Pilot-less Massive MIMO TDD System with Blind Channel Estimation using Non-coherent DMPSK

Manuel Jose Lopez Morales
Universidad Carlos III de Madrid
Madrid, Spain
mjlopez@tsc.uc3m.es

Kun Chen-Hu
Universidad Carlos III de Madrid
Madrid, Spain
kchen@tsc.uc3m.es

Ana Garcia Armada
Universidad Carlos III de Madrid
Madrid, Spain
agarcia@tsc.uc3m.es

***Abstract*—A novel time division duplex massive MIMO approach based on performing a blind channel estimation in the uplink using differentially encoded data and a precoding in the downlink, also with differentially encoded data, is proposed. In this system, the use of any type of explicit pilot symbol is completely avoided while maintaining spatial multiplexing capabilities in the downlink. We perform an analysis of the full system in terms or signal-to-interference-and-noise ratio (SINR) for the uplink and the downlink. The performance of the channel estimation using differentially encoded data is also analyzed, since it affects the performance of the downlink data transmission. A multi-user allocation strategy in an OFDM grid is proposed. The analysis is corroborated via numerical results and the proposed scheme is shown to outperform its coherent counterpart.**

***Index Terms*—Non-coherent, massive MIMO, differential modulation, spatial multiplexing, TDD.**

I. Introduction

A key ingredient for the advancement of wireless communications is massive multiple-input-multiple-output (MIMO), a technology where a base station (BS) is equipped with a large number of antennas [1]. Typically, time division duplex (TDD) is used in coherent massive MIMO as the strategy to organize the uplink (UL) and downlink (DL) usage of the channel resources [2]. In TDD, each user sends its orthogonal reference signals, also called pilots, multiplexed with data in the uplink, and the channel is estimated to obtain the so-called channel state information (CSI) using these pilots. With the obtained CSI, the uplink data are coherently demodulated thanks to the estimated channel response. Additionally, using the same CSI, the downlink data are precoded to spatially separate the users making use of channel reciprocity.

There are challenging scenarios of interest, such as those with high mobility and/ or low SNR, where the use of accurate enough CSI is problematic, since many pilots may be needed [3], [4]. Moreover, in scenarios with many users, this problem aggravates, since each user needs to transmit orthogonal pilots to obtain their respective CSI without interference. Noteworthily, additional pilot symbols must be added in each data stream in the downlink of each user since the channel precoding may not perfectly separate the users or compensate the channel effects, and additional processing is needed making use of these pilots, which are often called demodulation pilots or downlink pilots [5]. The overhead due to uplink and downlink pilots may become excessive in the above mentioned circumstances, causing an intolerable decrease of the actual data throughput.

A technique that allows to receive data without using CSI in the uplink, while benefiting from the use of many antennas at the BS and also reducing the complexity of the receivers, is called non-coherent (NC) massive single-input-mutiple-output (SIMO) [6]. It was shown in [3] that NC detection based on differential M-ary phase shift keying (DMPSK) [6] can detect data without CSI and that the channel can be estimated after reconstructing the received NC data in the uplink.

While the literature is very focused on the uplink, not much has been proposed for the downlink [7]. Then, a natural question arises: can we leverage the channel estimation method proposed in [3] to design a whole pilot-less (uplink and downlink) communications system? To answer this question, in this paper, we propose and analyze an innovative approach for TDD massive MIMO in which the uplink is based on the NC massive SIMO scheme where each user sends orthogonal differentially encoded data streams, which are used in the receiver to both receive data and estimate the channel, as shown in [3]. The downlink is based on precoding the data streams of all simultaneously served users with this CSI estimated using NC uplink data. Users are spatially separated through the precoder and their data is differentially encoded to avoid the need for downlink demodulation pilots. Even though these pilots are shared by many users, their main limitation comes from the fact that for each user, the resources available for the downlink may be very limited and it could still impact the performance in the downlink, as shown in Fig. 1. We consider in the performance analysis an orthogonal frequency division multiplexing (OFDM) waveform [8] and channels that are correlated in time.

The remainder of this paper is organized as follows. We introduce the system model in Sec. II. We analyze the mean square error (MSE) of the channel estimation using NC data with time variability and characterize the theoretical expression of the symbol-error-rate (SER) of the precoded downlink data in Sec. III. Numerical results in Sec. IV confirm our derivations and show a performance comparison between the proposed scheme and the coherent counterpart. Some conclusions and future work are pointed out in Sec. V.

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 $\mathbf{A}$, where a subindex $\forall$ indexes all the elements in that dimension. $[\mathbf{a}]_n$ represents the n-th element of vector $\mathbf{a}$. The superscripts $(\cdot)^H$, $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^{-1}$ and $*$ denote Hermitian, complex conjugate, matrix transpose, matrix inverse and convolution, respectively. $\mathbb{E}\{\cdot\}$ represents the expected value. $\mathcal{CN}(0,\sigma^2)$ represents the circularly-symmetric and zero-mean complex normal distribution with variance σ^2, $U(a,b)$ represents the uniform distribution between a and b and $VG(\mu,\alpha,\beta,\lambda)$ refers to the variance gamma distribution where μ is the location, α and β are the asymmetry parameters and λ is the shape parameter. $\Re\{(\cdot)\}$ and $\Im\{(\cdot)\}$ refer, respectively, to the real and imaginary parts of a complex number. $||x||_2$ denotes the euclidean norm of x. $\mathbf{1}$ and $\mathbf{0}$ indicate a column vector of all ones and all zeros, respectively. $Q(\cdot)$ is the Q-function and $\mathbf{I}$ the identity matrix. $\angle(\cdot)$ is the angle function.

