CHANNEL ESTIMATION AND PAPR REDUCTION IN OFDM BASED ON DUAL LAYERS-SUPERIMPOSED TRAINING

Kun Chen-Hu[1,2], M. Julia Fernández-Getino García[2], Ana García Armada[2]
[1]Department of Electronic Systems, Aalborg University, Denmark. E-mail: kchenhu@es.aau.dk, [2]Department of Signal Theory and Communications of Universidad Carlos III of Madrid, Spain. E-mails: {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 6G. The data and pilot symbols are sharing the same time and frequency resources, and hence, the overhead is significantly reduced. Moreover, the superimposed pilots can be also used for the reduction of the Peak-to-Average Power Ratio (PAPR). However, a joint channel estimation and PAPR reduction procedure has not been addressed yet. 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 a second layer is designed for PAPR reduction and it 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 and achievable rate are improved while significant PAPR reductions are attained with negligible complexity.*

Keywords – Averaging, channel estimation, PAPR, superimposed training

1. INTRODUCTION

The Peak-to-Average Power Ratio (PAPR) [1, 2] is one of the most critical issues in multicarrier 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 a more expensive hardware, which increases the cost, or decreases their average output 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]. At the transmitter, a peak search and cancellation algorithm is typically executed for each OFDM symbol, which corresponds to a data-dependent optimization problem. Consequently, 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. Note that the implementation of this side-link comes at the expense of reducing the available bandwidth for transmitting the data-link. Simultaneously, Superimposed Training (ST) [4, 5, 6] is considered as one of the most appealing channel estimation techniques for OFDM in order to replace the well-known Pilot Symbol Assisted Modulation (PSAM) [7], used in 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) [8], 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 [9]. 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 multicarrier waveforms in 6G [10].

Lately, ST has not only been exploited for channel estimation purposes, but it is also proposed to reduce the PAPR of the OFDM signal [11, 12, 13]. On one hand, some works have combined the existing PAPR reduction techniques with ST [11, 12], but however the complexity issues described before are still present. On the other hand, ST based on using a Constant Amplitude TS (CA-TS) that is added in the time domain is proposed to jointly estimate the channel and reduce the PAPR with a tiny complexity [13]. This work showed that the PAPR reduction is proportional to the energy allocated to this CA-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 CA-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 CA-TS [13]. The superimposed TS is built as the sum of two CA-TSs, where each 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 a negligible complexity as compared to the classical ones. On one hand, it takes advantage of the CA-TS which is capable of reducing the PAPR without any additional operations. On the other hand, the TS-2L is a modified version of a TR scheme [14, 15, 16], whose optimization problem can be solved by a low-complexity codebook search.

The main contributions are summarized as follows:

- Firstly, the phase dimension of a CA-TS in ST is theoretically analyzed, showing that the PAPR of the OFDM symbol can be reduced even further if the phase component of the TS is also tailored given the samples of each OFDM symbol. This property will be exploited in the design of the TS-2L.
- A novel DL-ST scheme is proposed in order to jointly estimate the channel and reduce the PAPR of an OFDM symbol, with a negligible additional complexity as compared to the existing solutions. The TS is composed of two CA-TSs, and each of them is designed by its own layer. The TS-1L will provide a CA-TS capable of obtaining an accurate channel estimation. Then, the TS-2L is able to effectively reduce the PAPR of each OFDM symbol by mitigating further its strongest peaks. Note that the design of the TS-2L must be also transparent from the perspective of the channel estimation process, otherwise it would be an additional undesirable interference. Furthermore, the TS-2L should be also easily canceled by the receiver without the need for deploying an additional side-link. Considering these new constraints and inspired by the classical TR scheme [14, 15, 16], the proposed solution for the TS-2L corresponds to the exploitation of a phase shifted single-TR, which is a low-complexity and low-overhead alternative, and it is also a CA-TS.
- 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 on the system performance. Moreover, a complexity comparison, in terms of the required amount of products, among the proposed TS-2L and the existing techniques is given. It is shown that the complexity of our proposed scheme is hugely lower with a significant PAPR reduction, making it especially suitable for realistic low-latency communication links.
- The benefit of DL-ST is also evaluated via numerical results, providing the simulations of PAPR, MSE and achievable rates. These simulations point out the advantages of this approach against the existing solutions and validate the theoretical analysis.

The remainder of the paper is organized as follows. Section 2 introduces the system model of an OFDM system. Section 3 theoretically analyzes the phase dimension of a CA-TS for PAPR reduction. Section 4 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 5 presents several numerical results for the proposed scheme, providing an assessment of the achieved performance. Finally, in Section 6, 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 $(M \times N)$. $\mathbf{A} = \text{diag}\,(\mathbf{a})$ is a diagonal matrix whose diagonal elements are formed by the elements of vector $\mathbf{a}$. $\text{tr}\,(\cdot)$ corresponds to the matrix trace operation. $\circledast$ is the circular convolution operation. $\otimes$ corresponds to the Kronecker product of two matrices. $\mathbb{E}\left\{\cdot\right\}$ 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.

2. 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 multipath 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 complex constellation. The time domain OFDM symbol ($\mathbf{s}$) can be obtained as

$$\mathbf{s} = \mathbf{F}_K^H \tilde{\mathbf{s}} \in \mathbb{C}^K, \quad \mathbb{E}\left\{\left|[\tilde{\mathbf{s}}]_k\right|^2\right\} = 1, \quad 1 \leq k \leq K, \tag{1}$$

$$[\mathbf{F}_K]_{k_1+1,k_2+1} = \frac{1}{\sqrt{K}} \exp\left(-j\frac{2\pi}{K}k_1k_2\right), \tag{2}$$

where $0 \leq k_1, k_2 \leq K-1$, $\mathbf{F}_K \in \mathbb{C}^{K\times K}$ 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 $\mathbf{s}$.

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 modeled as a circular convolution of K samples between the wireless channel and the transmitted signal, whose expression is given by

$$\mathbf{y} = \mathbf{h} \circledast \mathbf{s} + \mathbf{v}, \quad (3)$$

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 modeled as an independent complex random variable distributed as $\mathcal{CN}(0, \sigma_\tau^2)$ $1 \leq \tau \leq L_{CH}$, 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 ($T_c \geq (K + L_{CP}) T_s$), 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

$$\tilde{\mathbf{y}} = \mathbf{F}_K \mathbf{y} = \mathbf{H}\tilde{\mathbf{s}} + \tilde{\mathbf{v}} \in \mathbb{C}^K, \quad \tilde{\mathbf{v}} = \mathbf{F}_K \mathbf{v} \in \mathbb{C}^K, \quad (4)$$

$$\mathbf{H} = \text{diag}\left(\tilde{\mathbf{h}}\right) \in \mathbb{C}^{K\times K}, \quad \tilde{\mathbf{h}} = \mathbf{F}_K \mathbf{h} \in \mathbb{C}^K, \quad (5)$$

where $1 \leq n \leq K$, $\tilde{\mathbf{h}}$ is the channel frequency response and $\tilde{\mathbf{v}}$ accounts for the noise in 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

$$\hat{\tilde{\mathbf{s}}} = \mathbf{Q}\tilde{\mathbf{y}} = \mathbf{QH}\tilde{\mathbf{s}} + \mathbf{Q}\tilde{\mathbf{v}} \in \mathbb{C}^K, \quad (6)$$

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 [8].

