DOI: 10.1109/tbc.2018.2874542

Partial-Data Superimposed Training with Data Precoding for OFDM systems

Juan Carlos Estrada-Jiménez, *Student Member, IEEE*, and M. Julia Fernández-Getino García, *Member, IEEE*

***Abstract*—Superimposed Training (ST) is a recently addressed technique used for channel estimation where known training sequences are arithmetically added to data symbols, avoiding the use of dedicated pilot subcarriers, and thus, increasing the available bandwidth compared with traditional Pilot Symbol Assisted Modulation (PSAM) schemes. However, the system handles data interference over channel estimation as a result of the ST process; also, data detection is degraded by pilot interference. Recent ST methods have analyzed the data interference and presented schemes that deal with it. We propose a novel superimposed model over a precoded data scheme, named Partial-Data Superimposed Training (PDST), where an interference control factor assigns the adequate information level to be added to the training sequence in Orthogonal Frequency Division Multiplexing (OFDM) systems. Also, a data detection method is introduced to improve the Symbol Error Rate (SER) performance. Moreover, a capacity analysis of the system has been derived. Finally, simulation results confirm that performance of PDST is superior to previous proposals.**

***Index Terms*—Channel Estimation, Data Interference, OFDM, Superimposed Training.**

I. Introduction

In Orthogonal Frequency Division Multiplexing (OFDM) systems, channel estimation is a crucial process because it influences the overall performance [1] [2]. Pilot Symbol Assisted Modulation (PSAM), where pilots occupy certain time/slots subcarriers which waste valuable bandwidth [3], is commonly used. In broadcasting applications as Digital Video Broadcasting (DVB) scattered and continual pilots are considered for the framing structure in channel estimation [4].

Other methods like Superimposed Training (ST) arithmetically sum the data sequence with a known training sequence later used for channel estimation [5] [6]. Unlike PSAM, ST saves bandwidth because no resource is exclusively used for channel estimation. These ST schemes are very attractive for most wireless systems, especially for future next-generation 5G systems, where very high data rates are expected. New 5G technologies use ST to overcome existing boundaries like interference due to pilot contamination in massive Multiple Input Multiple Output (MIMO) [7], or mixed with novel techniques such as Non-Orthogonal Multiple Access (NOMA) [8], but using only simple methods of superimposition. In broadcasting scenarios, considering the significant length of OFDM symbols used in these applications, ST becomes an attractive choice in comparison with PSAM techniques [9]. However, ST causes interference in channel estimation process and a less accurate data detection is obtained [10].

Several ST proposals have been already proposed. Data-Dependent Superimposed Training (DDST) [11] [12], retires some periodical frequency components and adds a known sequence to estimate the channel. Nevertheless, this method cannot be applied to OFDM because those frequency components would have been filled with information symbols. This problem was solved by Gayosso *et. al.* in [13] with a Symbol-Blanking Superimposed Training (SBST) technique. The removed frequency-domain symbols will be spread over the next group of symbols, causing data interference and then performance degradation. Therefore, at subcarrier positions where data symbols are removed, pilot symbols are included. Inspired in [11], Data Nulling Superimposed Training (DNST) [14] by Dou *et. al.* precodes the data sequence and then cancels data at certain subcarriers where the pilot sequence will be inserted. The main drawback of [14] is that the cancelled data are not incorporated in the transmission what causes Symbol Error Rate (SER) degradation especially in high order constellations. Thus, an iterative method is used to overcome the distortion generated by the nulling process. In fact [13] and [14] are similar in the sense that, at the end, pilots are placed at subcarrier positions where data are not transmitted, so there is no superimposing in practice. Indeed, in both methods, there is a separation between pilots and data, as in PSAM, with the aim of achieving a channel estimation as accurate as possible. In principle this will lead to a better data detection process. However, in ST systems, a better channel estimation accuracy is obtained at the cost of a poorer data detection due to data interference.

In the literature, there are similar methods that also allow adding sequences for channel estimation by using spread codes [14] [15]. In [16], a turbo receiver is proposed for DNST, using a Bit Interleaved Coded Modulation (BICM) coder and an interference cancellation method, but with a prohibitive complexity. The Peak-to-Average power Ratio (PAR) in ST-OFDM is investigated in [17] and [18] for semiblind channel estimation in MIMO OFDM systems. Also, Optimal Channel Independent (OCI) sequences are introduced in [11]–[13] to address this problem.

In this paper, we propose a novel superimposed framework over a precoded data scheme, named Partial-Data Superimposed Training (PDST), where an interference control factor is incorporated, apart from the power allocation factor, to assign the adequate information level to be added to the training sequence sent at certain subcarriers in OFDM systems. A closed form expression for the data interference and for the

This work has been partially funded by the National Secretary of Science, Technology and Innovation (SENESCYT) in Ecuador and by the Spanish National Projects ELISA (TEC2014-59255-C3-3-R) and TERESA-ADA (TEC2017-90093-C3-2-R) (MINECO/AEI/FEDER, UE).

The authors are with the Department of Signal Theory and Communications of the University Carlos III of Madrid, Spain. Email: {jestrada, mjulia}@tsc.uc3m.es.

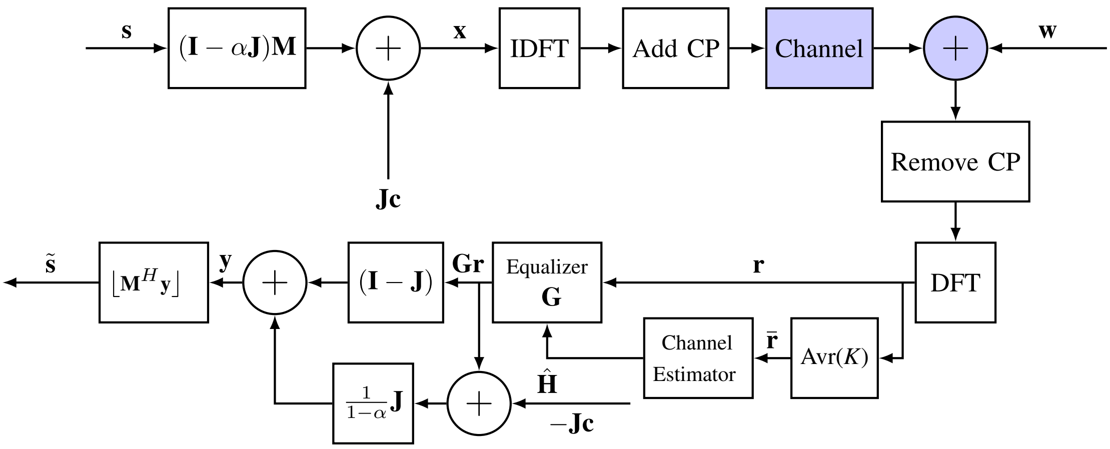

Fig. 1. System Architecture of the ST scheme proposed: Partial-Data Superimposed Training (PDST), where *Avr(K)* block denotes an averaging operation over K OFDM symbols.

channel estimation error is introduced, and based on this, an analysis of the system capacity is presented. Simulation results show that, with the appropriate value of the interference control factor, this scheme overcomes previous techniques. Additionally, a simple yet effective data detection method is proposed to improve SER performance.

This paper is organized as follows: Section II describes the system model and the proposed scheme. A Least Squares (LS) channel estimator and a data detection mechanism are provided. Section III describes the system capacity analysis of the proposal under the LS channel estimator previously derived. Section IV shows the simulation results. Finally, conclusions are discussed in section V.

II. System Model and Proposed Scheme

The architecture of the system, depicted in Fig. 1, is composed of several blocks, at the transmitter and receiver, detailed in the following sections. The OFDM modulation is carried out at the transmitter, then, the signal goes through the channel and noise is later added. Next, at the receiver the system recovers the information after performing a demodulation of the OFDM signal, channel estimation, equalization and data detection.

A. Signal Model

Let us consider an OFDM transmission in a multipath frequency selective channel. Let N be the number of subcarriers where P pilots will be inserted $(P < N)$ and $k = 0, ..., N-1$. The frequency-domain transmitted signal is represented by the $N \times 1$ vector $\mathbf{x}$, as a result of applying the proposed scheme to the data signal at the transmitter,

$$\mathbf{x} = (\mathbf{I} - \alpha\mathbf{J})\mathbf{M}\mathbf{s} + \mathbf{J}\mathbf{c} \tag{1}$$

where α is the interference control factor $0 < \alpha < 1$, $\mathbf{s}$ is the $N\times1$ data symbol vector in frequency-domain, $\mathbf{M}$ is the $N\times N$ precoding matrix, $\mathbf{J}$ is a diagonal matrix with ones at the P equispaced pilot positions, $\mathbf{I}$ is the $N\times N$ identity matrix and $\mathbf{c}$ is the $N\times1$ known training sequence vector with P equipowered and equispaced pilots that will be used for channel estimation process, while zeros are inserted in the remaining $(N - P)$ positions.

The precoding matrix $\mathbf{M}$ is formed by orthogonal codes, e.g. Hadamard, and, the matrix satisfies the condition $\mathbf{M}^H\mathbf{M} = \mathbf{I}$. Since it does not imply redundancy, each data symbol will be spread over N subcarriers, as in [14].

Note that $(\mathbf{I} - \alpha\mathbf{J})$ can be expressed as $(\mathbf{I} - \mathbf{J}) + (1 - \alpha)\mathbf{J}$. Then (1) is expanded to

$$\mathbf{x} = (\mathbf{I} - \mathbf{J})\mathbf{M}\mathbf{s} + (1 - \alpha)\mathbf{J}\mathbf{M}\mathbf{s} + \mathbf{J}\mathbf{c}. \tag{2}$$

As observed in the first term in (2), the $(\mathbf{I} - \mathbf{J})$ subcarrier positions, with cardinality $(N - P)$, will carry the *non-superimposed information signal*. These subcarriers will be denoted as *data* subcarriers.

