

# Differential Data-Aided Channel Estimation for Up-Link Massive SIMO-OFDM

MANUEL J. LOPEZ-MORALES<sup>ib</sup> (Student Member, IEEE),  
KUN CHEN-HU<sup>ib</sup> (Graduate Student Member, IEEE),  
AND ANA GARCIA-ARMADA<sup>ib</sup> (Senior Member, IEEE)

Department of Signal Theory and Communications, Universidad Carlos III de Madrid, 28911 Leganes, Spain

CORRESPONDING AUTHOR: M. J. LOPEZ-MORALES (e-mail: mjlopez@tsc.uc3m.es)

This work was supported in part by the European Union (EU) Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie ETN TeamUp5G, under Grant 813391, and in part by the Spanish National Project TERESA-ADA under Grant TEC2017-90093-C3-2-R (MINECO/AEI/FEDER, UE).

**ABSTRACT** Pilot symbol assisted modulation (PSAM) is widely used to obtain the channel state information (CSI) needed for coherent demodulation. It allows the density of pilot symbols to be dynamically chosen depending on the channel conditions. However, the insertion of pilots reduces the spectral efficiency, more severely when the channel is highly time-variant and/or frequency-selective. In these cases a significant amount of pilots is required to properly track the channel variations in both time and frequency dimensions. Alternatively, non-coherent demodulation does not require any CSI for the demodulation independently of the channel conditions. For the particular case of up-link (UL) based on massive single input - multiple output (SIMO) combined with orthogonal frequency division multiplexing (OFDM), we propose to replace the traditional reference signals of PSAM by a new differentially-encoded data stream that can be non-coherently detected. The latter can be demodulated without the knowledge of the CSI and subsequently used for the channel estimation. We denote our proposal as hybrid demodulation scheme (HDS) because it exploits both the benefits of a coherent demodulation scheme (CDS) and a non-coherent demodulation scheme (NCDS) to increase the spectral efficiency. The mean squared error (MSE) of the channel estimation, bit error rate (BER), achieved throughput and complexity are analyzed to highlight the benefits of this differential data-aided channel estimation as compared to other approaches. We show that the channel estimation is almost as good as PSAM, while the BER performance and throughput are improved for different channel conditions with a very small complexity increase.

**INDEX TERMS** Channel estimation, differential modulation, high-mobility, massive SIMO, non-coherent.

## I. INTRODUCTION

NEW DEMANDING applications have emerged in the last years, which require a tremendous increase in the capabilities of the communications networks. In this context, the Third Generation Partnership Project (3GPP) proposed the Fifth Generation (5G) of mobile communications, commonly known as the New-Radio (NR) [1]. Different scenarios are considered, from indoor to outdoor and from low to high mobility, where the latter has attracted much attention from industry [2] due to the commercial potential of new services related to the area of transportation, such as

autonomous vehicles, wide-band communications in high speed trains, etc.

Orthogonal frequency division multiplexing (OFDM) [3] and massive multiple input - multiple output (MIMO) [4] are key radio technologies adopted by 5G for the physical layer. They can provide a significant improvement in terms of spectral and energy efficiency with a constrained complexity. Thanks to the use of OFDM, the multi-path channel is converted into a set of independent flat-fading channels, and the MIMO pre/post-coding techniques can be applied to each subcarrier to obtain the equalized symbols. Moreover,

when the number of antennas is large enough, linear pre/post-coding techniques have a close to optimum performance [5], avoiding the use of other higher-complexity techniques.

Coherent demodulation is generally used at the receiver to compensate the propagation channel effects. Pilot symbol assisted modulation (PSAM) [6] is one of the most-frequently employed techniques to obtain the accurate channel state information (CSI) needed for the coherent processing with a reduced complexity. In PSAM several time-frequency resources are exclusively reserved to transmit some known reference signals, typically called pilots, where the density and position of these pilots in the resource grid can be dynamically chosen according to the channel conditions. However, even with this flexibility, the insertion of the PSAM pilots causes a transmit overhead, especially when the channel is fast-varying and/or very frequency-selective. In these cases an important quantity of pilots must be transmitted for periodically tracking the channel variations in each dimension.

Alternatives to PSAM exist, such as blind channel estimation [7], [8]. It requires no reference signals and consequently the spectral efficiency is increased, at the expense of sacrificing the complexity and/or the quality of the CSI. Blind methods mainly exploit the separability of the noise and signal sub-spaces by applying an eigenvalue decomposition (EVD) to the correlation matrix of the received signals. However, the CSI cannot be always obtained due to the ambiguity problem caused from the computation of EVD. Then, semi-blind techniques are proposed to solve the ambiguity issue and increase the quality of the channel estimates, now transmitting a reduced number of reference signals [9], [10]. Then, the spectral efficiency is again decreased. Moreover, the computational complexity of blind and semi-blind techniques is generally higher than PSAM's due to their iterative structure. Another alternative is superimposed training (ST) [11], [12], which is able to effectively alleviate the complexity of blind techniques, where the data and pilots share the same time-frequency resources. Hence, the spectral efficiency is increased at the expense of degrading the quality of the channel estimates due to the presence of data self-interference. All these techniques, blind, semi-blind or ST, are appropriate for time-invariant and/or flat-fading environments, since they need to average a significant number of samples to obtain an acceptable quality of the CSI.

Another interesting alternative is a non-coherent detection scheme (NCDS), capable of obtaining the transmitted data information without CSI knowledge. Though traditionally dismissed due to the SNR loss with respect to coherent approaches, the combination of non-coherent detection with massive MIMO has recently gained attention, since this SNR loss may become negligible when considering the gains in spectral efficiency and complexity reduction. A significant number of works in the literature focus on the up-link (UL) [13]–[16], where single antenna users transmit to a base station (BS) equipped with an array of a very

large number of elements, which corresponds to a multi-user single-input multiple-output (SIMO) scenario. The large number of elements at the receiver can leverage spatial diversity, improving the link performance. Reference [13] proposed a non-coherent scheme based on energy detection, that provides an insightful asymptotical performance analysis, while it practically requires an excessive number of antennas in order to provide an acceptable performance. References [14], [15] proposed the use of differential phase shift keying (DPSK), where the BS receiver non-coherently combines two contiguous symbols for each antenna and performs an averaging process over the antennas to mitigate the channel effects. Later, the combination of [14] with OFDM is studied in [16] assuming a doubly dispersive channel model. Moreover, the differential modulation showed great robustness against extremely high Doppler shifts, significantly outperforming the traditional coherent schemes in these scenarios. However, the system performance is compromised when the number of users simultaneously served in the same time and frequency resources is increased.

Taking into account the benefits and limitations of the existing techniques, we propose to take the best of both approaches. The idea is to maintain the SNR efficiency of coherent demodulation schemes (CDS) while not wasting resources for pilot transmission thanks to NCDS, even for those scenarios which suffer from high values of Doppler and/or delay spread. We focus on the particular case of UL since the recent literature has shown that the NCDS can make use of the large number of antennas at the BS to improve its performance [14]–[16]. Besides, most massive MIMO systems are based on a time division duplex (TDD) where training is performed in the UL, making the UL CSI estimation indispensable. Our proposal is a differential data-aided channel estimation scheme, where the reference symbols of traditional PSAM are replaced by differentially-encoded data that are non-coherently detected in a very robust manner thanks to the massive number of antennas. NCDS and CDS data are multiplexed in the same way as pilots and data in traditional PSAM and in most wireless communication standards based on OFDM. Then, these non-coherently detected symbols are used for channel estimation to produce the CSI to coherently demodulate the main data stream.

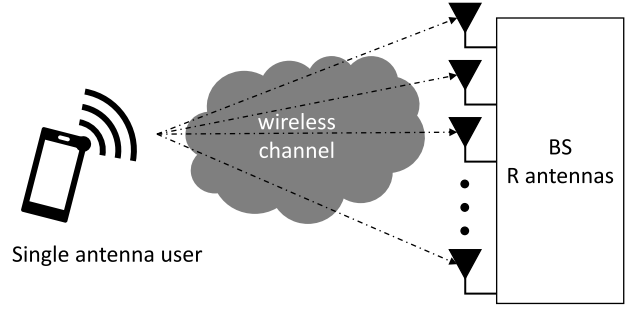
We denote our proposal as hybrid demodulation scheme (HDS), and we provide a complete description of how to multiplex coherent and differential data in the OFDM frame in a flexible way that may be adapted to different channel conditions. By combining two data streams, denoted as CDS and NCDS, we are able to exploit the benefits of the traditional CDS for different channel conditions without sacrificing the data-rate usually spent in pilot overhead. However, this combination is not straightforward. If not carefully designed, the CSI obtained with the NCDS stream may not have the required quality and decrease the achievable rate of the CDS stream with also a poor achievable rate of NCDS, so nothing is gained. Then, an analysis of the influence of the system parameters, particularly the constellation size and

the packet length of the NCDS stream, is crucial in order to understand their effects on the CSI accuracy, bit-error-rate (BER) and throughput, and provide a good design. The contributions of this paper are summarized as follows:

- An analytical derivation of the BER for the NCDS stream is provided, which differs from that of the CDS and is not available in the literature for non-coherent massive MIMO.
- An analysis of the MSE of the channel estimation is provided when the CSI is obtained using a NCDS data stream (unknown to the receiver). The analysis is based on the previous derivation of the BER and it allows us to determine the proper value of the parameters to ensure that the increment of the MSE with respect to the classical PSAM-based estimation is negligible.
- A comparative analysis of the complexity and achievable throughput of the CDS and HDS is provided. An alternative scheme such as ST is also evaluated for comparison. It is shown that HDS outperforms both ST and CDS with a very reduced increment in complexity. The analytical expressions of the throughput can be also used to choose the best constellation sizes for the HDS scheme in a certain scenario without the need of setting up a simulation environment.
- This approach is validated by evaluating the BER and throughput that are obtained using a realistic geometric channel model with link-level simulations. The numerical results offer the same conclusions as the analytical approach.

The remainder of the paper is organized as follows. Section II provides the system model based on SIMO-OFDM. It also provides an overview of the traditional CDS, PSAM and NCDS, as basic building blocks of our proposal. Section III describes the proposed HDS, which is based on the combination of CDS and NCDS, explaining the details of the proposed channel estimation and demodulation. The MSE obtained with the channel estimation is also analyzed. Section IV offers the throughput and complexity evaluation. Section V presents some numerical results to verify the analysis and compare the performance of our proposal with respect to other alternatives. Finally, in Section VI, some conclusions are pointed out.

