Phase Noise Cancelation Based on Polarization Modulation for Massive MIMO-OFDM Systems

Yao Nie,1,2 Chunyan Feng,2 Huan Wu,1,2 Fangfang Liu,2 Caili Guo,2 and Maosheng Fu1

1School of Electronics and Information Engineering, West Anhui University, Lu’an 237012, China
2Beijing Key Laboratory of Network System Architecture and Convergence, Beijing Laboratory of Advanced Information Networks, Beijing University of Posts and Telecommunications, Beijing 100876, China

Correspondence should be addressed to Huan Wu; wuhuan@bupt.edu.cn

Received 13 May 2018; Revised 22 September 2018; Accepted 30 September 2018; Published 13 March 2019

Copyright © 2019 Yao Nie et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In massive multiple-input multiple-output (MIMO) systems, phase noise introduced by oscillators can cause severe performance loss. It leads to common phase error and intercarrier interference in massive MIMO-OFDM uplink. To solve the issue, a novel phase noise cancelation scheme based on polarization modulation for the massive MIMO-OFDM system is proposed. We first introduce the polarization modulation (PM) exploited in massive MIMO-OFDM uplink. Then, by exploiting the zero-forcing detection, we analyze the asymptotically ICI and the distribution of the transformed noise under different XPD values. Furthermore, we demonstrate that phase noise can be asymptotically canceled and only the transformed additive white Gaussian noise exists as the number of antennas at the base station is very large. Moreover, we derive the instantaneous signal-to-noise ratio (SNR) on each subcarrier and analyze the ergodic capacity. To increase the ergodic capacity performance further, a joint modulation scheme combining the PM and 2PSK is proposed and the ergodic capacity performance of the joint modulation is discussed. The simulation results show that the proposed scheme can effectively mitigate phase noise and achieve a higher ergodic capacity.

1. Introduction

Massive multiple-input multiple-output (MIMO) is a promising technology recently for next-generation wireless communication systems and for green communications [1–7] due to its potential to provide significant spectral and energy efficiency [8–11]. By exploiting a large number of antennas at the transmit or receive side, simple detection schemes, such as maximum ratio combining (MRC), zero-forcing (ZF), and minimum mean-square error (MMSE), are asymptotically optimum and can increase the network throughput while reducing the required transmitting power [8]. However, the practical performance of massive MIMO is severely limited by hardware impairments [12], especially the phase noise (PN) introduced by the nonideal local oscillators (LO).

Usually, massive MIMO is used with orthogonal frequency-division multiplexing (OFDM) to counteract the effect of frequency-selective fading. However, massive MIMO-OFDM is sensitive to phase noise. The effect of phase noise is larger in the high-frequency band due to mmWave and THz bands used to achieve bit rates of 10 Gbps and greater for 5G. The impairment of phase noise is characterized by the common phase error (CPE) term and the intercarrier interference (ICI) term [13, 14]. Thus, the estimated channel during the pilot period is different from the channel during the data transmission period due to phase noise which can severely cause the ergodic capacity performance degradation significantly [15].

There have been some PN cancelation schemes for the MIMO-OFDM system, such as Bayesian-based [16] and maximum likelihood- (ML-) based [17]. However, the computational complexity of these schemes increases with the number of antennas, which is not suitable for the massive MIMO-OFDM system where the number of antennas is very large [18]. The effect of phase noise has been investigated for massive MIMO uplink in [19]. Different performance metrics, such as SINR and achievable rate, have been derived for MRC [20] and ZF [21] schemes. In [15, 22], the impact
of phase noise on massive MIMO-OFDM uplink system performance has been analyzed for the single-user (SU) and multiuser (MU) cases. The MU massive MIMO-OFDM downlink system with ZF and maximal ratio transmission (MRT) precoding has been compared [23, 24]. Recently, some phase noise cancelation schemes have been proposed for massive MIMO uplink systems [25, 26] and for downlink systems [27].

In wireless communication, the polarization modulation (PM) is a novel modulation technique to improve the performance, such as BER [28] and energy efficiency [29–32]. Moreover, it has been found in [33] that the polarization state (PS) of the signal is insensitive to phase noise. Therefore, the PM, which transmits information through PSs, has the potential to cancel the impact of phase noise. We have proposed the differential polarization shift keying, which is a special PM, to cancel phase noise for the OFDM system in our prior work [34]. In a realistic wireless channel, the polarized signals are distorted by depolarization, such as cross-polarization discrimination (XPD) [35, 36], polarization-dependent loss (PDL), and polarization mode dispersion (PMD) [37]; some papers propose the channel compensation methods to improve the PM performance [29–32, 38–40]. In this paper, we extend the phase noise cancelation scheme for a larger number of antenna scenarios under XPD effect.

In this paper, we proposed a novel phase noise cancelation scheme based on the PM for the massive MIMO-OFDM uplink system. The proposed scheme cancels phase noise since the PSs used by the PM are insensitive to phase noise. The ZF detector is used to receive the PM signal and reduce the XPD effect. Then, with the help of the polarization signal processing, we transform the multiplicative phase noise into the additive noise and demonstrate that the additive noise multiplied by ZF factor is with Gaussian statistical characteristics under perfect CSI. After analyzing asymptotical ICI power, we investigate the phase noise cancelation performance in terms of the instantaneous SNR and the ergodic capacity. Furthermore, the joint modulated scheme combining the PM and 2PSK is proposed to improve the ergodic capacity. The ergodic capacity of the joint modulated scheme is also discussed.

This rest of paper is organized as follows. The system model is introduced in Section 2. Then, the proposed phase noise cancelation scheme is derived in Section 3. In Section 4, we investigate the phase noise cancelation performance and analyze the SNR and ergodic capacity. Moreover, the joint modulated scheme is proposed to improve the ergodic capacity performance in this section. The numerical results are provided in Section 5. Finally, conclusions are drawn in Section 6.