In the second term in (2), the $\mathbf{J}$ subcarrier positions, with cardinality P, will carry the *superimposed information signal*, with $(1 - \alpha)$ factor premultiplying $\mathbf{M}\mathbf{s}$. That signal is added with the pilot signal $\mathbf{J}\mathbf{c}$, given by the third term. Thus, pilot and data are simultaneously transmitted at these subcarriers, denoted as *pilot-data* subcarriers.

The total signal power $\mathcal{P}$ is composed of two terms: the power of the information signal $\mathcal{P}_{\mathbf{s}}$ and the power of the training sequence $\mathcal{P}_{\mathbf{c}}$. It is given by

$$\mathcal{P} = \mathcal{P}_{\mathbf{s}} + \mathcal{P}_{\mathbf{c}}, \tag{3}$$

$$\mathcal{P}_{\mathbf{s}} = (1 - \beta)\mathcal{P}, \tag{4}$$

$$\mathcal{P}_{\mathbf{c}} = \beta\mathcal{P} \tag{5}$$

where β is the power allocation factor, i.e. the ratio between the power assigned to the training signal and the total power $\beta = \frac{\mathcal{P}_c}{\mathcal{P}}$. The information signal power is divided into

$$\mathcal{P}_{\mathbf{s}} = \mathcal{P}_{s1} + \mathcal{P}_{s2} \tag{6}$$

where $\mathcal{P}_{s1}$ and $\mathcal{P}_{s2}$ are the power of the *non-superimposed information signal* and the power of the *superimposed information signal*, respectively. For easiness, we represent $\tilde{\mathbf{x}} = (\mathbf{I} - \alpha\mathbf{J})\mathbf{M}\mathbf{s}$ and assume that $\mathcal{P}_{s1}$ is equally distributed over $N - P$ subcarriers,

$$E[|\tilde{\mathbf{x}}|^2]_k = \frac{\mathcal{P}_{s1}}{N - P}, k \in \mathcal{I}_0, \tag{7}$$

and, $\mathcal{P}_{s2}$ is equally distributed over P subcarriers,

$$E[|\tilde{\mathbf{x}}|^2]_k = \frac{\mathcal{P}_{s2}}{P}, k \in \mathcal{I}_0^{\perp} \tag{8}$$

where $E[.]$ denotes the expected value, $\mathcal{I}_0$ is the *data* subcarriers index set, with cardinality $|\mathcal{I}_0| = N - P$, and, $\mathcal{I}_0^{\perp}$ is the complementary set, with *pilot-data* subcarriers and cardinality $|\mathcal{I}_0^{\perp}| = P$. Since data symbols at pilot positions are scaled with $(1-\alpha)$ factor, the relationship between powers is:

$$E[|\tilde{\mathbf{x}}|^2]_k = (1-\alpha)^2 E[|\mathbf{Ms}|^2]_k, k \in \mathcal{I}_0^{\perp}, \tag{9}$$

while at *data* subcarriers

$$E[|\tilde{\mathbf{x}}|^2]_k = E[|\mathbf{Ms}|^2]_k, k \in \mathcal{I}_0. \tag{10}$$

Considering (8), we obtain

$$\frac{\mathcal{P}_{s2}}{P} = (1-\alpha)^2 E[|\mathbf{Ms}|^2]_k, k \in \mathcal{I}_0^{\perp}. \tag{11}$$

Also rewriting (7) it is obtained

$$\mathcal{P}_{s1} = (N-P) E[|\mathbf{Ms}|^2]_k, k \in \mathcal{I}_0. \tag{12}$$

Finally, substituting (11) and (12) into (6),

$$\mathcal{P}_{\mathbf{s}} = [(N - P + P(1-\alpha)^2)] E[|\mathbf{Ms}|^2]_k. \tag{13}$$

More useful expressions are obtained for the equally distributed power assumptions depending on $\mathcal{P}_s$, N, P, and α. For ST subcarriers, using (9) and (13)

$$E[|\tilde{\mathbf{x}}|^2]_k = \frac{\mathcal{P}_s (1-\alpha)^2}{N - P + P(1-\alpha)^2}, k \in \mathcal{I}_0^{\perp}, \tag{14}$$

and then, for non-ST subcarriers, using (10) and (13)

$$E[|\tilde{\mathbf{x}}|^2]_k = \frac{\mathcal{P}_s}{N - P + P(1-\alpha)^2}, k \in \mathcal{I}_0. \tag{15}$$

The resulting $N\times 1$ signal $\mathbf{x}$ is modulated via Inverse Discrete Fourier Transform (IDFT) yielding $\mathbf{x}_t = \frac{1}{\sqrt{N}} \mathbf{F}^H \mathbf{x}$, where $\mathbf{x}_t$ is the $N\times 1$ time-domain transmitted signal vector and $\mathbf{F}$ corresponds to the $N \times N$ normalized DFT matrix where each (p,q)th element has the form $[\mathbf{F}]_{p,q} = \mathrm{e}^{\frac{-j2\pi pq}{N}}$, and $\mathbf{F}^H$ is its Hermitian transpose. A cyclic prefix (CP) is inserted to avoid the inter-block interference (IBI). Let the column-vector $\mathbf{h}$ be the discrete time channel impulse response with L coefficients, $\{h_\ell\}_{\ell=0}^{L-1}$, that are assumed to be Rayleigh distributed with zero mean and variance ${\sigma_{h_\ell}}^2$. To minimize the channel estimation error P should satisfy the $P \geq L$ condition. The channel is quasi-stationary, and, varies every K OFDM symbols. Assuming a perfect synchronization at the receiver, the CP is eliminated. The received signal can be expressed as

$$\mathbf{r} = \mathbf{Hx} + \mathbf{w} \tag{16}$$

$$\mathbf{r} = \mathbf{H}(\mathbf{I} - \alpha\mathbf{J})\mathbf{Ms} + \mathbf{HJc} + \mathbf{w} \tag{17}$$

where $\mathbf{H}$ is an $N\times N$ diagonal matrix, whose diagonal is the $N\times 1$ channel's frequency response $\mathbf{h}_f = \mathbf{F}_L\mathbf{h}$ ($\mathbf{F}_L$ is the front $N\times L$ submatrix of $\mathbf{F}$), and $\mathbf{w}$ is $N\times 1$ Additive White Gaussian Noise (AWGN) vector with zero mean and $\sigma_{\mathbf{w}}^2\mathbf{I}$ covariance matrix. Finally, each element of $\mathbf{h_f}$ is a random variable with zero mean and $\sigma_H^2 = \sum_{\ell=0}^{L-1} \sigma_{h_\ell}^2$ variance. The received signal to noise ratio (SNR) is defined as

$$\rho_{snr} = \frac{\mathcal{P}}{N} \frac{\sigma_H^2}{\sigma_W^2}. \tag{18}$$

TABLE I
CHANNEL ESTIMATION ERROR FOR ST AND PSAM IN OFDM

CE OFDM technique	CE	$\sigma^2_{\Delta_h}$
Huang et. al. [19]	LS	$\frac{L}{K\beta\mathcal{P}}\left(\frac{\mathcal{P}_S\sigma_H^2}{N}+\sigma_W^2\right)$
Dou et. al. [14]	LS	$\frac{L\sigma_W^2}{K\beta\mathcal{P}}$
Gayosso et. al. [13]	LS	$\frac{L\sigma_W^2}{K\beta\mathcal{P}}$
PDST	LS	$\frac{L}{K\beta\mathcal{P}}\frac{1}{K}\left[\frac{\mathcal{P}_{\mathbf{s}}(1-\alpha)^2\sigma_H^2}{N-P+P(1-\alpha)^2}+\sigma_{\mathbf{w}}^2\right]$
PSAM	LS	$\frac{L\sigma_W^2}{K\beta\mathcal{P}}$

B. Channel Estimation

The LS time-domain channel estimator can be derived as [19]

$$\begin{aligned} \hat{\mathbf{h}}_{LS} &= (\tilde{\mathbf{F}}_L^H \tilde{\mathbf{F}}_L)^{-1} \tilde{\mathbf{F}}_L^H \mathbf{\Lambda} \mathbf{r}_{\mathcal{I}_0^{\perp}} \\ &= \frac{1}{P} \tilde{\mathbf{F}}_L^H \mathbf{\Lambda} \mathbf{r}_{\mathcal{I}_0^{\perp}} \end{aligned} \tag{19}$$

where $\tilde{\mathbf{F}}_L$ is the $P\times L$ submatrix of $\mathbf{F}_L$ (P-pilot rows of $\mathbf{F}_L$), $\mathbf{\Lambda}$ is a $P\times P$ diagonal matrix diag$\{\mathbf{\Lambda}\} = [\mathbf{c}_{\mathcal{I}_0^{\perp}}]^{-1}$, where $\mathbf{c}_{\mathcal{I}_0^{\perp}}$ is a $P\times 1$ vector with the nonzero elements of the training sequence $\mathbf{c}$, and $\mathbf{r}_{\mathcal{I}_0^{\perp}}$ corresponds to a $P\times 1$ vector formed by the *pilot-data* subcarriers of $\mathbf{r}$. According to the quasi-stationary condition, the channel estimate is improved with an average over K OFDM symbols, and then,

$$\hat{\mathbf{h}}_{LS} = \frac{1}{P} \tilde{\mathbf{F}}_L^H \mathbf{\Lambda} \frac{1}{K} \sum_{t=0}^{K-1} \mathbf{r}^t_{\mathcal{I}_0^{\perp}} \tag{20}$$

where $\mathbf{r}^t_{\mathcal{I}_0^{\perp}}$ is the received signal at tth OFDM symbol, and replacing $\bar{\mathbf{r}} = \sum_{t=0}^{K-1} \mathbf{r}^t_{\mathcal{I}_0^{\perp}}$ in (20)

$$\hat{\mathbf{h}}_{LS} = \frac{1}{P}\frac{1}{K} \tilde{\mathbf{F}}_L^H \mathbf{\Lambda} \bar{\mathbf{r}}. \tag{21}$$