**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  $\mathbf{a}$ .  $\angle(x)$  and  $\lceil x \rceil$  represent the angle and closest superior integer number of  $x$ , respectively. The superscripts  $(\cdot)^H$ ,  $(\cdot)^*$  and  $(\cdot)^\dagger$  denote Hermitian, complex conjugate and pseudoinverse of Moore-Penrose, respectively.  $\mathbb{E}\{\cdot\}$  represents the expected value.  $\text{Var}\{\cdot\}$  denotes variance.  $\mathcal{CN}(0, \sigma^2)$  represents the circularly-symmetric and zero-mean complex normal distribution with variance  $\sigma^2$ .  $f(a|b)$  and  $P(a|b)$  is the conditional probability density function (PDF) of  $a$  conditioned to  $b$  for continuous and discrete random variables, respectively.  $\Gamma(k, \xi)$  is the Gamma distribution with shape



**FIGURE 1.** SIMO-OFDM scenario, where a BS equipped with  $R$  antennas is serving a particular single-antenna user.

parameter  $k$  and scale parameter  $\xi$ , with mean  $\mu = k/\xi$  and variance  $\sigma = k/\xi^2$ .  $\Re$  and  $\Im$  refer, respectively, to the real and imaginary parts of a complex number.  $\mathbb{N}$  indicates the set of integer numbers.

## II. SYSTEM MODEL FOR CDS AND NCDS

In this section we first introduce the system model of an UL based on SIMO-OFDM. Focusing on this system model, we detail the channel estimation and equalization procedures of a classical CDS with PSAM. Then, a NCDS based on differential modulation is also presented. CDS and NCDS are building blocks of the HDS proposed in Section III.

### A. MASSIVE SIMO-OFDM

We consider one BS equipped with an array of  $R$  antenna elements which is serving several single-antenna users. For the purpose of channel estimation, we focus on the UL of a particular user (see Fig. 1) that transmits  $N$  consecutive OFDM symbols to the BS. The OFDM signal has  $K$  subcarriers and the length of the cyclic prefix ( $L_{CP}$ ) is long enough to absorb the effects of the multi-path channel ( $L_{CP} \geq L_{CH} - 1$ ). The number of channel taps is denoted by  $L_{CH}$ . At the BS, after removing the cyclic prefix and performing a fast-Fourier transform (FFT), we can process each subcarrier as one of a set of  $K$  independent sub-channels. Moreover, the whole time-frequency resource grid ( $K \times N$  resource elements) is split into unit blocks of  $K_B$  consecutive subcarriers and  $N_B$  contiguous OFDM symbols. Each unit block can be independently processed for channel estimation and detection. This model has practical interest since it is used, for example, in the Fourth Generation (4G) and 5G, where these generic unit blocks are denoted as physical resource blocks [1]. Hence, for the sake of conciseness and without loss of generality, we focus on a particular unit block.

The signal received at the BS on the  $k$ -th subcarrier and  $n$ -th OFDM symbol is denoted by the vector  $\mathbf{y}_k^n$  ( $R \times 1$ ) given by

$$\mathbf{y}_k^n = \mathbf{h}_k^n x_k^n + \mathbf{v}_k^n, \quad 1 \leq k \leq K_B, \quad 1 \leq n \leq N_B, \quad (1)$$

where  $x_k^n$  denotes the symbol transmitted by the user on the  $k$ -th subcarrier and  $n$ -th OFDM symbol,  $\mathbf{v}_k^n$  ( $R \times 1$ ) represents the additive white Gaussian noise vector, whose

elements are distributed according to  $\mathcal{CN}(0, \sigma_v^2)$ , and  $\mathbf{h}_k^n$  is a  $(R \times 1)$  vector that corresponds to the frequency channel response experienced at the  $R$  antennas, whose elements are distributed as  $\mathcal{CN}(0, 1)$ . Following [17], the channel coefficients are spatially uncorrelated but suffer from time variability according to

$$\mathbb{E}\left\{\left|([\mathbf{h}_k^n]_r)^*[\mathbf{h}_k^{n'}]_r\right|\right\} = \left|J_0\left(2\pi f_d \frac{\Delta n}{\Delta f} \left(1 + \frac{L_{CP}}{K}\right)\right)\right|, \quad \Delta n = |n' - n|, \quad 1 \leq r \leq R, \quad (2)$$

where  $\Delta f$  is the subcarrier spacing in Hertz,  $J_0(\cdot)$  denotes the zero-th order Bessel function of the first kind and  $f_d$  represents the maximum Doppler shift experienced by the transmitted signal, also in Hertz. Finally, assuming unit-power transmitted symbols, the SNR is defined as  $\rho = 1/\sigma_v^2$ .

### B. PILOT SYMBOL ASSISTED MODULATION (PSAM)

In PSAM, some of the time-frequency resources of the unit block are exclusively reserved for the transmission of pilot symbols while the others are dedicated to data symbols. Firstly, let  $\mathcal{A}$  denote the set of pairs of frequency and time indexes that describe the resources of the unit block, defined as

$$\mathcal{A} = \{(k, n) | k \in \{1, \dots, K_B\}, n \in \{1, \dots, N_B\}\}, \quad (3)$$

with cardinality  $|\mathcal{A}| = K_B \times N_B$ . Let  $\mathcal{A}_d$  denote the subset of the pairs of frequency and time indexes intended for the data transmission, and  $\mathcal{A}_p$  denote the subset that contains the pairs of frequency and time indexes reserved for the pilot symbols. These subsets must satisfy

$$\mathcal{A} = \mathcal{A}_d \cup \mathcal{A}_p, \quad \emptyset = \mathcal{A}_d \cap \mathcal{A}_p, \quad (4)$$

and their cardinality is given by

$$|\mathcal{A}_p| = K_p \times N_p, \quad |\mathcal{A}_d| = K_B N_B - K_p N_p. \quad (5)$$

In this equation  $K_p$  and  $N_p$  refer to the number of pilots in the frequency and time domain, respectively, in a unit block. According to [18], these pilot symbols must be equally spaced over the available resources. The distance between two contiguous pilot symbols in the frequency and time dimensions is denoted by  $L_K$  and  $L_N$ , measured in number of resources, and they must satisfy

$$K_B = L_K \times K_p, \quad N_B = L_N \times N_p. \quad (6)$$

Hence,  $x_k^n$  can be built in the following way

$$x_k^n = \begin{cases} s_k^n, & (k, n) \in \mathcal{A}_d \\ p_k^n, & (k, n) \in \mathcal{A}_p, \end{cases} \quad (7)$$

from the data symbol denoted by  $s_k^n$  that belongs to a quadrature amplitude modulation (QAM) constellation of size  $M_C$  and the pilot symbol at the  $k$ -th subcarrier and  $n$ -th OFDM symbol denoted by  $p_k^n$ .

At the BS, the channel estimation is performed at the resources occupied by the reference signals and obtained as

$$\hat{\mathbf{h}}_k^n = \min_{\mathbf{h}_k^n} \|\mathbf{y}_k^n - \mathbf{h}_k^n x_k^n\|^2, \quad (k, n) \in \mathcal{A}_p. \quad (8)$$

By applying the LS criterion [19], the channel estimate is

$$\hat{\mathbf{h}}_k^n = (x_k^n)^{-1} \mathbf{y}_k^n = \mathbf{h}_k^n + (x_k^n)^{-1} \mathbf{v}_k^n \quad (k, n) \in \mathcal{A}_p, \quad (9)$$

and its corresponding MSE is  $\sigma_e^2 = \sigma_v^2$ .

Finally, these channel estimates  $\hat{\mathbf{h}}_k^n$  are interpolated to obtain the channel estimates at the data symbol positions ( $\mathcal{A}_d$ ) by using some interpolation function. The channel estimation error after interpolation ( $\sigma_d^2$ ) depends on the channel variations in time and frequency and on the chosen interpolation function [20], whose detailed analysis is out of scope of this work.

### C. COHERENT DEMODULATION SCHEME (CDS)

Once the CSI is available, the post-coding vectors can be computed to equalize the received data symbols. Using a zero-forcing (ZF) criterion, they are obtained as

$$\mathbf{w}_k^n = (\hat{\mathbf{h}}_k^n)^\dagger, \quad (k, n) \in \mathcal{A}_d. \quad (10)$$

Then, the following decision variable is obtained in at the receiver to recover the transmitted data symbols

$$z_k^n = \mathbf{w}_k^n \mathbf{y}_k^n = \mathbf{w}_k^n \mathbf{h}_k^n x_k^n + \mathbf{w}_k^n \mathbf{v}_k^n, \quad (k, n) \in \mathcal{A}_d. \quad (11)$$

According to [21], the BER for a  $M_C$ -QAM constellation using a ZF post-equalization may be approximated as

$$P_{b,C} \approx \frac{2}{\sqrt{M_C} g_C} \sum_{i_1=1}^{g_C} \sum_{i_2=0}^{(1-2^{-i_1})\sqrt{M_C}-1} (-1)^{\lfloor \frac{i_2 2^{i_1-1}}{\sqrt{M_C}} \rfloor} \left( 2^{i_1-1} - \left\lfloor \frac{i_2 2^{i_1-1}}{\sqrt{M_C}} + \frac{1}{2} \right\rfloor \right) \times \left[ \frac{1-\mu_{i_2}}{2} \right]^R \sum_{i_3=0}^{R-1} \binom{R-1+i_3}{i_3} \left[ \frac{1+\mu_{i_2}}{2} \right]^{i_3}, \quad (12)$$

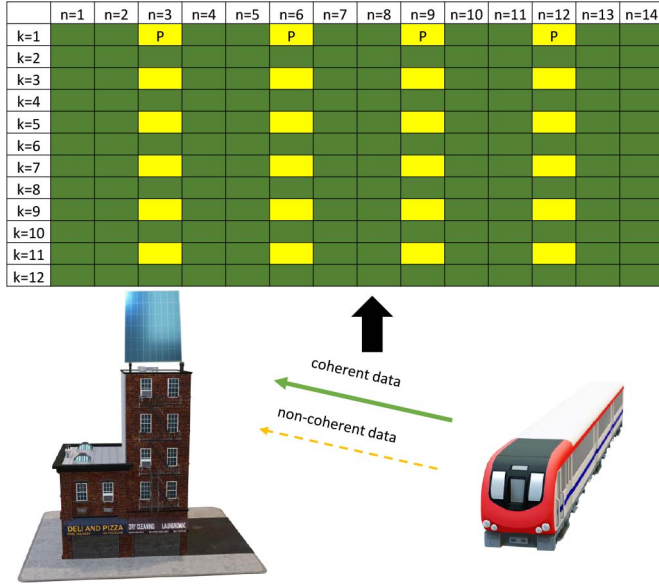
$$\mu_{i_2} = \sqrt{\frac{3(2i_2+1)^2 \gamma_s}{2(M_C-1) + 3(2i_2+1)^2 \gamma_s}}, \quad \gamma_s = \frac{\rho}{1 + \rho \sigma_d^2}, \quad (13)$$

where  $g_C = \log_2(\sqrt{M_C})$  and we assume the use of a rectangular pulse.

