Dual Layers-Superimposed Training for Joint Channel Estimation and PAPR reduction in OFDM

Chen Hu, Kun; Fernández-Getino García, Maria Julia; García Armada, Ana

Published in:
GLOBECOM 2023 - 2023 IEEE Global Communications Conference

DOI (link to publication from Publisher):
10.1109/GLOBECOM54140.2023.10437185

Publication date:
2023

Document Version
Accepted author manuscript, peer reviewed version

Dual Layers-Superimposed Training for Joint Channel Estimation and PAPR reduction in OFDM

Kun Chen-Hu[1,2], M. Julia Fernández-Getino García[1], and Ana García Armada[1]
[1]Department of Signal Theory and Communications, Universidad Carlos III de Madrid, Spain
[2]Department of Electronic Systems, Aalborg University, Denmark
E-mails: {kchen, mjulia, agarcia}@tsc.uc3m.es

***Abstract*—Superimposed training (ST) is one of the most appealing channel estimation techniques for orthogonal frequency division multiplexing (OFDM), to be possibly exploited in the 6G. The data and pilot symbols are sharing the same time and frequency resources, and the overhead is significantly reduced. Besides, the superimposed pilots can be also used for the reduction of the peak-to-average power ratio (PAPR). However, joint channel estimation and PAPR reduction procedures have never been addressed. In this work, a novel scheme denoted as dual layers-superimposed training (DL-ST) is proposed for this joint purpose. The training sequence (TS) of the first layer is targeted to perform channel estimation, while the TS of the second layer is designed for PAPR reduction and is made transparent to the first one. Both layers can be independently processed, which implies a reduced complexity. To verify the performance of the proposed technique, the analytical expression of the channel estimation mean squared error (MSE) is derived. Finally, several numerical results further illustrate the performance of the proposal, showing how the MSE is improved while significant PAPR reductions are attained with negligible complexity.**

***Index Terms*—channel estimation, layers, OFDM, PAPR, ST.**

I. Introduction

The peak-to-average power ratio (PAPR) [1], [2] is one of the most critical issues in multi-carrier waveforms, such as orthogonal frequency division multiplexing (OFDM) and its variants. The high peaks of the signal are frequently cut-off by the non-linear region of the power amplifier (PA) [3], and therefore, the transmitted signal is significantly distorted. This inefficiency in the amplification stage, which worsens as we use higher frequencies according to the trend for future wireless communications, can be partially solved by either using more expensive hardware, which increments the cost or decreasing the average input power, which reduces the coverage. Traditionally, several PAPR reduction techniques have been presented for OFDM, such as clipping, partial transmit sequences, selective mapping, tone reservation (TR), etc. [1], [2]. However, the main challenge of the PAPR reduction is characterized by the required computational complexity to remove the strong peaks of each OFDM symbol without distorting the transmitted information. Moreover, some of these existing techniques may require an additional side-link to transmit the modifications performed to the data symbols at the transmitter, and then, the receiver can undo these modifications performed by the PAPR cancellation algorithm.

Simultaneously, superimposed training (ST) [4]–[6] is considered one of the most appealing channel estimation techniques for OFDM in order to replace the well-known pilot symbol assisted modulation (PSAM), used in the 5G. In ST, data symbols and training sequence (TS) share the same time/frequency resources, and hence, the efficiency of the system is significantly increased since the TS is no longer exclusively occupying resources as in PSAM. Additionally, this efficiency is even higher in some scenarios such as large-scale multi-antenna systems, typically known as massive multiple-input multiple-output (MIMO), and high-speed communications which require a significant amount of reference symbols to periodically track the channel. At the receiver, an averaging process is performed to remove the self-interference induced by the data symbols, and the channel estimation can be obtained by applying a simple least squares (LS) criterion [7]. Enabling powerful amplification and compensating for channel effects in a variety of demanding scenarios are significant challenges that must be solved to continue with the advantageous use of multi-carrier waveforms in the 6G.

Lately, ST has not only been exploited for channel estimation purposes but it is also proposed to reduce the PAPR of the OFDM signal [8]–[10]. On one hand, some works have combined the existing PAPR reduction techniques with ST [8], [9], but however, the complexity issues described before are still present. On the other hand, ST based on using a constant amplitude TS is proposed to jointly estimate the channel and reduce the PAPR with a tiny complexity [10]. This work showed that the PAPR reduction is proportional to the energy allocated to this constant amplitude TS, and it does not require any additional complex products. Even though this technique exhibits a negligible complexity, it can only reduce the PAPR by 2 dB, and this moderate performance is due to the fact that the phase dimension of the constant amplitude TS is specifically tailored to increase the quality of the channel estimates, and it is not designed to further reduce the PAPR.