II. System Model

A. Baseline TDD coherent massive MIMO

We consider a massive MIMO base-station (BS), with R antennas and U single-antenna users. An OFDM signal composed of K subcarriers is used for transmission in UL and DL, with cyclic prefix length (L_{CP}), which is designed to be long enough to absorb the effects of the multipath channel. At each receiving side, after removing the cyclic prefix and performing a fast Fourier Transform (FFT), it is possible to process each subcarrier independently.

In the UL, each user u transmits a symbol x_n^k, containing either data or a pilot, placed in an orthogonal subcarrier k a time instant n. The propagation channel between the user u an the BS at time instant n and subcarrier k is represented by $\mathbf{h}$ of size $(R\times 1)$ where $\mathbf{h}_n^k \sim \mathcal{CN}(\mathbf{0},\mathbf{I})$ is modeling a Rayleig distributed channel, as defined in [9]. Following [10], [11 the channel coefficients suffer from time variability and a autoregressive model approximates the temporally correlate fading channel coefficients of subcarrier k at time instant n a

$$\mathbf{h}_{n'}^k = \alpha_d \mathbf{h}_n^k + \mathbf{w}_{n'}^h, \qquad (1$$

where n' refers to a time instant in the future with respe to n ($d = |n'-n|$ time difference in OFDM symbols $\alpha_d = J_0\left(2\pi d f_D\left(\frac{K+L_{CP}}{K\Delta f}\right)\right) < 1$ is the temporal correla tion parameter, Δf 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. Also, we assume that $\mathbf{w}_{n'}^h$ is uncorrelated with $\mathbf{h}_n^k$ since it represents a random channel change component and is modeled [12] as a Gaussian random process with i.i.d. entries and distribution $\mathbf{w}_n^h \sim \mathcal{CN}\left(\mathbf{0},(1-\alpha_d^2)\mathbf{I}\right)$.

The received signal in the UL for each user u at the k-th subcarrier and n-th time instant is given by

$$\mathbf{y}_n^k = \mathbf{h}_n^k x_n^k + \boldsymbol{\nu}_n^k, \qquad (2)$$

where $\boldsymbol{\nu}_n^k$ $(R\times 1)$ denotes the additive white Gaussian noise (AWGN) in the UL distributed according to $\boldsymbol{\nu}_n^k \sim \mathcal{CN}(0,\sigma_{\nu,u}^2)$. In the BS, $\mathbf{h}_n^k$ is estimated from $\mathbf{y}_n^k$. In case x_n^k is a known pilot, the channel for each user u is estimated following pilot-symbol-assisted modulation (PSAM) [13] and denoted as $[\hat{\mathbf{h}}_n^k]_u$. The estimated channels of all the users are stacked as $\hat{\mathbf{H}}_n^k = [[\hat{\mathbf{h}}_n^k]_1,\cdots,[\hat{\mathbf{h}}_n^k]_U]^T$.

In the DL, the symbols of all the users are stacked in $\mathbf{x}_n^k$ of size $(U\times 1)$ for each time instant n and subcarrier k and are precoded before transmission using the precoding matrix $\mathbf{B}_n^k = (\hat{\mathbf{H}}_n^k)^H$ for maximum ratio transmission (MRT). The DL channel is composed as $\mathbf{H}_n^k = [[\mathbf{h}_n^k]_1,\cdots,[\mathbf{h}_n^k]_U]^T$, where the DL channels of all users are stacked. Thus, in the DL the received signal is

$$\mathbf{y}_n^k = \mathbf{H}_n^k \mathbf{B}_n^k \mathbf{x}_n^k + \boldsymbol{\nu}_n^k, \qquad (3)$$

where $\boldsymbol{\nu}_n^k$ is a vector $(U\times 1)$ where each element is the noise at the receiver of each user u and is distributed as $[\boldsymbol{\nu}_n^k]_u \sim \mathcal{CN}(0,\sigma_k^2)$. When MRT is applied in the DL of the BS, we can extend (3) by separating the matrix between the desired user and the rest of the users, so (3) can be rewritten as

$$[\mathbf{y}_n^k]_u = [\mathbf{H}_n^k]_{u,\forall}[\mathbf{B}_n^k]_{\forall,u}[\mathbf{x}_n^k]_u + \sum_{u'\neq u}[\mathbf{H}_n^k]_{u,\forall}[\mathbf{B}_n^k]_{\forall,u'}[\mathbf{x}_n^k]_{u'} + \boldsymbol{\nu}_n^k. \qquad (4)$$

For coherent demodulation, in this manuscript it is assumed that the coherence time of the channel is at least four OFDM symbols (which is the most pessimistic case) as shown in Fig. 1. Thus, due to the reciprocity of the channel, we can ensure that $\mathbf{h}_{n_u}^k \approx (\mathbf{h}_{n_d}^k)^T$ so long as $|n_u - n_d| < n_c$, where $n_c =$ $0.15 f_D^{-1} \Delta f$ [10] denotes the coherence time in OFDM symbol

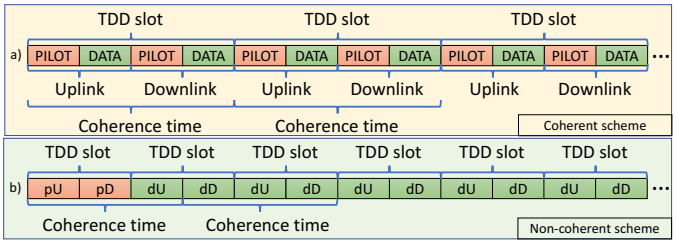

Fig. 1. TDD frame comparison for coherent (a) and non-coherent (b).