Throughout this paper, $\mathcal{C}$,$\mathcal{N}$ denotes the complex Gaussian distribution, $\mathcal{N}$ denotes the Gaussian distribution, $\cdot^*$ denotes the complex conjugate, $\cdot^H$ denotes the conjugate transpose, $\cdot^{-T}$ denotes the inverse transpose, $\cdot^\dagger$ denotes the matrix transpose, and $\mathbb{E}(\cdot)$ denotes the expectation operator. $\odot$ denotes the Hadamard product. $\text{var}(\mathbf{A})$ and $\text{cov}(\mathbf{A})$ denote variance and covariance of the matrix $\mathbf{A}$, respectively.

2. System Model

We consider a single-user MIMO-OFDM system with a BS and a UE in the uplink. At the receiver side, the BS is equipped with $M$ antennas. In massive MIMO, it is common to have a single high-frequency oscillator (possibly at millimeter waves) driving the up- or downconversion for all the antennas. Thus, we consider the common oscillator at the BS or the UE, respectively. OFDM is with over $N$ subcarriers and spacing $1/T_f$. An OFDM frame consists of the training symbols (the pilot symbols) and uplink data symbols in nonoverlapping time periods, and the duration of a frame is shorter than the channel coherence time. In the training period, a pilot OFDM symbol is sent from the UE to the BS to estimate the channel state information. Then, the information symbols present in all the subcarriers are known to the BS.

The channel coefficient for the $n$th subcarrier and between the UE and the $m$th BS antenna is denoted by $h_{mn}^r = \sqrt{\beta} g_{mn}^r$ in the $i$th time instant, where $\beta$ and $g_{mn}^r$ are the slow-fading and the fast-fading coefficients, respectively. $\beta$ is caused by the geometric attenuation and the shadow fading and fixed during many frames so that it is independent of $m$ and $i$. $g_{mn}^r$ is assumed to be independent and identically distributed (i.i.d.) complex Gaussian random variables with zero mean and unit variance, $g_{mn}^r \sim \mathcal{CN}(0, \sigma_n^2)$.

Phase noise is introduced by the oscillator impairment and as a multiplicative term. Free-running oscillators are assumed; hence, the phase noise realizations follow Wiener processes [13, 41, 42]. Let $\phi_i^f$ be the transmitter phase noise and $\phi_i^r$ be the receiver phase noise in the $i$th time instant, where we assume that $\phi_i^f$ and $\phi_i^r$ are independent since the users and the BS use different oscillators:

$$\phi_i^f = \phi_{i-1}^f + \Delta_i^f,$$
$$\phi_i^r = \phi_{i-1}^r + \Delta_i^r,$$

where $\Delta_i^f$ and $\Delta_i^r$ are independent and identically distributed real Gaussian random variables, $\Delta_i^f \sim (0, \sigma_f^2)$ and $\Delta_i^r \sim (0, \sigma_r^2)$.

The total phase noise increment variance is $\sigma^2 = \sigma_f^2 + \sigma_r^2$.

According to [15], in the SU MIMO-OFDM uplink system, the frequency domain representation of the received signal
signal vector at the BS after FFT in the $i$th time instant and the $n$th subcarrier can be expressed as

$$Y_{i,n} = \sqrt{P_{\text{CPE}}} \Theta_{i,n} H_{i,n} S_{i,n} + \sqrt{P} \sum_{l=0, l \neq n}^{N} \Theta_{i,l} H_{i,l} S_{i,l} + W_{i,n},$$

where $\Theta_{i,n} = \text{diag}\{\theta_{i,n}, \ldots, \theta_{i,n}\}$ is the $M \times M$ phase noise matrix after FFT with $\theta_{i,n} = (1/N) \sum_{m=0}^{N-1} e^{j\theta_{m}+j\phi_{m}}$ being the $n$th frequency coefficient of the $N$ discrete phase noise samples affecting an OFDM symbol. $P$ denotes the transmit power per subcarrier and the additive white Gaussian noise (AWGN) affecting the subcarriers is denoted by the $M \times 1$ vector $W_{i,n} = [W_{i,n}^{1}, \ldots, W_{i,n}^{M}]^T$ with $W_{i,n}^{m} \sim \mathcal{C}(0, \sigma_{w}^2)$. It can be seen that the received signals are affected by CPE and ICI. CPE rotates the constellation of all the received signals equally, and ICI spreads the constellation.

### 3. Proposed Scheme

In this section, the PM-based phase noise cancelation scheme is proposed for massive MIMO-OFDM systems to cancel phase noise since the PS is insensitive to phase noise. The PM scheme for massive MIMO-OFDM uplink and massive MIMO-OFDM dual-polarized channel is derived, and the pilot and data transmission periods are investigated.

#### 3.1. Polarization Modulation and Dual-Polarized Massive MIMO Channel

Figure 1 shows the polarization modulation (PM) scheme in the massive MIMO-OFDM system. As shown in the figure, the UE exploits a pair of orthogonal dual-polarized antennas (ODPA), polarized horizontally and vertically. And at the receiver side, the BS is equipped with $M/2$ horizontally polarized antennas and $M/2$ vertically polarized antennas.

The polarization modulation (PM) exploits the PS of a radiated electromagnetic wave as the information-bearing parameter, which can be realized by controlling the ratio of the power and the difference of the phase of two components in the baseband [33–30]. The ODPAs, equipped both at the transmitter and receiver, are to transmit and receive the PM signal.

The PM unit is shown in Figure 2. The transmitting sequence is $S_{i,n}$, whose elements are all “1.” On the lower branch, the source information is firstly mapping to polarization constellation $\{Q_v\}_{v=1}^{V}$ which is the constellation points of $V$ order PM by the polarization mapping constellation unit (PCMU). Meanwhile, $S_{i}$ divides into the two same components. Then, the PCMU exploits the power division unit (PDU) and phase shifting unit (PSU) to modify the amplitude ratio and phase difference of the two components, respectively. These two components are the orthogonal components of the PM signal. Then, the orthogonal component outputs from PSU, denoted as $S_{H,n}$ and $S_{V,n}$, are modulated to a OFDM symbol by a OFDM modulator.