In this paper, a novel technique denoted as dual layers-superimposed training (DL-ST) is proposed to jointly estimate the channel and reduce the PAPR significantly improving the performance provided by the constant amplitude TS [10], [11]. The superimposed TS is built as the sum of two constant amplitude TSs, where each of the TS belongs to one layer. The TS of the first layer (TS-1L) is specifically tailored to obtain accurate channel estimates, while the TS of the second layer (TS-2L) is specifically designed to reduce the PAPR of each OFDM symbol. This proposal exhibits negligible complexity as compared to the classical ones. Firstly, it takes advantage of the constant amplitude TS which is capable of reducing the PAPR without any additional operations. Then, the TS-2L is a modified version of TR scheme [12]–[14], whose optimization problem can be solved by a low-complexity codebook search. The mean squared error (MSE) of the channel estimation based on the DL-ST is theoretically analyzed, showing that the effect of the TS-2L designed for PAPR reduction implies a negligible degradation in the system performance. The benefit of DL-ST is also evaluated via numerical results, providing the simulations of PAPR and MSE. This point out the advantages of this approach against the existing techniques and validates the theoretical analysis.

The remainder of the paper is organized as follows. Section II introduces the system model of an OFDM system. Section III explains the proposed DL-ST by detailing the design of the TSs for the two layers. Moreover, it provides the channel estimation procedure and the theoretical analysis of MSE. Section IV presents several numerical results for the proposed scheme, providing an assessment of the achieved performance. Finally, in Section V, the conclusions are reported.

Notation: matrices, vectors and scalar quantities are denoted by boldface uppercase, boldface lowercase, and normal letters, respectively. $[\mathbf{A}]_{m,n}$ denotes the element in the m-th row and n-th column of $\mathbf{A}$. $[\mathbf{a}]_n$ represents the n-th element of vector $\mathbf{a}$. $\mathbf{I}_M$ is the identity matrix of size $(M \times M)$. $\mathbf{0}_{M,N}$ is the zero matrix of size $(M \times N)$. $\mathbf{1}_{(M\times N)}$ denotes a matrix of ones of size $(M \times N)$. $\mathbf{A} = \operatorname{diag}(\mathbf{a})$ is a diagonal matrix whose diagonal elements are formed by the elements of vector $\mathbf{a}$. $\operatorname{tr}(\cdot)$ corresponds to the matrix trace operation. $\circledast$ is the circular convolution operation. $\otimes$ corresponds to the Kronecker product of two matrices. $\mathbb{E}\{\cdot\}$ represents the expected value. $\mathcal{CN}(0,\sigma^2)$ represents the circularly-symmetric and zero-mean complex normal distribution with variance σ^2. $|\cdot|$ represents the absolute value of a complex number. $\mathbb{C}^K$ and $\mathbb{R}^K$ are K-dimensional complex and real spaces, respectively, and $\mathbb{C}^{K\times K}$ is the $K \times K$-dimensional complex space.

II. System Model

Two wireless transceiver units are transmitting to each other using a bidirectional link. Both terminals can transmit and receive OFDM symbols. Each OFDM symbol is composed of K subcarriers with a subcarrier spacing of Δf Hz and a cyclic prefix (CP), whose length is measured in samples (L_{CP}), to mitigate the multi-path effects of the channel.

Let us define the complex data vector $\tilde{\mathbf{s}} \in \mathbb{C}^K$ to be transmitted at an OFDM symbol. Its elements belong to a constellation with energy normalized to one ($\mathbb{E}\left\{\left|[\tilde{\mathbf{s}}]_k\right|^2\right\} = 1$, $1 \le k \le K$). The OFDM symbol ($\mathbf{s}$) can be obtained as

$$\mathbf{s} = \mathbf{F}_K^H \tilde{\mathbf{s}} \in \mathbb{C}^K, \quad [\mathbf{F}_K]_{k_1+1,k_2+1} = \frac{1}{\sqrt{K}} \exp\left(-j\frac{2\pi}{K}k_1 k_2\right), \quad (1)$$

where $0 \le k_1, k_2 \le K-1$, $\mathbf{F}_K \in \mathbb{C}^{K\times K}$ is the normalized discrete Fourier transform (DFT) matrix of K-points. Then, a CP, whose length is given by L_{CP} is appended to each OFDM symbol $\mathbf{s}$ in order to absorb the multi-path effect.

At the receiver, after discarding the CP and assuming that the length of the CP is long enough to absorb all the taps of the channel, the received signal is modelled as a circular convolution of K samples between the wireless channel and the transmitted signal, whose expression is given by

$$\mathbf{y} = \mathbf{h} \circledast \mathbf{s} + \mathbf{v}, \quad (2)$$

where $\mathbf{v} \in \mathbb{C}^{K+L_{CP}}$ is the additive white Gaussian noise (AWGN) vector, each element is distributed as $\mathcal{CN}(0,\sigma_v^2)$ and $\mathbf{h} \in \mathbb{C}^{L_{CH}}$ corresponds to the channel impulse response of L_{CH} taps. Each tap is modelled as an independent complex random variable distributed as $\mathcal{CN}(0,\sigma_\tau^2)$ $1 \le \tau \le L_{CH}$, and the channel gain is normalized to unity. Considering a high mobility case and choosing a block-fading model, it is assumed that the channel coherence time (T_c) remains quasi-static during, at least, one OFDM symbol ($T_c \ge (K + L_{CP})\, T_s$), where $T_s = 1/(K\Delta f)$ denotes the sampling period.