### D. NON-COHERENT DEMODULATION SCHEME (NCDS)

We will use the frequency domain implementation of the NCDS presented in [16]. This method has the advantage of involving a lower processing delay. In turn, only for strongly frequency-selective channels, an additional phase shift needs to be estimated and compensated. This phase shift remains quasi-static, even for extremely high Doppler scenarios and/or strongly frequency selective channels, when the number of antennas at the BS is large enough. Hence, if needed, this phase error can be estimated by using only two





**FIGURE 2.** Example of a unit block for the proposed HDS, where  $K_p = 6$  and  $N_p = 4$ . The green boxes are data modulated by CDS, the yellow boxes are data modulated by NCDS and “p” denotes a reference symbol.

pilots for the whole time-frequency resource grid ( $K \times N$ ), which has a minimum impact in terms of data rate.

For a given unit block of size  $\mathcal{A}$ , the transmitted symbols  $x_k^n$  are given by

$$x_k^n = \begin{cases} p_k^n, & k = 1 \\ x_{k-1}^n q_{k-1}^n, & \text{otherwise} \end{cases}, \quad (k, n) \in \mathcal{A}. \quad (14)$$

In this equation  $p_k^n$  is a reference symbol known by both ends of the communication link placed at the first subcarrier and each OFDM symbol (equivalent to one pilot in PSAM) and  $q_k^n$  is the complex data symbol to be transmitted in subcarrier  $k$  and symbol  $n$ , which belongs to a DPSK constellation of size  $M_N$ .

At the receiver, two symbols contiguously transmitted in the frequency domain are non-coherently combined as

$$z_k^n = \frac{1}{R} (\mathbf{y}_{k-1}^n)^H \mathbf{y}_k^n, \quad 2 \leq k \leq K_B, \quad 1 \leq n \leq N_B. \quad (15)$$

After some manipulations, (15) can be expanded as

$$z_k^n = \frac{1}{R} (\mathbf{h}_{k-1}^n x_{k-1}^n)^H \mathbf{h}_k^n x_k^n + \frac{1}{R} (\mathbf{v}_{k-1}^n)^H \mathbf{v}_k^n + \frac{1}{R} (\mathbf{v}_{k-1}^n)^H \mathbf{h}_k^n x_k^n + \frac{1}{R} (\mathbf{h}_{k-1}^n x_{k-1}^n)^H \mathbf{v}_k^n. \quad (16)$$

Assuming that the number of antenna elements at the BS is large enough, and making use of the Law of Large Numbers [22], the following is satisfied

$$(\mathbf{h}_{k-1}^n)^H \mathbf{h}_k^n \xrightarrow{R \rightarrow \infty} R, \quad (17)$$

$$(\mathbf{v}_{k-1}^n)^H \mathbf{h}_k^n, (\mathbf{h}_{k-1}^n)^H \mathbf{v}_k^n, (\mathbf{v}_{k-1}^n)^H \mathbf{v}_k^n \xrightarrow{R \rightarrow \infty} 0, \quad (18)$$

and since  $(x_{k-1}^n)^H x_k^n = q_k^n$ , (16) asymptotically tends to

$$z_k^n \xrightarrow{R \rightarrow \infty} q_k^n. \quad (19)$$

Otherwise, when the number of antennas is not large enough, there are some distortion and noise terms that are computed in [14]. Additionally, Appendix A shows the development of the analytical expressions for the symbol error rate (SER) and BER of NCDS combined with massive SIMO. The latter is given by

$$P_{b,N} = \frac{2}{\log_2(M_N)} - \frac{1}{\sigma_x \sigma_y \pi \log_2(M_N)} \times \int_{\mathcal{D}} \int_0^\infty e^{-\left(\frac{r \cos(\gamma) - \mu_x}{\sigma_x \sqrt{2}}\right)^2 - \left(\frac{r \sin(\gamma) - \mu_y}{\sigma_y \sqrt{2}}\right)^2} r dr d\gamma, \quad (20)$$

where  $\mathcal{D} \in [-\pi/M_N, \pi/M_N]$  denotes the decision region for the correspondent symbol of interest and  $(\mu_x, \mu_y) = (1, 0)$  (set for the DPSK symbol placed in the positive part of the real axis of the complex plane without loss of generality). The parameters  $\sigma_x$  and  $\sigma_y$  are as follows

$$\sigma_x = \sqrt{\frac{2 + 2\sigma_v^2 + \sigma_v^4}{2R}} \quad \text{and} \quad \sigma_y = \sqrt{\frac{2\sigma_v^2 + \sigma_v^4}{2R}}.$$

### III. PROPOSED DIFFERENTIAL DATA-AIDED CHANNEL ESTIMATION

In the previous section we showed that the CDS is limited by the additional overhead produced by the transmission of pilots. This inefficiency is more severe in scenarios with a high variability in time and/or frequency since the pilot density needs to be increased. We propose a new scheme, that we denote as the HDS, based on the combination of the CDS and the NCDS techniques to alleviate the overhead problem while maintaining a good performance. Our proposal consists of replacing the pilot symbols of the traditional PSAM-based CDS by an additional data stream following the NCDS approach, as shown in Fig. 2. At the BS, the data transmitted on the NCDS stream is firstly demodulated, without CSI knowledge, and then used to obtain the channel estimates. The diversity introduced by the large number of antennas at the BS and a careful selection of the parameters of the NCDS stream ensure a good performance, so it can provide an accurate channel estimation for the subsequent demodulation of the CDS data. Based on this channel estimates, the post-coding matrices are computed and the data transmitted on the CDS stream is equalized. Since the whole time-frequency grid is occupied by data symbols, to be demodulated either non-coherently (NCDS) or coherently (CDS), the system efficiency is increased.

In the following we provide the details of how the two data streams are multiplexed in the time-frequency grid of the unit block, the proposed procedure to estimate the CSI from the received differential data, and the analysis of the MSE that is obtained.

#### A. MULTIPLEXING CDS AND NCDS DATA STREAMS

The proposed HDS consists of multiplexing two data streams, namely a data stream transmitted following the traditional CDS ( $s_k^n$ ), and another data stream transmitted by the

NCDS ( $q_k^n$ ) (see Fig. 2). Comparing the HDS with the traditional NCDS, the traditional reference symbols are replaced by a new data stream to be non-coherently detected. Hence, the distribution of resources expressed in (7) is modified as

$$x_k^n = \begin{cases} s_k^n, & (k, n) \in \mathcal{A}_d \\ p_k^n, & (k, n) \in \mathcal{A}_p, k = k_{min}^n \\ x_{k-L_K}^n q_k^n, & (k, n) \in \mathcal{A}_p, k \neq k_{min}^n, \end{cases} \quad (21)$$

where  $k_{min}^n$  is the minimum value of the index  $k$  for the  $n$ -th OFDM symbol belonging to the subset  $\mathcal{A}_p$  (that is, the first sub-carrier of each OFDM symbol that carries the NCDS stream). For example,  $k_{min}^n = 1$  in Fig. 2. Interestingly, our proposed differential data-aided system can be adapted to any existing pilot distribution based on PSAM, by simply replacing the pilots by a NCDS data stream.

### B. DETECTION OF THE NCDS STREAM AND CHANNEL ESTIMATION

As described in (15), for those resources that carry the NCDS data stream,  $((k, n) \in \mathcal{A}_p)$ , the non-coherently combined symbols ( $z_k^n$ ) can be obtained at the BS receiver. Then, the transmitted data symbols ( $\hat{q}_k^n$ ) can be recovered from  $z_k^n$  by using the maximum-likelihood (ML) criterion [23] as

$$\hat{q}_k^n = \arg \max_{q_k^n} \{f_{z|q}(z_k^n | q_k^n)\}, \quad (k, n) \in \mathcal{A}_p, \quad (22)$$

where  $f_{z|q}(z_k^n | q_k^n)$  is the conditional probability that  $z_k^n$  is received having transmitted  $q_k^n$ .

Once the decision is taken on the NCDS data symbols, they can be used for the channel estimation as follows. Firstly, the differential symbols are reconstructed ( $\hat{q}_k^n \rightarrow \hat{x}_k^n$ ) by applying (21). Then, the LS criterion is used to estimate the channel at the  $(k, n) \in \mathcal{A}_p$  resources as

$$\hat{\mathbf{h}}_k^n = \frac{1}{\hat{x}_k^n} \mathbf{y}_k^n = \frac{x_k^n}{\hat{x}_k^n} \mathbf{h}_k^n + \frac{1}{\hat{x}_k^n} \mathbf{v}_k^n, \quad (k, n) \in \mathcal{A}_p. \quad (23)$$

Inspecting (23), we can see that there is no noise enhancement since  $|\hat{x}_k^n| = 1$ . However, unlike in traditional PSAM, there is an additional error term in the channel estimation produced by the possible mismatch between the transmitted and the reconstructed differential symbol. This error in the channel estimation is studied in the next subsection.

### C. ANALYSIS OF THE MSE OF THE CHANNEL ESTIMATION

The two terms given in (23) are independent to each other, which can be verified as

$$\begin{aligned} \mathbb{E} \left\{ \frac{x_k^n}{\hat{x}_k^n} [\mathbf{h}_k^n]_r \left( \frac{1}{\hat{x}_k^n} [\mathbf{v}_k^n]_r \right)^* \right\} &= \mathbb{E} \left\{ \frac{x_k^n}{|\hat{x}_k^n|^2} [\mathbf{h}_k^n]_r ([\mathbf{v}_k^n]_r)^* \right\} \\ &= \mathbb{E} \{x_k^n\} \mathbb{E} \{[\mathbf{h}_k^n]_r\} \mathbb{E} \{([\mathbf{v}_k^n]_r)^*\} = 0, \\ &\quad \times (k, n) \in \mathcal{A}_p, \end{aligned} \quad (24)$$

due to the fact that the transmitted symbol, channel and noise are uncorrelated random variables. Hence, the channel estimation error incurred when using (23) is given by

$$\sigma_e^2 = \mathbb{E} \left\{ |\hat{\mathbf{h}}_k^n - \mathbf{h}_k^n|^2 \right\} = \sigma_q^2 + \sigma_v^2 \quad (k, n) \in \mathcal{A}_p, \quad (25)$$

where the expectation is performed over both time and frequency dimensions. Recall that  $\sigma_v^2$  is the noise variance, and the first term can be obtained as

$$\begin{aligned} \sigma_q^2 &= \mathbb{E} \left\{ \left| \frac{x_k^n}{\hat{x}_k^n} \mathbf{h}_k^n - \mathbf{h}_k^n \right|^2 \right\} = \mathbb{E} \left\{ \left| \frac{x_k^n}{\hat{x}_k^n} - 1 \right|^2 \right\} = \\ &= 1 + \mathbb{E} \left\{ \frac{|x_k^n|^2}{|\hat{x}_k^n|^2} - \frac{x_k^n}{\hat{x}_k^n} - \frac{(x_k^n)^*}{(\hat{x}_k^n)^*} \right\} = \\ &= 2 - 2\mathbb{E} \{ \cos(\angle(x_k^n) - \angle(\hat{x}_k^n)) \} = 2\alpha. \end{aligned} \quad (26)$$