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

$$\text{PAPR}(\mathbf{s}) = \frac{\max\limits_{1\leq n\leq K} \left|[\mathbf{s}]_n\right|^2}{\mathbb{E}\left\{\|\mathbf{s}\|_2^2\right\}/K} \propto \max_{1\leq n\leq K} \left|[\mathbf{s}]_n\right|^2 \quad (7)$$

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 *(7)* 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).

3. ANALYSIS OF THE PHASE DIMENSION OF A SUPERIMPOSED CA-TS

In this section, the phase dimension of the superimposed CA-TS is theoretically analyzed. According to [13], a CA-TS is the best superimposed TS in terms of reducing the high peaks of an OFDM symbol. However, the phase dimension of the CA-TS has been only tailored until now to improve the quality of the channel estimates. In this section, these phase values are analytically studied to show that an appropriate choice for these phases is also able to further reduce the PAPR without the need of degrading the channel estimates. This analytical result will be exploited in the proposed DL-ST, shown in Section 4.

After the IDFT operation given in *(1)*, the CA-TS is superimposed in the time domain to the data symbols as

$$\mathbf{x} = \sqrt{\beta_s}\mathbf{s} + \sqrt{1-\beta_s}\mathbf{p} \in \mathbb{C}^K, \quad 0 < \beta_s < 1, \quad (8)$$

where β_s is the power allocated to the data symbols, $\mathbf{p}$ is the superimposed CA-TS and $\mathbf{x}$ denotes the transmitted symbol vector. According to [13], the superimposed TS for channel estimation should have a constant amplitude ($\left|[\mathbf{p}]_n\right|^2 = 1, 1 \leq n \leq K$) in order to decrease the PAPR of the OFDM symbol. In order to ease the notation of the theoretical analysis, let us redefine the data symbols ($\mathbf{s}$) and the superimposed CA-TS ($\mathbf{p}$) as

$$[\mathbf{s}]_n = s_{r,n} + js_{i,n} = r_{s,n} \exp\left(j\theta_{s,n}\right), \quad (9)$$

$$[\mathbf{p}]_n = p_{r,n} + jp_{i,n} = \exp\left(j\theta_{p,n}\right), \quad -\pi \leq \theta_{p,n} < \pi, \quad (10)$$

respectively, where $1 \leq n \leq K$. Note that the distribution of the amplitude and phase components of the data symbols follow a Rayleigh and uniform distribution, respectively, whose expressions are given by

$$r_{s,n} \sim \mathcal{R}\left(1/\sqrt{2}\right), \quad \theta_{s,n} \sim \mathcal{U}\left[-\pi, \pi\right]. \quad (11)$$

After superimposing a known CA-TS ($\mathbf{p}$), let us define the two following probabilities as

$$F_{r,n}\left(r = r_{s,n}\right) = Pr\left(\left|[\mathbf{x}]_n\right|^2 \leq \left|[\mathbf{s}]_n\right|^2 \mid r = r_{s,n}, [\mathbf{p}]_n\right), \quad (12)$$

$$\begin{aligned} &F_{t,n}\left(r_{min} \leq r_{s,n}\right) = \\ &Pr\left(\left|[\mathbf{x}]_n\right|^2 \leq \left|[\mathbf{s}]_n\right|^2 \mid r_{min} \leq r_{s,n}, [\mathbf{p}]_n\right), \end{aligned} \quad (13)$$

$\forall n \in \{1, \cdots, K\}$, where both *(12)* and *(13)* are conditional probabilities to the chosen CA-TS ($\mathbf{p}$). However, *(12)* is also conditioned to a particular amplitude of the data symbol ($r = r_{s,n}$), while *(13)* is considering all those samples in the time domain whose amplitude is equal to or higher than a minimum chosen threshold ($r_{min} \leq r_{s,n}$). All these chosen samples correspond to those potential symbols associated with a high energy peak responsible for raising the PAPR.

Inspecting the argument of *(13)*, and making use of the definition of ST given in *(8)*, it can be developed as

$$\begin{aligned} &\sqrt{\beta_s(1-\beta_s)}\left([\mathbf{s}]_n [\mathbf{p}]_n^* + [\mathbf{s}]_n^* [\mathbf{p}]_n\right) \\ &+ (1-\beta_s)\left|[\mathbf{p}]_n\right|^2 + \beta_s \left|[\mathbf{s}]_n\right|^2 \leq \left|[\mathbf{s}]_n\right|^2, \end{aligned} \quad (14)$$

Considering that the TS has a constant amplitude, *(14)* can be simplified as

$$2\sqrt{\beta_s(1-\beta_s)}\Re\left\{[\mathbf{s}]_n [\mathbf{p}]_n^*\right\} \leq \left|[\mathbf{s}]_n\right|^2 (1-\beta_s) - (1-\beta_s), \quad (15)$$

$$\begin{aligned}\Re\left\{[\mathbf{s}]_n[\mathbf{p}]_n^*\right\} &= s_{r,n}p_{r,n} + s_{i,n}p_{i,n} \\ &\leq \frac{(1-\beta_s)\left(\left|[\mathbf{s}]_n\right|^2 - 1\right)}{2\sqrt{\beta_s(1-\beta_s)}},\end{aligned} \tag{16}$$

Then, taking into account the definitions given in *(9)* and *(10)*, *(16)* can be rewritten as

$$\left(\cos\left(\theta_{s,n}\right)\cos\left(\theta_{p,n}\right) + \sin\left(\theta_{s,n}\right)\sin\left(\theta_{p,n}\right)\right) \leq \rho_{s,n}, \tag{17}$$

$$\rho_{s,n} = \frac{\sqrt{1-\beta_s}\left(r_{s,n}^2 - 1\right)}{2\sqrt{\beta_s}r_{s,n}}, \tag{18}$$

where ρ_s is a positive real value which is proportional to the amplitude of the data samples ($r_{s,n}$). Finally, making use of some trigonometric identities, *(17)* can be simplified to

$$\cos\left(\theta_{s,n} - \theta_{p,n}\right) \leq \rho_{s,n}, \quad \forall n \in \{1, \cdots, K\}, \tag{19}$$

where it is shown that the PAPR of the OFDM symbols with ST given in *(8)* depends on the phase alignment between the data symbols (**s**) and the chosen CA-TS (**p**). If both phases ($\theta_{s,n}$ and $\theta_{p,n}$) have the same sign, the PAPR of the ST symbol of interest will be enhanced ($\cos(0) = 1$). On the contrary, the PAPR will be reduced if the phases have an opposite sign ($\cos(\pi) = -1$). Consequently, *(19)* is an alternative way to evaluate the PAPR of an OFDM symbol given in *(7)*. Note that *(19)* does not require the use of any complex product operation, and this low-complexity property will be exploited by the DL-ST, described in Section 4.4.