The received OFDM symbol must be equalized in order to remove the effects of the channel. The frequency-domain output of the DFT computed over $\mathbf{y}$ can be decomposed as

$$\tilde{\mathbf{y}} = \mathbf{F}_K\mathbf{y} = \mathbf{H}\tilde{\mathbf{s}} + \tilde{\mathbf{v}} \in \mathbb{C}^K, \quad \tilde{\mathbf{v}} = \mathbf{F}_K\mathbf{v} \in \mathbb{C}^K, \quad (3)$$

$$\mathbf{H} = \operatorname{diag}\left(\tilde{\mathbf{h}}\right) \in \mathbb{C}^{K\times K}, \quad \tilde{\mathbf{h}} = \mathbf{F}_K\mathbf{h} \in \mathbb{C}^K, \quad 1 \le n \le K, \quad (4)$$

where $\tilde{\mathbf{h}}$ is the channel frequency response and $\tilde{\mathbf{v}}$ accounts for the noise in the frequency domain at the received OFDM symbol. Then, the channel effects can be equalized by using a one-tap equalizer in the frequency domain as

$$\widehat{\tilde{\mathbf{s}}} = \mathbf{Q}\tilde{\mathbf{y}} = \mathbf{Q}\mathbf{H}\tilde{\mathbf{s}} + \mathbf{Q}\tilde{\mathbf{v}} \in \mathbb{C}^K, \quad (5)$$

where $\mathbf{Q} \in \mathbb{C}^{K\times K}$ is the diagonal equalization matrix in the frequency domain, which is typically computed using the zero-forcing (ZF) criterion [15].

In order to evaluate the peaks of each OFDM symbol ($\mathbf{s}$), its PAPR can be defined as

$$\operatorname{PAPR}(\mathbf{s}) = \frac{\max\limits_{1\le n\le K} \left|[\mathbf{s}]_n\right|^2}{\mathbb{E}\left\{\|\mathbf{s}\|_2^2\right\}/K} \propto \max_{1\le n\le K} \left|[\mathbf{s}]_n\right|^2 \quad (6)$$

where the proportionality is accurate since the average power of each OFDM symbol is the same value when the number of subcarriers is large enough, and hence, the denominator of (6) can be discarded for comparison purposes. Since the PAPR is a random variable, it is typically represented by its cumulative distribution function (CDF) (peak value versus probability).

III. Proposed Dual Layers-Superimposed Training (DL-ST)

The joint optimization problem for channel estimation and PAPR reduction will be stated in this section in order to show that it is intractable in terms of complexity. Hence, a low-complexity DL-ST is proposed as a realistic alternative. Each layer is responsible for providing the suitable constant amplitude TS in order to either estimate the channel or remove the high peaks. The TS-1L is in charge of estimating the channel coefficients and equalizing the data symbols, and it is obtained by an offline optimization given in [4], [6]. The TS-2L is designed to remove the high peaks of each OFDM symbol, and it can be obtained by using a low-overhead single-TR with an additional phase shift. The best value of the phase shift is a data-dependent value and it can be obtained by using a low-complexity solver based on codebook search. Additionally, the TS-2L not only must be transparent from the perspective of the channel estimation process in order to avoid degrading its accuracy but it must be easily detected and removed at the receiver.

A. Joint Optimization Problem

After IDFT operation given in (1), the TS is superimposed to the data symbols as

$$\mathbf{x} = \sqrt{\beta_s}\mathbf{s} + \sqrt{1-\beta_s}\mathbf{p} \in \mathbb{C}^K, \quad 0 < \beta_s < 1, \tag{7}$$

where β_s is the power allocated to the data symbols and $\mathbf{p}$ is the superimposed TS.

Inspecting (7) and according to [10], a proper choice of the superimposed constant amplitude TS ($\mathbf{p}$) is capable of improving the overall performance of the link. The general optimization problem to be solved can be expressed as

$$\max_{\mathbf{p}} \ \text{SINR}(\mathbf{p}), \quad \text{s.t.} \left|[\mathbf{p}]_n\right| = 1, \quad 1 \le n \le K, \tag{8}$$

$$\text{SINR}(\mathbf{p}) \approx \frac{\beta_s}{\sigma^2_{\Delta h}(\mathbf{p}) + \sigma^2_c(\mathbf{p}) + \sigma^2_w}, \tag{9}$$

where the average performance, measured by the signal to interference and noise ratio (SINR) of the received data symbols ($\widehat{\mathbf{s}}$), relies on the quality of the channel estimates obtained at the receiver ($\sigma^2_{\Delta h}$) [16] and the interference produced as the result of the clipping effect caused to the OFDM symbol by the non-linear region of the PA [3]. Consequently, the choice of a suitable superimposed TS is relevant to provide a significant impact on the overall performance by minimizing the channel estimation error and reducing the occurrence of high peaks. The problem described by (8) has not been solved yet in real-time since it is a non-convex problem. Note that it may be solved by using artificial intelligence, specifically designed for these kinds of non-convex problems, at the expense of significantly increasing the processing time and/or resources that are not affordable, especially for low-latency communication and/or high-mobility scenarios.