Here  $\alpha$  accounts for the channel estimation error produced by the use of an unknown data sequence. It depends on the error probability of the decision of ( $\hat{q}_k^n$ ). Its value is derived in Appendix B and shown to be

$$\alpha \leq \begin{cases} 0, & P_{s,N} = 0 \\ 1 - \frac{1 - P_{s,N} - (1 - P_{s,N})^{K_p}}{(K_p - 1)P_{s,N}}, & 0 < P_{s,N} \leq 1. \end{cases} \quad (27)$$

It depends on the NCDS stream length ( $K_p$ ) and the error probability of the symbols  $\hat{q}_k^n$ , denoted as  $P_{s,N}$ . Moreover,  $\alpha$  is bounded by  $0 \leq \alpha \leq 1$ . The case for  $\alpha = 0$  is equivalent to PSAM, where all symbol decisions are correct, while  $\alpha = 1$  means that all decisions are wrong, degrading the channel estimates. Therefore, the channel estimation error for our proposed scheme  $\sigma_e$  is bounded by

$$\sigma_v^2 \leq \sigma_e^2 \leq \sigma_v^2 + 2. \quad (28)$$

For low values of SNR, the dominant term is given by the noise since  $\sigma_v^2 \gg 1$ . For high and moderate SNR values, both terms in (28) decrease. Moreover, according to (26) and (27),  $\sigma_q^2$  can be minimized by properly setting  $P_{s,N}$  and  $K_p$ , thus assuring a performance similar to the traditional PSAM with the advantage of being capable of transmitting a higher data rate. The performance of  $P_{s,N}$  depends on the constellation size ( $M_N$ ), number of antennas ( $R$ ) and the SNR ( $\rho$ ). In order to improve the error performance of the symbols transmitted by the NCDS and the quality of the channel estimation,  $M_N$  should be specifically chosen for a particular scenario. On the other hand, the length of the sequence  $K_p$  should be also constrained. However, this is not restrictive since  $K_p$  is already constrained by using the unit block.

Finally, the estimated channel ( $\hat{\mathbf{h}}_k^n$ ,  $(k, n) \in \mathcal{A}_p$ ) is interpolated to provide the CSI for the whole resources of the unit block ( $\mathcal{A}_d$ ), the post-coding matrices based on ZF are computed, and the CDS data symbols are equalized, following the procedure explained in the previous section.

### IV. ANALYSIS OF THE THROUGHPUT AND COMPLEXITY

In this section we analyze and compare in terms of throughput and complexity the traditional CDS and our proposed HDS. We show that HDS can increase the throughput with a negligible increment of the complexity.

### A. THROUGHPUT ANALYSIS

For a typical packet-based transmission, let us define the total throughput of the HDS as

$$T = T_C + T_N[\text{packet/s}], \quad (29)$$

where  $T_C$  and  $T_N$  refer to the throughput of the CDS and NCDS streams, respectively, for each unit block. The throughput of each scheme can be found as

$$T_C = \frac{\eta_C \Delta f K_B}{L_P} (1 - P_{b,C})^{L_P} \log_2(M_C), \quad (30)$$

$$T_N = \frac{\eta_N \Delta f K_B}{L_P} (1 - P_{b,N})^{L_P} \log_2(M_N). \quad (31)$$

In these equations,  $L_P$  denotes the number of bits in one packet,  $P_{b,N}$  is given by (20) and  $P_{b,C}$  by (12). The efficiencies of CDS and NCDS in terms of the occupied resources in the unit block are

$$\eta_C = 1 - \frac{K_p N_p}{K_B N_B}, \quad \eta_N = \frac{N_p}{K_B N_B} (K_p - 1). \quad (32)$$

Inspecting (29)-(31), we can see that the total throughput of our proposed HDS is the sum of the throughput of each scheme, CDS and NCDS given in (30) and (31), respectively. If some parameters are given by the system design, such as the number of antennas, SNR, bandwidth and length of the packet, the throughput of the NCDS only depends on the selected modulation scheme ( $M_N$ ) and its corresponding BER ( $P_{b,C}$  given in (20)). In turn, the BER if CDS ( $P_{b,C}$  given in (12)) not only depends on the constellation size ( $M_C$ ), but it also relies on the performance of the channel estimation provided by the NCDS stream. Hence, the proposed HDS outperforms the traditional CDS due to the additional throughput provided by the NCDS if the channel estimation error ( $\sigma_e^2$ ) is properly constrained in order to avoid the increment of  $P_{b,C}$ .

In order to obtain the best pair of constellation sizes ( $M_C$  and  $M_N$ ) for a given scenario with some specific parameters, the throughput given in (29)-(31) should be maximized. Due to the difficulty to find closed-form expressions from (12) and (20) and motivated by the fact that the search space is very small ( $\log_2(\max(M_N)) \times \log_2(\max(M_C))$  combinations), we propose to resort to an exhaustive search over these few options. The throughput expressions are evaluated for a particular scenario and the values that provide the maximum throughput given in (29) are chosen. As we show in the next section, the search is limited to a small number of options since the possible values of  $M_C$  and  $M_N$  are usually constrained to a few, e.g.,  $8 \times 6$ .

### B. COMPLEXITY

In general, non-coherent processing is much less complex than coherent processing. Anyway, as the NCDS stream is an addition, its demodulation will incur some extra complexity that would not exist in a CDS system alone. In order to show that the proposed HDS does not significantly

TABLE 1. Simulation parameters.

$K_p$	1, 2, 3, 4, 6, 12	$K_B, N_B$	12, 14	$K$	1024
$N_p$	1, 2, 4, 7, 14	$M_C$	4, 16, 64	$L_P$	20
$R$	16, 64, 256	$M_N$	4, 8, 16, 64	$\Delta f$	30 KHz

increase the complexity of the system, the number of complex multiplications (NCM) required for each scheme is accounted as follows.

For the particular case of CDS, the expression of the NCM is given by

$$D_C = K_B N_B (3R + 1) - 3RK_p N_p, \quad (33)$$

where the LS channel estimation given in (23) is considered, as well as the post-coding matrix computation given in (10) and the equalization process given in (11). The complexity introduced by the interpolation process is not taken into account, so this is an optimistic evaluation of the complexity of CDS. If we consider it, the additional relative complexity of NCDS is even lower.

For the case of HDS, the NCM can be expressed as

$$\begin{aligned} D_H &= D_C + D_N = D_C + RN_p(K_p - 1) = \\ &= K_B N_B (3R + 1) - RN_p(2K_p + 1), \end{aligned} \quad (34)$$

where  $D_N$  accounts for the differential decoding given in (15), required for the NCDS. Note that the differential encoding required for computing the differential symbols  $q_k^n$  at the transmitter, and for the reconstruction of the differential symbols  $\hat{q}_k^n$  at the receiver are considered negligible. The reason for this is the fact that the phase difference can be computed in polar coordinates, which corresponds to adding/subtracting phases, and the conversion from polar to binomial coordinates can be implemented by using look-up tables, since the differential symbols ( $q_k^n$ ) belong to a DPSK constellation, which is a finite set. On the contrary, these simplifications cannot be implemented for the differential decoding, since the received symbols, which are modified by the channel and noise effects, no longer belong to a finite set.

Comparing (33) and (34), we can see that the dominant term in (34) is the first one, which corresponds to the CDS, and therefore, both techniques have the same complexity order  $O(3RK_B M_B)$ , rendering the additional complexity of the NCDS stream negligible, as predicted.

### V. NUMERICAL RESULTS

In this section, we provide some numerical results to verify our analysis and show the validity of the proposed HDS, which is capable of outperforming the traditional CDS for several scenarios of interest, as well as other alternative schemes such as ST. In Table 1, we provide a summary of some numerical values for the different system parameters, where they have been chosen taking into account the numerology given in 5G [1]. Besides, the size of the unit block has been set  $K_B = 12$ ,  $N_B = 14$  and the interpolation method to ‘spline’ [24].

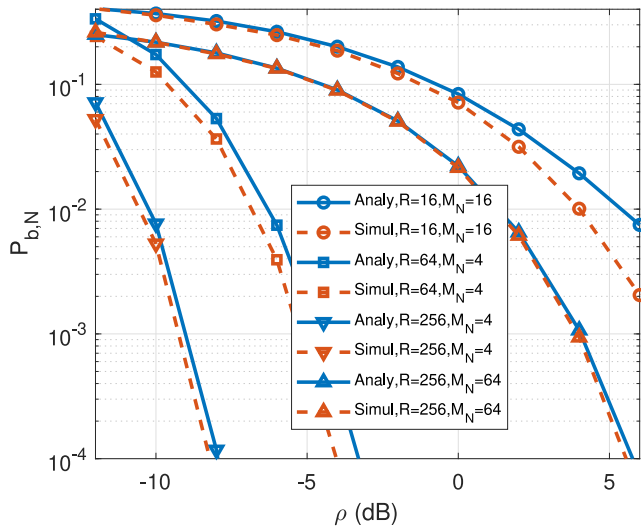


FIGURE 3. BER of NCDS with different values of  $R$  and  $M_N$ .

The number of pilots placed in the unit block (given by  $K_p$  and  $N_p$ ) must be incremented as the variations of the channel in both dimensions increase. Note that the time variability of the channel is modelled by using (2). According to [18], the number of pilots placed in the frequency dimension should be at least twice the number of taps of the multi-path channel, and in the time dimension there must be at least two pilots within the coherence time [17]. However, realistic communication links increase the number of pilots beyond the minimum, especially in the frequency dimension, to improve the quality of the estimators. Then, by changing the values of  $K_p$  and  $N_p$  we illustrate different scenarios of time and frequency variability of the channel.

The ST scheme of [12] is also taken into account for the comparison of throughput. This scheme has a parameter denoted as  $\beta$  that indicates the power allocated for the superimposed pilot symbols, while  $1 - \beta$  corresponds to the power of the data symbols. According to [12], we set  $\beta = 0.2$  which is the most frequently used value in the literature. ST requires an averaging process, which is implemented in both time and frequency dimensions and the number of resources to average is dynamically adapted for different values of Doppler and delay spreads to avoid the degradation of the channel estimates due to these effects. Hence, considering the definitions of (6), for the ST the number of averaged samples has been set equal to the amount of resources between two contiguous pilot symbols for each dimension ( $L_K \times L_N$ ).