Considering the inequality given in *(19)*, the probability described in *(13)* can be rewritten as

$$\begin{aligned}&F_{t,n}\left(r_{min} \leq r_{s,n}\right) = \frac{1}{Pr\left(r_{s,n} \geq r_{min}\right)} \\ &\times \int_{r_{s,n}=r_{min}}^{\infty} F_{r,n}\left(r = r_{s,n}\right) f_r\left(r_{s,n}\right) \mathrm{d}r_{s,n},\end{aligned} \tag{20}$$

$$f_r\left(r_{s,n}\right) = 2r_{s,n}\exp\left(-r_{s,n}^2\right), \quad r_{s,n} \geq 0, \tag{21}$$

$$\begin{aligned}Pr\left(r_{s,n} \geq r_{min}\right) &= \int_{r_{s,n}=r_{min}}^{\infty} f_r\left(r_{s,n}\right) \mathrm{d}r_{s,n} \\ &= \exp\left(-r_{min}^2\right),\end{aligned} \tag{22}$$

where $Pr(r_s)$ is the probability density function of the Rayleigh distribution given in *(11)* and $Pr(r_s \geq r_{min})$ is the normalization factor that guarantees that $0 \leq F_{t,n}\left(r_{min} \leq r_{s,n}\right) \leq 1$ for the chosen r_{min}.

Then, $F_{r,n}\left(r = r_{s,n}\right)$ can be expressed as

$$\begin{aligned}&F_{r,n}\left(r = r_{s,n}\right) = \frac{1}{2\pi} \\ &\times \int_{\theta_{s,n}=-\pi}^{\pi} Pr\left(\cos\left(\theta_{s,n} - \theta_{p,n}\right) \leq \rho_{s,n}\right) \mathrm{d}\theta_{s,n}.\end{aligned} \tag{23}$$

By applying some trigonometric properties, the close-form expression of $F_{r,n}\left(r = r_{s,n}\right)$ is given by

$$\begin{aligned}F_{r,n}\left(r = r_{s,n}\right) &= 1 - \frac{\arccos\left(\rho_{s,n}\right)}{\pi} \\ &= \left\{\begin{array}{ll} < 0.5 & r_{s,n} < 1 \\ \geq 0.5 & r_{s,n} \geq 1 \end{array}\right\},\end{aligned} \tag{24}$$

where it is shown that the probability of reducing the amplitude of the peaks (the phases are misaligned) is always greater than 50% when the amplitude of the peak is higher than the unity, which corresponds to the interesting region for PAPR reduction. Additionally, this probability is further enhanced as the amplitude of the peaks is also increased. Note that *(24)* is an alternative way to show that a CA-TS is capable of reducing the highest peaks of an OFDM symbol as in [13]. However, this analysis is also providing the relation of the phase component between the data symbols and a CA-TS in order to reduce the PAPR even further. This property will be exploited by the second layer of the proposed TL-ST, as shown in the following sections.

Finally, substituting *(21)*-*(24)* in *(20)*, its analytical expression can be given by

$$\begin{aligned}&F_{t,n}\left(r_{min} \leq r_{s,n}\right) = 1 - \frac{2}{\pi}\exp\left(r_{min}^2\right) \\ &\times \int_{r_{s,n}=r_{min}}^{\infty} r_{s,n}\exp\left(-r_{s,n}^2\right)\arccos\left(\rho_{s,n}\right) \mathrm{d}r_{s,n}.\end{aligned} \tag{25}$$

4. 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 CA-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 a 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.

4.1 Joint optimization problem

Inspecting *(8)* and according to [13], an appropriate choice of the superimposed CA-TS (**p**) should be capable

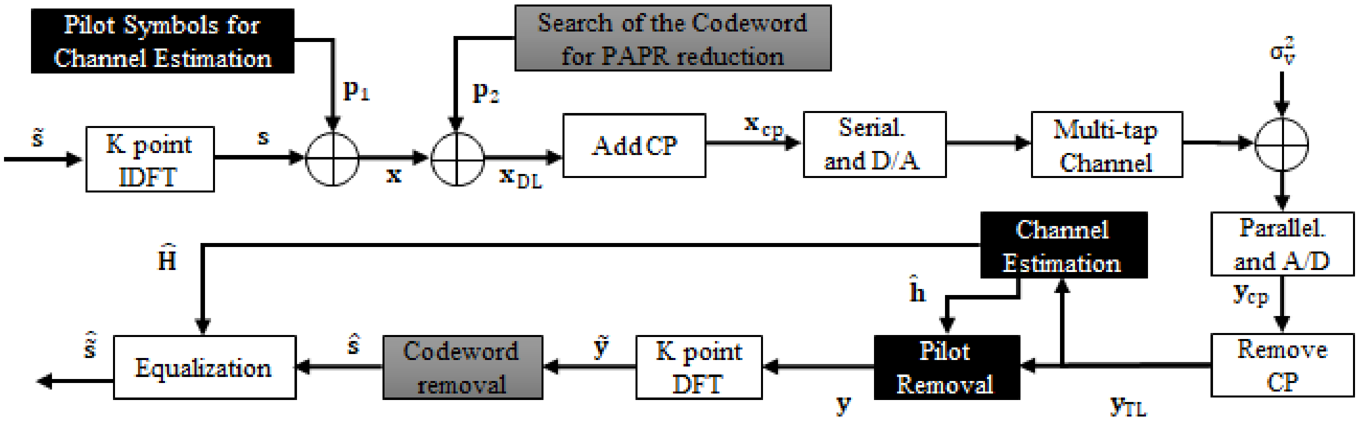


Fig. 1 – Block diagram of DL-ST in OFDM .The two superimposed TSs are designed for jointly performing the channel estimation and PAPR reduction.

of improving the overall performance of the link. The general optimization problem [4, 11, 12, 14] to be solved can be expressed as

$$\max_{\mathbf{p}} \ \mathrm{SINR}\left(\mathbf{p}\right), \quad \text{s.t.} \left|\left[\mathbf{p}\right]_n\right| = 1, \quad 1 \leq n \leq K, \qquad (26)$$

$$\mathrm{SINR}\left(\mathbf{p}\right) \approx \frac{\beta_s}{\sigma_{\Delta h}^2\left(\mathbf{p}\right) + \sigma_c^2\left(\mathbf{p}\right) + \sigma_w^2}, \qquad (27)$$

where the average performance [4, 14], measured by the Signal to Interference and Noise Ratio (SINR) of the received data symbols $(\tilde{\mathbf{s}})$, should be maximized. Its performance directly relies on the quality of the channel estimates obtained at the receiver $(\sigma_{\Delta h}^2)$ [17] 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]. Note that if either the channel estimates are not accurate enough or the PAPR is not properly reduced, the overall performance will be compromised since the received signal will be significantly degraded. 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 *(26)* 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, such as evolutionary computation algorithms [18], specifically designed for these kind 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.