B. DL-ST

In order to circumvent the high complexity issue given by the joint optimization problem given in (8), the DL-ST is proposed. At the transmitter, two TSs are superimposed on the data symbols as

$$\mathbf{x}_{TL} = \left(\sqrt{\beta_s}\mathbf{s} + \sqrt{\beta_1}\mathbf{p}_1\right) + \sqrt{\beta_2}\mathbf{p}_2 = \mathbf{x} + \sqrt{\beta_2}\mathbf{p}_2 \in \mathbb{C}^K, \tag{10}$$

$$0 < \beta_s, \beta_1, \beta_2 < 1, \quad \beta_s + \beta_1 + \beta_2 = 1, \tag{11}$$

where β_1 and β_2 are the power allocated to the TSs of the first and second layers, respectively. Note that, (10) can be seen as the traditional ST ($\mathbf{x}$), defined in (7), plus an additional second layer. Albeit (10) and (11) are pointing out that the additional power spent for the second layer may represent an inefficiency to the system, however, this second layer will be capable of reducing the PAPR further with a reduced amount of allocated energy (e.g. $\beta_2 = 0.05 - 0.1$), as it will be shown.

Taking into account (10) and (11), the original optimization problem described in (8) can be transformed into the following two independent optimization problems

$$\min_{\mathbf{p}_1} \ \sigma^2_{\Delta h}(\mathbf{p}_1), \ \text{s.t.} \left|[\mathbf{p}_1]_n\right| = 1, \ 1 \le n \le K, \tag{12}$$

$$\min_{\mathbf{p}_2} \ \sigma^2_c(\mathbf{p}_2), \ \text{s.t.} \left|[\mathbf{p}_2]_n\right| = 1, \ 1 \le n \le K, \ \text{given } \mathbf{s} \ \& \ \mathbf{p}_1, \tag{13}$$

where (12) is seeking for a constant amplitude TS-1L ($\mathbf{p}_1$) capable of minimizing the channel estimation error by assuming that the TS-2L is transparent, while (13) is searching for the best constant amplitude TS-2L ($\mathbf{p}_2$) which is able to reduce PAPR for a given OFDM symbol ($\mathbf{s}$) and the chosen TS-1L ($\mathbf{p}_1$). The complexity of solving these two optimization problems is much lower than (8) due to the fact that only the PAPR reduction optimization, given in (13), requires to be solved in real-time, while the channel estimation optimization problem given in (12) is data-independent and can be executed offline [4], [6].

C. Design of the TS-1L for Channel Estimation

Let us define a short constant amplitude TS for channel estimation purposes $\mathbf{p}_b \in \mathbb{C}^{L_p}$ as

$$[\mathbf{p}_b]_n = \exp(j\phi_n), \quad 1 \le n \le L_p, \quad L_{CH} \le L_p \le L_{CP} \tag{14}$$

where $\phi_n \in \mathbb{R}$ is the phase of the n-th unit modulus pilot symbol and L_p corresponds to the sequence length. Then, the TS-1L is a cyclic repetition of $\mathbf{p}_b$, given by

$$\mathbf{p}_1 = \mathbf{1}_{(N_p \times 1)} \otimes \mathbf{p}_b \in \mathbb{C}^K, \quad K = N_p L_p. \tag{15}$$

where N_p is the number of times that each ϕ_n appears in the constant amplitude TS. Note that the number of different phases should have, at least, the same number of taps of the channel ($L_p \ge L_{CH}$) in order to allow the estimation of all of them by using just a simple LS criterion [7]. Moreover, the N_p repetitions of a block of L_p different phases within the OFDM symbol will allow to average over these N_p blocks within one single OFDM symbol.

Given the TS-1L which is a block repetition of $\mathbf{p}_b$ and assuming that the TS-2L is transparent in the channel estimation process, a block averaging process over the N_p blocks within the OFDM symbol is performed to obtain $\overline{\mathbf{y}} \in \mathbb{C}^{L_p}$ as

$$\overline{\mathbf{y}} = \frac{1}{N_p}\left(\mathbf{1}_{(1\times N_p)} \otimes \mathbf{I}_{L_p}\right)\mathbf{y} = \mathbf{h} \circledast \left(\sqrt{\beta_s}\overline{\mathbf{s}} + \sqrt{\beta_1}\mathbf{p}_1\right) + \overline{\mathbf{v}}, \tag{16}$$

$$[\overline{\mathbf{s}}]_n = \frac{1}{N_p}\sum_{u=0}^{N_p-1}[\mathbf{s}]_{\text{mod}\left(n+uN_p-\tau, K\right)+1}, \tag{17}$$

$$[\overline{\mathbf{v}}]_n = \frac{1}{N_p}\sum_{u=0}^{N_p-1}[\mathbf{v}]_{uN_p+n}, \quad 1 \le n \le L_p, \tag{18}$$

where $\bar{\mathbf{s}}$ is the averaged self-interference induced by the data symbols ($\mathbf{s}$) and $\bar{\mathbf{v}}$ corresponds to the averaged noise samples. Their corresponding variances are given by $\sigma_{\bar{s}}^2 = \beta_s/N_p$ $\sigma_{\bar{v}}^2 = \sigma_v^2/N_p$, respectively. Note that, the TS-2L will be specifically designed to be transparent for the channel estimation process, and therefore, the interference source produced by TS-2L will be nullified, as will be shown in Subsection III-D.

