receivers are trained using Eq. (6), thus resulting in $O(M_p\tilde{N}^2N_p)$, while the M_d SW-ELM receivers of the data subcarriers are first linearly interpolated and then used in Eq. (8) to detect the transmitted data of the training phase, resulting in $O(M_d\tilde{N}N_p)$. Hence, the overall complexity can be given by $O(M_p\tilde{N}^2N_p)$. This means that the proposed interpolated SW-ELM receiver can use $\frac{M_p}{M} \times 100\%$ of the computational and spectral resources of the conventional block-trained SW-ELM receiver, while having almost the same performance accuracy. In fact, this depends on the channel frequency selectivity and the chosen pilot spacing scheme, as can be seen in table II which provides the relative gains of the interpolated SW-ELM w.r.t. the conventional block-trained SW-ELM, for different values of the RMS delay spread and pilot spacing schemes. The results show that, for an RMS of 200 ns, a gain of up to 75% can be obtained by adopting the pilot spacing scheme of 4, while incurring a negligible loss in BER. For an RMS of 500 ns, a gain of at least 50% can be achieved.

V. CONCLUSION

In this paper, the subcarrier-wise ELM receiver was revisited to take advantage of its superiority over the block-based ELM, mainly in terms of its robustness against the number of used OFDM subcarriers and its fast training and global optimization capabilities. The first contribution of our work is the extension of SW-ELM receivers to exploit adjacent subcarrier channels so as to achieve a high detection performance with simple architectures, while necessitating a low number of training symbols. Indeed, we showed that $3\times4\times1$ and $5\times6\times1$ ELM structures that exploit up to 4 adjacent subcarriers are sufficient to yield a high detection accuracy. Based on the extended SW-ELM receivers, we then proposed a novel interpolated training concept that saves more than 50% of the spectral and computational resources of conventional SW-ELM receivers. We showed that these gains can be increased when the channel frequency selectivity decreases. In addition, we conducted an indoor measurement campaign that validated the efficiency of the interpolated training concept in different transmission scenarios. By incorporating the prediction phase of data symbols within the training phase, the new ML structures avoid the need for the traditional long offline training phases, and therefore allow their real-time implementation in low-latency applications.

REFERENCES

[1] H. Ye, G. Y. Li, and B.-H. Juang, "Power of deep learning for channel estimation and signal detection in OFDM systems," *IEEE Wireless Communications Letters*, vol. 7, no. 1, pp. 114–117, 2017.

[2] J. Zhang, C.-K. Wen, S. Jin, and G. Y. Li, "Artificial intelligence-aided receiver for a CP-free OFDM system: Design, simulation, and experimental test," *IEEE Access*, vol. 7, pp. 58 901–58 914, 2019.

[3] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1-3, pp. 489–501, 2006.

[4] L. Yang, Q. Zhao, and Y. Jing, "Channel equalization and detection with ELM-based regressors for OFDM systems," *IEEE Communications Letters*, vol. 24, no. 1, pp. 86–89, 2020.

[5] M. Saideh, E. P. Simon, J. Farah, J. Villain, A. Fleury, V. Deniau, and C. Gransart, "Ensemble extreme learning machine based equalizers for OFDM systems," in *2020 14th International Conference on Signal Processing and Communication Systems (ICSPCS)*. IEEE, 2020, pp. 1–6.

[6] M.-B. Li, G.-B. Huang, P. Saratchandran, and N. Sundararajan, "Fully complex extreme learning machine," *Neurocomputing*, vol. 68, pp. 306–314, 2005.

[7] I. G. Muhammad, K. E. Tepe, and E. Abdel-Raheem, "QAM equalization and symbol detection in OFDM systems using extreme learning machine," *Neural Computing and Applications*, vol. 22, no. 3-4, pp. 491–500, 2013.

[8] J. Liu, K. Mei, X. Zhang, D. Ma, and J. Wei, "Online extreme learning machine-based channel estimation and equalization for OFDM systems," *IEEE Communications Letters*, vol. 23, no. 7, pp. 1276–1279, 2019.

[9] S. S. Haykin, *Adaptive filter theory*. Pearson Education India, 2005.

[10] 3GPP, ETSI TR 138 900 V14.2.0 (2017-06), "Study on channel model for frequency spectrum above 6 ghz."