#### A. BER FOR NCDS COMBINED WITH MASSIVE SIMO

We verify the validity of (20) to obtain the BER of the DPSK symbols received using the NCDS with a massive number of antennas at the BS ( $P_{b,N}$ ). In Fig. 3, we provide the simulation results and the analytical approximation of the BER for different values of  $M_N$  and  $R$ . The chosen power delay profile for the channel impulse response corresponds to the type B, given in [25], with a spatially uncorrelated channel of

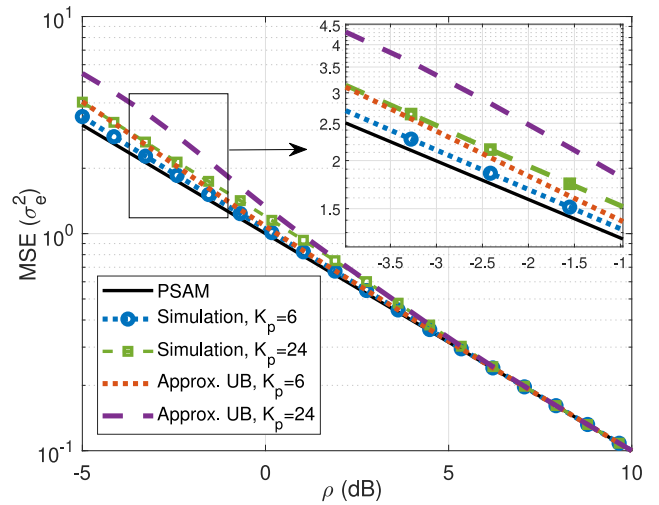


FIGURE 4. MSE of the channel estimation for  $R = 64$ ,  $M_N = 16$ ,  $N_p = 4$  and different values of  $K_p$ .

delay spread 363 ns. It can be confirmed that the difference between the simulation and the analytical approximation is negligible, in particular when the number of antennas is large enough.

#### B. MSE OF THE DIFFERENTIAL DATA-AIDED CHANNEL ESTIMATION

We check the performance of the proposed channel estimation using a differential data-aided sequence as compared to the traditional PSAM. The number of antennas is constrained to  $R = 64$  for feasibility reasons, due to the fact that this is a typical number of antennas for the deployment of 5G [25]. In Fig. 4, we show the MSE of the channel estimation for different values of  $K_p$ , given in (25), when  $M_N = 16$  and  $N_p = 4$ . Note that generally we process the NCDS streams divided in unit blocks, so  $K_p$  is lower than or equal to  $K_B$ . The motivation for this choice is to show the feasibility of our proposed scheme for a realistic implementation, since the unit block (or resource block) is the minimum unit of processing in the frequency domain in 4G or 5G [1]. Only for this subsection, both  $K_p$  and  $K_B$  are exceptionally allowed to be increased up to 24 subcarriers for illustrative purposes and discussion of the performance in other cases. The chosen power profile for the channel impulse response is the same as the one defined in the previous subsection. For medium and high SNR, our scheme and the PSAM have the same performance. For very low SNR, our proposed scheme is only slightly worse than PSAM, where the relative degradation increases with  $K_p$ . However, we can properly adjust  $K_p$  and  $M_N$  to obtain a required performance even for very low SNR regimes, as the subsequent BER and throughput analysis show. Furthermore, we can see that the analytical approximation (denoted upper-bound (UB)) is very close to the results obtained by simulation, saturating to  $\sigma_v^2 + 2$  ( $\alpha = 1$ ) for low SNRs, verifying the validity of the analysis shown in Appendix B.



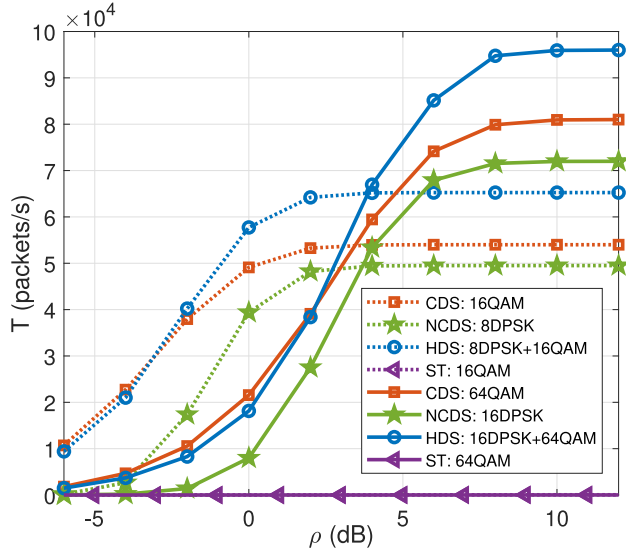


FIGURE 5. Throughput comparison of CDS, HDS, ST and NCDS for different constellation sizes,  $R = 64$ ,  $K_p = 6$  and  $N_p = 7$ .

### C. THROUGHPUT EVALUATION FOR DIFFERENT CHANNEL SCENARIOS

In Fig. 5, we show the throughput comparison for the CDS, NCDS, HDS and ST for a unit block where  $R = 64$ ,  $K_p = 6$  and  $N_p = 7$ , which is an interesting case in terms of medium to high delay and Doppler spreads. There is a substantial improvement of performance for the HDS with respect to the CDS, since the latter is highly penalized by the great amount of necessary pilots. The throughput of the HDS is the highest, not only due to a low average BER, but also to the additional throughput provided by the NCDS stream (recall (30) and (31)). Moreover, we also show the performance of the pure NCDS, where the DPSK symbols are occupying the entire unit block. We can see that NCDS is not able to outperform neither CDS nor HDS in this scenario. This is due to the fact that a QAM constellation has always a better performance than DPSK in terms of BER when  $M_C = M_N$ , provided that the channel can be adequately estimated. Also, it is known that the differential detection process increases the noise [14]. On the other hand, ST has a very poor performance since in this scenario there are not enough resources to perform the required averaging process in order to avoid the self-interference produced by the data, degrading its performance in terms of throughput.

According to the throughput analysis, for each particular scenario of interest imposed by the system design (number of antennas, length of the packet, delay and Doppler spreads, etc.), an optimum value of  $M_N$  and  $M_C$  can be chosen for each case of SNR ( $\rho$ ) in order to obtain the maximum throughput. Tables 2 and 3 show the throughput evaluation of the HDS calculated for  $M_C = 16$  and  $M_C = 64$ , respectively, where  $R = 64$ ,  $K_p = 6$  and  $N_p = 7$ . The maximum throughput values are highlighted in bold letters, and their corresponding optimal values of  $M_N$  and  $M_C$  depend on the

TABLE 2. Throughput for HDS ( $10^3 \times$  packets/second) for increasing  $\rho$  and increasing  $M_N$  for  $R = 64$ ,  $K_p = 6$ ,  $N_p = 7$  and  $M_C = 16$ .

$\rho \backslash M_N$	4	8	16	32	64	128
5 dB	61.5	<b>65</b>	64.6	55.6	47.3	44.9
10 dB	61.5	65	69	<b>69.9</b>	57.9	51
15 dB	61.5	65	69	<b>72.7</b>	71.8	58.4
20 dB	61.5	65	69	72.7	<b>76.4</b>	72.9

TABLE 3. Throughput for HDS ( $10^3 \times$  packets/second) for increasing  $\rho$  and increasing  $M_N$  for  $R = 64$ ,  $K_p = 6$ ,  $N_p = 7$  and  $M_C = 64$ .

$\rho \backslash M_N$	4	8	16	32	64	128
5 dB	75.5	<b>79.2</b>	76.7	33.8	18.4	15.7
10 dB	88.4	92.1	<b>95.9</b>	95.7	44.8	24.6
15 dB	88.5	92.2	96	<b>99.7</b>	98.4	45.3
20 dB	88.5	92.2	96	99.7	<b>103.4</b>	98.4

SNR ( $\rho$ ). Both constellation sizes can be increased when the SNR is high enough, maximizing the data-rate of the system. Note that the optimum value of  $M_N$  is not only chosen to increase the throughput of the NCDS stream, but it also has to constrain the channel estimation error, guaranteeing the performance of the CDS stream ( $M_C$ ). Hence, given the expressions of the throughput, we can easily choose the best configuration for a given case. For example  $M_C = 16$  and  $M_N = 8$  should be chosen for  $\rho = 5$  dB, while  $M_C = 64$  and  $M_N = 32$  are best for  $\rho = 15$  dB.

Table 4 shows a throughput comparison among the CDS, HDS and ST for different values of  $K_p$  and  $N_p$  where  $R = 64$ ,  $M_C = 16$ ,  $M_N = 8$  and  $\rho = 5$  dB. NCDS is not shown in this Table since its throughput remains constant for any of the considered values of delay and Doppler spread ( $K_p$  and  $N_p$  respectively). Its throughput is  $49.5 \times 10^3$  packets/s for  $M_N = 8$  and  $\rho = 5$ . This value is the same given in Table 4 for HDS when  $K_p = 12$  and  $N_p = 14$ , since in this extreme case all the symbols in the HDS are differentially encoded and HDS boils down to NCDS. Again, the different values of  $K_p$  and  $N_p$  correspond to different maximum supported values of delay and Doppler spreads. Once more, the use of ST does not bring in general an adequate performance. When there are enough resources for averaging ( $K_p$  and  $M_p$  low), the results are acceptable, but they degrade fast when  $K_p$  and  $M_p$  increase. In Table 5, for the same scenario, we provide the percentage of throughput increment of the proposed HDS with respect to the CDS, which is defined as

$$\Delta T_H(\%) = \frac{T - T_C}{T_C} \times 100. \quad (35)$$

We can see in both Tables 4 and 5 that for low mobility scenarios with a low or medium delay spread (low  $K_p$  and  $N_p$ ), the NCDS does not provide a significant increase of the throughput (approximately 0 – 5%) over the HDS. However, when either the Doppler or delay spreads are significantly increased, the throughput of the CDS is decreased while the throughput of NCDS is increased, improving  $\Delta T_H$  (approximately 6 – 70%). For the extreme case of extremely high Doppler and delay spread ( $K_p = 12$  and  $N_p = 14$ ),

**TABLE 4.** Throughput comparison among CDS, HDS and ST ( $10^3 \times$  packets/second) for  $\rho = 5$ ,  $R = 64$ ,  $M_N = 8$  and  $M_C = 16$ . Increasing  $K_p$  and  $N_p$  means increasing the maximum supported values of delay and Doppler spread.

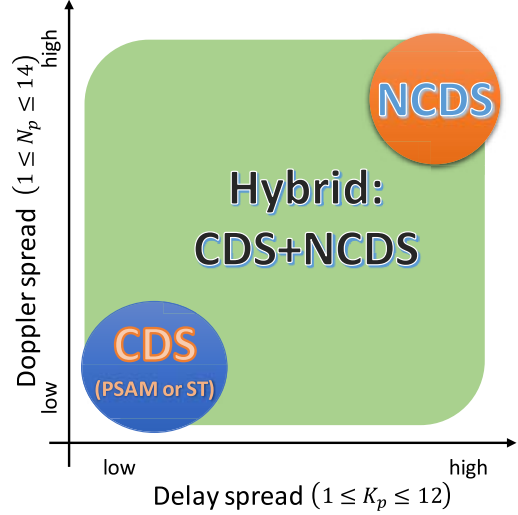
$N_p \backslash K_p$	1			2			3			4			6			12		
	CDS	HDS	ST	CDS	HDS	ST	CDS	HDS	ST	CDS	HDS	ST	CDS	HDS	ST	CDS	HDS	ST
1	72	71.6	62	71	71.5	42	70.7	71	31	70	71	19	69.4	71	9.6	67	70.4	2.7
2	71	71	45	70	71	20	69.4	71	10	69	70.5	7	67	70	2.3	62	69	0.3
4	70.3	70.3	20	67	67	5.6	67	69.4	2	65	69	1	62	68	0.4	51.4	65.6	0.1
7	69	69	7	66	68	1.5	63	67.5	0.5	60	67	0.2	54	65	0.1	36	61	0
14	66	66	2	60	64.5	0.2	54	63	0	48	61.5	0	36	58.5	0	0	49.5	0

**TABLE 5.** Percentage improvement of the throughput for the HDS with respect to the CDS; same parameters as in Table 4.