In order to ease the notation, (16) can be rewritten in matrix form as

$$\bar{\mathbf{y}} = \left(\sqrt{\beta_s}\mathbf{S} + \sqrt{\beta_1}\mathbf{P}_1\right)\mathbf{h} + \bar{\mathbf{v}}, \tag{19}$$

where $\mathbf{P}_1 \in \mathbb{R}^{L_p \times L_p}$ and $\mathbf{S} \in \mathbb{R}^{L_p \times L_p}$ are Toeplitz matrices defined by

$$\mathbf{P}_1 = \begin{bmatrix} \exp(j\phi_1) & \exp(j\phi_2) & \cdots & \exp\left(j\phi_{L_p}\right) \\ \exp\left(j\phi_{L_p}\right) & j\exp(j\phi_1) & \cdots & j\phi_{L_p-1} \\ \vdots & \vdots & \ddots & \vdots \\ \exp(j\phi_2) & \cdots & \exp\left(j\phi_{L_p}\right) & \exp(j\phi_1) \end{bmatrix}, \tag{20}$$

and $\mathbf{S}$ are built by using (20) and replacing the training sequence by $\bar{\mathbf{s}}$.

Considering (19), LS estimation [7] can be applied in the time domain. Assuming that the matrix $\mathbf{P}_1$ is known at the receiver, the estimated channel can be decomposed as

$$\widehat{\mathbf{h}} = \frac{1}{\sqrt{\beta_1}}\mathbf{P}_1^{-1}\bar{\mathbf{y}} = \mathbf{h} + \frac{1}{\sqrt{\beta_1}}\mathbf{P}_1^{-1}\left(\sqrt{\beta_s}\mathbf{S}\mathbf{h} + \bar{\mathbf{v}}\right), \tag{21}$$

The estimator given in (21) is unbiased since the averaged noise ($\bar{\mathbf{v}}$) and data interference ($\mathbf{s}$) are zero mean random variables.

The MSE of the channel estimation can be derived as

$$\begin{aligned}\sigma_{\Delta h}^2 &= \mathbb{E}\left\{\left|\widehat{\mathbf{h}} - \mathbf{h}\right|^2\right\} = \mathbb{E}\left\{\left|\frac{1}{\sqrt{\beta_1}}\mathbf{P}_1^{-1}\left(\sqrt{\beta_s}\mathbf{S}\mathbf{h} + \bar{\mathbf{v}}\right)\right|^2\right\} \\ &= \frac{\sigma_v^2}{N_p\beta_1}\mathrm{tr}\left(\left(\mathbf{P}_1^H\mathbf{P}_1\right)^{-1}\right),\end{aligned} \tag{22}$$

where it is shown that the MSE of the channel estimates is only proportional to the noise power (σ_v^2) and the effect of the TS-2L is not included since it is cancelled by the averaging process as will be shown in the next subsection. Additionally, it is inversely proportional to the power allocated to the TS in the first layer (β_1) and the number of blocks to be averaged (N_p). The best TS-1L capable of minimizing (22), which corresponds to a non-convex optimization problem, has already been given in [4], [6].

D. Design of the TS-2L for PAPR Reduction

Inspecting (10) from a PAPR reduction perspective, it resembles the optimization problem described by the well-known TR technique [12]–[14]. However, the TR techniques reported in the literature are not appealing for deploying the next 6G mobile system, since they not only possess a high latency and complexity, but they also waste a lot of data subcarriers reserved for the PAPR reduction. Taking into account the literature and the constraint imposed by the channel estimation process, a low-complexity single-TR with an additional phase shift is proposed to find out the TS-2L capable of reducing the PAPR with a tiny complexity.

The proposed TS-2L only wastes a single subcarrier out of K, which is given by

$$\begin{gathered}[\mathbf{p}_2]_n = \exp\left(j\left(2\pi n\frac{k_a}{K} + \varphi\right)\right), \quad 1 \le n \le K, \\ 0 \le k_a \le K-1, \quad -\pi \le \varphi < \pi,\end{gathered} \tag{23}$$

where k_a and φ are the reserved subcarrier and the additional phase shift, respectively, for the PAPR reduction purpose.

Firstly, (23) must satisfy the condition of transparency required by the channel estimation process when the block averaging is executed, given in (16), as

$$\left(\mathbf{1}_{(1\times N_p)} \otimes \mathbf{I}_{L_p}\right)\mathbf{p}_2 = \mathbf{0}_{(L_p\times 1)}, \quad k_a \ne 0. \tag{24}$$

This property can be easily analyzed by substituting (23) in the block averaging process described in (24), and hence it is obtained that

$$\sum_{u=0}^{N_p-1}[\mathbf{p}_2]_{uL_p+n} = \left\{\begin{array}{cc} 0 & k_a \ne 0 \\ N_p & k_a = 0 \end{array}\right\}, \quad 0 \le n < L_p, \tag{25}$$

where it is shown that summing up multiple periods of an equally-sampled exponential function equals zero. Consequently, the MSE of the channel estimation given in (22) is not degraded by the TS-2L, since it is nullified by the block averaging process.