Let PM signal for the $i$th time instant and the $n$th subcarrier be $S_{i,n}$. Using Jones vector [33], the PM signal can be written as

$$S_{i,n} = \begin{bmatrix} S_{H,i,n} \\ S_{V,i,n} \end{bmatrix},$$

where $S_{H,i,n}$ and $S_{V,i,n}$ are a pair of orthogonal components of $S_{i,n}$. The amplitude ratio and phase difference are

$$\psi_{i,n} = \begin{bmatrix} S_{H,i,n} \\ S_{V,i,n} \end{bmatrix},$$

$$\varphi_{i,n} = \arctan\left(S_{V,i,n}\right) - \arctan\left(S_{H,i,n}\right).$$

The PDU and PSU are adjusted to modify the amplitude and phase of the orthogonal components according to $\psi_{i,n}$ and $\varphi_{i,n}$. After the above processing, the PM signal is sent by the horizontally and vertically polarized antennas.

Considering the ODPAs deployed in massive MIMO systems, the channel propagation between BS and UE is modeled by large-scale fading, small-scale fast fading, and depolarization effect of the ODPAs. For a dual-polarized MIMO channel, some depolarization factors are considered: cross-polarization discrimination (XPD), copolar correlation, polarization mismatch, etc. [44]. To facilitate the analysis, the XPD effect caused by wireless channel is only considered, which is denoted as the antennas’ ability to distinguish the orthogonal components of the PM signals [45]. Let $\chi$ be the inverse of XPD, $0 \leq \chi \leq 1$. The polarization fading effects can be modeled by the following matrix:

$$X = \begin{bmatrix} 1 & \sqrt{\chi} \cdot 1 \times (M/2) \times (K/2) \\ \sqrt{\chi} \cdot 1 \times (M/2) \times (K/2) & 1 \times (M/2) \times (K/2) \end{bmatrix},$$

where the matrix $1_{a \times b}$ is an $a \times b$ matrix with all ones, $M$ and $K$ are the antenna’s numbers at the BS and UE, respectively. Then, the polarized uplink channel matrix for the $n$th subcarrier and in the $i$th time instant is denoted by

$$H_{i,n} = X \circ \tilde{h}_{i,n},$$

where $\circ$ stands for the Hadamard product. $H_{i,n}$ is only the massive MIMO channel modeled by large-scale fading and small-scale fast fading, which is given by

$$H_{i,n} = \begin{bmatrix} h_{i,n}^{(1)T} \\ h_{i,n}^{(2)T} \end{bmatrix} \in \mathbb{C}^{M \times 2},$$

where $k = 1, 2$ which is the antenna’s numbers of the UE, and the channel vector $h_{i,n}^{(k)T}$ is denoted by

$$h_{i,n}^{(k)} = \sqrt{\beta_{k}} g_{i,n}^{(k)},$$

where $\beta_{k}$ and $g_{i,n}^{(k)}$ are the large-scale fading and small-scale fast fading coefficients, respectively. $g_{i,n}^{(k)}$ is assumed to be
independent and identically distributed (i.i.d.) complex Gaussian random variables with zero mean and unit variance, \( g_{k,m} \sim \mathcal{CN}(0, \sigma^2) \).

Usually, massive MIMO-OFDM systems first send the pilot in training period to estimate the channel state information (CSI). Due to the duration of a frame including the training and data transmission which is less than the coherence time, the estimated CSI can be used to suppress the channel effect. Thus, the proposed scheme can be divided into two stages: pilot transmission and data transmission.

3.2. Pilot Transmission and Channel Estimation. As in most of wireless communication systems, pilot symbols are first sent to estimate the channel state information (CSI). Assume that the user transmits one pilot OFDM symbol at time instant \( i = 0 \) (the training period) which is used to estimate the CSI between BS antennas and the user for all subcarriers.

Following (3), at the 0th subcarrier, the received signal (pilot) can be expressed as

\[
Y_{0,0} = \sqrt{P} \Theta_{0,0} H_{0,0} S_{0,0} + \sqrt{P} \sum_{n=1}^{N-1} \Theta_{0,n} H_{0,n} S_{0,n} + W_{0,0} = \sqrt{P} \Theta_{0,0} H_{0,0} S_{0,0} + W'_{0,0}.
\]  

Note that the transmitted and received signals are polarized signals. After the ZF detection, we will estimate the CSI by the methods in [15]. In this paper, the channel is estimated by \( \hat{G}_{0,0} = Y_{0,0} = \sqrt{P} \Theta_{0,0} H_{0,0} \) by the maximum likelihood estimation at the BS, due to that ICI adds \( w'_{0,0} \) as a complex Gaussian distribution. Thus, we have

\[
G_{0,0} = \left( G_{0,0}^H \hat{G}_{0,0} \right)^{-1} G_{0,0}^H = \left( \sqrt{P} H_{0,0}^H \Theta_{0,0}^H \sqrt{P} H_{0,0} \right)^{-1} \sqrt{P} H_{0,0} \Theta_{0,0}^H, \tag{11}
\]

where \( G_{0,0}' = (1/\sqrt{P}) (H_{0,0}^H H_{0,0})^{-1} H_{0,0}^H \) is the ZF detection without phase noise.