It must be taken into account that $\bar{\mathbf{r}} = \mathbf{H}_{\mathcal{I}_0^{\perp}}(1-\alpha)\overline{\mathbf{Ms}} + \mathbf{H}_{\mathcal{I}_0^{\perp}}\bar{\mathbf{c}} + \bar{\mathbf{w}}$, from (17), where $\mathbf{H}_{\mathcal{I}_0^{\perp}}$ is a $P \times P$ diagonal matrix formed as a submatrix of $\mathbf{H}$ with the P elements of $\mathcal{I}_0^{\perp}$ index set; $\overline{\mathbf{Ms}}$, $\bar{\mathbf{c}}$ and $\bar{\mathbf{w}}$ are $P\times 1$ vectors defined similarly to $\bar{\mathbf{r}}$. Then,

$$\hat{\mathbf{h}}_{LS} = \mathbf{h} + \frac{1}{P}\frac{1}{K} \tilde{\mathbf{F}}_L^H \mathbf{\Lambda} \left[(1-\alpha)\mathbf{H}_{\mathcal{I}_0^{\perp}}\overline{\mathbf{Ms}} + \bar{\mathbf{w}}\right] \tag{22}$$

where a new factor $(1-\alpha)$ and a precoding scheme have been included in this proposal compared to Huang *et. al.* [19], which helps to reduce the data interference. It is important to realize that the channel estimation accuracy depends on α and K. Having that $\mathbf{v} = \frac{1}{K}[(1-\alpha)\mathbf{H}_{\mathcal{I}_0^{\perp}}\overline{\mathbf{Ms}} + \bar{\mathbf{w}}]$ is the frequency-domain interference and noise vector with zero mean, the $R_{\mathbf{v}} = E[\mathbf{v}\mathbf{v}^H]$ covariance matrix has the following form:

$$R_{\mathbf{v}} = \frac{1}{K}\left[\frac{\mathcal{P}_{\mathbf{s}}(1-\alpha)^2\sigma_H^2}{N-P+P(1-\alpha)^2} + \sigma_{\mathbf{w}}^2\right]\mathbf{I}_P \tag{23}$$

where $\mathbf{I}_P$ is a $P \times P$ identity matrix. Taking into account that the interference has zero mean as well as the AWGN noise we can say that the channel estimator is an unbiased estimator [20]. Finally, the variance of the LS channel estimator can be expressed as $\sigma_{\hat{h}_{LS}} = L\frac{\sigma_{\mathbf{v}}^2}{\beta\mathcal{P}}$. Finally, a comparison of the channel estimation error for every method can be observed in Table I.

TABLE II
COMPLEXITY ORDER OF THE PDST RECEIVER

Operation	Expression	Times	CM+
Op. 1	$\mathbf{G}$	1	N
Op. 2	$\mathbf{M}^H(\mathbf{I}-\mathbf{J})\mathbf{y}$, where $\mathbf{y}=(\mathbf{I}-\mathbf{J})\mathbf{Gr}+\frac{1}{1-\alpha}\mathbf{J}(\mathbf{Gr}-\mathbf{Jc})$	1	N^2
Op. 3	$\mathbf{M}^H\mathbf{JM}\hat{\mathbf{s}}^{i-1}$	#it	0

+ A similar complexity analysis for DNST was done in [16]. Note that CM refers to complex multiplications.

C. Data Detection

An MMSE and LS equalizer are used for the *data* and *pilot-data* subcarriers, respectively, similarly as in [14]. Then, the received signal in (17) is equalized with a diagonal matrix diag$\{\mathbf{G}\}=[g_0,...,g_{N-1}]$ whose diagonal components are filled with a ZF scheme $g_k=\frac{1}{([\hat{\mathbf{h}}_f]_k)}, k\in\mathcal{I}_0^{\perp}$ for *pilot-data* subcarriers, and an MMSE scheme $g_k=\frac{([\hat{\mathbf{h}}_f]_k)}{([\hat{\mathbf{h}}_f]_k)^2+\sigma_{\mathbf{w}}^2}, k\in\mathcal{I}_0$ for *data* subcarriers, where $\hat{\mathbf{h}}_f$ is the estimated channel in frequency-domain $\hat{\mathbf{h}}_f=\mathbf{F}_L\hat{\mathbf{h}}$. The equalization process yields

$$\mathbf{Gr}=\mathbf{GH}(\mathbf{I}-\alpha\mathbf{J})\mathbf{Ms}+\mathbf{GHJc}+\mathbf{Gw}. \quad (24)$$

For easiness of theoretical analysis of data detection mechanism, $\mathbf{GH}\cong\mathbf{I}$ is assumed, and expanding $(\mathbf{I}-\alpha\mathbf{J})\mathbf{Ms}$ to $(\mathbf{I}-\mathbf{J})\mathbf{Ms}+(1-\alpha)\mathbf{JMs}$, we obtain

$$\mathbf{Gr}=(\mathbf{I}-\mathbf{J})\mathbf{Ms}+(1-\alpha)\mathbf{JMs}+\mathbf{Jc}+\mathbf{w} \quad (25)$$

and reordering the terms,

$$\mathbf{Gr}-\mathbf{Jc}=(\mathbf{I}-\mathbf{J})\mathbf{Ms}+(1-\alpha)\mathbf{JMs}+\mathbf{w}. \quad (26)$$

As we can see in (26), the detector receives $(1-\alpha)$ part of the data symbols that superimpose the P pilots. To recover the superimposed data signal, a $\frac{\mathbf{J}}{1-\alpha}$ factor is inserted to premultiply (26), and taking into account that $\mathbf{J}^2=\mathbf{J}$ and $\mathbf{J}(\mathbf{I}-\mathbf{J})=\mathbf{0}_{N\times N}$

$$\frac{1}{1-\alpha}\mathbf{J}(\mathbf{Gr}-\mathbf{Jc})=\mathbf{JMs}+\frac{\mathbf{J}}{1-\alpha}\mathbf{w}. \quad (27)$$

Thus with (27), the *superimposed information signal* can be recovered. Next, to recover the *non-superimposed one* we premultiply (25) by $(\mathbf{I}-\mathbf{J})$

$$\begin{aligned}(\mathbf{I}-\mathbf{J})\mathbf{Gr}=&(\mathbf{I}-\mathbf{J})(\mathbf{I}-\mathbf{J})\mathbf{Ms}+(\mathbf{I}-\mathbf{J})(1-\alpha)\mathbf{JMs}\\&+(\mathbf{I}-\mathbf{J})\mathbf{Jc}+(\mathbf{I}-\mathbf{J})\mathbf{w}.\end{aligned} \quad (28)$$

Then, considering $(\mathbf{I}-\mathbf{J})^2=(\mathbf{I}-\mathbf{J})$,

$$(\mathbf{I}-\mathbf{J})\mathbf{Gr}=(\mathbf{I}-\mathbf{J})\mathbf{Ms}+(\mathbf{I}-\mathbf{J})\mathbf{w} \quad (29)$$

where the *non-superimposed information signal* can be recovered. Adding (27) and (29), we obtain,

$$(\mathbf{I}-\mathbf{J})\mathbf{Gr}+\frac{1}{1-\alpha}\mathbf{J}(\mathbf{Gr}-\mathbf{Jc})=\mathbf{Ms}+(\mathbf{I}-\mathbf{J})\mathbf{w}+\mathbf{J}\frac{\mathbf{w}}{1-\alpha} \quad (30)$$

yielding, at the left-hand part, our recovering data model, that will be denoted with $\mathbf{y}$, and, at the right-hand part, the desired signal with the noise that affects the model

$$\mathbf{y}=\mathbf{Ms}+(\mathbf{I}-\mathbf{J})\mathbf{w}+\mathbf{J}\frac{\mathbf{w}}{1-\alpha}. \quad (31)$$

TABLE III
DATA DETECTION ALGORITHMS FOR ST OFDM TECHNIQUES

ST OFDM tech.	Data detection	Rec. complexity
Dou et. al. [14]	see (10) (11) from [14]	$N{+}N(N-P)$
Gayosso et. al. [13]	see (19)-(22) from [13]	N^2
PDST	$\hat{\mathbf{s}}^i{=}\lfloor\mathbf{M}^H(\mathbf{I}{-}\mathbf{J})\mathbf{y}{+}\mathbf{M}^H\mathbf{JM}\hat{\mathbf{s}}^{i-1}\rfloor$	$N{+}N^2$

Taking into account that $\mathbf{M}^H\mathbf{M}=\mathbf{I}$ and since the information signal was precoded, the unspread recovered data signal can be obtained with $\tilde{\mathbf{s}}=\mathbf{M}^H\mathbf{y}$. Then,

$$\tilde{\mathbf{s}}=\mathbf{s}+\mathbf{M}^H(\mathbf{I}-\mathbf{J})\mathbf{w}+\mathbf{M}^H\mathbf{J}\frac{\mathbf{w}}{1-\alpha}. \quad (32)$$

The data detection process consists in a hard detector aided by an iterative process, which initiates with hard decision $\hat{\mathbf{s}}^0=\lfloor\tilde{\mathbf{s}}\rfloor$, where $\lfloor.\rfloor$ corresponds to the minimum distance detector. Then, this proposal defines the next ith iterations as

$$\hat{\mathbf{s}}^i=\lfloor\mathbf{M}^H(\mathbf{I}-\mathbf{J})\mathbf{y}+\mathbf{M}^H\mathbf{JM}\hat{\mathbf{s}}^{i-1}\rfloor \quad (33)$$

where the $(\mathbf{I}-\mathbf{J})$ positions of $\mathbf{y}$ are unspread and added to the unspread $\mathbf{J}$ positions of the previous iteration $\hat{\mathbf{s}}^{i-1}$.