4.2 DL-ST

In order to circumvent the high complexity issue given by the joint optimization problem in (*(26)*), the DL-ST is proposed. At the transmitter, two TSs are superimposed to the data symbols as

$$\begin{aligned} \mathbf{x}_{DL} &= \left(\sqrt{\beta_s}\mathbf{s} + \sqrt{\beta_1}\mathbf{p}_1\right) + \sqrt{\beta_2}\mathbf{p}_2 \\ &= \mathbf{x} + \sqrt{\beta_2}\mathbf{p}_2 \in \mathbb{C}^K, \end{aligned} \qquad (28)$$

$$0 < \beta_s, \beta_1, \beta_2 < 1, \quad \beta_s + \beta_1 + \beta_2 = 1, \qquad (29)$$

where β_1 and β_2 are the power allocated to the TSs of the first and second layers, respectively. Note that, *(28)* can be seen as the traditional ST $(\mathbf{x})$, defined in *(8)*, plus an additional second layer $(\sqrt{\beta_2}\mathbf{p}_2)$. Albeit *(28)* and *(29)* is 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. Note that the TS-1L designed for channel estimation is firstly added, while the TS-2L tailored for the PAPR is added later. It is crucial that the TS-2L should be the last sequence to be superimposed, since it is capable of reducing any new potential raised high peaks produced by superimposing the TS-1L. Additionally, the TS-1L does not depend on either the data symbols or channel coefficients, and hence, its optimization problem *(30)* can be executed offline [4, 6] if the TS-2L is properly designed to be transparent to the first layer. Hence, the obtained TS for the first layer can be stored in the memory of the transceiver and be selected according to the higher layers, and hence the complexity is significantly reduced. Taking into account *(28)* and *(29)*, the original optimization problem described in *(26)* can be transformed into the following two independent optimization problems

$$\min_{\mathbf{p}_1} \ \sigma_{\Delta h}^2\left(\mathbf{p}_1\right), \ \text{s.t.} \left|\left[\mathbf{p}_1\right]_n\right| = 1, \qquad (30)$$

$$\min_{\mathbf{p}_2} \ \sigma_c^2\left(\mathbf{p}_2\right), \ \text{s.t.} \left|\left[\mathbf{p}_2\right]_n\right| = 1, \ \text{given } \mathbf{s} \ \& \ \mathbf{p}_1, \qquad (31)$$

where $1 \leq n \leq K$. Eq. *(30)* is the optimization problem responsible for seeking a CA-TS-1L $(\mathbf{p}_1)$ capable of minimizing the channel estimation error by assuming that the TS-2L is fully transparent, while *(31)* accounts for another optimization problem focusing on searching for the best CA-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 *(26)* due to the fact that only the PAPR reduction optimization, given in *(31)*, requires solving in real time, while the channel estimation optimization problem given in *(30)* is data-independent and can be executed offline [4, 6].

4.3 Design of the TS-1L for channel estimation

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

$$[\mathbf{p}_b]_n = \exp\left(j\phi_n\right),\, 1 \leq n \leq L_p,\, L_{CH} \leq L_p \leq L_{CP}, \tag{32}$$

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

$$\mathbf{p}_1 = \mathbf{1}_{(N_p \times 1)} \otimes \mathbf{p}_b \in \mathbb{C}^K, \quad K = N_p L_p. \tag{33}$$

where N_p is the number of times that each ϕ_n appears in the CA-TS. Note that the number of different phases should have, at least, the same number of taps of the channel ($L_p \geq L_{CH}$) in order to allow the estimation of all of them by using just a simple LS criterion [9]. Moreover, the N_p repetitions of a block of L_p different phases within the OFDM symbol will allow averaging 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 $\bar{\mathbf{y}} \in \mathbb{C}^{L_p}$ as

$$\bar{\mathbf{y}} = \frac{1}{N_p}\left(\mathbf{1}_{(1\times N_p)} \otimes \mathbf{I}_{L_p}\right)\mathbf{y} = \mathbf{h} \circledast \left(\sqrt{\beta_s}\bar{\mathbf{s}} + \sqrt{\beta_1}\mathbf{p}_1\right) + \bar{\mathbf{v}}, \tag{34}$$

$$[\bar{\mathbf{s}}]_n = \frac{1}{N_p}\sum_{u=0}^{N_p-1}[\mathbf{s}]_{\mathrm{mod}(n+uN_p-\tau,K)+1}, \tag{35}$$

$$[\bar{\mathbf{v}}]_n = \frac{1}{N_p}\sum_{u=0}^{N_p-1}[\mathbf{v}]_{uN_p+n}, \quad 1 \leq n \leq L_p, \tag{36}$$

where $\bar{\mathbf{s}}$ is the averaged self-interference induced by the data symbols ($\mathbf{s}$) and $\bar{\mathbf{v}}$ which correspond to the averaged noise samples. Their corresponding variances are given by $\sigma^2_{\bar{s}} = \beta_s/N_p$ $\sigma^2_{\bar{v}} = \sigma^2_v/N_p$, respectively. Note that the TS-2L ($\mathbf{p}_2$) 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 Section 4.4.

In order to ease the notation, *(34)* can be rewritten in matrix form as

$$\bar{\mathbf{y}} = \left(\sqrt{\beta_s}\mathbf{S} + \sqrt{\beta_1}\mathbf{P}_1\right)\mathbf{h} + \bar{\mathbf{v}}, \tag{37}$$

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

$$\mathbf{P}_1 = \exp\left(j\begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_{L_p} \\ \phi_{L_p} & \phi_1 & \cdots & \phi_{L_p-1} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_2 & \cdots & \phi_{L_p} & \phi_1 \end{bmatrix}\right), \tag{38}$$

and $\mathbf{S}$ are built by using *(38)* and replacing the training sequence by $\bar{\mathbf{s}}$.