**Lemma 1.** \( G_{0,0} \) can be written as

\[
G_{0,0} = \frac{1}{\sqrt{P}} G_{0,0}' \Theta_{0,0}^H = \frac{1}{\sqrt{P}} \left( H_{0,0}^H \Theta_{0,0} \right)^{-1} H_{0,0}^H \Theta_{0,0}^H, \tag{12}
\]

**Proof.** For the system model, the common oscillator is exploited at both the BS and UE. The phase noise matrix \( \Theta_{0,0} \) is the diagonal matrix and has the same diagonal elements, \( \Theta_{0,0}^H \Theta_{0,0} = \theta_0^H \theta_0 \cdot \text{diag} \{1, \ldots, 1\} \). Therefore, \( \Theta_{0,0}^H \Theta_{0,0} \) can be replaced by \( \theta_0^H \theta_0 \):

\[
\theta_0^H \theta_0 = \sum_{l=0}^{N-1} e^{-j\phi(lT)} \sum_{n=0}^{N} e^{j\phi(\tau T)}.
\]

Since phase noise is a Wiener process, we know

\[
\phi_K = \phi_k + \sum_{l=k+1}^{K} \Delta_l, \tag{14}
\]

where \( \phi_k \) is the \( K \) time sample phase noise. Then, we have

\[
e^{j\phi} = e^{j\phi_k} + \sum_{l=k+1}^{K} \Delta_l. \tag{15}
\]
Thus, $\theta_{0,0}$ and $\theta_{0,0}^{H}$ can be rewritten as

$$\theta_{0,0} = \frac{1}{N} \sum_{n=0}^{N-1} e^{j\phi_n} = \frac{1}{N} \left( e^{j\phi_0} + e^{j(\phi_0 + \Delta_1)} + \ldots + e^{j \sum_{l=1}^{N-1} \Delta_l} \right)$$

and

$$\theta_{0,0}^{H} = \frac{1}{N} e^{-j\phi_0} \left( 1 + e^{-j\phi_1} + \ldots + e^{-j \sum_{l=1}^{N-1} \Delta_l} \right).$$ (16)

According to (16) and (17), we obtain

$$\theta_{0,0}^{H} \theta_{0,0} = \frac{1}{N^2} \sum_{n=0}^{N-1} e^{-j\phi_n} \cdot \frac{1}{N} \sum_{n=0}^{N-1} e^{j\phi_n}$$

$$= \frac{1}{N^2} \left( 1 + e^{j\phi_1} + \ldots + e^{\sum_{l=1}^{N-1} \Delta_l} \right) \left( 1 + e^{-j\phi_1} + \ldots + e^{-\sum_{l=1}^{N-1} \Delta_l} \right)$$

$$= \frac{1}{N^2} \left( N + e^{j\phi_1} + \ldots + e^{\sum_{l=1}^{N-1} \Delta_l} \right) \left( N + e^{-j\phi_1} + \ldots + e^{-\sum_{l=1}^{N-1} \Delta_l} \right)$$

$$= \frac{1}{N^2} \sum_{l=1}^{N-1} \sum_{j=0}^{N-1} \Delta_l \Delta_j + \sum_{j=1}^{N-1} \sum_{l=1}^{N-2} \Delta_l \Delta_j + \sum_{l=1}^{N-1} \sum_{j=1}^{N-2} \Delta_l \Delta_j.$$. (18)

Since the values of the phase noise increments are small, we have the following approximation:

$$\frac{1}{N} \sum_{l=k+1}^{N} \Delta_l \approx 1 + j \sum_{l=k+1}^{K} \Delta_l,$$ (19)

Accordingly, (17) can be rewritten as

$$\theta_{0,0}^{H} \theta_{0,0} = \frac{1}{N^2} \sum_{n=0}^{N-1} e^{j\phi_n} \cdot \frac{1}{N^2} \sum_{n=0}^{N-1} e^{-j\phi_n} = \frac{1}{N^2} \left( N^2 + j\Delta_1 + \ldots + j \sum_{l=1}^{N-1} \Delta_l - j \sum_{l=1}^{N-1} \Delta_l + j \sum_{l=1}^{N-1} \Delta_l \right) + \ldots + \frac{1}{N^2} \left( j \sum_{l=1}^{N-1} \Delta_l - j \sum_{l=1}^{N-1} \Delta_l \right) = 1.$$. (20)

Thus, (12) follows that $\Theta_{0,0}^{H} \Theta_{0,0} = \text{diag}(1, \ldots, 1)$. Then, the ZF detection is obtained in the presence of phase noise, which is used to compensate for the channel distortion during the data transmission period.

3.3. Data Transmission. We consider the data transmission from the UE at the $D$th time instant after the training period, where $D$ is less than the coherence time of the channel, implying that $G_{D,0} = G_{0,0}$. After polarization matching receiving, the received PM signal is given by

$$Y_{D,0} = \sqrt{P} \Theta_{D,0}^H H_{D,0} S_{D,0} + \sqrt{P} \sum_{n=1}^{N-1} \Theta_{D,n} H_{D,n} S_{D,n} + W_{D,0}.$$ (21)

Upon performing ZF on the signal received in the 0th subcarrier and the $D$th time instant, the transmitted PM signal is detected after the ZF detection which can be written as

$$\bar{S}_{D,0} = G_{0,0} Y_{D,0} = \sqrt{P} G_{0,0} \Theta_{D,0}^H H_{D,0} S_{D,0} + G_{0,0} W_{D,0}.$$ (22)

According to (12), (22) also can be rewritten as

$$\bar{S}_{D,0} = \theta_{0,0}^{H} \theta_{0,0} G_{0,0} \Theta_{D,0}^H H_{D,0} S_{D,0} + \theta_{0,0}^{H} \theta_{0,0} G_{0,0} \sqrt{P} \sum_{n=1}^{N-1} \Theta_{D,n} H_{D,n} S_{D,n} + \theta_{0,0}^{H} G_{0,0} W_{D,0}.$$ (23)

where $\gamma = \theta_{0,0}^{H} \theta_{0,0}$. Since $H_{D,0} = H_{0,0}$ due to that an OFDM system time duration is less than the channel coherence time:

$$I = \left( H_{0,0}^H H_{0,0} \right)^{-1} H_{0,0}^H H_{D,0} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$ (24)

According to [33], the polarization is insensitive to phase noise subject to the Gaussian noise. Moreover, the ZF detector output signal $\bar{S}_{D,0}$ consists of the desired signal, ICI, and the additive noise. Therefore, the ICI and distribution of the additive noise are investigated to verify the proposed scheme performance.
4. Performance Analysis

In this section, we use random matrix theory to analyze the received signal in (23) as $M \to \infty$. Importantly, the ICI and AWGN of the received PM signal after ZF detection are asymptotically analyzed to derive the phase noise cancelation performance. Furthermore, we analyze the instantaneous SNR and discuss the ergodic capacity performance.