To analyze the effect of imperfect channel estimation for the proposed scheme in the next Section in the downlink transmission, we assume the following definition [14]

$$\hat{\mathbf{H}}_n^k = \sqrt{1-e_d^2}\mathbf{H}_n^k + \mathbf{H}_e^k, \qquad (5)$$

where $\mathbf{H}_e^k \sim \mathcal{CN}\left(\mathbf{0},e_d^2\mathbf{I}\right)$ is an error component which is uncorrelated with $\mathbf{H}_n^k$.

B. Proposed pilot-less TDD massive MIMO

We propose a system in which each UL and DL data symbol of each user s_n^k is differentially encoded as

$$x_n^k = x_{n-1}^k s_n^k \quad 1\le n\le N, \qquad (6)$$

where N denotes the length of the differentially encoded data stream which only needs one reference symbol or pilot s_0^k

known at the BS (for the UL) and at each user (for the DL), and s_n^k belongs to a DMPSK constellation. Because the information is only encoded in the phase of unit modulus symbols, we know that $(x_n^k)^* x_{n-1}^k = s_n^k$.

The reception is performed via differential detection of two consecutive received signals (2) in time, as shown in [15],

$$z_n^k = \left(\mathbf{y}_{n-1}^k\right)^H \mathbf{y}_n^k = \left(\mathbf{h}_{n-1}^k\right)^H \mathbf{h}_n^k s_n^k + \eta_n^k, \tag{7}$$

with the noise terms (η_n^k) as

$$\eta_n^k = \left(\mathbf{h}_{n-1}^k\right)^H \left(\mathbf{x}_{n-1}^k\right)^H \boldsymbol{\nu}_n^k + \left(\boldsymbol{\nu}_{n-1}^k\right)^H \mathbf{h}_n^k \mathbf{x}_n^k + \left(\boldsymbol{\nu}_{n-1}^k\right)^H \boldsymbol{\nu}_n^k, \tag{8}$$

and the transmitted symbol for each user is estimated in the BS (for the uplink) and in each user terminal (for the downlink) according to [6] as

$$\hat{s}_n^k = \arg\min_{s_n^k} \left\{ \left| s_n^k - z_n^k \right|,\ s_n^k \in \mathfrak{M} \right\}, \tag{9}$$

where $\mathfrak{M}$ indicates the DMPSK constellation set, of size M, either for the uplink or downlink.

For the downlink, spatial multiplexing is utilized, via the use of MRT. Once the decision is taken on the UL NC data symbols, the channel can be estimated following [3, Eq. (21)]. First, we create the differential symbols from the detected ones as $\hat{x}_n^k = \hat{x}_{n-1}^k \hat{s}_n^k$, where $\hat{x}_0^k$ is known a priori in the receiver. Then, the LS criterion is applied to estimate the channel as

$$\hat{\mathbf{h}}_n^k = \frac{\mathbf{y}_n^k}{\hat{x}_n^k} = \mathbf{h}_n^k \frac{x_n^k}{\hat{x}_n^k} + \frac{\boldsymbol{\nu}_n^k}{\hat{x}_n^k}, \tag{10}$$

where there is no noise enhancement since $|\hat{x}_n^k| = 1$. A maximum length in the uplink stream of each user is defined so that the probability that an erroneous detection ($\hat{x}_n^k \neq x_n^k$) in some of the uplink NC symbols affecting the data detection process is controlled. For this purpose, K_p is defined as the length of the differential data in the UL used for channel estimation before a new stream starting with a reference pilot is inserted, as defined in [3]. With this estimated channel, the DL is precoded following (3).

The frequency of the needed channel estimation is related to the link adaptation problem [16], [17], which is not broadly discussed in this paper. A first approximation is to perform the estimation at least once every coherence time period, and this is supported by the numerical results in Sec. IV.

III. Analysis of Channel Estimation and Error Probability

In this section, we first characterize the effect of time correlated channels on the mean square error (MSE) of the channel estimation, which depends on the symbol error rate (SER) of the differentially encoded symbol in the UL [3] and the time variability, which was not analyzed in [3]. The main difference with the cited work in this regard comes from the fact that now we are considering channel time variability, so the estimated channel will be used in another time instant n' different from the one where the channel has been estimated. Second, we characterize the symbol-error-rate (SER) of the differentially encoded symbol of the DL of each user.