$N_p \backslash K_p$	1	2	3	4	6	12
1	0%	0.5%	0.9%	1.4%	2.3%	5.3%
2	0%	0.9%	1.9%	2.8%	4.8%	11.5%
4	0%	1.9%	3.8%	5.9%	10.4%	27.5%
7	0%	3.4%	7.1%	11.2%	20.8%	68.7%
14	0%	7.5%	16.7%	28.1%	62.5%	$\infty$

only NCDS can provide an acceptable performance (HDS collapses to pure NCDS), while CDS cannot be used. In summary, for increasing  $K_p$  and  $N_p$  values the performance of CDS worsens faster than that of HDS. ST can only provide an acceptable performance for the particular case of  $K_p = N_p = 1$ , where the channel has very mild variations in both dimensions and the averaging processing can be performed satisfactorily.

According to this numerical evaluation of the throughput, we provide a graphical summary of the advisable technique for different scenarios in Fig. 6. CDS is not recommendable in scenarios with high delay and/or Doppler spreads, where an excessive number of pilots must be transmitted to provide a continuous tracking of the channel at both time and frequency dimensions ( $K_p \uparrow$  and  $N_p \uparrow$ ), so that an acceptable quality of the channel estimates ( $\sigma_{e,p}^2 \rightarrow 0$  given in (25)) is provided. In these cases,  $T_C = 0$  since  $\eta_C \rightarrow 0$ . On the other hand, NCDS is suitable for these extreme cases where resources are not wasted to transmit reference signals. Indeed, [14]–[16] showed that NCDS has a great robustness against high Doppler scenarios, no matter the delay spread when OFDM is used. On the contrary, for low mobility or fixed communication scenarios with low or moderate values of delay spread, the number of reference signals can be greatly reduced ( $K_p \downarrow$  and  $N_p \downarrow$ ) since the channel remains quasi-static in both dimensions. Therefore, the throughput increment provided by the addition of the NCDS stream is not significant as compared to the throughput of the CDS ( $T_C \gg T_N$ ). Finally, in those scenarios where the channel variations are not excessively high, at least in one dimension, the combination of CDS and NCDS, that is the proposed HDS, outperforms the existing solutions in terms of throughput. In these scenarios, the proportion of resources allocated to the reference signals in CDS is significant and the effective data-rate of the link is reduced if only the CDS stream is sent. In the example provided in Fig. 2, which is a typical configuration in 5G-NR [1], the pilot symbols correspond to 28.6% of the grid, which is an important overhead, and our proposal can take advantage of this overhead by transmitting an additional data stream using NCDS.



**FIGURE 6.** Summary of the chosen scheme for different scenarios.

#### D. COMPLEXITY EVALUATION FOR DIFFERENT CHANNEL SCENARIOS

In order to compare the complexity of CDS and HDS, we define the complexity increment of HDS with respect to CDS as

$$\Delta D_H(\%) = \frac{D_H - D_C}{D_C} \times 100. \quad (36)$$

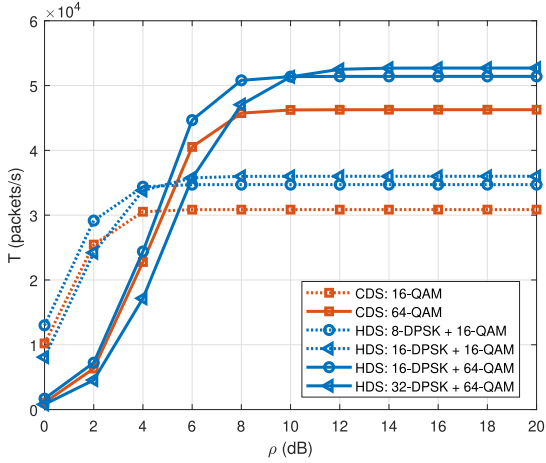
This complexity increment is presented in Table 6 where we can see that it is proportional to the number of symbols transmitted in the NCDS stream (given by  $K_p$  and  $N_p$ ). This complexity increment is below 10% for a low to medium amount of resources dedicated to NCDS, which corresponds to a throughput increment of up to 21% as shown in Table 5. For the extreme case of very fast time-varying and/or strongly frequency-selective channels, many more resources can be dedicated to the NCDS stream and the additional complexity can be increased up to 30%, while the throughput increase is almost 70%. In summary, the additional complexity produced by the proposed scheme is always much lower than the additional amount of throughput that it can provide, showing that HDS is more efficient than the traditional PSAM-based CDS.

#### E. BER AND THROUGHPUT EVALUATION WITH A GEOMETRIC CHANNEL MODEL

In this subsection we provide some simulation results with a more realistic channel model to compare CDS and HDS. Particularly, we adopt a geometric wide-band channel model

**TABLE 6.** Complexity increment of HDS with respect to CDS for different  $K_P$  and  $N_P$  values.

$N_P \backslash K_P$	1	2	3	4	6	12
1	0.0%	0.2%	0.4%	0.6%	1.0%	2.3%
2	0.0%	0.4%	0.8%	1.2%	2.1%	5.1%
4	0.0%	0.8%	1.7%	2.6%	4.6%	12.1%
7	0.0%	1.5%	3.2%	5.0%	9.2%	30.2%
14	0.0%	3.3%	7.4%	12.4%	27.5%	—

**FIGURE 7.** Throughput comparison of CDS and HDS for different constellation sizes,  $K_P = 6$  and  $N_P = 7$ .

which enables the characterization of the effects of the propagation channel and the antenna arrays [26]. It is characterized by the geometric superposition of several separate clusters, where each of them has a different value of delay and gain. Moreover, each cluster is made of a certain number of rays with different angle of arrival and departure. The chosen array configuration corresponds to a uniform linear array (ULA), where the distance of two contiguous elements is half the wavelength. The delay spread is set to 363 ns and the angular spread is set to 5 degrees, which are example values defined in [25]. Additionally, with interest in high mobility scenarios, we set a Doppler frequency of 1.6 KHz, which corresponds to a carrier frequency of 3.5 GHz and an approximated speed of 500 km/h. Moreover, we set  $K_P = 6$  and  $N_P = 4$ , which is the configuration given in Fig. 2.

In Fig. 7, we provide the throughput comparison for the different values of constellation size. We can see that the HDS provides a significant additional throughput as compared to CDS. Under this particular realistic channel model, the achieved throughput of both CDS and HDS is lower than in the previous case, due to the effect of the spatial correlation produced by the chosen array configuration. However, HDS still outperforms the traditional CDS by approximately a 11% of throughput increment, which is similar to what was obtained in the equivalent case with spatially uncorrelated channels. Hence, these results also show the advantages of our proposed scheme in a more realistic environment.

## VI. CONCLUSION

For an UL massive SIMO-OFDM system, we have proposed a differential data-aided channel estimation scheme, where

the traditional reference signals are replaced by a differential data stream. This data stream is demodulated using non-coherent detection at the BS and is used to perform the channel estimation. Therefore, by proposing a hybrid scheme denoted as HDS, the channel can be accurately estimated to perform coherent demodulation of the CDS stream, while the resources typically occupied by pilots are now leveraged for transmission of the NCDS data stream. The benefits of our proposal have been evaluated in terms of the channel estimation MSE, BER, throughput and complexity.

We have provided the analytical expressions of the channel estimation MSE and we have shown that it has very close performance to PSAM-based estimation. We have provided analytical expressions of the BER and evaluated the throughput of HDS for different configurations of the resources in the time-frequency grid, which correspond to different values of Doppler and delay spreads, showing that it outperforms the CDS with up to a 75% of throughput increment, which is obtained for high mobility scenarios. The different theoretical derivations given in this work are shown to match the numerical results, showing the accuracy of the analysis, and facilitating an optimization of the system parameters. Illustrative values of the number of antennas and the length of the NCDS data stream were chosen to show the feasibility of this solution in the frame of the current mobile communications standards and deployments. With the analytical tools provided in this work, the same optimization can be performed for any other values of these parameters.

In summary, with the HDS scheme we are able to replace the reference signals with a differential data stream, which in most of the analyzed scenarios provides a higher throughput than what can be achieved with a coherent system based on PSAM. Other alternatives such as ST perform worse when there is a mild frequency selectivity or channel variability. Then, this work contributes to the improvement of the spectral efficiency of massive MIMO-OFDM systems, which is crucial for the evolution of wireless communications.

It is worth noting that the NCDS stream can be used, for example, to establish a low latency service in parallel with the regular high data rate application provided by the CDS, since the detected NCDS data stream can be quickly forwarded to upper protocol layers while the channel is estimated for its use to equalize the CDS data stream. Therefore, it is possible to use this system for a single service with enhanced throughput or to multiplex two parallel services at the physical level.