4.1. Phase Noise Cancelation Performance. After ZF detection, we have the received signal in the presence of phase noise in the frequency domain. The CPE and ICI are analyzed asymptotically to verify the phase noise cancelation.

**Proposition 1.** According to (22), the ICI after ZF detection is given by

$$S_{ICL} = (PH_0^H H_0^H)^{-1} P H_0^H \Theta_0^H \sum_{n=1}^{N-1} \Theta_{D,n}^H H_{D,n}^H S_{D,n}$$

$$= \left( \frac{1}{M} H_0^H H_0^H \right)^{-1} \frac{1}{M} H_0^H \Theta_0^H \sum_{n=1}^{N-1} \Theta_{D,n}^H H_{D,n}^H S_{D,n}. \tag{25}$$

$S_{ICL}$ tends asymptotically to 0.

**Proof.** The large-scale fading coefficient $\sqrt{\beta_k}$ caused by the geometric attenuation and the shadow fading, is independent of BS’s antenna and fixed during many frames, so that it is set $\beta_k = 1$. Moreover, the variance of the small-scale fading coefficient $\sigma_{\epsilon_k}^2$ is equal to 1, $\sigma^2 = 1$. Furthermore, the XPD effect is also considered due to that the PM signal is distorted by depolarization in the wireless channel. According to (6) and (7), the BS has $M/2$ horizontal and $M/2$ vertical polarized antennas so that the XPD effect modifies the variance of $H_{\text{x,r}}$.

For the convenience of analysis, we first analyze the channel model without XPD effect. Since $H_{0,0} = \{h_{mk}\}_{M \times 2}$ and $h_{mk} \sim \mathcal{CN}(0, 1)$ without considering the XPD effect, as $M \to \infty$, by the law of large numbers, we have

$$\frac{1}{M} H_0^H H_0^H \xrightarrow{a.s.} \left( \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right), \tag{26}$$

where $\xrightarrow{a.s.}$ denotes the almost sure convergence. When considering the CO scenario, $\Theta_0^H$ can be expressed as $\theta_0^H$ which can be put at the head of the right hand term of (25). Thus, we only analyze asymptotically ICI after ZF detection.

According to [15], $\sum_{n=1}^{N-1} \Theta_{D,n}^H H_{D,n}^H S_{D,n}$ is approximately Gaussian distributed whose elements are i.i.d. random variables with zero mean and unit variance $\sigma^2_{\text{ICI}}$. Also, by the law of large numbers, we have

$$\frac{1}{M} \bar{H}_{0,0}^H \xrightarrow{a.s.} 0. \tag{27}$$

Thus, after ZF detection, the ICI is asymptotically given by

$$S_{ICL} = \theta_0^H \left( \frac{1}{M} \bar{H}_{0,0}^H \right)^{-1} \frac{1}{M} H_0^H \sum_{n=1}^{N-1} \Theta_{D,n}^H H_{D,n}^H S_{D,n} = 0. \tag{28}$$

The derivation of (26) and (27) is outlined in Appendix A. The XPD effect is not considered while the ICI tends to 0. The ICI power comparison before and after ZF detection is shown in Figure 3. With different variances and antenna numbers, the ICI power after ZF detection decreases faster than that without ZF detection, which means that the ZF detection in the proposed scheme can cancel the ICI.

The polarization state which is insensitive to phase noise is based on the additive noise subject to Gaussian distribution [33]. However, after ZF detection, the additive noise is the product of the AWGN and ZF. We derive the distribution of the noise after ZF detection in (23).

**Proposition 2.** Let $W_{D,0}^\prime$ be the additive noise after ZF detection without phase noise $\theta_{0,0}$. $W_{D,0}^\prime = G_{0,0} W_{D,0}^\prime$, according to (23). $W_{D,0}^\prime$ is the random variable with Gaussian distribution.

**Proof.** From (23), the additive noise $W_{D,0}^\prime$ is given by

$$W_{D,0}^\prime = G_{0,0} W_{D,0} = \left( \frac{1}{M} H_0^H H_0^H \right)^{-1} \frac{1}{\sqrt{M}} H_0^H W_{D,0}. \tag{29}$$

The same as Proposition 1, the XPD effect is dropped. $W_{D,0}$ is the AWGN vector, $W_{D,0}^\prime \sim \mathcal{CN}(0, 1)$, and independent of $H_{0,0}$. Thus, as $M \to \infty$ and by the central limit theorem, we have

$$\frac{1}{\sqrt{M}} H_{0,0} W_{D,0} \xrightarrow{d} \mathcal{CN} \left( \left( \begin{array}{c} 0 \\ 0 \end{array} \right), \left( \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right) \right), \tag{30}$$

where $\xrightarrow{d}$ denotes convergence in distribution. The derivation of (30) is outlined in Appendix B. According to (26), $W_{D,0}^\prime$ can be expressed as

$$W_{D,0}^\prime \xrightarrow{d} \left( \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right)^{-1} \mathcal{CN} \left( \left( \begin{array}{c} 0 \\ \end{array} \right), \left( \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right) \right) \tag{31}$$

where $\mathcal{CN}$ denotes the zero mean and unit variance Gaussian distribution.
\[ W_{D,0}' \overset{d}{\rightarrow} \text{CN} \left( \begin{bmatrix} 0 & \left( \frac{2}{1 + \chi} \right) \\ 0 & 0 \end{bmatrix} \right) \]  

As the range of \( \chi \) is \( 0 \leq \chi \leq 1 \), the variance of \( H_{D,0} \) is modified from 1 to 4 which increases the variance of the AWGN.