A. Mean square error of the channel estimation

An additional error term in the channel estimation is produced by a possible mismatch between transmitted and reconstructed differential symbols. This error was characterized in [3], without considering channel time variability, so following the same approach, we calculate the MSE of the channel estimation in the UL of a NC stream, now considering channel time variability. The estimated channel will be used in another time instant n' different from the one where the channel has been estimated. Hence, the channel estimation error is composed of two independent components as shown below

$$e_d^2 = \mathbb{E}\left\{ \left| \hat{\mathbf{h}}_n^k - \mathbf{h}_{n+d}^k \right|^2 \right\} = \sigma_{x,d}^2 + \sigma_{\nu,u}^2, \tag{11}$$

where $\sigma_{x,d}^2$ is the channel estimation error that comes from compensation and estimation in different time instants with a possible mismatch between transmitted and reconstructed differential symbol. We compute

$$\begin{aligned} \sigma_{x,d}^2 &= \mathbb{E}\left\{ \left| \mathbf{h}_n^k \frac{x_n^k}{\hat{x}_n^k} - \mathbf{h}_{n+d}^k \right|^2 \right\} = \mathbb{E}\left\{ \left| \mathbf{h}_n^k \frac{x_n^k}{\hat{x}_n^k} - \alpha_d \mathbf{h}_n^k - \mathbf{w}_d^h \right|^2 \right\} = \\ &= \mathbb{E}\left\{ \left| \mathbf{h}_n^k \right|^2 \left| \frac{x_n^k}{\hat{x}_n^k} - \alpha_d \right|^2 + \left| \mathbf{w}_d^h \right|^2 \right\} = \mathbb{E}\left\{ \left| \frac{x_n^k}{\hat{x}_n^k} - \alpha_d \right|^2 \right\} + \sigma_w^2, \end{aligned} \tag{12}$$

where $\mathbb{E}\left\{ \left| \mathbf{h}_n^k \right|^2 \right\} = 1$ since the channel is normalized, $\left| \mathbf{h}_n^k \right|^2$ is uncorrelated and independent of $\left| x_n^k / \hat{x}_n^k - \alpha_d - \mathbf{w}_d^h \right|^2$ and $x_n^k / \hat{x}_n^k - \alpha_d$ and $\mathbf{w}_d^h$ are uncorrelated and independent. In this case, $\sigma_w^2 = (1 - \alpha_d^2)$, since the power of (1) is normalized. Developing the first term of the last part of (12), we have

$$\begin{aligned} \mathbb{E}\left\{ \left| \frac{x_n^k}{\hat{x}_n^k} - \alpha_d \right|^2 \right\} &= \mathbb{E}\left\{ \left| \frac{x_n^k}{\hat{x}_n^k} \right|^2 - 2\alpha_d \Re\left\{ \frac{x_n^k}{\hat{x}_n^k} \right\} + \alpha_d^2 \right\} \\ &= 1 + \alpha_d^2 - 2\alpha_d \mathbb{E}\left\{ \cos\left(\angle(x_n^k) - \angle(\hat{x}_n^k) \right) \right\}. \end{aligned} \tag{13}$$

We define $\beta_n^k = \mathbb{E}\{\cos \angle(x_n^k) - \angle(\hat{x}_n^k)\}$ and by following the same approach as that of [3, Appendix B] (omitted here for the sake of conciseness), we can define

$$\beta_n^k \approx \begin{cases} 1, & P_n^k = 0 \\ \frac{1 - P_n^k - \left(1 - P_n^k\right)^{K_p}}{(K_p - 1) P_n^k}, & 0 < P_n^k \leq 1 \end{cases}, \tag{14}$$

that will serve as an upper bound (UB) for e_d^2, as stated in [3, Appendix B], and we can write $\sigma_{x,d}^2$ and e_d^2 as

$$\sigma_{x,d}^2 = 1 + \alpha_d^2 - 2\alpha_d \beta_n^k + 1 - \alpha_d^2 = 2(1 - \alpha_d \beta_n^k), \tag{15}$$

$$e_d^2 = \mathbb{E}\left\{ \left| \hat{\mathbf{h}}_n^k - \mathbf{h}_{n'}^k \right|^2 \right\} = 2(1 - \alpha_d \beta_n^k) + \sigma_{\nu,u}^2. \tag{16}$$

Analyzing (16), it can be seen that in case either α_d or β_n^k is zero, the channel error estimation is the highest, while both need to be 1 to avoid any increment in the channel estimation error, with respect to the classical estimation with pilots.

B. SER of the received symbol in the DL

In order to analyze the effect of an imperfectly estimated channel in the SER of the DL, we extend (4) following the definition in (5) as

$$[\mathbf{H}_n^k]_{u,\forall}[\mathbf{B}_n^k]_{\forall,u'} = [\mathbf{H}_n^k]_{u,\forall}\left(\sqrt{1-e_d^2}[\mathbf{H}_n^k]_{u',\forall} + \mathbf{H}_e^k\right)^H = \sqrt{1-e_d^2}[\mathbf{H}_n^k]_{u,\forall}[\mathbf{H}_n^k]_{u',\forall}^H + [\mathbf{H}_n^k]_{u,\forall}(\mathbf{H}_e^k)^H, \tag{17}$$

and $[\mathbf{x}_n^k]_u$ is just a phase rotation. Thus, following the properties of the product of normal variables [18] and the properties of VG distributions [19], [20], we have that

$$\sqrt{1-e_d^2}[\mathbf{H}_n^k]_{u,\forall}[\mathbf{H}_n^k]_{u,\forall}^H \sim \mathcal{N}\left(R\sqrt{1-e_d^2}, R(1-e_d^2)\right), \tag{18}$$

$$\sqrt{1-e_d^2}[\mathbf{H}_n^k]_{u,\forall}[\mathbf{H}_n^k]_{u',\forall}^H \sim \mathcal{CN}\left(0, R(1-e_d^2)\right), \tag{19}$$

$$[\mathbf{H}_n^k]_{u',\forall}([\mathbf{H}_e^k])^H \sim \mathcal{CN}\left(0, Re_d^2\right). \tag{20}$$