III. ANALYSIS OF THE SYSTEM CAPACITY

We have from [15] [19] [21] that the averaged channel capacity is given by

$$C=\frac{1}{N}\sum_{k=0}^{N-1}E\{\log_2(1+|g|^2\rho_k)\} \quad (34)$$

where ρ_k is the SNR of the kth subcarrier, and the expectation is taken with respect to a zero mean and unitary variance random variable $g\sim(0,1)$. We have that the capacity depends on the interference control and power allocation factor, α and β, respectively, where

$$C(\alpha,\beta)=\frac{1}{N}E\left\{\sum_{k\in\mathcal{I}_0}log_2(1+|g|^2\rho_k)+\sum_{k\in\mathcal{I}_0^{\perp}}log_2(1+|g|^2\rho_k)\right\}. \quad (35)$$

Applying the approximation used in [(22)- [19]] we can obtain

$$\begin{aligned}C(\alpha,\beta)=&\frac{N-P}{N}E\{log_2(1+|g|^2\rho_{NST})\}+\\&\frac{P}{N}E\{log_2(1+|g|^2\rho_{ST})\}\end{aligned} \quad (36)$$

where ρ_k, $k\in\mathcal{I}_0$, is denoted as ρ_{NST}. Similarly, ρ_k, $k\in\mathcal{I}_0^{\perp}$, will be denoted as ρ_{ST}.

For *pilot-data* subcarriers, $k\in\mathcal{I}_0^{\perp}$, where ST is applied, using [(22) [19]] ρ_{ST} is given by

$$\rho_{ST}=\frac{\sigma_H^2-\sigma_{\Delta_h}^2}{\sigma_{\Delta_h}^2\frac{E[|\tilde{\mathbf{x}}|^2]+\frac{\beta\mathcal{P}}{P}}{E[|\tilde{\mathbf{x}}|^2]}+\frac{\sigma_W^2}{E[|\tilde{\mathbf{x}}|^2]}} \quad (37)$$

where $E[|\tilde{\mathbf{x}}|^2]$ and $\sigma_{\Delta_h}^2$ are defined as in (14) and (23), respectively. Using (14), (18) and (23) the relation between the channel estimation error and the channel variance can be written as

$$\frac{\sigma_{\Delta_h}^2}{\sigma_H^2}=\frac{L}{\beta K}\left[\frac{(1-\beta)(1-\alpha)^2}{N-P+P(1-\alpha)^2}+\frac{1}{N\rho_{snr}}\right] \quad (38)$$

and the relation between the noise variance and the channel variance as

$$\frac{\sigma_W^2}{\sigma_H^2}\frac{1}{E[|\tilde{\mathbf{x}}|^2]}=\frac{N-P+P(1-\alpha)^2}{\rho_{snr}N(1-\beta)(1-\alpha)^2}, \quad (39)$$

TABLE IV
SYSTEM PARAMETERS

Parameter	Value
Number of subcarriers (N)	64, 128, 256
Number of pilots (P)	8, 16, 32
Modulation	QPSK, 16QAM
Number of OFDM symbols (K)	10, 20
Channel distribution	Rayleigh
Number of taps (L)	8
Cyclic prefix (CP) length	8

then, considering (37), (38) and (39) ρ_{ST} can be written as

$$\rho_{ST} = \frac{(1-\alpha)^2(1-\beta)\left[1-\frac{L}{\beta K}\left(\frac{(1-\beta)(1-\alpha)^2}{N-P+P(1-\alpha)^2}+\frac{1}{N\rho_{snr}}\right)\right]}{\frac{L}{\beta K}\left[(1-\alpha)^2+\beta(\frac{N}{P}-1)\right]\left[\frac{(1-\beta)(1-\alpha)^2}{N-P+P(1-\alpha)^2}+\frac{1}{N\rho_{snr}}\right]+\frac{N-P+P(1-\alpha)^2}{\rho_{snr}N}}. \tag{40}$$

For *data* subcarriers, $k \in \mathcal{I}_0$, where non-ST is applied, using [(27)- [15]] and (15), ρ_{NST} can be viewed as

$$\rho_{NST} = \frac{\sigma_H^2 - \sigma_{\Delta_h}^2}{\sigma_{\Delta_h}^2 + \sigma_W^2 \frac{1}{E[|\tilde{\mathbf{x}}|^2]}}. \tag{41}$$

The relation $\sigma_W^2/\sigma_H^2 E[|\tilde{\mathbf{x}}|^2]$ for *data* subcarriers is

$$\frac{\sigma_W^2}{\sigma_H^2}\frac{1}{E[|\tilde{\mathbf{x}}|^2]} = \frac{N-P+P(1-\alpha)^2}{(1-\beta)N\rho_{snr}}. \tag{42}$$

Then, the channel estimation is obtained from the ST subcarriers and the relation $\sigma_{\Delta_h}^2/\sigma_W^2$ is taken from (38). Considering (38), (41) and (42) ρ_{NST} can be expressed as

$$\rho_{NST} = \frac{N\rho_{snr}(1-\beta)\left[1-\frac{L}{\beta K}\left(\frac{(1-\beta)(1-\alpha)^2}{N-P+P(1-\alpha)^2}+\frac{1}{N\rho_{snr}}\right)\right]}{\frac{L(1-\beta)}{\beta K}\frac{N\rho_{snr}(1-\beta)(1-\alpha)^2}{N-P+P(1-\alpha)^2}+1+N-P+P(1-\alpha)^2}. \tag{43}$$

The adequate parameters α and β can be found searching the maximum capacity of the proposed PDST scheme with (36), (40) and (43).

$$\begin{aligned} &\underset{\alpha,\beta}{\text{maximize}} \quad C(\alpha,\beta), \\ &\text{subject to} \quad 0<\alpha<1, \qquad 0<\beta<1. \end{aligned} \tag{44}$$

A. Complexity analysis

1) Transmitter: At the transmitter side, it is required to search the parameters that maximize the capacity. To find the values of α and β from (44), exhaustive search can be carried out. Let us assume that the size of the range of $0<\alpha<1$ and $0<\beta<1$ is n. Then, using nested loops, n^2 comparisons are required, and thus complexity order is $O(n^2)$. The use of exhaustive search in an off-line mode and a pre-computing table can reduce the complexity.

2) Receiver: This proposal differs from previous related works in data detection process. Its complexity lies in the number of complex multiplications (CM), which we calculate similarly as in [16]. The analysis is provided in Table II and Table III. In first operation Op. 1, the equalization matrix is constructed using N CM, one per subcarrier. In Op. 2, the recovering data model in (30), where $N-P$ CM are used to obtain the information of the *data* subcarriers and P CM to get the data from *pilot-data* subcarriers, requires N times the Op. 1 to construct $\mathbf{y}$. Finally, no complex multiplications are performed in Op. 3. Note that the use of selection matrices and products by 0 and 1 are not computationally demanding so they can be neglected. Thus, N^2+N CM are required at the receiver side.

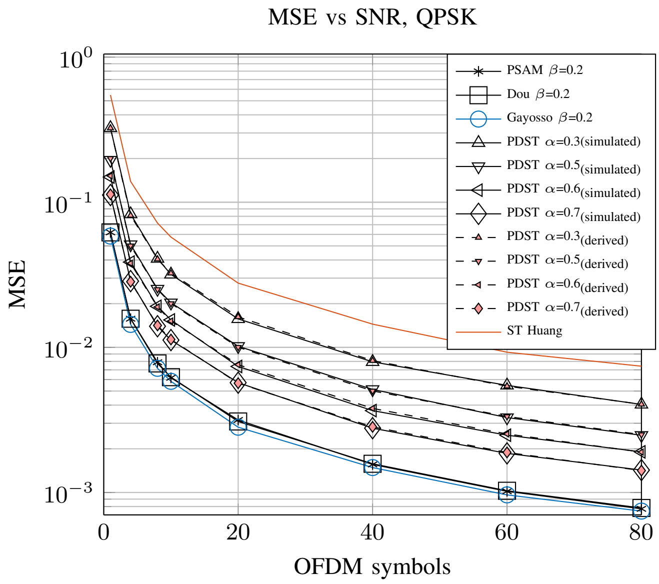

Fig. 2. MSE of channel estimation for various schemes.

IV. SIMULATION RESULTS

We consider an OFDM transmission system with $N = \{64, 128, 256\}$ subcarriers using Quadrature Phase-Shift Keying (QPSK) constellations and Quadrature Amplitude Modulation (16QAM). The channel is assumed to be Rayleigh distributed with L=8 and a power delay profile $E[|h_\ell|^2]=e^{-\frac{L}{10}\ell}$. The length of the CP is greater or equal to L. The known training sequence is chosen as in [11]–[13] to be optimal. The precoder is an $N\times N$ Hadamard matrix. The methods have been tested when the channel varies every K OFDM symbols. The system parameters are described in Table IV.

A. Channel Estimator performance

Fig. 2 shows the MSE performance of the channel estimation methods with a fixed SNR=10dB N=64, P=8 and β=0.2. The compared techniques in this figure consider equipowered and equispaced pilot subcarriers. For PDST, theoretical curves analytically derived in (23) perfectly match with simulation results what confirms the validity of our theoretical expression. These results are provided for different values of α. Then, it is observed how our proposal overcomes in MSE to classical ST approach proposed by Huang. *et. al.* [19]. This improvement attained by PDST is due to the interference control factor, α, introduced in this work. Also, PSAM technique is used as lower bound for MSE of channel estimation.