Based on the additive noise after ZF detection which is subject to Gaussian distribution, the phase noise effect on the desired signal can be canceled. According to (28) and (32), the multiplicative phase noises in (23) can be transformed to additive as

\[ \hat{S}_{D,0} = YI S_{D,0} + U\theta_{0,0}' G_{0,0} W_{D,0}' = S_{D,0} + U\theta_{0,0}' W_{D,0}' \]  

where \( U\theta_{0,0}' \) is a real unitary stochastic matrix as

\[ U\theta_{0,0}' = \begin{bmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{bmatrix}, \]

where \( \phi = (1/N^2) \sum_{n=0}^{N-1} e^{-j\phi_n} \sum_{n=0}^{N-1} e^{j\phi_{n+1}} \). In the following, we will demonstrate that the multiplicative phase noises above can be transformed to the new additive Gaussian noise [46].

**Lemma 2.** Let \( U = U\theta_{0,0}' \): \( U \) is a stochastic unitary matrix.

**Proof.** \( U^H U = (U\theta_{0,0}')^H U\theta_{0,0}' = \theta_{0,0}' U\theta_{0,0}' \). According to (20), we have \( \theta_{0,0}' \theta_{0,0}' = 1 \). Also, \( U \) is the real unitary stochastic matrix. Therefore, we have \( U^H U = I \).

Indeed, the real unitary stochastic matrix \( U \) rotates the polarization constellation as a rigid body, which means that both the constellation structure and the power of the PM signal remain the same. The ZF-detected PM signal components in (33) can be rewritten according to Euler’s formula as

\[ \hat{S}_{D,0} = \begin{bmatrix} |S_{D,0,x}'| \left( \cos (\phi_{D,0,x} + \tilde{\phi}) + j \sin (\phi_{D,0,x} + \tilde{\phi}) \right) \\ |S_{D,0,y}'| \left( \cos (\phi_{D,0,y} + \tilde{\phi}) + j \sin (\phi_{D,0,y} + \tilde{\phi}) \right) \\ W_{D,0,xc}' + j W_{D,0,xc}' \\ W_{D,0,yc}' + j W_{D,0,yc}' \end{bmatrix}, \]

where \( |S_{D,0,x}'| \) and \( |S_{D,0,y}'| \) denote the magnitudes for polarization \( x \) and \( y \) and \( \phi_{D,0,x} \) and \( \phi_{D,0,y} \) are the phases of \( S_{D,0,x} \) and \( S_{D,0,y} \), respectively. \( W_{D,0,xc}' \) and \( W_{D,0,yc}' \) are the additive noise components. Note that polarization \( x \) and \( y \) are orthogonal.

\begin{align*}
S_1 &= \left( |S_{D,0,x}'|^2 - |S_{D,0,y}'|^2 \right) = \left( W_{D,0,x}' \right)^2 - \left( W_{D,0,y}' \right)^2 \\
&+ 2 |S_{D,0,x}'| \left[ \cos \phi_{D,0,x} \sin \phi_{D,0,x} \right] U \left[ W_{D,0,xc}' \right] \\
&- 2 |S_{D,0,y}'| \left[ \cos \phi_{D,0,y} \sin \phi_{D,0,y} \right] U \left[ W_{D,0,yc}' \right] \\
S_2 &= 2 |S_{D,0,x}'| \left| S_{D,0,y}' \right| \cos (\phi_{D,0,y} - \phi_{D,0,x}) \\
&+ 2 \left( W_{D,0,xc}' + W_{D,0,yc}' + W_{D,0,yc}' \right) \\
&+ 2 |S_{D,0,x}'| \left[ \cos \phi_{D,0,y} \sin \phi_{D,0,y} \right] U \left[ W_{D,0,yc}' \right] \\
&+ 2 |S_{D,0,y}'| \left[ \cos \phi_{D,0,y} \sin \phi_{D,0,y} \right] U \left[ W_{D,0,yc}' \right] \\
S_3 &= 2 |S_{D,0,x}'| \left| S_{D,0,y}' \right| \sin (\phi_{D,0,y} - \phi_{D,0,x}) \\
&+ 2 \left( W_{D,0,xc}' + W_{D,0,yc}' + W_{D,0,yc}' \right) \\
&+ 2 |S_{D,0,x}'| \left[ \cos \phi_{D,0,y} \sin \phi_{D,0,y} \right] U^T \left[ W_{D,0,yc}' \right] \\
&+ 2 |S_{D,0,y}'| \left[ \cos \phi_{D,0,y} \sin \phi_{D,0,y} \right] U^T \left[ W_{D,0,yc}' \right] \\
&+ 2 |S_{D,0,y}'| \left[ \cos \phi_{D,0,y} \sin \phi_{D,0,y} \right] U^T \left[ W_{D,0,xc}' \right],
\end{align*}
Stokes vector is another polarization presentation method, where $S_1$, $S_2$, and $S_3$ are the parameters of the Stokes vector [47, 48]. Stokes vector is readily connected to Poincare sphere, which is used to represent the polarization constellation. Moreover, the Stokes vector can readily represent magnitudes, phase of the PM signal, and phase noise as in (36).

Thus, we can get (33) which indicates that multiplicative phase noise has been transformed to the AWGN, while the received PM signal after ZF detection is not alerted. Then, the question is how the real unitary stochastic matrix $U$ with phase noise affects the AWGN.