Considering *(37)*, LS estimation [9] 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

$$\hat{\mathbf{h}} = \frac{1}{\sqrt{\beta_1}}\mathbf{P}_1^{-1}\bar{\mathbf{y}} = \mathbf{h} + \frac{1}{\sqrt{\beta_1}}\mathbf{P}_1^{-1}\left(\sqrt{\beta_s}\mathbf{S}\mathbf{h} + \bar{\mathbf{v}}\right), \tag{39}$$

The estimator given in *(39)* 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

$$\sigma^2_{\Delta h} = \mathbb{E}\left\{\left|\hat{\mathbf{h}} - \mathbf{h}\right|^2\right\} = \mathbb{E}\left\{\left|\frac{1}{\sqrt{\beta_1}}\mathbf{P}_1^{-1}\left(\sqrt{\beta_s}\mathbf{S}\mathbf{h} + \bar{\mathbf{v}}\right)\right|^2\right\}, \tag{40}$$

$$\left|\hat{\mathbf{h}} - \mathbf{h}\right|^2 = \frac{1}{\sqrt{\beta_1}} \times \mathrm{tr}\left(\left(\mathbf{P}_1^{-1}\left(\sqrt{\beta_s}\mathbf{S}\mathbf{h} + \bar{\mathbf{v}}\right)\right)^H\left(\mathbf{P}_1^{-1}\left(\sqrt{\beta_s}\mathbf{S}\mathbf{h} + \bar{\mathbf{v}}\right)\right)\right). \tag{41}$$

Making use of the fact that the data symbols ($\mathbf{s}$) and noise ($\mathbf{n}$) are independent random variables, and the channel gain is normalized to one, *(40)* can be simplified to

$$\sigma^2_{\Delta h} = \frac{\beta_s + \sigma^2_v}{N_p\beta_1}\mathrm{tr}\left(\left(\mathbf{P}_1^H\mathbf{P}_1\right)^{-1}\right), \tag{42}$$

where it is shown that the MSE of the channel estimates is only proportional to the noise power (σ^2_v) and the effect of the TS-2L is not included since it is canceled 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 *(40)*, which corresponds to a non-convex optimization problem, has already been given in [4, 6].

4.4 Design of the TS-2L for PAPR reduction

Inspecting *(28)* from a PAPR reduction perspective, it resembles the optimization problem described by the well-known TR technique [14, 15, 16]. However, the TR techniques reported in the literature are not appealing for deploying the next 6G mobile system, since they not only posses 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

$$[\mathbf{p}_2]_n = \exp\left(j\left(2\pi n\frac{k_a}{K} + \varphi\right)\right), \quad 1 \leq n \leq K, \tag{43}$$

$$0 \leq k_a \leq K-1, \quad -\pi \leq \varphi < \pi,$$

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

Firstly, *(43)* must satisfy the condition of transparency required by the channel estimation process when the block averaging is executed, given in *(34)*, as

$$\left(\mathbf{1}_{(1\times N_p)} \otimes \mathbf{I}_{L_p}\right) \mathbf{p}_2 = \mathbf{0}_{(L_p\times 1)}, \quad k_a \neq 0. \tag{44}$$

This property can be easily analyzed by substituting *(43)* in the block averaging process described in *(44)*, and hence it is obtained that

$$\sum_{u=0}^{N_p-1} [\mathbf{p}_2]_{uL_p+n} = \left\{ \begin{array}{cc} 0 & k_a \neq 0 \\ N_p & k_a = 0 \end{array} \right\}, \quad 0 \leq n < L_p, \tag{45}$$

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 *(40)* 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 anything in the range $1 \leq k_a \leq K-1$, excluding the first subcarrier $k_a = 0$. Making use of the Parseval's relation, all the energy of the TS-2L is concentrated in the k_a-th subcarrier as

$$\sqrt{\beta_2}\tilde{\mathbf{p}}_2 = \sqrt{\beta_2}\mathbf{F}_k\mathbf{p}_2 = \left\{ \begin{array}{cc} 0 & k \neq k_a \\ \sqrt{K\beta_2}\,\text{exp}\left(j\varphi\right) & k = k_a \end{array} \right\}. \tag{46}$$

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 multicarrier 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 *(31)* can be transformed as

$$\min_{\varphi} \quad \text{PAPR}\left(\mathbf{x}_{TL}\right), \quad \text{s.t.} \quad (43), \tag{47}$$

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

$$\varphi \in \mathcal{B} = \left\{\varphi_b = b\frac{2\pi}{B}, \quad 0 \leq b \leq B-1,\right\}, \tag{48}$$

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

Algorithm 1 Search algorithm based on codebook sweeping

```
Inputs: x, 𝓑, k_a, ρ_r, ρ_i
Initialization: ν = ∞, φ = 0
Compute the set 𝒩 using (49)
Compute the phases θ_{x,n} = arctan([x]_n), ∀n ∈ 𝒩
Compute the vector 2πnk_a/K, ∀n ∈ 𝒩
for all 1 ≤ b ≤ B do
    Select the phase: φ_b ∈ 𝓑
    Compute θ_{a,n} = φ_b + 2πnk_a/K, ∀n ∈ 𝒩
    Compute ν_b = Σ_{n∈𝒩} cos(θ_{x,n} − θ_{a,n})
    if ν_b < ν then
        ν = ν_b and φ = φ_b
return φ
```

In order to keep reducing the complexity further and shorten the processing time, *(47)* can be significantly simplified. Firstly, instead of analyzing all the K samples of each OFDM symbol, given in *(47)*, only a subset with the highest peaks is considered in the optimization problem. This approximation will not induce a degradation in the performance, since only the high peaks samples of the OFDM symbol are being cut off by the high power amplifier. Hence, before executing *(47)*, a selection is performed as

$$\mathcal{N} = \left\{n \ / \ \Re\left\{[\mathbf{x}]_n\right\} > \rho_r \ \text{ OR } \ \Im\left\{[\mathbf{x}]_n\right\} > \rho_i\right\}, \tag{49}$$

where the set $\mathcal{N}$ contains all those sample indexes of $\mathbf{x}$ where either the real or imaginary part are higher than their respectively thresholds (ρ_r and ρ_i).