## APPENDIX A

### DERIVATION OF SER AND BER OF NON-COHERENTLY DETECTED $M_N$ -DPSK COMBINED WITH MASSIVE MIMO

From Eq. (16), we can characterize the statistical distribution of the received symbol. Without loss of generality and to ease the notation, we focus on a particular subcarrier  $k$  and time instant  $n$ . According to [15], each term of (16) is independent

to each other and (16) can be redefined as

$$z_k^n = (z_k^n)_g + (z_k^n)_r = \sum_{l=1}^3 (z_k^n)_{g,l} + (z_k^n)_r, \quad (37)$$

where

$$(z_k^n)_{g,1} = \frac{1}{R} \sum_{r=1}^R [\mathbf{v}_{k-1}^m]_r^* [\mathbf{v}_k^n]_r \quad (38)$$

$$(z_k^n)_{g,2} = \exp(j\angle(q_k^n)) \frac{1}{R} \sum_{r=1}^R [\mathbf{v}_{k-1}^m]_r^* [\mathbf{h}_k^n]_r, \quad (39)$$

$$(z_k^n)_{g,3} = \exp(-j\angle(q_{k-1}^m)) \frac{1}{R} \sum_{r=1}^R [\mathbf{v}_k^n]_r ([\mathbf{h}_{k-1}^m]_r)^*, \quad (40)$$

$$(z_k^n)_r = q_k^n \frac{1}{R} \sum_{r=1}^R |[\mathbf{h}_k^n]_r|^2. \quad (41)$$

The three terms  $(z_k^n)_{g,l}$  described in (37) are composed of the sum of the product of independent complex circularly symmetric Gaussian variables. Focusing on the first one, each term of  $(z_k^n)_{g,1}$  can be decomposed as

$$[\mathbf{v}_{k-1}^m]_r^* [\mathbf{v}_k^n]_r = A_v + jB_v, \quad (42)$$

$$A_v = \Re\{[\mathbf{v}_{k-1}^m]_r^* \Re\{[\mathbf{v}_k^n]_r\} - \Im\{[\mathbf{v}_{k-1}^m]_r^* \Im\{[\mathbf{v}_k^n]_r\}\} \quad (43)$$

$$B_v = \Im\{[\mathbf{v}_{k-1}^m]_r^* \Re\{[\mathbf{v}_k^n]_r\} + \Im\{[\mathbf{v}_{k-1}^m]_r^* \Im\{[\mathbf{v}_k^n]_r\}\}. \quad (44)$$

According to [27], the PDF of the product of two Gaussian random variables with the same parameters is proportional to the Modified Bessel function of the second kind and zero-th order [28]. Hence, both real and imaginary parts of (42) are a sum of  $2R$  independent product terms, and assuming the number of antennas at the BS ( $R$ ) is large enough, the central limit theorem (CLT) [29] can be used. As shown in [30], and due to the independence among the different terms of  $(z_k^n)_g$ , its distribution can be approximated by

$$(z_k^n)_g \sim \mathcal{CN}\left(0, \frac{1}{R}(\sigma_v^4 + 2\sigma_v^2)\right). \quad (45)$$

According to [31], the second term of (37) is a one-dimensional random variable that follows a Gamma distribution, and it can also be approximated as an one-dimensional Gaussian distribution when the number of terms in the sum is large enough given by

$$\frac{1}{R} \sum_{r=1}^R |[\mathbf{h}_k^n]_r|^2 \sim \mathcal{N}\left(1, \frac{1}{R}\right). \quad (46)$$

Hence, the distribution of  $(z_k^n)_r$  is given by

$$(z_k^n)_r \sim \mathcal{N}\left(\exp(j\angle(q_k^n)), \frac{1}{R} \exp(j\angle(q_k^n))\right). \quad (47)$$

Note that the distortion produced by  $(z_k^n)_r$  only affects the amplitude of the received signal. Then, the interference and

noise terms can be characterized as the sum of a complex circularly symmetric Gaussian variable affecting both real and imaginary parts, and a unidimensional Gaussian variable affecting only the amplitude.

To calculate the SER for the  $M_N$ -DPSK, it is sufficient to calculate the error probability for one symbol, due to the symmetry in the distribution of the symbols over the complex plane. Therefore, and without loss of generality, we take the symbol placed over the positive part of the real axis  $(x, y) = (1, 0)$ . We now use two variables  $x$  and  $y$  to define the real and imaginary axis of the complex plane  $\mathcal{C}$ . Then, we can approximate the distribution of  $z_k^n$  with real and imaginary parts as

$$\Re\{z_k^n\} \sim \mathcal{N}(\mu_x, \sigma_x^2) = \mathcal{N}\left(1, \frac{2 + 2\sigma_v^2 + \sigma_v^4}{2R}\right), \quad (48)$$

$$\Im\{z_k^n\} \sim \mathcal{N}(\mu_y, \sigma_y^2) = \mathcal{N}\left(0, \frac{\sigma_v^4 + 2\sigma_v^2}{2R}\right). \quad (49)$$

Due to the independence of the real and imaginary parts of  $z_k^n$ , the bidimensional PDF of  $z_k^n$  ( $f(\Re\{z_k^n\}, \Im\{z_k^n\})$ ) can be computed as the product of both parts, distributed according to (48) and (49), resulting in

$$f(\Re\{z_k^n\}, \Im\{z_k^n\}) = f(x, y) = \frac{e^{-\left(\frac{x-\mu_x}{\sigma_x\sqrt{2}}\right)^2 - \left(\frac{y-\mu_y}{\sigma_y\sqrt{2}}\right)^2}}{\sigma_x\sigma_y2\pi}. \quad (50)$$

To simplify the integration to calculate the SER, we integrate the PDF in polar coordinates, for which we have to make the following change of variables ( $x = r\cos(\gamma)$  and  $y = r\sin(\gamma)$ ) as

$$f(r, \phi) = \frac{e^{-\left(\frac{r\cos(\gamma)-\mu_x}{\sigma_x\sqrt{2}}\right)^2 - \left(\frac{r\sin(\gamma)-\mu_y}{\sigma_y\sqrt{2}}\right)^2}}{\sigma_x\sigma_y2\pi}. \quad (51)$$

Hence, the SER probability of the  $q_k^n$  is given by

$$P_{s,N} = 1 - \frac{\int_{\mathcal{D}} \int_0^\infty e^{-\left(\frac{r\cos(\gamma)-\mu_x}{\sigma_x\sqrt{2}}\right)^2 - \left(\frac{r\sin(\gamma)-\mu_y}{\sigma_y\sqrt{2}}\right)^2} r dr d\gamma}{\sigma_x\sigma_y2\pi}, \quad (52)$$

where  $\mathcal{D}$  denotes the decision region for the particular symbol of interest, which in the selected case  $(x, y) = (1, 0)$  is defined as  $\mathcal{D} \in [-\pi/M_N, \pi/M_N]$ .

Finally, according to [21], the BER of a  $M_N$ -DPSK symbol in NCDS can be approximated with a lower bound as

$$P_{b,N} \approx \frac{2}{\log_2(M_N)} P_{s,N}. \quad (53)$$

## APPENDIX B

### ANALYSIS OF THE MSE OF THE DIFFERENTIAL DATA-AIDED CHANNEL ESTIMATION

Given (21), the obtained differential symbol at the  $k$ -th position is given by

$$\hat{x}_k^n = p_{k_{min}}^n \prod_{i=2}^k \hat{q}_{i-L_K}^n, \quad (k, n) \in \mathcal{A}_p, \quad k < K_p. \quad (54)$$



The first differential symbol of the block has the same performance as PSAM since the reference symbol is known. Moreover, the differential symbols  $x_k^n$  belong to the same constellation as the complex data symbols  $\hat{q}_k^n$  with possibly an additional phase rotation produced by the phase of the first reference symbol.

The channel estimation error due to the wrong decision on the symbols transmitted by the NCDS was shown in (26) to depend on  $\alpha$  and can be defined as

$$\begin{aligned}\alpha &= 1 - \mathbb{E}\{\cos(\angle(x_k^n) - \angle(\hat{x}_k^n))\} = \\ &= 1 - \frac{1}{K_p} \sum_{k=2}^{K_p} \sum_{i_t=1}^{M_N} \sum_{i_r=1}^{M_N} \cos(\angle(x_k^n(i_t)) - \angle(\hat{x}_k^n(i_r))) \times \\ &\quad \times P_{x_r|x_t}(x_r = \hat{x}_k^n(i_r)|x_t = x_k^n(i_t))P_{x_t}(x_t = x_k^n(i_t)), \quad (55)\end{aligned}$$

where  $x_k^n(i_t)$  and  $\hat{x}_k^n(i_r)$  denote that the transmitted and decided differential symbols correspond to the  $i_t$ -th and  $i_r$ -th element of the constellation, respectively. According to (54) and inspecting (55), we realize that each differential symbol of the constellation ( $\hat{x}_k^n(i_r)$ ) can be produced by different combinations of the decided symbols ( $\hat{q}_k^n$ ). Hence, let us define the following transmission and reception vector indices as

$$\mathbf{c}_{t,k} = [c_{r,1} \quad \cdots \quad c_{r,k-1}] \in C_{t,k}, \quad 1 \leq c_{r,i} \leq M_N, \quad (56)$$

$$\mathbf{c}_{r,k} = [c_{r,1} \quad \cdots \quad c_{r,k-1}] \in C_{r,k}, \quad 1 \leq c_{t,i} \leq M_N, \quad (57)$$

and they must satisfy that

$$x_k^n(i_t) = p_{k_{\min}}^n \prod_{i=2}^k q_{i-L_K}^n(c_{t,i-1}), \quad (k, n) \in \mathcal{A}_p, \quad (58)$$

$$\hat{x}_k^n(i_r) = p_{k_{\min}}^n \prod_{i=2}^k \hat{q}_{i-L_K}^n(c_{r,i-1}), \quad (k, n) \in \mathcal{A}_p. \quad (59)$$

Therefore, the second term of the product given in (55) can be expressed as

$$\begin{aligned}P_{x_2|x_1}(x_2 = \hat{x}_k^n(i_r)|x_1 = x_k^n(i_t))P_{x_t}(x_t = x_k^n(i_t)) &= \\ &= \frac{1}{(M_N)^{k-1}} \sum_{\mathbf{c}_{t,k} \in C_{t,k}} \sum_{\mathbf{c}_{r,k} \in C_{r,k}} \prod_{i=2}^k P_{q_1}(q_{i-L_K}^n(c_{t,i-1})) \\ &\quad \times P_{q_2|q_1}(\hat{q}_{i-L_K}^n(c_{r,i-1})|q_{i-L_K}^n(c_{t,i-1})). \quad (60)\end{aligned}$$

Given the analytical expression of  $\alpha$ , it can be solved by numerical computation. However, we provide the upper and lower bounds of  $\alpha$  in order to provide some insights. Given (55), we can split it as

$$\alpha = 1 - \alpha_1 - \alpha_2, \quad (61)$$

$$\begin{aligned}\alpha_1 &= \frac{1}{K_p} \sum_{k=2}^{K_p} \sum_{i=1}^{M_N} \times \\ &\quad \times P_{x_r|x_t}(x_r = x_k^n(i)|x_t = x_k^n(i))P_{x_t}(x_t = x_k^n(i)), \quad (62)\end{aligned}$$

$$\alpha_2 = \frac{1}{K_p} \sum_{k=2}^{K_p} \sum_{i_t=1}^{M_N} \sum_{\substack{i_r=1 \\ i_r \neq i_t}}^{M_N} \cos(\angle(x_k^n(i_t)) - \angle(\hat{x}_k^n(i_r))) \times$$

$$\times P_{x_r|x_t}(x_r = \hat{x}_k^n(i_r)|x_t = x_k^n(i_t))P_{x_t}(x_t = x_k^n(i_t)), \quad (63)$$

where  $\alpha_1$  accounts for the probability of taking correct decisions in a block of  $K_p$  differential symbols, and  $\alpha_2$  corresponds to the average probability of taking wrong decisions weighted by the cosine of the angular distance of the symbols in one block of  $K_p$  differential symbols. Hence, the lower-bound of  $\alpha$  corresponds to taking correct decisions of all transmitted symbols by the NCDS, which is the same case as PSAM. Hence

$$\alpha_1 = 1, \quad \alpha_2 = 0 \rightarrow \alpha = 0. \quad (64)$$

Inspecting (63) and taking into account that the constellation is DPSK, we have that  $\alpha_2 > 0$  since it is more likely to wrongly choose those symbols that belong to the contiguous decision regions from the correct one (which corresponds to low phase difference). Therefore, after performing the sum over all the possible values of the differential symbol, the positive terms ( $\cos(\angle(x_k^n(i_t)) - \angle(\hat{x}_k^n(i_r))) > 0$ ) have a higher probability than the negative ones. Hence, the upper-bound of  $\alpha$  is given by

$$\alpha = (1 - \alpha_1 - \alpha_2) \leq (1 - \alpha_1), \quad (65)$$

and  $\alpha_1$  can be obtained as

$$\begin{aligned}\alpha_1 &= \frac{1}{K_p - 1} \sum_{k=2}^{K_p} (1 - P_{s,N})^{k-1} = \\ &= \begin{cases} 1, & P_{s,N} = 0 \\ \frac{1 - P_{s,N} - (1 - P_{s,N})^{K_p}}{(K_p - 1)P_{s,N}}, & 0 < P_{s,N} \leq 1, \end{cases} \quad (66)\end{aligned}$$

where  $P_{s,N}$  is the probability of the symbol error of the NCDS, which is computed in the Appendix A. Hence,  $\alpha$  can be bounded as

$$\alpha \leq \begin{cases} 0, & P_{s,N} = 0 \\ 1 - \frac{1 - P_{s,N} - (1 - P_{s,N})^{K_p}}{(K_p - 1)P_{s,N}}, & 0 < P_{s,N} \leq 1. \end{cases} \quad (67)$$