It must be noted that the proposals of Dou *et. al.* [14] and Gayosso *et. al.* [13] perform as PSAM in terms of channel estimation MSE. It can be observed that PDST performance

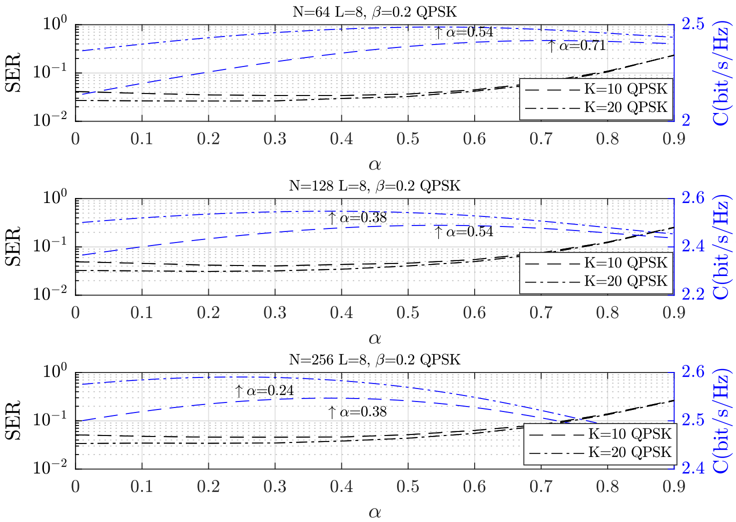

Fig. 3. Comparison of system capacity (right-axis) with SER performance (left-axis) as a function of α for QPSK and a fixed SNR=10dB.

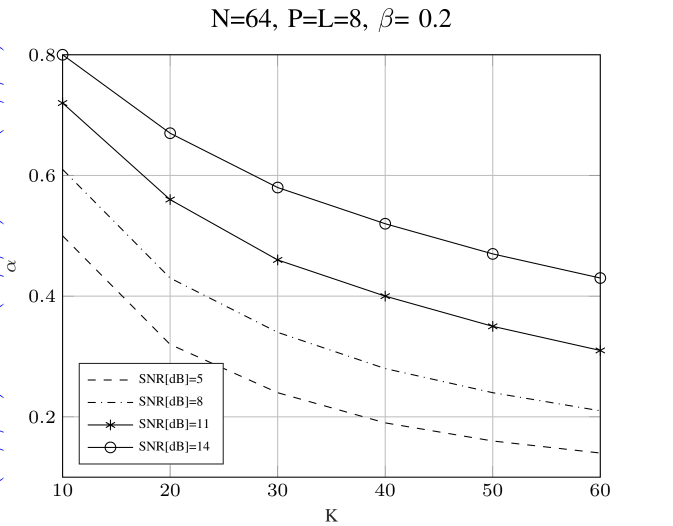

Fig. 5. Optimal interference control factor (α) varying the number of averaged OFDM symbols, K.

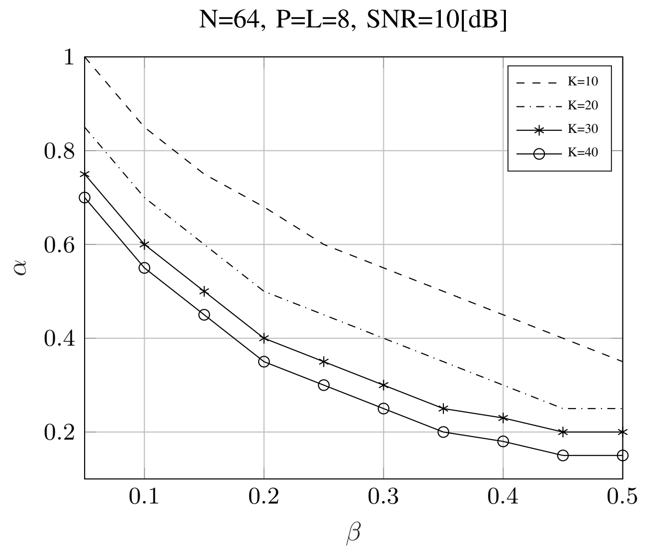

Fig. 4. Interference control factor, α, as a function of the power allocation factor, β, for different values of K.

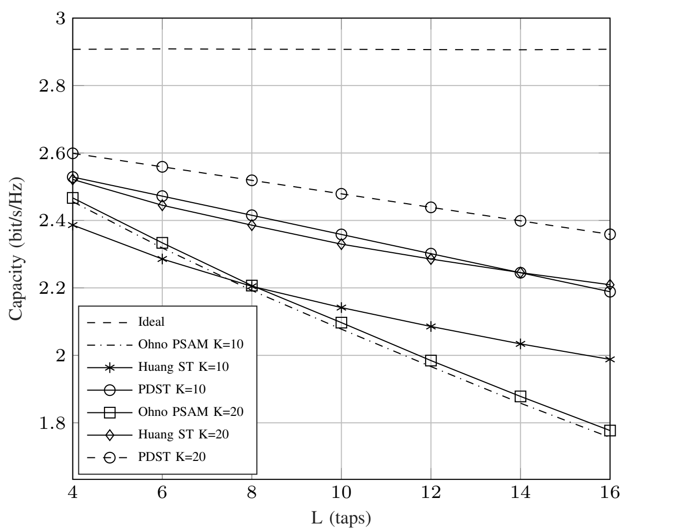

Fig. 6. Average channel capacity of the ST and PSAM schemes in channels with an i.i.d. power delay profile with $N = 64$ subcarriers (SNR=10dB, K=10 and K=20).

tends to PSAM curve as the parameter α is tuned. However, SER performance for [13] [14] is worse than the proposed PDST technique due to data interference. It must be considered that SER is the key parameter for communication systems rather than MSE of channel estimator.

B. Capacity performance

To evaluate the average channel capacity we have considered (36) and the references [(32), (34) [15]] and [(26), (29) [19]]. The channel capacity for the ideal case can be expressed as

$$C_{ideal} = \frac{N-P}{N} E\{\log_2(1+\rho_{ideal}|g|^2)\} \quad (45)$$

where this is ideal as P and the channel estimation error $\sigma^2_{\Delta_h}$ tend to zero. Then $\rho_{ideal} = \sigma^2_H \mathcal{P}_S / N\sigma^2_W$ where the total power is used for data transmission.

Optimal values of α for the capacity expression in (36) for a given β can be found as illustrated in Fig. 3. Once these values have been determined, they will be used later for SER analysis.

The three sub-figures within Fig. 3 present results for N=$\{64, 128, 256\}$, respectively. Each sub-figure contains curves for K=10 and K=20. The value of K, which is a necessary parameter in our proposal, is selected according to practical standard considerations, *e.g.* LTE system establishes slots with several OFDM symbols [22], and also taking into account that other related-works use even larger values of K [19]. The optimal value of α, that provides the maximum capacity accordingly to (44), is highlighted for every curve. It is important to realize that the optimal values in SER and in capacity metrics do not exactly match because SER

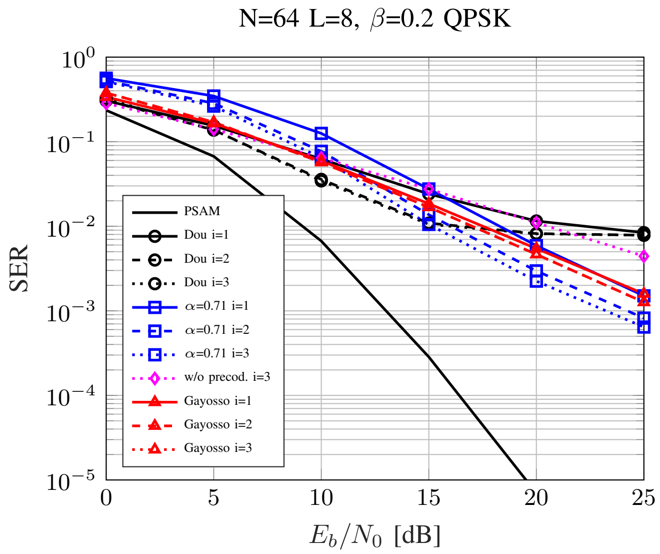

Fig. 7. Symbol Error Rate for K=10 OFDM symbols and QPSK modulation with $N = 64$ subcarriers and $P = 8$. The parameter α that maximizes the average channel capacity is $\alpha = 0.71$.

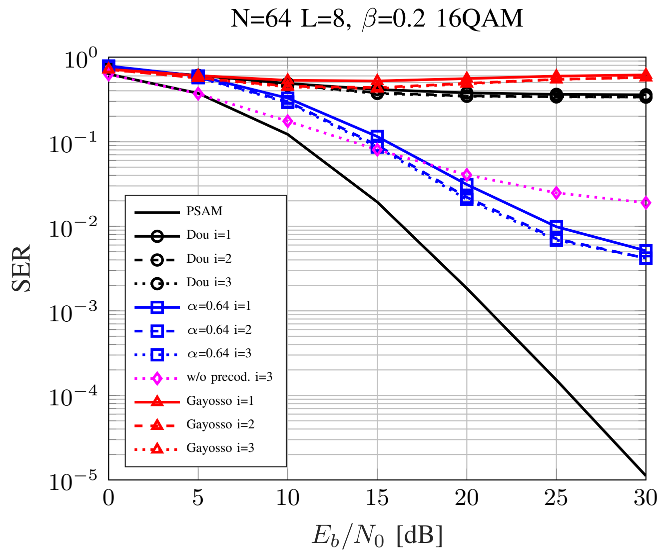

Fig. 9. Symbol Error Rate for K=20 OFDM symbols and 16QAM modulation with N=64 subcarriers and $P = 8$. The parameter α that maximizes the average channel capacity is $\alpha = 0.64$.