Then, making use of the theoretical analysis provided in Section 3, the PAPR evaluation described in *(7)* can be substituted by its low-complexity alternative given by *(19)* as

$$\min_{\varphi} \quad \nu = \sum_{n\in\mathcal{N}} \cos\left(\theta_{x,n} - \theta_{2,n}\right), \quad \text{s.t.} \quad (43), \tag{50}$$

$$\theta_{2,n} = \varphi + 2\pi n k_a/K, \quad \forall n \in \mathcal{N}, \tag{51}$$

where ν is the objective function to be minimized. Consequently, the optimization problem can be simplified as a computation of the cosine distance among the angle of the data samples of an OFDM symbol and the superimposed TS. Algorithm 1 provides a pseudo-code of the proposed low-complexity algorithm, detailing the operations disclosed in *(47)*-*(51)*.

Table 1 shows the complexity evaluated for the proposal as compared to some of the existing approaches in the literature, in terms of number of required complex products, number of reserved subcarriers for TR (K_a) to highlight the overhead, and the amount of PAPR reduction at the probabilities of $Pr = 10^{-2}$ and 10^{-3} to point out the performance. Since the proposed DL-ST scheme is similar to the TR, the complexity comparison is performed against some existing solutions based on this approach [14, 15, 16]. The proposed techniques in the literature

Table 1 – Complexity comparison in terms of number of complex products and the amount of subcarriers used for a PAPR reduction at two probabilities

Scheme	# Products	Reduction $Pr = 10^{-2}$	Reduction $Pr = 10^{-3}$	Subc. used
Kernel Matrix [14]	$2KK_aI_t$	2.9 dB	4 dB	5%
Peak-Window. [15]	$K^2K_aI_t$	5 dB	6 dB	5%
MS-SCR [16]	$K\left(\log_2(K)+I_t\right)$	3.8 dB	5 dB	12%
Const. Amp. [13]	0	1.5 dB	2.5 dB	0
DL-ST	0	4 dB	5 dB	1

are wasting between $K_a/K \propto 5-12\%$ of the subcarriers for a PAPR reduction of typically $4-6$ dB. Moreover, all these techniques have a considerable associated complexity since they firstly 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. Furthermore, this process is typically iterated several times (I_t) until the PAPR is lower than an established threshold. According to Algorithm 1, the proposed DL-ST does not require any complex products in order to perform the PAPR reduction, which is similar to the traditional ST with the exploitation of a CA-TS [13]. 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 [19], whose accuracy of the results depends on the chosen number of operations involved in the computation. 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 CA-TS at the expense of only sacrificing one single subcarrier, which constitutes a negligible loss in terms of data-rate.

5. 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 [7] and ST based on CA-TS (ST-CA) schemes [13]. Note that those PAPR reduction techniques based on TR [14, 15, 16] are not considered in this section due to the fact that they possess an excessive complexity as compared to our proposal (see Table 1). A summary of the simulation parameters is given in Table 2.

Table 2 – Simulation parameters

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

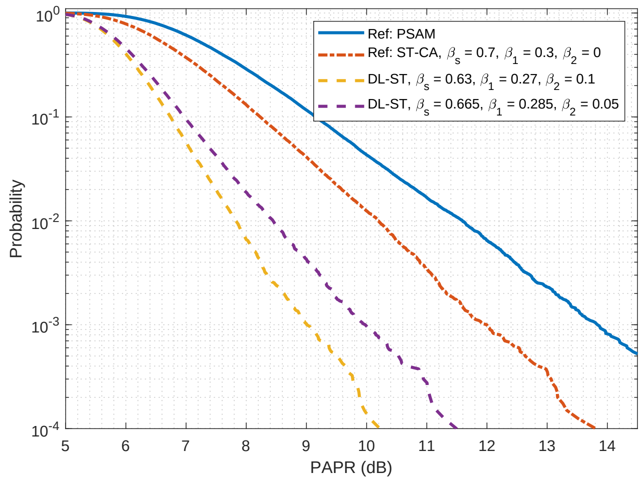
The channel model adopted for the simulation corresponds to the Tapped Delay Line (TDL) model proposed by the 3rd Generation Partnership Project (3GPP) to evaluate 5G performance [20], specifically the Urban Microcell (UMi) scenario is chosen. Moreover, the channel coefficients, given in *(5)*, are obtained by using the TDL-A power-delay profile whose delay spread is $\tau = 105$ ns, which corresponds to a Non-Line-of-Sight (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].

5.1 PAPR performance

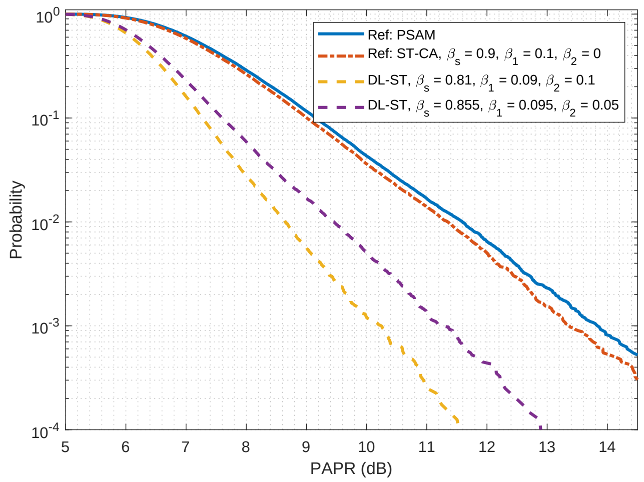
A comparison of the PAPR performance between the proposed DL-ST, PSAM and ST-CA is given in figures 2a and 2b, which correspond to different values of β_2 and β_1. In both figures, the proposed DL-ST significantly outperforms the existing techniques. In Fig. 2a, the improvement is about 2.5 and 3.4 dB 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. 2b, 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. 2a than in Fig. 2b 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 [13], regardless of 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.

5.2 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. 3. Note that PSAM is discarded in this comparison since its MSE performance is significantly worse due to the absence of a noise averaging block [11, 12, 13]. 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 *(40)*, 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.



(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. 2 – Comparison for PAPR for different values of $\beta_2 = 0 - 0.1$.

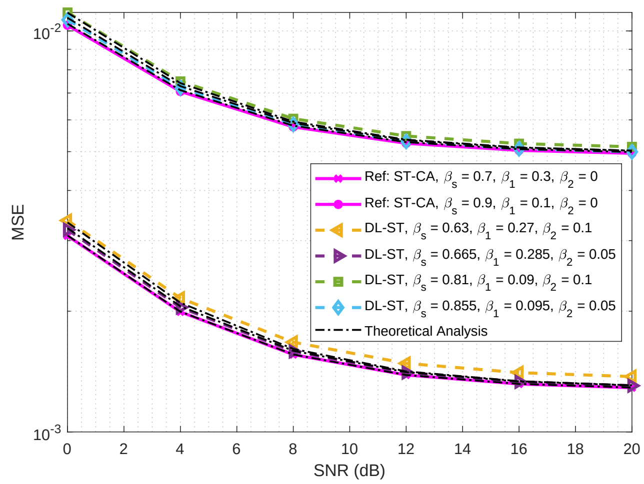


Fig. 3 – Comparison for MSE performance for different values of $\beta_2 = 0 - 0.1$ and the theoretical analysis verification.