We define a useful component (UF) of the received signal

$$\mathrm{UF}_n^u = \left(\sqrt{1-e_d^2}[\mathbf{H}_n^k]_{u,\forall}[\mathbf{H}_n^k]_{u,\forall}^H\right)[\mathbf{x}_n^k]_u \tag{21}$$

and the rest of the terms of (4) are defined as non-useful (NUF) terms. Using the properties of the product of normal variables and the properties of VG distributions, the distribution of UF_n^u is defined as (18). The distribution of NUF_n^u is composed of the sum of several complex circularly symmetric Gaussian variables, including (20) from the desired user and $U-1$ of (19) and (20) from the interfering users over the desired user. Since $[\mathbf{x}_n^k]_{u'}$ only rotates the distributions, NUF_n^u is

$$\mathrm{NUF}_n^u \sim \mathcal{CN}\left(0, R(U+e_d^2-1)+\sigma_{n,k}^2\right). \tag{22}$$

Assuming $[\mathbf{x}_n^k]_u = 1$ without loss of generality[1], we have $\Re\{[\mathbf{y}_n^k]_u\} \sim \mu_\Re + \mathcal{N}\left(0, \sigma_\Re^2\right)$ and $\Im\{[\mathbf{y}_n^k]_u\} \sim \mathcal{N}\left(\mu_\Im, \sigma_\Im^2\right)$, so

$$\Re\{[\mathbf{y}_n^k]_u\} \sim R\sqrt{1-e_d^2} + \mathcal{N}\left(0, \frac{R(U-e_d^2+1)+\sigma_{n,k}^2}{2}\right) \tag{23}$$

$$\Im\{[\mathbf{y}_n^k]_u\} \sim \mathcal{N}\left(0, \frac{R(U+e_d^2-1)+\sigma_{n,k}^2}{2}\right). \tag{24}$$

The differential decoding performed in reception for the received signal at each user (7) results in the product of complex normally distributed variables, where in order to find the distribution of the received symbol, we have to consider the product of two complex variables $x = a+jb$ and $y = c+jd$, so the product $(x)^*y = (ac+bd)+j(ad-bc)$. Using again the properties of the product of normal variables and VG distributions, we have that

$$\Re\{[\mathbf{z}_n^k]_u\} = \Re\{[\mathbf{y}_{n-1}^k]_u\}\Re\{[\mathbf{y}_n^k]_u\} + \Im\{[\mathbf{y}_{n-1}^k]_u\}\Im\{[\mathbf{y}_n^k]_u\}, \tag{25}$$

[1]The error is computed for $[\mathbf{x}_n^k]_u = 1$ for simplicity but is the same for the rest of the symbols.

$$\Im\{[\mathbf{z}_n^k]_u\} = \Re\{[\mathbf{y}_{n-1}^k]_u\}\Im\{[\mathbf{y}_n^k]_u\} - \Im\{[\mathbf{y}_{n-1}^k]_u\}\Re\{[\mathbf{y}_n^k]_u\}. \tag{26}$$

Thus, the first term of (25) is composed of three terms as

$$\Re_1\{[\mathbf{z}_n^k]_u\} = R^2(1-e_d^2), \tag{27}$$

$$\Re_2\{[\mathbf{z}_n^k]_u\} \sim R\sqrt{1-e_d^2}\mathcal{N}\left(0, 2\sigma_\Re^2\right), \tag{28}$$

$$\Re_3\{[\mathbf{z}_n^k]_u\} \sim VG\left(1, 0, \sigma_\Re^2, 0\right), \tag{29}$$

while the second term of (25) is distributed as

$$\Im\{[\mathbf{y}_{n-1}^k]_u\}\Im\{[\mathbf{y}_n^k]_u\} \sim VG\left(1, 0, \sigma_\Im^2, 0\right). \tag{30}$$

Each term of (26) is composed of two terms, them being

$$\Im_1\{[\mathbf{z}_n^k]_u\} = R\sqrt{1-e_d^2}\mathcal{N}\left(0, \sigma_\Im^2\right), \tag{31}$$

$$\Im_2\{[\mathbf{z}_n^k]_u\} \sim VG\left(1, 0, \sigma_\Re\sigma_\Im, 0\right). \tag{32}$$

The VG distributions can be approximated to normal distributions as $VG\left(1,0,\sigma_\Re^2,0\right) \approx \mathcal{N}\left(0,\sigma_\Re^4\right)$, $VG\left(1,0,\sigma_\Im^2,0\right) \approx \mathcal{N}\left(0,\sigma_\Im^4\right)$ and $VG\left(1,0,\sigma_\Re\sigma_\Im,0\right) \approx \mathcal{N}\left(0,\sigma_\Re^2\sigma_\Im^2\right)$.