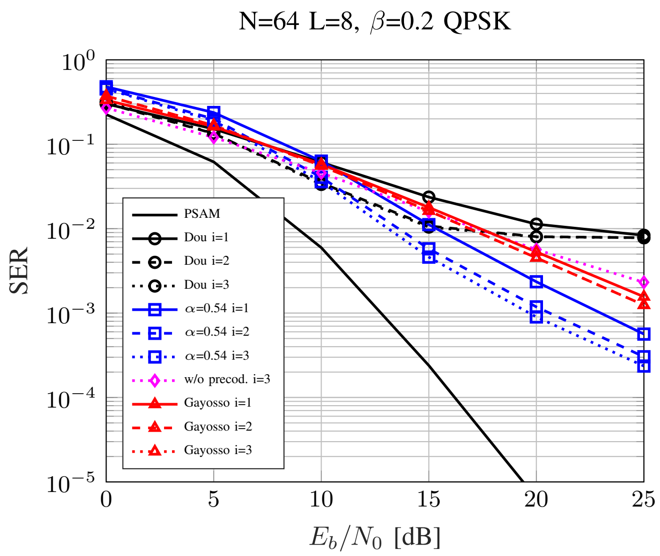

Fig. 8. Symbol Error Rate for K=20 OFDM symbols and QPSK modulation with $N = 64$ subcarriers and $P = 8$. The parameter α that maximizes the average channel capacity is $\alpha = 0.54$.

simulations additionally evaluate equalization and detection processes which improve its performance and modify the value of α. Also, it can be observed that as K increases, the value of α is reduced as expected.

In Fig. 4, we can see the optimal interference control factor α plotted with respect to β with a fixed SNR=10dB and different values of K parameter. This implies that the more power is assigned to the pilots, β, the better the channel estimation is and thus the more data information can be superimposed at *pilot-data* subcarriers, i. e. $(1{-}\alpha)$ increases.

In Fig. 5, it is observed how α decreases when K is incremented, under fixed SNR and β values. This means that when the channel is more stable, α diminishes or similarly $(1{-}\alpha)$ is incremented, i.e., more data information can be superimposed.

In Fig. 6, we can see the average channel capacity of PDST evaluated for various channel orders L under $K = 10$ and $K = 20$ scenarios. The ideal case (45), the capacity for PSAM-OFDM given by Ohno *et. al.* [(32), (34) [15]] and for ST given by Huang *et. al.* [(26), (29) [19]] are plotted for comparison purposes. Each point is evaluated using the optimal value of α for $P = L$ taps. We can see the performance of the proposal under scenarios where the channel changes every $K = 10$ and $K = 20$ consecutive OFDM symbols. PSAM falls faster than PDST because of wasting exclusive subcarriers with pilots. ST presents a lower performance than PDST because in ST [19] the data interference is still significant for these values of K. Note that data interference in ST, given by [(26), (29) [19]], is controlled by K, so the choice of quite high values for this parameter is a requirement so that the channel estimation error $\sigma^2_{\Delta_h}$ gets reduced. On the contrary, in PDST data interference is controlled by K and α, what allows a better capacity with lower values of K. Note that in practice, the average channel capacity depends on the CP length, since it influences the signal-to-noise ratio. To analyze its effect ρ_{snr} can be expressed as $\rho_{snr}{=}\rho_{snr\_\mathrm{cp}} \cdot N/(N{+}\mathrm{CP})$ where ρ_{snr_cp} refers to the signal-to-noise ratio including the CP length. Then to observe the evolution of the capacity as a function of the CP length, the above equation for ρ_{snr} should be replaced in (40) and (43).

C. SER performance

For the evaluation of SER performance, the parameter α was determined by using exhaustive search to maximize capacity in (35). For comparison purposes, each ST proposal has been simulated with its own iterative detection algorithm, employing 3 iterations, under the same conditions of stationarity, channel selectivity and noise.

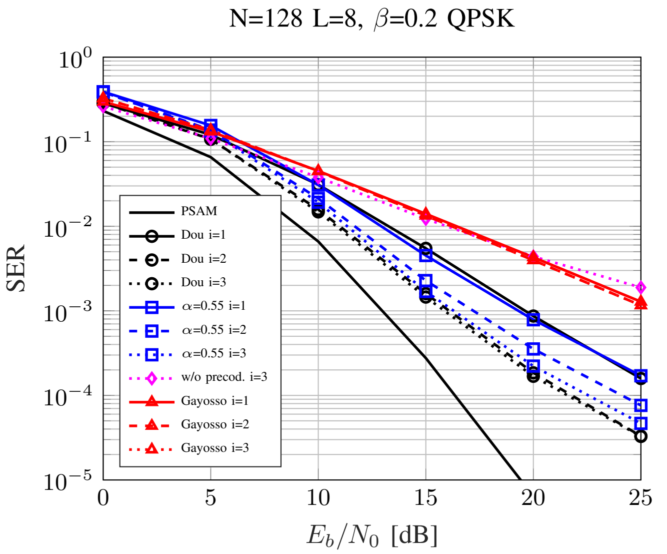

Fig. 10. Symbol Error Rate for K=20 OFDM symbols and QPSK modulation with N=128 subcarriers and $P = 8$. The parameter α that maximizes the average channel capacity is $\alpha = 0.55$.

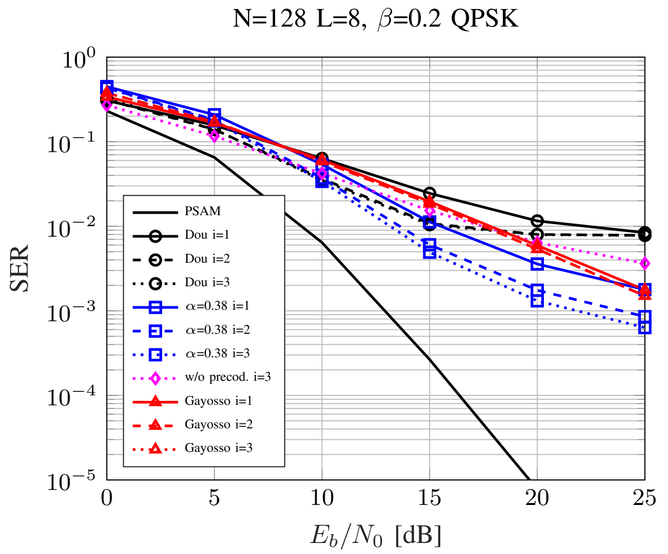

Fig. 12. Symbol Error Rate for K=20 OFDM symbols and QPSK modulation with N=128 subcarriers and $P = 16$. The parameter α that maximizes the average channel capacity is $\alpha = 0.38$.

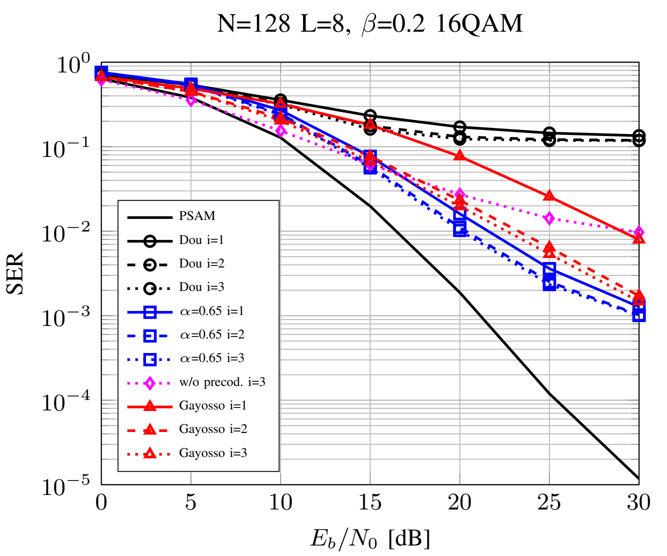

Fig. 11. Symbol Error Rate for K=20 OFDM symbols and 16QAM modulation with $N = 128$ subcarriers and $P = 8$. The parameter α that maximizes the average channel capacity is $\alpha = 0.65$.

In Figs. 7-15 results for QAM constellation, considering $N = \{64, 128, 256\}$ and $K = \{10, 20\}$ are provided. In these figures, SER performance of PDST is compared with previous proposals in literature: Gayosso *et. al.*, SBST [13], Dou *et. al.*, DNST [14], Tao *et. al.*, PSAM [3]. Fig. 7 evaluates the SER performance for K=10, N=64, P=8 and QPSK constellation scenario, while Fig. 8 evaluates the SER performance for the same parameters, excepting K=20. In both cases we can see that the proposal obtains a better performance than the other ST methods. It can be observed that if we increase K, the interference control factor selected will be reduced what indicates that an increased superimposed level is allowed.

In Fig. 9, the SER performance is evaluated for a 16QAM constellation for $K = 20$, $N = 64$ and $P = 8$. It can be seen that our method, compared to previous works, can face a dense constellation scenario as 16QAM while the other ST proposals remain useless.

In Figs. 10 and 11, the SER performance is depicted employing N=128 with P=8 using QPSK and 16QAM modulation, respectively. It can be observed that DNST obtains a similar performance to PDST using QPSK. However, our proposal is clearly better in a 16QAM scenario. Then, Fig. 12 shows the SER performance for N=128 with P=16 using QPSK.

Figs. 13 and 14 show the SER performance employing N=256 with P=8 using QPSK and 16QAM modulation, respectively. The robustness of our proposal using 16QAM is confirmed in this scenario.

In Fig. 15 we can see the SER performance for N=256 and P=32 using QPSK. Thus, comparing results previously depicted in Figs. 12 and 15, it can be seen that PDST behaves better than [13] and [14] with a growing number of pilots.