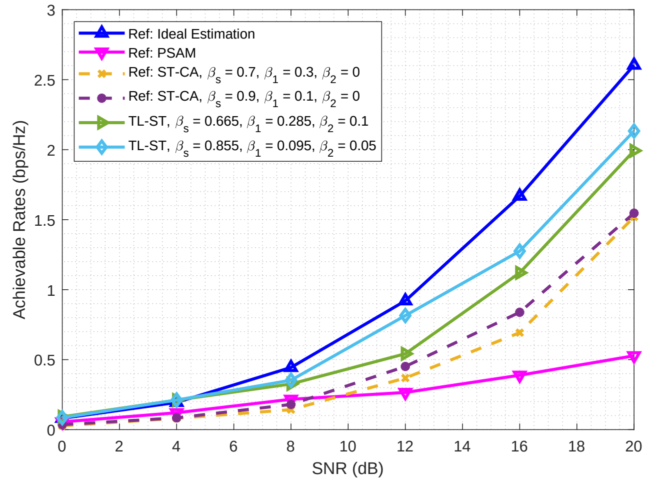


Fig. 4 – Comparison of achievable rates for different values of $\beta_2 = 0 - 0.1$ and using PA whose IBO = 8 dB.

5.3 Achievable rate performance with the PA effect

An achievable rate comparison of DL-ST, PSAM and ST-CA is given. Assuming Gaussian signaling, the achievable rate measured in bps/Hz can be obtained as

$$R = \mu \log_2 (1 + \mathrm{SNR}), \tag{52}$$

where μ accounts for the efficiency of the system due to the transmission of the reference symbols. For PSAM, the efficiency is given by $\mu = (1 - 1/3)$, while for DL-ST and ST-CA, the efficiency is $\mu = 1$ since no reference symbols are exclusively occupying a time/frequency resource.

In Fig. 4, a comparison of achievable rates among DL-ST, ST-CA and PSAM is plotted. The saturation point is set such that the input back-off corresponds to IBO $= 8$ dB and the order is set to $\lambda = 2$. Note that an upper-bound is also shown as a reference, which corresponds to the idealistic case where the channel estimates are perfectly given (dark blue line). The worst case is PSAM due to the fact that it possesses the highest PAPR and MSE while the lowest efficiency, which strongly penalizes the performance of the system. On the contrary, ST-CA is better than PSAM, since the constant amplitude reduces the PAPR and a lower number of data samples is clipped by the PA, and the MSE is better thanks to the averaging factor capable of reducing the undesirable noise and self-interference effects, and the efficiency is not penalized ($\mu = 1$). However, the proposed DL-ST is even better than ST-CA, since it not only inherits all the benefits of ST-CA, but it is also able to reduce the PAPR further thanks to the second layer, and hence its overall performance is the best as compared to the other schemes.

6. CONCLUSIONS AND FUTURE DIRECTIONS

A DL-ST scheme is proposed in this work for a joint channel estimation and PAPR reduction in OFDM. DL-ST is a novel contribution since the TS is made by two independent CA-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. Additionally, the latter has been especially designed to be transparent for the channel estimation process done with the first layer. By splitting the TS in 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 services (augmented reality, tactile Internet) and high-mobility scenarios (vehicular and satellite communications).

The PAPR reduction performance of DL-ST still relies on the amount of allocated power to the superimposed TSs, and hence the spectral efficiency may be compromised for some cases, as shown in the numerical results. Novel techniques based on artificial intelligence could find a new set of TSs for this purpose, such as evolutionary computation. These algorithms are specifically designed for solving non-convex optimization problems, and can be executed offline to seek new TSs for different aims, such as channel estimation, PAPR reduction, phase-noise compensation, among others.

ACKNOWLEDGMENT

This work has been partially funded by the Spanish National project IRENE-EARTH (PID2020-115323RB-C33AEI /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] Seung Hee Han and Jae Hong Lee. "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission". In: *IEEE Wirel. Commun.* 12.2 (Apr. 2005), pp. 56–65.

[2] Yasir Rahmatallah and Seshadri Mohan. "Peak-To-Average Power Ratio Reduction in OFDM Systems: A Survey And Taxonomy". In: *IEEE Commun. Surv. Tutorials* 15.4 (2013), pp. 1567–1592. ISSN:1553-877X.

[3] E. Costa, M. Midrio, and S. Pupolin. "Impact of amplifier nonlinearities on OFDM transmission system performance". In: *IEEE Commun. Lett.* 3.2 (Feb. 1999), pp. 37–39. ISSN:2373-7891.

[4] A.G. Orozco-Lugo, M.M. Lara, and D.C. McLernon. "Channel estimation using implicit training". In: *IEEE Trans. Signal Process.* 52.1 (Jan. 2004), pp. 240–254. ISSN:1941-0476.

[5] Kun Chen-Hu, M. Julia Fernández-Getino García, Andrea M. Tonello, and Ana García Armada. "Pilot Pouring in Superimposed Training for Channel Estimation in CB-FMT". In: *IEEE Trans. Wirel. Commun.* 20.6 (June 2021), pp. 3366–3380.

[6] Kun Chen-Hu, M. Julia Fernández-Getino García, Andrea M. Tonello, and Ana García Armada. "Phase-domain Injected Training for Channel Estimation in Constant Envelope OFDM". In: *IEEE Trans. Wirel. Commun.* (2022). ISSN:1558-2248.

[7] M. K. Tsatsanis and Zhengyuan Xu. "Pilot symbol assisted modulation in frequency selective fading wireless channels". In: *IEEE Trans. Signal Process.* 48.8 (Aug. 2000), pp. 2353–2365. ISSN:1941-0476.

[8] A. Hedayat and A. Nosratinia. "Outage and Diversity of Linear Receivers in Flat-Fading MIMO Channels". In: *IEEE Trans. Signal Process.* 55.12 (Dec. 2007), pp. 5868–5873. ISSN:1941-0476.

[9] J. Lin. "Least-Squares Channel Estimation for Mobile OFDM Communication on Time-Varying Frequency-Selective Fading Channels". In: *IEEE Trans. Veh. Technol.* 57.6 (Nov. 2008), pp. 3538–3550. ISSN:1939-9359.

[10] *5G-NR; Physical channels and modulation (Release 17)*. France: 3GPP, 2022.

[11] Guangtian Tan, Zhuyi Li, Jinli Su, and Han 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.