The choice of the reserved subcarrier k_a can be anyone in the range $1 \le k_a \le K-1$, excluding the first subcarrier $k_a = 0$. Making use of Parseval's relation, all the energy of the TS-2L is concentrated in the k_a-th subcarrier as

$$\sqrt{\beta_2}\tilde{\mathbf{p}}_2 = \sqrt{\beta_2}\mathbf{F}_k\mathbf{p}_2 = \left\{\begin{array}{cc} 0 & k \ne k_a \\ \sqrt{K\beta_2}\exp(j\varphi) & k = k_a \end{array}\right\}. \tag{26}$$

Although β_2 is typically small, the product of the power allocated to the TS-2L and the number of subcarriers is much higher than the energy allocated to the data symbols ($K\beta_2 >> 1$) since K is typically very large in multi-carrier waveforms. Therefore, it does not matter which subcarrier is selected by the transmitter for the PAPR cancellation since the receiver can easily measure and avoid it by using a simple amplitude detector in the frequency domain without using any additional side-link to convey this information.

Finally, the value of φ can be selected for the PAPR reduction. The optimization problem given in (13) can be transformed as

$$\min_{\varphi} \quad \mathrm{PAPR}(\mathbf{x}_{TL}), \quad \text{s.t.} \quad (23), \tag{27}$$

where it means looking for the best additional phase shift capable of reducing the PAPR. To solve this problem, a low-complexity codebook search can be adopted, where the value of φ is discretized and stored at the transceiver as

$$\varphi \in \mathcal{B} = \left\{\varphi_b = b\frac{2\pi}{B}, \quad 0 \le b \le B-1,\right\}, \tag{28}$$

TABLE I

COMPLEXITY COMPARISON IN TERMS OF THE NUMBER OF COMPLEX PRODUCTS AND THE NUMBER OF SUBCARRIERS USED FOR A PAPR REDUCTION AT THE PROBABILITY OF $Pr = 10^{-3}$.

Scheme	# Products	Reduct.	Subc. used
Kernel Matrix [12]	$2KK_aI_t$	4 dB	5%
Peak-Window. [13]	$K^2K_aI_t$	6 dB	5%
MS-SCR [14]	$K\left(\log_2(K)+I_t\right)$	5 dB	12%
Const. Amp. [10]	0	2.5 dB	0
DL-ST	0	5 dB	1

where $\mathcal{B}$ is the set that contains all the phases to be tested (φ_b) and B denotes the cardinality of the set.

Table I shows the complexity evaluated for the proposal as compared to some of the existing approaches, in terms of the number of complex products, the number of reserved subcarriers for TR (K_a) and the amount of PAPR reduction at the probability of $Pr = 10^{-3}$. Since the DL-ST scheme is similar to the TR, the complexity comparison is performed against some solutions based on this approach [12]–[14]. The techniques in the literature are wasting between $K_a/K \propto 5-12\%$ of the subcarriers for a PAPR reduction of typically $4-6$ dB. Besides, all these techniques have a considerable associated complexity since they first obtain the peaks in the time domain, and then, these peaks are projected in the frequency domain in order to find out what complex values should be allocated in the reserved subcarriers to avoid them. Moreover, this process is typically iterated several times (I_t) until the PAPR is lower than an established threshold. The proposed DL-ST does not require any complex products in order to perform the PAPR reduction, similarly to the traditional ST with the exploitation of a constant amplitude TS [10]. Note that all the trigonometric evaluations, such as phase rotation operations, can be easily performed by using the well-known low-complexity coordinate rotation digital computer (CORDIC) algorithm. Moreover, as it will be seen in the next section of numerical results, the proposed DL-ST has a better PAPR reduction as compared to the ST based on constant amplitude TS at the expense of only sacrificing one single subcarrier, what constitutes a negligible loss in terms of data-rate.

IV. PERFORMANCE EVALUATION

In this section, several numerical results are provided in order to show the performance of the proposed DL-ST for channel estimation and PAPR reduction as compared to the classical PSAM [17] and ST based on constant amplitude TS (ST-CA) schemes [10]. Note that those PAPR reduction techniques based on TR [12]–[14] are not considered in this section due to the fact that they possess an excessive complexity as compared to our proposal (see Table I).

A summary of the simulation parameters is given in Table II. The channel model adopted for the simulation corresponds to the tapped delay line (TDL) model proposed by the 3rd Generation Partnership Project (3GPP) to evaluate 5G performance [18], specifically the urban micro-cell (UMi) scenario is chosen. Moreover, the channel coefficients, given in (4), are obtained by using the TDL-A power-delay profile, which corresponds to a non-line-of-sigh (NLOS) channel model. Note that the length of the channel is ensured to be shorter than or equal to the length of the CP ($L_{CH} \leq L_{CP}$). The SNR can be defined as SNR $= \sigma_v^{-2}$ since the transmitted power and the channel gain are both normalized to one. The TS-1L for channel estimation is chosen according to [4], [6].