The PSAM method is included as a graphical bound taking into account that uses extra resources to transmit the training sequence; N subcarriers for data transmission and P additional subcarriers for explicit training sequence, respectively.

For comparison purposes, the case of PDST without precoding is included. It can be seen how precoding improves the SER performance specially in high SNR values.

Regarding system synchronization, there are numerous proposals in the literature for OFDM, based on specific headers or training sequences, that attain an adequate synchronization with negligible residual errors, independently if the system employs PSAM or ST-based channel estimation [23]–[26]. It seems that ST schemes may offer tracking capabilities, with timing and frequency synchronization as mentioned in [19].

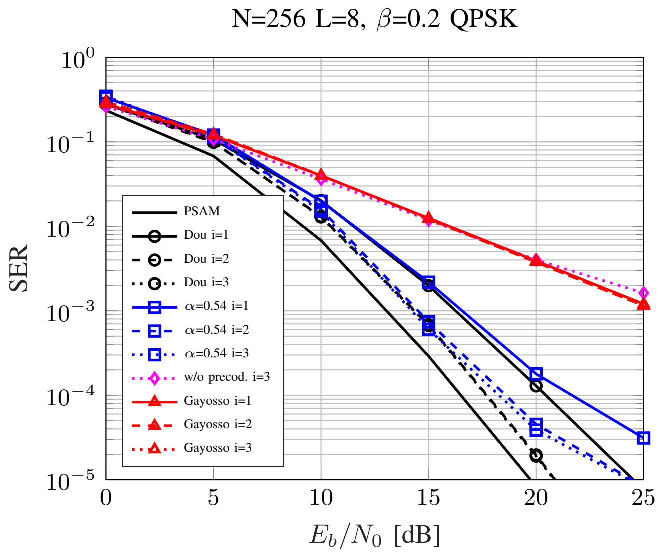

Fig. 13. Symbol Error Rate for K=20 OFDM symbols and QPSK modulation with $N = 256$ subcarriers and $P = 8$. The parameter α that maximizes the average channel capacity is $\alpha = 0.54$.

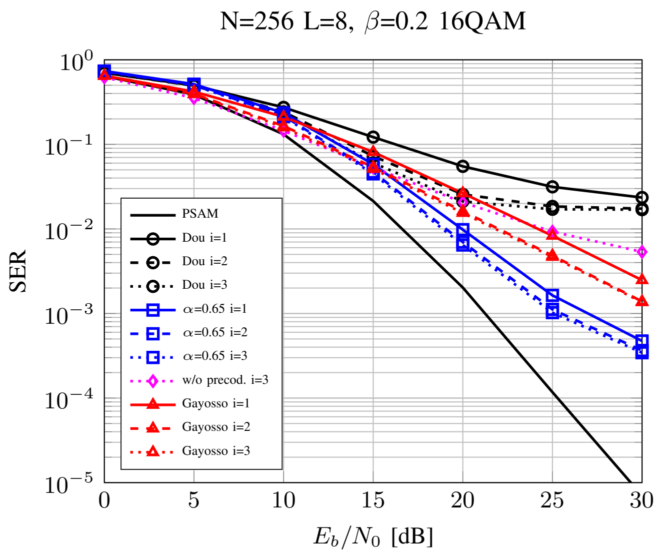

Fig. 14. Symbol Error Rate for K=20 OFDM symbols and 16QAM modulation with $N = 256$ subcarriers and $P = 8$. The parameter α that maximizes the average channel capacity is $\alpha = 0.65$.

D. Throughput Analysis

In order to clarify which technique, PSAM or PDST, has a better performance based on SER results, a throughput comparison is required. The throughput is usually defined as the number of correct bits per unit of time. Then, according to [27], it can be expressed as

$$R_t = \frac{N-P}{N}\frac{N_t\,(1-\mathrm{BER})}{T}. \tag{46}$$

If Gray encoding is assumed, then $\mathrm{BER} \approx \frac{\mathrm{SER}}{N_t}$, where N_t is the number of transmitted bits per symbol and T is the transmission time that has been assumed as $T = 1s$.

For illustration purposes, the throughput per subcarrier for $N = 64$ scenarios is depicted in Fig. 16. In the first and

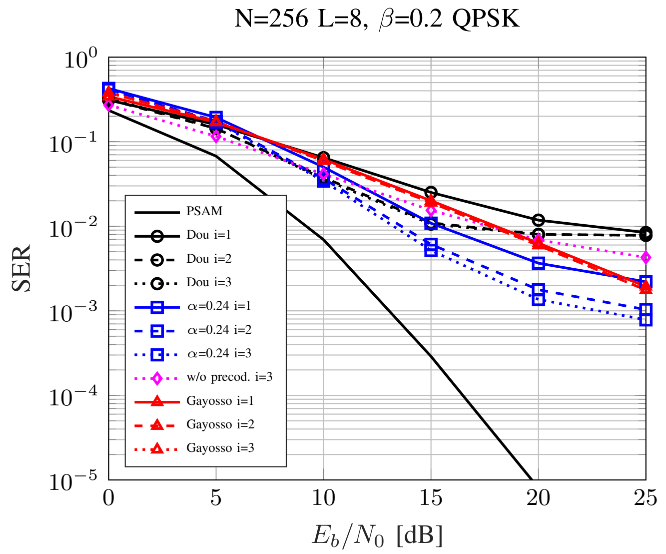

Fig. 15. Symbol Error Rate for K=20 OFDM symbols and QPSK modulation with $N = 256$ subcarriers and $P = 32$. The parameter α that maximizes the average channel capacity is $\alpha = 0.24$.

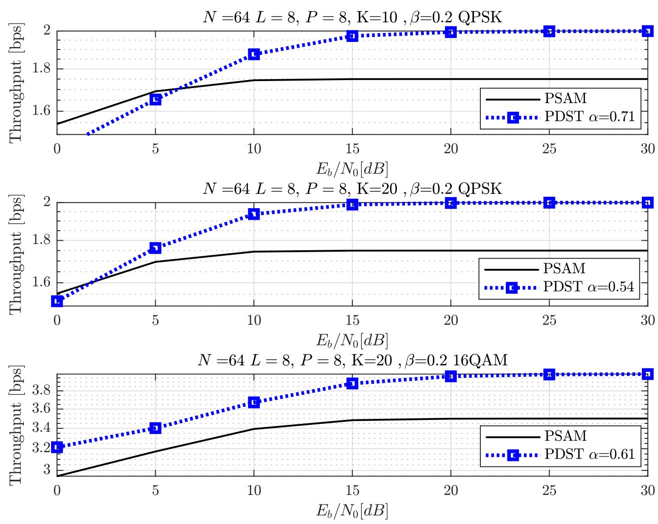

Fig. 16. Comparison of throughput for PSAM and PDST ($i = 3$).

second graphics, using QPSK, the effect of K variation can be observed. If K increases, a higher throughput is obtained.

Finally, the throughput for a 16QAM scenario is depicted in the third graphic. Note that for ST $\frac{N-P}{N}$ is assumed to be 1. The performance in the remainder scenarios follows the same behavior as the results illustrated in Fig. 16.

E. PAR performance

To illustrate the PAR characteristics of PDST, the complementary cumulative distribution function (CCDF) is used. This can be defined as

$$\mathrm{CCDF}(\gamma) = \Pr(\mathrm{PAR} > \gamma) \tag{47}$$

where PAR is defined as in [17]. In Figs. 17 and 18, the PAR is compared to ST and PSAM schemes. It is observed that

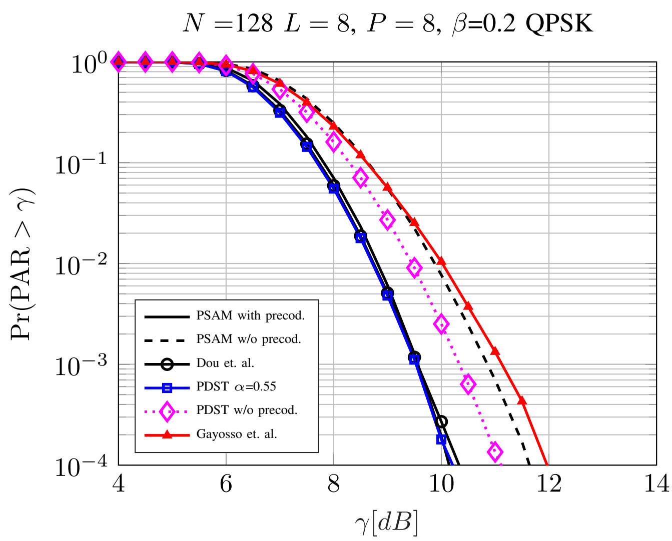

Fig. 17. CCDF comparison of PAR performance of the ST analyzed techniques for $N = 128$ and $P = 8$.

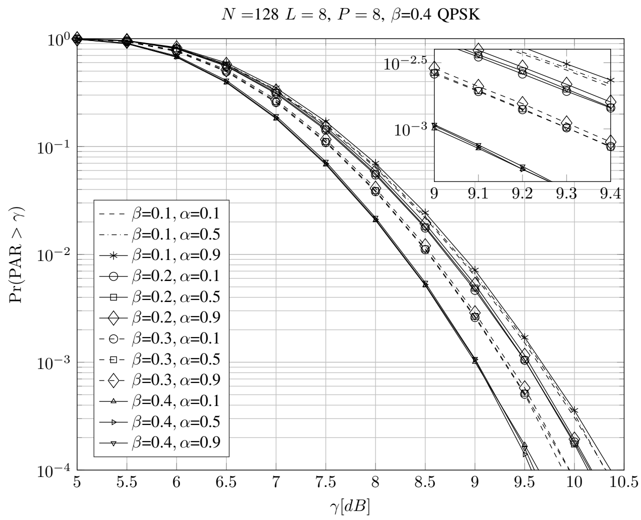

Fig. 19. CCDF comparison of PAR performance varying α and β for $N = 128$ and $P = 8$.