[12] Radhashyam Patra, Arunanshu Mahapatro, Himanshu B. Mishra, Prem Singh, and Sonali Panda. "PAPR and CCDF Analysis of Superimposed Training Sequence-based MIMO-FBMC OQAM Systems". In: *2019 TENCON*. Oct. 2019, pp. 1489–1493.

[13] Ning Chen and G.T. Zhou. "Superimposed training for OFDM: a peak-to-average power ratio analysis". In: *IEEE Trans. Signal Process.* 54.6 (June 2006), pp. 2277–2287. ISSN:1941-0476.

[14] Pingyuan Yu and Shubo Jin. "A Low Complexity Tone Reservation Scheme Based on Time-Domain Kernel Matrix for PAPR Reduction in OFDM Systems". In: *IEEE Trans. Broadcast.* 61.4 (Dec. 2015), pp. 710–716. ISSN:1557-9611.

[15] Jun Hou, Jianhua Ge, and Fengkui Gong. "Tone Reservation Technique Based on Peak-Windowing Residual Noise for PAPR Reduction in OFDM Systems". In: *IEEE Trans. Veh. Technol.* 64.11 (Nov. 2015), pp. 5373–5378. ISSN:1939-9359.

[16] Jingqi Wang, Xin Lv, and Wen Wu. “SCR-Based Tone Reservation Schemes With Fast Convergence for PAPR Reduction in OFDM System”. In: *IEEE Wirel. Commun. Lett.* 8.2 (Apr. 2019), pp. 624–627. ISSN: 2162-2345.

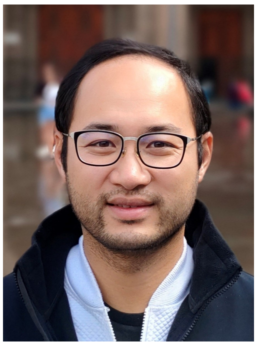
[17] S. Ohno and G.B. Giannakis. “Average-rate optimal PSAM transmissions over time-selective fading channels”. In: *IEEE Trans. Wirel. Commun.* 1.4 (Oct. 2002), pp. 712–720. ISSN:1558-2248.

[18] Andrew N. Sloss and Steven Gustafson. “2019 Evolutionary Algorithms Review”. In: *Genetic Programming Theory and Practice XVII*. Ed. by Wolfgang Banzhaf, Erik Goodman, Leigh Sheneman, Leonardo Trujillo, and Bill Worzel. Cham: Springer International Publishing, 2020, pp. 307–344. ISBN: 978-3-030-39958-0.

[19] Jack E. Volder. “The Birst of Cordic”. In: *J. VLSI Sig. Proc. Syst.* 25 (June 2000), pp. 101–105. ISSN:0922-5773.

[20] *Study on channel model for frequencies from 0.5 to 100 GHz (Release 17)*. France: 3GPP, 2022.

AUTHORS



Kun Chen-Hu received his Ph.D. degree in multimedia and communications in 2019 from Universidad Carlos III de Madrid (Spain). Currently, he is a postdoctoral researcher in Aalborg University (Denmark). He was awarded by UC3M in 2019 in recognition of his outstanding professional career after graduation. He visited Eurecom (France) and Vodafone Chair TU Dresden (Germany), both as guest researcher. He also participated in different research projects in collaboration with several top companies in the area of mobile communications. He was the Web Chair for Globecom 2021, Madrid (Spain), and online content editor for IEEE ComSoc. His research interests are related to signal processing techniques, such as waveforms design, reconfigurable intelligent surfaces, non-coherent massive MIMO and channel estimation.

M. Julia Fernández-Getino García received M. Eng. and Ph.D. degrees in telecommunication engineering from the Polytechnic University of Madrid, Spain, in 1996 and 2001, respectively. She is currently with the Department of Signal Theory and Communications, Carlos III University of Madrid, Spain, as an associate professor. From 1996 to 2001, she held a research position with the Department of Signals, Systems and Radiocommunications, Polytechnic University of Madrid. She visited Bell Laboratories, Murray Hill, NJ, USA, in 1998; visited Lund University, Sweden, during two periods in 1999 and 2000; visited Politecnico di Torino, Italy, in 2003 and 2004; and visited Aveiro University, Portugal, in 2009 and 2010. Her research interests include multicarrier communications, coding and signal processing for wireless systems. She received the best ”Master Thesis” and ”Ph.D. Thesis” awards from the Professional Association of Telecommunication Engineers of Spain in 1998 and 2003, respectively; the ”Student Paper Award” at the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC) in 1999; the ”Certificate of Appreciation” at the IEEE Vehicular Technology Conference (VTC) in 2000; the ”Ph.D. Extraordinary Award” from the Polytechnic University of Madrid in 2004; the ”Juan de la Cierva National Award” from AENA Foundation in 2004; and the ”Excellence Award” from Carlos III University of Madrid in 2012 for her research career.

Ana García Armada received a Ph.D. degree in electrical engineering from the Polytechnic University of Madrid in February 1998. She is currently a professor at Universidad Carlos III de Madrid, Spain. She is leading the Communications Research Group at this university. She has participated in more than 30 national and 10 international research projects, as well as having 20 contracts with the industry. Her research has resulted in nine book chapters, and more than 150 publications in prestigious international journals and conferences, as well as five patents. She has also contributed to standardization organizations (ITU, ETSI) and is a member of the European 5GPPP Group of Experts, as well as the Spanish representative on the committee of the ESA Joint Board on Communication Satellite Programs 5G Advisory Committee (5JAC). She has been editor (2016–2019, Exemplary Editor Award 2017 and 2018) and area editor (2019-2020, Exemplary Edi- tor Award 2020) of IEEE Communication Letters. She has been the editor of IEEE Transactions on Communications since 2019, area editor of IEEE Open Journal of the Communications Society since 2019, editor of the ITU Journal on Future and Evolving Technologies and is a regular member of the technical program committees of the most relevant international conferences in their field. She has formed / is part of the organizing committee of the IEEE Globecom 2019 and 2021 (General Chair), IEEE Vehicular Technology Conference Spring 2018, 2019 and Fall 2018, IEEE 5G Summit 2017, among others. She is Secretary of the IEEE ComSoc Signal Processing and Computing for Communications Committee, has been Secretary and Chair of the IEEE ComSoc Women in Communications Engineering Standing Committee. Since January 2020 she is Director of Online Content of the IEEE Communications Society. She has received the Award of Excellence from the Social Council and the Award for Best Teaching Practices from Universidad Carlos II de Madrid, as well as the third place Bell Labs Prize 2014, the Outstanding Service Award 2019 from the SPCE committee of the IEEE Communications Society and the Outstanding Service Award 2020 from the Women in Communications Engineering (WICE) standing committee.