TABLE II

SIMULATION PARAMETERS

K	1024	β_s	$0.665-0.9$	$\mathbf{N_p}$	64	Δf	15 KHz
$\mathbf{L_{CP}}$	16	β_1	$0.285-0.3$	**B**	8	β_2	$0-0.1$

A. PAPR Performance

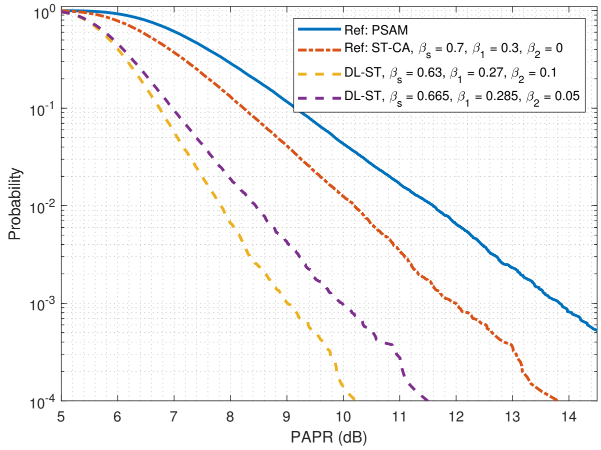
A comparison of the PAPR performance among the proposed DL-ST, PSAM and ST-CA is given in Figs. 1a and 1b, which correspond to different values of β_2 and β_1. In both figures, the proposed DL-ST significantly outperforms the existing techniques. In Fig. 1a, the improvement is about 2.5 and 3.4 dB with respect to ST-CA and 4 and 5 dB with respect to PSAM for $\beta_2 = 0.05$ and $\beta_2 = 0.1$ at the probability of $Pr = 10^{-3}$, respectively. In Fig. 1b, the improvement is 2.2 and 3.2 dB with respect to ST-CA and PSAM for $\beta_2 = 0.05$ and $\beta_2 = 0.1$, respectively. The improvement is higher in Fig. 1a than in Fig. 1b, because the power allocated to the TS-1L for channel estimation is significantly higher, verifying that the constant envelope sequence is capable of further reducing the PAPR as shown in [10], regardless the chosen phase components. Moreover, it can be seen that a tiny amount of power allocated to TS-2L ($\beta_2 = 0-0.1$) is capable of significantly reducing the PAPR.

B. Verification of MSE Performance

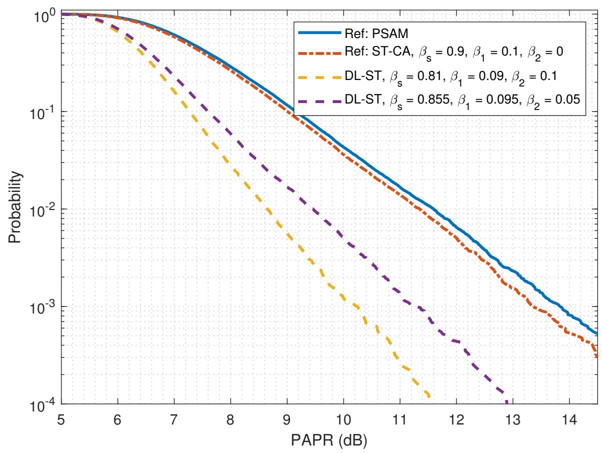
A comparison of the MSE obtained on the channel estimation between DL-ST and ST-CA techniques is illustrated in Fig. 2. Note that PSAM is discarded in this comparison since its MSE performance is significantly worse due to the absence of a noise averaging block [8]–[10]. Firstly, it must be remarked that the MSE of the channel estimates worsens as the power allocation to the TS-2L is increased. However, this degradation is negligible for the realistic values given in the numerical results ($\beta_2 = 0-0.1$). Hence, it can be concluded that the TS-2L not only is effectively reducing the PAPR, but it also has a negligible impact on the channel estimation performance. Additionally, the analytical results of the MSE for channel estimation based on the proposed DL-ST, given in (22), match with the simulation results, showing the accuracy of the theoretical expressions and how the design of TS-2L is fully transparent from the perspective of the first layer.

V. CONCLUSIONS

A DL-ST scheme is proposed in this work for joint channel estimation and PAPR reduction in OFDM. DL-ST is a novel contribution since the TS is made by two independent constant amplitude TSs, and each of them is generated for its specific layer. The TS-1L is responsible for performing the channel estimation, while the TS-2L is tailored for PAPR reduction.



(a) $\beta_s = 0.665 - 0.7$ and $\beta_1 = 0.285 - 0.3$.



(b) $\beta_s = 0.855 - 0.9$ and $\beta_1 = 0.095 - 0.1$.

Fig. 1. Comparison for PAPR for different values of $\beta_2 = 0 - 0.1$.

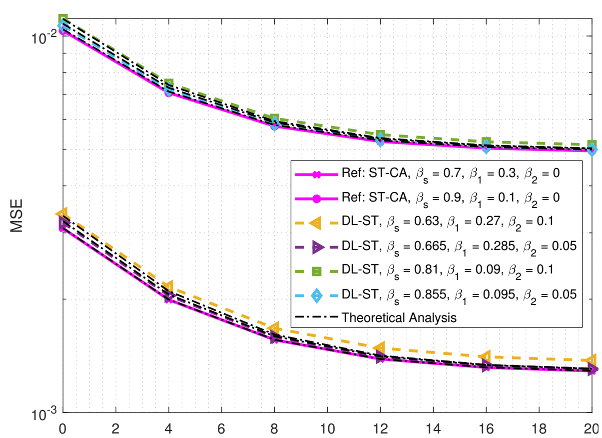


Fig. 2. Comparison for MSE performance for different values of $\beta_2 = 0 - 0.1$ and the theoretical analysis verification.