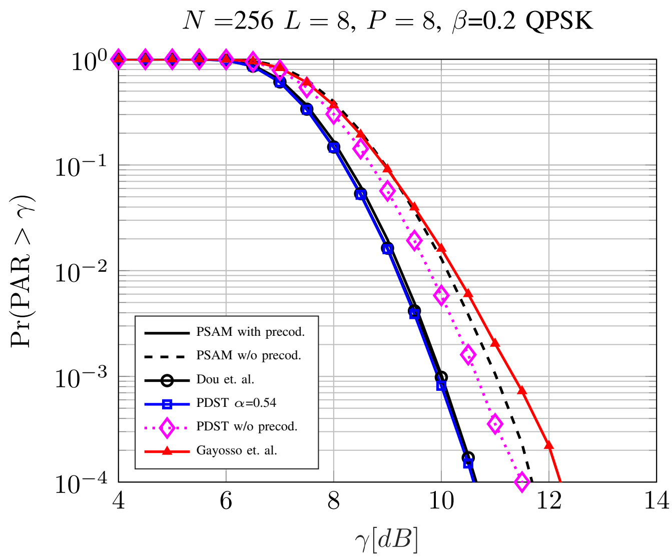

Fig. 18. CCDF comparison of PAR performance of the ST analyzed techniques for $N = 256$ and $P = 8$.

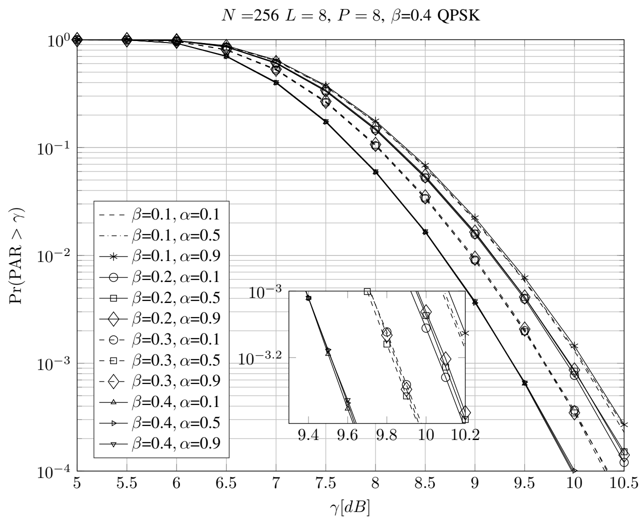

Fig. 20. CCDF comparison of PAR performance varying α and β for N=256 and P=8.

PDST and DNST behave similarly and better than [13] and PDST w/o precoding.

PAR depends on the pilot sequence; then, in this work, the employed pilot signal is composed of optimal equipowered and equispaced tones as in [11].

In Figs. 19 and 20, it can be seen that the CCDF function lowers with the increase of β. In the same figures, a slight PAR deterioration with the change of α can be seen when β is low. Then, when increasing the pilot signal power, *i.e.* β, this degradation is almost negligible. Similar results for the variation of β were reported in [17].

V. Conclusions

We present a novel ST method for OFDM systems named Partial-Data Superimposed Training (PDST). In order to face the degradation produced by the interference of the information signal on channel estimation, we propose a *superimposed information signal* scheme at certain subcarriers that overcomes the previous methods proposed in the literature. We show that a certain level of interference on channel estimation can be tolerated in order to preserve a level of *superimposed information signal* that, after being recovered, can reduce the error probability and increase the average capacity. A data detection method was also proposed to improve the Symbol Error Rate (SER) performance. The validity of our proposal is confirmed with a capacity analysis of the system that allows to determine the adequate level of superimposition.

References

[1] Y. Li and G. L. Stüber, *Orthogonal Frequency Division Multiplexing for Wireless Communications*. Springer, Boston, MA, 2006.

[2] W.-G. Song and J.-T. Lim, "Pilot-symbol aided channel estimation for OFDM with fast fading channels," *IEEE Trans. Broadcast.*, vol. 49, no. 4, pp. 398–402, Dec 2003.

[3] T. Cui and C. Tellambura, "Pilot symbols for channel estimation in OFDM systems," in *GLOBECOM '05. IEEE Global Telecommunications Conference*, vol. 4, Dec 2005, pp. 2229–2233.

[4] Z. Gao, C. Zhang, and Z. Wang, "Robust Preamble Design for Synchronization, Signaling Transmission, and Channel Estimation," *IEEE Trans. Broadcast.*, vol. 61, no. 1, pp. 98–104, March 2015.

[5] F. Mazzenga, "Channel estimation and equalization for M-QAM transmission with a hidden pilot sequence," *IEEE Trans. Broadcast.*, vol. 46, no. 2, pp. 170–176, Jun 2000.

[6] W. W. Hu and C. P. Li, "Super-Imposed training scheme for timing and frequency synchronization in OFDM systems," in *2007 IEEE 65th Vehicular Technology Conference*, April 2007, pp. 1718–1722.

[7] K. Upadhya, S. A. Vorobyov, and M. Vehkapera, "Superimposed pilots are superior for mitigating pilot contamination in massive MIMO," *IEEE Trans. Signal Process.*, vol. 65, no. 11, pp. 2917–2932, June 2017.

[8] J. Ma, C. Liang, C. Xu, and L. Ping, "On orthogonal and superimposed pilot schemes in massive MIMO NOMA systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 12, pp. 2696–2707, Dec 2017.

[9] A. Goljahani, N. Benvenuto, S. Tomasin, and L. Vangelista, "Superimposed sequence versus pilot aided channel estimations for next generation DVB-T systems," *IEEE Trans. Broadcast.*, vol. 55, no. 1, pp. 140–144, March 2009.

[10] N. Chen and G. T. Zhou, "What is the price paid for superimposed training in OFDM?" in *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, May 2004, pp. 421–424.

[11] M. Ghogho, D. McLernon, E. Alameda-Hernandez, and A. Swami, "Channel estimation and symbol detection for block transmission using data-dependent superimposed training," *IEEE Signal Process. Lett.*, vol. 12, no. 3, pp. 226–229, March 2005.

[12] M. Ghogho, D. McLernon, E. Alameda-Hemandez, and A. Swami, "SISO and MIMO channel estimation and symbol detection using data-dependent superimposed training," in *Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, March 2005, pp. 461–464.

[13] E. Gayosso-Rios, M. M. Lara, A. G. Orozco-Lugo, and D. McLernon, "Symbol-Blanking Superimposed Training for Orthogonal Frequency Division Multiplexing systems," in *2010 7th International Symposium on Wireless Communication Systems*, Sept 2010, pp. 204–208.

[14] G. Dou, C. He, C. Li, and J. Gao, "Channel estimation and symbol detection for OFDM systems using data-nulling superimposed pilots," *Electronics Letters*, vol. 50, no. 3, pp. 179–180, January 2014.

[15] S. Ohno and G. B. Giannakis, "Capacity maximizing MMSE-optimal pilots for wireless OFDM over frequency-selective block Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 2138–2145, Sept 2004.

[16] F. Zaarour, E. Simon, M. Zwingelstein-Colin, and I. Dayoub, "A low complexity turbo receiver for data nulling superimposed pilots in OFDM," in *2015 Fifth International Conference on Digital Information and Communication Technology and its Applications*, April 2015, pp. 32–37.

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

[18] Myung Sun Baek, M.-J. Kim, Y.-H. You, and H.-K. Song, "Semi-blind channel estimation and PAR reduction for MIMO-OFDM system with multiple antennas," *IEEE Trans. Broadcast.*, vol. 50, no. 4, pp. 414–424, Dec 2004.

[19] W. C. Huang, C. P. Li, and H. J. Li, "On the power allocation and system capacity of OFDM systems using superimposed training schemes," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1731–1740, May 2009.

[20] S. M. Kay, *Fundamentals of statistical signal processing*. Englewood Cliffs, NJ: Prentice Hall, 1993.

[21] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 933–946, May 2000.

[22] *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation*, 3GPP TS, 1 2011, 36.211 version 10.0.0 Release 10.

[23] E. Alameda-Hernandez, D. C. McLernon, A. G. Orozco-Lugo, M. M. Lara, and M. Ghogho, "Frame/Training sequence synchronization and DC-offset removal for (data-dependent) superimposed training based channel estimation," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2557–2569, June 2007.

[24] C. P. Li and W. W. Hu, "Super-imposed training scheme for timing and frequency synchronization in OFDM systems," *IEEE Trans. Broadcast.*, vol. 53, no. 2, pp. 574–583, June 2007.

[25] M. Ahmadi and A. S. Mehr, "Superimposed training aided carrier frequency offset estimation in ofdm systems," in *2007 IEEE International Conference on Electro/Information Technology*, May 2007, pp. 296–299.

[26] X. Xie, M. Peng, Y. Li, W. Wang, and H. V. Poor, "Channel estimation for two-way relay networks in the presence of synchronization errors," *IEEE Trans. Signal Process.*, vol. 62, no. 23, pp. 6235–6248, Dec 2014.

[27] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

Juan Carlos Estrada-Jiménez received his B.E. degree from Escuela Politcnica Nacional (EPN), Quito, Ecuador, in 2009 and the M.Sc. degree in multimedia and communications in the Signal Theory and Communications department from Carlos III University in Madrid (UC3M), Spain in 2013. He joined the Communications Research Group of UC3M in 2015. Currently, he is pursuing the Ph.D. degree at the same place. He has been involved in ELISA and TERESA-ADA national projects. His research is focused in channel estimation strategies for new generation wireless communications.

M. Julia Fernndez-Getino Garca (S'99 - AM'02 - M'03) received the 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.