The proposed scheme exploits the PM signal, which can be decomposed to a pair of orthogonal components by using Jones vector. Thus, the new additive noise $W_{D,0}''$, comprising the real unitary matrix $U$ and the AWGN, can be denoted as

$$ W_{D,0}'' = \begin{bmatrix} W_{D,0,x}'' \\ W_{D,0,y}'' \end{bmatrix} = U \begin{bmatrix} W_{D,0,x}' \\ W_{D,0,y}' \end{bmatrix}, \quad (37) $$

where $W_{D,0,x}''$ and $W_{D,0,y}''$ are the mutually orthogonal components of the new additive noise. According to [33], it is proved that the phase noise does not alter the statistical characteristics of the AWGN.

Proposition 3. For the transformation in (37), it is given that $W_{D,0,x}''$ and $W_{D,0,y}''$ are two independent identically distributed as stationary complex Gaussian random processes with zero mean and subject to

$$ \mathbb{E}\left\{ W_{0,x}'(t + \tau)W_{0,x}'(t) \right\} = \mathbb{E}\left\{ W_{0,y}'(t + \tau)W_{0,y}'(t) \right\} = 0, \quad \forall \tau. \quad (38) $$

In the sequel, the distribution of the AWGN does not vary with time so that we will drop the time index $m$ for notational convenience. The processes $W_{D,0,x}''$ and $W_{D,0,y}''$ are still independent zero-mean complex Gaussian random processes with the same distributions as $W_{D,0,x}'$ and $W_{D,0,y}'$.

Proof. The joint probability density function (PDF) of the random process $W_{D,0,x}'$ and $W_{D,0,y}'$ can be

$$ f(z) = \frac{1}{\pi^p \text{Det}\mathbf{D}} \exp(-z^H \mathbf{D}^{-1} z), \quad (39) $$

where $z = [W_{0,x}', \ldots, W_{0,xq}', W_{0,y}', \ldots, W_{0,yq}']$, $q$ is the variable order, and $\mathbf{D} = \mathbb{E}\{zz^H\}$. Then, the transformed new additive noise is given by

$$ \mathbb{E}\left\{ \begin{bmatrix} W_{0,x}'(t + \tau) \\ W_{0,y}'(t + \tau) \end{bmatrix} \begin{bmatrix} W_{0,x}''(t) \\ W_{0,y}''(t) \end{bmatrix} \right\} = \mathbb{E}\left\{ \begin{bmatrix} U(t + \tau) \\ U(t + \tau) \end{bmatrix} \begin{bmatrix} W_{0,x}'(t) \\ W_{0,y}'(t) \end{bmatrix} \right\} \mathbf{U}^H(t) \right\} = U(t + \tau)0U^H(t) = 0. \quad (40) $$

Thus, the PDF of $W_{D,0,x}''$ and $W_{D,0,y}''$ are the same as that of $W_{D,0,x}'$ and $W_{D,0,y}'$ and independent of $U$. With respect to $U$, $W_{D,0}''$ leaves the PDF unchanged.

Using the results from (29), (32), and Proposition 3, we know that the multiplicative phase noise has been transformed to the additive noise. Moreover, we prove that the new additive noise has asymptotically the same distribution as the AWGN before the transformation. Phase noise only changes the polarization base of the AWGN leading to the distribution of the new additive noise that does not alter. Therefore, the Stokes vector in (36) is not affected by phase noise which is the same as that without phase noise. Furthermore, we prove that the ICI tends asymptotically to 0. Based on these conclusions, the proposed scheme can cancel phase noise for massive MIMO-OFDM uplink.

Figure 4 depicts the polarization constellation comparison of the proposed scheme with and without phase noise (PN) under the $M = 500$, $\sigma^2 = \sigma^2 = 0.1$, and SNR = 15 dB with different $\chi$. It can be seen that in Figures 4(a) and 4(b), the constellation points have the same distribution which means that phase noise cannot distort the PM signal. Moreover, the difference between Figures 4(a) and 4(b) is caused by the different XPD values. According to (32), the XPD effect increases the AWGN related to $\chi$. Therefore, the scattered area of constellation in Figure 4(b) is more than that in Figure 4(a).

4.2. The Instantaneous SNR and Ergodic Capacity. The massive MIMO-OFDM uplink performance is related to the SNR, where the interference consists of the ICI. The SNR is defined as the ratio of the power of the desired signal and the additive noise plus ICI. Moreover, as $M \to \infty$, we analyze the instantaneous SNR, which can be written as

$$ \lim_{M \to \infty} \text{SNR} = \lim_{M \to \infty} \frac{\mathbb{E}\{|S_{D,0}^2|\}}{\mathbb{E}\{|S_{IC}^2| + |S_{0,0}^H W_{0,0}^T|^2\}}. \quad (41) $$

According to (28), the power of ICI tends asymptotically to 0, $\lim_{M \to \infty} \mathbb{E}\{|S_{IC}^2|\} = 0$. Using (32), (33), and Proposition 3, we can see that the power of the new additive noise is $\lim_{M \to \infty} \mathbb{E}\{|S_{D,0}^2|\} = (2/(1 + x))^2 \sigma_w^2$, where $\sigma_w^2$ denotes the power of the AWGN before ZF detection. However, in other phase noise investigation for massive MIMO systems, for example, in [15], the SNR analysis must consider the phase noise power. In the proposed scheme, we
do not consider this distortion due to the PM signal exploited. Therefore, the instantaneous SNR is given by
\[
\lim_{M \to \infty} \text{SNR} = \frac{\sigma_s^2}{\sigma_w^2} \frac{1}{1 + x^2}.
\]
where \(\sigma_s^2 = \mathbb{E}(\hat{S}_{D,0}^2)\), which is the power of the design single.

Finally, the ergodic capacity per subcarrier for the proposed scheme in (33) is evaluated as
\[
C_{\text{erg}} = \log_2(1 + \text{SNR}).
\]

4.3. Complexity of the Proposed Scheme. In this section, we consider the computational complexities. For single-user MIMO systems, the phase noise cancelation schemes usually suffer high complexity leading to these schemes which may not be extended to massive MIMO-OFDM uplink systems especially when the BS antenna number \(M\) is large. For example, the complexities of the phase noise cancelation schemes in [18] are related to \(M^3\). In [25], the author proposed a simple phase noise cancelation scheme for the massive MIMO-OFDM system. The complexity of the phase noise cancelation scheme in [25] is \((2N_D)MK + N_DK\) multiplications accompanied by \(O(K^2M)\) based on LS channel estimation, where \(N_D\) is the frame length during the data transmission period.