Summarizing, the distributions for $[\mathbf{s}_n^k]_u = 1$ are the following $\Re\{[\mathbf{z}_n^k]_u\} \sim \mathcal{N}\left(\mu_{\Re\{z_d\}}, \sigma_{\Re\{z_d\}}^2\right)$, $\Im\{[\mathbf{z}_n^k]_u\} \sim \mathcal{N}\left(\mu_{\Im\{z_d\}}, \sigma_{\Im\{z_d\}}^2\right)$ and defined as

$$\Re\{[\mathbf{z}_n^k]_u\} \sim \mathcal{N}\left(R^2(1-e_d^2), 2R^2(1-e_d^2)\sigma_\Re^2 + \sigma_\Re^4 + \sigma_\Im^4\right), \tag{33}$$

$$\Im\{[\mathbf{z}_n^k]_u\} \sim \mathcal{N}\left(0, 2\sigma_\Im^2\left(R^2(1-e_d^2)+\sigma_\Re^2\right)\right), \tag{34}$$

so the SER for the DL of user u can be computed following the approach in [3, Appendix A] as

$$P_n^k \approx 1 - \frac{\int_{-\pi/M}^{\pi/M}\int_0^\infty e^{-\left(\frac{r\cos(\gamma)-\mu_{\Re\{z_d\}}}{\sqrt{2}\sigma_{\Re\{z_d\}}}\right)^2} e^{-\left(\frac{r\sin(\gamma)-\mu_{\Im\{z_d\}}}{\sqrt{2}\sigma_{\Im\{z_d\}}}\right)^2} r\,dr\,d\gamma}{2\pi\sigma_{\Re\{z_d\}}\sigma_{\Im\{z_d\}}}. \tag{35}$$

IV. Numerical Results

In order to compare the classical coherent scheme and the proposed pilot-less scheme, we will use the SER. Since the overhead is different in them, for a fair comparison we will ensure the same spectral efficiency in both systems. For such purpose, different constellation sizes are used. For instance, in a TDD slot of 4 OFDM symbols, where 2 are used for the uplink and the other 2 for the downlink, half of the OFDM symbols should be used for pilots (one for uplink and another one for downlink). In this case, for a fair comparison, the size of the constellation used in the coherent scheme should be the square of that used in the non-coherent scheme. Furthermore, the proposed pilot-less scheme can work with a coherence time of 3 symbol slots (1.5 TDD slots), as shown in Fig. 1, while the classical scheme is more restrictive and needs at least 4 symbols.

In this section, we first validate the theoretical analysis previously shown in Sec. III and afterwards we compare the

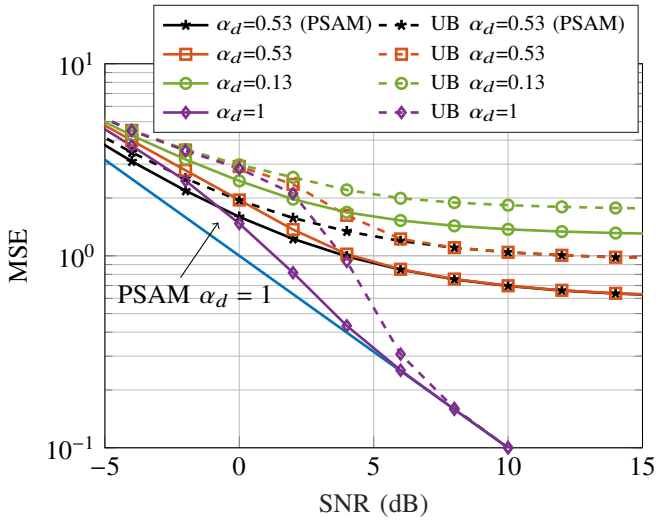

Fig. 2. MSE of channel estimation for $M_{UL} = 16$ and $R = 100$. Continuous line corresponds to the Monte Carlo simulation while dashed line corresponds to the theoretical upper bound. Blue line represents the PSAM without channel time variability, which is the best case.

classical scheme with the proposed scheme in a deployment scenario with ideal conditions. Unless otherwise stated, $R = 100$, $M_u = M_d = 4$. The SNR in the simulations is defined as the inverse of the noise power.

A. Corroboration of the theoretical analysis

It can be seen in Fig. 2, that the theoretical analysis for the channel estimation fits well with the Monte Carlo simulations, with the theoretical results being an upper bound of the simulations. The lower bound is given by the performance of the PSAM with a pilot in every coherence time, that is, without any degradation due to time variability (at the expense of a large pilot overhead). It can be observed that a correlation α_d caused by time difference between the estimated channel and the real one results in a MSE floor, caused by the time variability of the channel. When there is no time variability, the MSE is affected only by the error probability in the detection of (7) via (14). Last but not least, the case of PSAM with time variability is shown, were it can be seen that it is below the proposed channel estimation with channel time variability, but is equal to it for high SNR.