## REFERENCES

- [1] *Physical Channels and Modulation (Release 15)*, 3GPP Standard 38.211, 2018.
- [2] P. Fan, J. Zhao, and C. L., "5G high mobility wireless communications: Challenges and solutions," *China Commun.*, vol. 13, no. S2, pp. 1–13, 2016.
- [3] T. Hwang, C. Yang, G. Wu, S. Li, and G. Y. Li, "OFDM and its wireless applications: A survey," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1673–1694, May 2009.
- [4] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.
- [5] H. Yang and T. L. Marzetta, "Performance of conjugate and zero-forcing beamforming in large-scale antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 172–179, Feb. 2013.
- [6] F. Tufvesson and T. Maseng, "Pilot assisted channel estimation for OFDM in mobile cellular systems," in *Proc. IEEE 47th Veh. Technol. Conf. Motion*, vol. 3, 1997, pp. 1639–1643.
- [7] C. Tu and B. Champagne, "Subspace-based blind channel estimation for MIMO-OFDM systems with reduced time averaging," *IEEE Trans. Veh. Technol.*, vol. 59, no. 3, pp. 1539–1544, Mar. 2010.
- [8] S. A. Banani and R. G. Vaughan, "OFDM with iterative blind channel estimation," *IEEE Trans. Veh. Technol.*, vol. 59, no. 9, pp. 4298–4308, Nov. 2010.

- [9] J. Mirzaei, R. S. Adve, and S. Shahbazpanahi, "Semi-blind time-domain channel estimation for frequency-selective multiuser massive MIMO systems," *IEEE Trans. Commun.*, vol. 67, no. 2, pp. 1045–1058, Feb. 2019.
- [10] E. Nayeibi and B. D. Rao, "Semi-blind channel estimation for multiuser massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 66, no. 2, pp. 540–553, Jan. 2018.
- [11] W. Huang, C. Li, and H. 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.
- [12] J. C. Estrada-Jiménez, K. Chen-Hu, M. J. F. García, and A. G. Armada, "Power allocation and capacity analysis for FBMC-OQAM with superimposed training," *IEEE Access*, vol. 7, pp. 46968–46976, 2019.
- [13] A. Manolakis, M. Chowdhury, and A. Goldsmith, "Energy-based modulation for noncoherent massive SIMO systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7831–7846, Nov. 2016.
- [14] A. G. Armada and L. Hanzo, "A non-coherent multi-user large scale SIMO system relaying on M-ary DPSK," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 2517–2522.
- [15] V. M. Baeza and A. G. Armada, "Non-coherent massive SIMO system based on M-DPSK for Rician channels," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 2413–2426, Mar. 2019.
- [16] K. Chen-Hu and A. G. Armada, "Non-coherent multiuser massive MIMO-OFDM with differential modulation," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.
- [17] T. S. Rappaport *et al.*, *Wireless Communications: Principles and Practice*, vol. 2. Upper Saddle River, NJ, USA: Prentice-Hall, 1996.
- [18] R. Nilsson, O. Edfors, M. Sandell, and P. O. Borjesson, "An analysis of two-dimensional pilot-symbol assisted modulation for OFDM," in *Proc. IEEE Int. Conf. Pers. Wireless Commun.*, Dec. 1997, pp. 71–74.
- [19] Y. Qiao, S. Yu, P. Su, and L. Zhang, "Research on an iterative algorithm of LS channel estimation in MIMO-OFDM systems," *IEEE Trans. Broadcast.*, vol. 51, no. 1, pp. 149–153, Mar. 2005.
- [20] 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.
- [21] C. Wang, E. K. S. Au, R. D. Murch, W. H. Mow, R. S. Cheng, and V. Lau, "On the performance of the MIMO zero-forcing receiver in the presence of channel estimation error," *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 805–810, Feb. 2007.
- [22] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Found. Trends Commun. Inf. Theory*, vol. 1, no. 1, pp. 1–182, 2004.
- [23] J. L. Arbuckle, G. A. Marcoulides, and R. E. Schumacker, "Full information estimation in the presence of incomplete data," in *Advanced Structural Equation Modeling: Issues and Techniques*, vol. 243. Mahwah, NJ, USA: Lawrence Erlbaum, 1996, p. 277.
- [24] Z. Cao, G. Wen, L. Chen, Q. Shu, and L. Chen, "Comparison of interpolation methods for pilot aided estimation in direct-detection optical OFDM system," *Microw. Opt. Technol. Lett.*, vol. 55, no. 11, pp. 2604–2608, 2013.
- [25] *Study on Channel Model for Frequencies from 0.5 to 100 GHz (Release 15)*, 3GPP Standard 38.901, 2018.
- [26] A. Alkhateeb and R. W. Heath, "Frequency selective hybrid precoding for limited feedback millimeter wave systems," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 1801–1818, May 2016.
- [27] G. Cui, X. Yu, S. Iommelli, and L. Kong, "Exact distribution for the product of two correlated Gaussian random variables," *IEEE Signal Process. Lett.*, vol. 23, no. 11, pp. 1662–1666, Nov. 2016.
- [28] Z.-H. Yang and Y.-M. Chu, "On approximating the modified Bessel function of the second kind," *J. Inequalities Appl.*, vol. 2017, no. 1, p. 41, 2017. [Online]. Available: <https://europepmc.org/articles/PMC5306441>
- [29] J. A. Rice, *Mathematical Statistics and Data Analysis*, 3rd ed. Belmont, CA, USA: Duxbury, 2006.
- [30] R. Ware and F. Lad, "Approximating the distribution for sums of products of normal variables," University of Canterbury, Canterbury, U.K., Rep. UCDSMS, 2003.
- [31] C. Walck, "Hand-book on statistical distributions for experimentalists," Dept. Phys., Univ. Stockholm, Stockholm, Sweden, Rep. SUF-PFY/96-01, 1996.



**MANUEL J. LOPEZ-MORALES** (Student Member, IEEE) received the B.Eng. degree in electrical engineering from the Polytechnic University of Madrid in 2015 and the double M.Sc. degree (Hons.) in telecommunications engineering and wireless communications from the Polytechnic University of Madrid and Lund University, in 2018. He is currently pursuing the Ph.D. degree with the Universidad Carlos III de Madrid and working on the TeamUp5G Project with MSCA. He was a Visiting Student with the Microwave Vision Group, Satimo, Italy, since 2014, the Tokyo Institute of Technology, Japan, in 2016, Ericsson AB, Sweden, in 2017, and with the ISAE Supaero, France, in 2018. His research interests include signal processing, wireless communications, antenna technology, massive MIMO, machine learning, and optimization applied to communications.



**KUN CHEN-HU** (Graduate Student Member, IEEE) received the Ph.D. degree in multimedia and communications from the Universidad Carlos III de Madrid (UC3M), Spain, in 2019, where he is currently a Postdoctoral Researcher. He visited Eurecom, France, and Vodafone Chair TU Dresden, Germany, as a Guest Researcher. He also participated in different research projects in collaboration with several top companies in the area of mobile communications. His research interests are related to signal processing techniques, such as waveforms design, non-coherent massive MIMO, and channel estimation. In 2019, he received an award from UC3M recognizing his outstanding professional career after graduation.



**ANA GARCIA-ARMADA** (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the Polytechnical University of Madrid in February 1998. She is currently a Professor with the University Carlos III of Madrid, Spain, where she is leading the Communications Research Group. She has been a Visiting Scholar with Stanford University, Bell Labs, and the University of Southampton. She has participated (and coordinated most of them) in more than 30 national and 10 international research projects, as well as 20 contracts with the industry, all related to wireless communications. She has coauthored eight book chapters on wireless communications and signal processing. She has published around 150 papers in international journals and conference proceedings and she holds 4 patents. Her main interests are multicarrier and multiantenna techniques and signal processing applied to wireless communications. She has received the Young Researchers Excellence Award, the Award for Outstanding Achievement in Research, Teaching and Management and the Award for Best Practices in Teaching, all from University Carlos III of Madrid. She was awarded the third place Bell Labs Prize in 2014 for shaping the future of information and communications technology. She received the Outstanding Service Award from the IEEE ComSoc Signal Processing and Communications Electronics Technical Committee in 2019. She has contributed to international standards organizations, such as ITU and ETSI, and is member of the expert group of the European 5G PPP and of the advisory committee 5JAC of the ESA as an expert appointed by Spain on 5G. She has served on the Editorial Board of *Physical Communication* from 2008 to 2017, *IET Communications* from 2014 to 2017. She has been serving on the Editorial Board of *IEEE COMMUNICATIONS LETTERS* since 2016 (Editor until February 2019, Senior Editor from March 2019, Exemplary Editor Award in 2017 and 2018), and *IEEE TRANSACTIONS ON COMMUNICATIONS* since 2019. She has served on the TPC of more than 40 conferences and she has been member of the organizing committee of IEEE Globecom in 2019, IEEE Vehicular Technology Conference Fall in 2018, Spring 2018 and 2019, and IEEE 5G Summit 2017. She will be the General Chair of Globecom 2021. She was the Newsletter Editor of the IEEE ComSoc Signal Processing and Consumer Electronics Committee from 2017 to 2018, of which she has been Secretary since 2019. She was the Secretary of the IEEE ComSoc Women in Communications Engineering Standing Committee from 2016 to 2017, and the Chair from 2018 to 2019.