Additionally, the latter is specially designed to be transparent for the channel estimation process done with the first layer. By splitting the TS into two layers, the complexity of the system is significantly decreased and the performance is kept in terms of PAPR reduction, as compared to the existing techniques using OFDM. All these benefits have proven that the proposed DL-ST is a solid alternative to be implemented in communication systems for upcoming 6G, in situations such as low-latency links and high-mobility scenarios.

ACKNOWLEDGMENTS

This work has been partially funded by the Spanish National project IRENE-EARTH (PID2020-115323RB-C33/AEI /10.13039/501100011033) and the work of K. Chen-Hu was also, in part, supported by the Villum Investigator Grant "WATER" from the Velux Foundation, Denmark.

REFERENCES

[1] S. H. Han and J. H. Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission," *IEEE Wirel. Commun.*, vol. 12, no. 2, pp. 56–65, Apr. 2005.

[2] Y. Rahmatallah and S. Mohan, "Peak-to-average power ratio reduction in ofdm systems: A survey and taxonomy," *IEEE Commun. Surv. Tutorials*, vol. 15, no. 4, pp. 1567–1592, 2013.

[3] E. Costa, M. Midrio, and S. Pupolin, "Impact of amplifier nonlinearities on ofdm transmission system performance," *IEEE Commun. Lett.*, vol. 3, no. 2, pp. 37–39, Feb. 1999.

[4] A. Orozco-Lugo, M. Lara, and D. McLernon, "Channel estimation using implicit training," *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 240–254, Jan. 2004.

[5] K. Chen-Hu, M. J. F.-G. García, A. M. Tonello, and A. G. Armada, "Pilot pouring in superimposed training for channel estimation in CB-FMT," *IEEE Trans. Wirel. Commun.*, vol. 20, no. 6, Jun. 2021.

[6] K. Chen-Hu, M. Julia Fernández-Getino García, A. M. Tonello, and A. G. Armada, "Phase-domain injected training for channel estimation in constant envelope OFDM," *IEEE Trans. Wirel. Commun.*, vol. 22, no. 6, pp. 3869–3883, 2023.

[7] J. Lin, "Least-squares channel estimation for mobile OFDM communication on time-varying frequency-selective fading channels," *IEEE Trans. Veh. Technol.*, vol. 57, no. 6, pp. 3538–3550, Nov. 2008.

[8] G. Tan, Z. Li, J. Su, and H. Zhang, "Superimposed training for PTS-PAPR reduction in OFDM: A side information free data recovery scheme," in *2013 8th CHINACOM*, Aug. 2013, pp. 694–698.

[9] R. Patra, A. Mahapatro, H. B. Mishra, P. Singh, and S. Panda, "PAPR and CCDF analysis of superimposed training sequence-based MIMO-FBMC OQAM systems," in *2019 TENCON*, Oct. 2019, pp. 1489–1493.

[10] N. Chen and G. Zhou, "Superimposed training for OFDM: a peak-to-average power ratio analysis," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2277–2287, Jun. 2006.

[11] W. Liu, L. Zou, B. Bai, and T. Sun, "Low PAPR channel estimation for OTFS with scattered superimposed pilots," *China Commun.*, vol. 20, no. 1, pp. 79–87, 2023.

[12] P. Yu and S. Jin, "A low complexity tone reservation scheme based on time-domain kernel matrix for PAPR reduction in OFDM systems," *IEEE Trans. Broadcast.*, vol. 61, no. 4, pp. 710–716, Dec. 2015.

[13] J. Hou, J. Ge, and F. Gong, "Tone reservation technique based on peak-windowing residual noise for PAPR reduction in OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 64, no. 11, pp. 5373–5378, Nov. 2015.

[14] J. Wang, X. Lv, and W. Wu, "SCR-based tone reservation schemes with fast convergence for PAPR reduction in OFDM system," *IEEE Wirel. Commun. Lett.*, vol. 8, no. 2, pp. 624–627, Apr. 2019.

[15] A. Hedayat and A. Nosratinia, "Outage and diversity of linear receivers in flat-fading MIMO channels," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5868–5873, Dec. 2007.

[16] S. Ohno and G. Giannakis, "Average-rate optimal PSAM transmissions over time-selective fading channels," *IEEE Trans. Wirel. Commun.*, vol. 1, no. 4, pp. 712–720, Oct. 2002.

[17] M. K. Tsatsanis and Zhengyuan Xu, "Pilot symbol assisted modulation in frequency selective fading wireless channels," *IEEE Trans. Signal Process.*, vol. 48, no. 8, pp. 2353–2365, Aug. 2000.

[18] *Study on channel model for frequencies from 0.5 to 100 GHz (Release 17)*, 3GPP Std. 38.901, 2022.