In Fig. 3, we can see the histogram and the theoretical PDFs for the real and imaginary parts of the received symbol when $[\mathbf{z}_n^{k_d}]_u = 1$. A very good agreement can be seen for both cases, but for the real part, since more approximations were made, the agreement is slightly worse. This fact directly affects the accuracy of the SER performance for the DL of the proposed scheme, which is shown in Fig. 4. The theoretical SER (labeled 'TH' in the legend) is very similar to that of the Monte Carlo simulations even though there are little discrepancies, due to the left tail of the distribution of the real part.

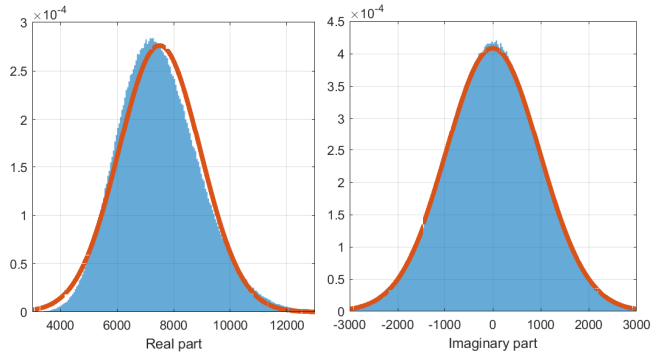

Fig. 3. Monte Carlo histogram (blue) versus theoretical PDF (red) of the real (left) and imaginary (right) part of $[\mathbf{s}_n^{k_d}]_u = 1$ for $R = 100$, $M_{UL} = 4$ and $e_d = 0.5$.

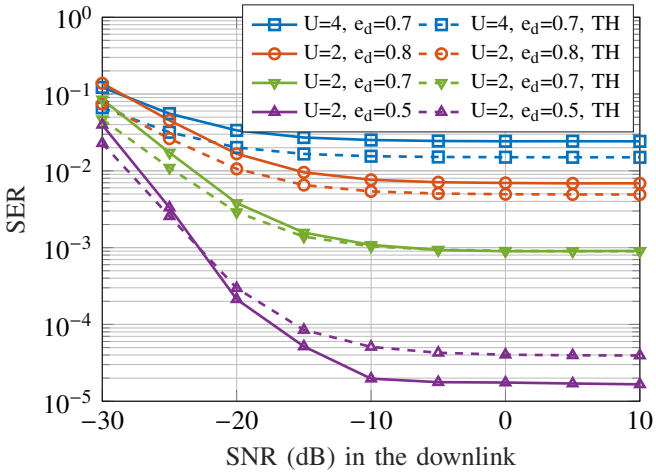

Fig. 4. DL performance with different U and e_d for the proposed scheme for $M_{DL} = 8$ and $R = 200$. Continuous line corresponds to the Monte Carlo simulation while dashed line corresponds to the theoretical upper bound.

B. Comparison with a classical pilot assisted system

The classical and the proposed schemes are compared with Monte Carlo simulations. The coherence time n_c is defined in number of OFDM symbols of the TDD scheme, according to the frame shown in Fig. 1, and the values of 2, 3 and 10 are considered in Fig. 5. The number of symbols in a TDD slot period for the UL τ_u and the DL τ_d are set to 1 and 2 to see the dependence with this parameter, and the constellation size in the DL M_d to 4 and 16. An SNR sufficiently large to avoid any errors is used in the UL and $U = 2$ are considered. The classical (coherent) and the proposed (non-coherent) schemes are referred in Fig. 5 as C and N, respectively.

It can be observed how, for large n_c (10 in Fig. 5), the scheme C is approximately the same as the scheme N, which reinforces the validity of the proposed scheme even for large n_c. For very fast varying channels ($n_c = 2$), the latter outperforms the former for the same spectral efficiency. More concretely, for $\tau_u = \tau_d = 2$, the scheme C utilizes one symbol pilot while the scheme N does not. Thus, the scheme C transmits a 16-QAM while the scheme N transmits a QPSK.

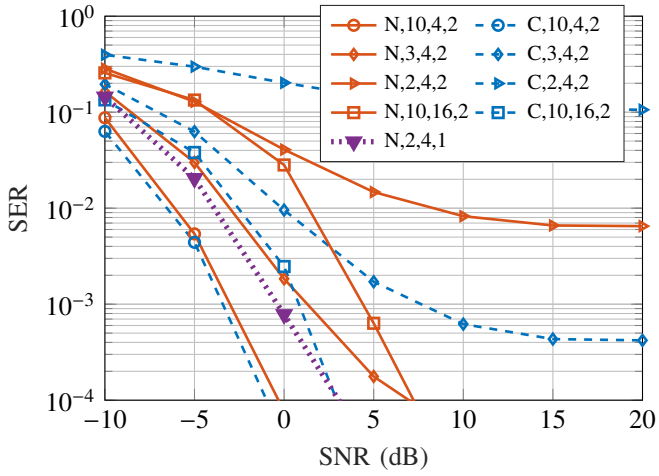

Fig. 5. Comparison between classical (C, dashed) and proposed (N, continuous) schemes, labelled from left to right with n_c, M_d, $\tau_u = \tau_d$, for $R = 100$.

For both schemes transmitting a QPSK (scheme C with a lower spectral efficiency than scheme N) and for $n_c = 3$, scheme N outperforms scheme C, which reinforces our proposal. In fact, scheme N can go down to $\tau_u = \tau_d = 1$, while C cannot since then, only pilots would be sent both in UL and DL. This is an advantage for the proposed scheme N which can work in extremely fast varying channels. As a clarification example, in a 5G system with carrier frequency $f_c = 3.6$GHz, carrier spacing $\Delta_f = 30$KHz at a speed of 500km/h, the coherence time is $n_c = 2.8$ OFDM symbols.

V. Conclusions

In this work, we propose a pilot-less massive MIMO TDD scheme where the UL data is detected via non-coherent processing, the channel is estimated using the data received in the UL and the data of several users in the DL is precoded to spatially separate them. To avoid the use of pilots in both the UL and DL, differential PSK is utilized. The MSE of the channel estimation for time-varying channels and the performance of the DL using precoding and differential encoding are characterized and validated with numerical simulations. The proposed pilot-less TDD scheme is compared with its coherent PSAM counterpart and it is shown that it largely outperforms it when the coherence time is very small, which is exemplified to be realistic in the context of 5G.

References

[1] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser mimo systems," *IEEE Transactions on Communications*, vol. 61, no. 4, pp. 1436–1449, 2013.

[2] O. Elijah, C. Y. Leow, T. A. Rahman, S. Nunoo, and S. Z. Iliya, "A Comprehensive Survey of Pilot Contamination in Massive MIMO—5G System," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 905–923, 2016.

[3] M. J. Lopez-Morales, K. Chen-Hu, and A. Garcia-Armada, "Differential Data-aided Channel Estimation for Up-link Massive SIMO-OFDM," *IEEE Open Journal of the Comms. Society*, vol. 1, pp. 976–989, 2020.

[4] M. Chowdhury, A. Manolakos, and A. J. Goldsmith, "Coherent versus noncoherent massive simo systems: Which has better performance?" in *2015 IEEE International Conf. on Comms. (ICC)*, 2015, pp. 1691–1696.

[5] S. Malkowsky, J. Vieira, L. Liu, P. Harris, K. Nieman, N. Kundargi, I. C. Wong, F. Tufvesson, V. Öwall, and O. Edfors, "The World's First Real-Time Testbed for Massive MIMO: Design, Implementation, and Validation," *IEEE Access*, vol. 5, pp. 9073–9088, 2017.

[6] A. G. Armada and L. Hanzo, "A Non-Coherent Multi-User Large Scale SIMO System Relaying on M-ary DPSK," in *2015 IEEE International Conf. on Comms. (ICC)*, June 2015, pp. 2517–2522.

[7] K. Chen-Hu, Y. Liu, and A. G. Armada, "Non-coherent massive mimo-ofdm down-link based on differential modulation," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 11 281–11 294, 2020.

[8] T. Hwang, C. Yang, G. Wu, S. Li, and G. Ye Li, "OFDM and Its Wireless Applications: A Survey," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1673–1694, May 2009.

[9] M. J. L. Morales, K. Chen-Hu, and A. G. Armada, "Effect of spatial correlation on the performance of non-coherent massive mimo based on dmpsk," in *2021 IEEE Global Comms. Conf.*, 2021, pp. 1–6.

[10] T. S. Rappaport *et al.*, *Wireless communications: principles and practice*. Prentice Hall PTR, 1996, vol. 2.

[11] K. T. Truong and R. W. Heath, "Effects of channel aging in massive MIMO systems," *Journal of Communications and Networks*, vol. 15, no. 4, pp. 338–351, 2013.

[12] M. Vu and A. Paulraj, "On the capacity of mimo wireless channels with dynamic csit," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 7, pp. 1269–1283, 2007.

[13] J. K. Cavers, "An Analysis of Pilot Symbol Assisted Modulation for Rayleigh Fading Channels (Mobile Radio)," *IEEE Trans. Veh. Technol.*, vol. 40, no. 4, pp. 686–693, 1991.

[14] D. Mi, M. Dianati, L. Zhang, S. Muhaidat, and R. Tafazolli, "Massive mimo performance with imperfect channel reciprocity and channel estimation error," *IEEE Transactions on Communications*, vol. 65, no. 9, pp. 3734–3749, 2017.

[15] K. Chen-Hu, Y. Liu, and A. G. Armada, "Non-coherent massive MIMO-OFDM for communications in high mobility scenarios," *ITU Journal on Fut. and Evolv. Tech.*, vol. 1, 2020.

[16] L. Toni and A. Conti, "Does Fast Adaptive Modulation Always Outperform Slow Adaptive Modulation?" *IEEE Transactions on Wireless Communications*, vol. 10, no. 5, pp. 1504–1513, 2011.

[17] A. Conti, M. Z. Win, and M. Chiani, "Slow Adaptive M-QAM With Diversity in Fast Fading and Shadowing," *IEEE Transactions on Communications*, vol. 55, no. 5, pp. 895–905, 2007.

[18] G. W. Bohrnstedt and A. S. Goldberger, "On the exact covariance of products of random variables," *Journal of the American Statistical Association*, vol. 64, no. 328, pp. 1439–1442, 1969.

[19] R. Gaunt, "Variance-gamma approximation via stein's method," *Electron. J. Probab.*, vol. 19, p. 33 pp., 2014.

[20] R. E. Gaunt, "A note on the distribution of the product of zero-mean correlated normal random variables," *Statistica Neerlandica*, vol. 73, no. 2, pp. 176–179, 2019.