In contrast, the complexity of the proposed scheme is low due to exploiting the phase noise insensitive property of the PS without high complexity phase noise estimation algorithm. The polarization modulation and demodulation are only the modulated technologies which do not need to be calculated. Therefore, calculating \(G_{0,0}\) in the channel estimation period requires only about \((2N_D)MK + N_DK\) based on LS channel estimation. Compared with the scheme in [25], the complexity of the proposed scheme is lower due to not calculating phase noise estimation, \(O(K^2M)\). The low complexity leads that the proposed scheme is suitable for massive MIMO-OFDM uplink systems, in which \(M\) is large.

5. The Joint Modulation Scheme

Note that the proposed scheme exploits ODPAs to transmit and receive the orthogonal components of a PM signal. Two autonomous single-antenna user (TU) schemes for massive MIMO-OFDM uplink systems also exploit two polarized antennas, and each user exploits one antenna with the horizontally or vertically polarized. Thus, the TU scheme for massive MIMO-OFDM uplink systems has the same architecture as the proposed scheme in this paper. Compared with the TU scheme, the proposed scheme only has the half ergodic capacity due to that the ODPAs are exploited to transmit the orthogonal components of the PM signal in the absence of phase noise. Fortunately, the PM is realized by controlling the relative amplitude and phase and does not depend on the original phases or amplitudes of the two components. Therefore, we can combine the PM and other amplitude or/and phase modulated technology, PSK, QAM, etc., which is named the joint modulation, to improve the ergodic capacity.

We modify the PM modulation unit in Figure 1 to realize the joint modulation as shown in Figure 5. In this section, 2PSK is exploited for an example to evaluate the joint modulation scheme. At the transmitter, 2PSK signal \(S_t[n]\) is generated by a 2PSK modulator and then sent to PDU and PSU to modify the amplitude and phase of the upper (or lower) components. In this paper, we denote that the amplitude and phase of the lower component do...
not modify, which is actually the 2PSK signal $S_i[n]$ written as follows:

$$S_i[n] = e^{i\theta_i[n]},$$  \hspace{1cm} (44)$$

where $\theta_i[n]$ is the phase information after 2PSK. Then, the PSU and PDU modify the amplitude and phase of the upper component to control the PS. Finally, the output signal of PSU conveys information including both the PS related to PM and the phase related to 2PSK. According to (2), the joint modulated signal for the $n$th subcarrier is given by

$$S[n] = S_{PSK} \begin{bmatrix} \cos \psi[n]e^{i\phi[n]} \\ \sin \psi[n]e^{i\phi[n]} \end{bmatrix},$$  \hspace{1cm} (45)$$

where $S_{PSK}$ is the normalization factor which guarantees that the power of $S[n]$ is equal to 1. To decide the information from 2PSK and PM, the receiver needs to add the 2PSK received branch. The received unit of the joint modulation scheme is shown in Figure 6. In Figure 6, the PM modulator is exploited to demodulate the PM signal and the component of the PM signal on the lower branch received by the horizontal antenna is demodulated by the 2PSK modulator. Based on the joint demodulation unit, after ZF detection, the Stokes vectors as mentioned in (36) and then remaps the constellation on the Poincare sphere to decide the PS. Meanwhile, the component on the lower branch is sent to the 2PSK modulator to decide the phase information. Finally, the information from PM and 2PSK is demodulated.

The PM and PSK are independent in the joint modulation scheme so that the ergodic capacity of the PM and 2PSK can be calculated separately. Thus, the ergodic capacity of the joint modulation scheme is the sum of that of the PM and 2PSK. Due to that the PM is insensitive to phase noise, the ergodic capacity of PM is given in (43) in the presence of phase noise. However, 2PSK is distorted by phase noise; the asymptotical ergodic capacity of 2PSK can be written as

$$C_{PSK} = \log_2(1 + \text{SNR}_{PSK}),$$  \hspace{1cm} (46)$$

where SNR_{PSK} is the instantaneous SNR of 2PSK, which is given by

$$\lim_{M \to \infty} \text{SNR}_{PSK} = \frac{\sigma_{PSK}^2}{\sigma_{PSK,\text{IC}}^2 + \sigma_{PSK,\text{AW}}^2},$$  \hspace{1cm} (47)$$

where $\sigma_{PSK}^2$, $\sigma_{PSK,\text{IC}}^2$ and $\sigma_{PSK,\text{AW}}^2$ are the power of the 2PSK desired signal and ICI caused by phase noise and AWGN, respectively. Note that the ICI of 2PSK cannot be canceled so that the ergodic capacity of 2PSK is decreased and only half of the TU scheme exploited 2PSK in the presence of phase noise. Therefore, the total ergodic capacity of the joint modulation scheme is

$$C_{\text{total}} = C_{\text{erg}} + C_{PSK}.$$  \hspace{1cm} (48)$$

6. Results

The performance of the proposed scheme is verified via computer simulations in this section. The ergodic capacity is found to discuss the system performance distorted by phase noise for the massive MIMO-OFDM uplink system. Moreover, the proposed scheme is simulated in comparison with the single-user (SU) and two-user (TU) schemes with the single antenna scheme in the presence of phase noise. Further, the XPD effect is also simulated to show the ergodic capacity performance with different $\chi$. The ergodic capacity of the joint modulation scheme and the TU scheme is also compared. The numbers of the OFDM subcarriers are 64. We consider that the total transmitted power is 1; $P = 1/M$, $M = 300$, $\sigma = \sigma_r + \sigma_p$, $\sigma_r = \sigma_p$, and $\sigma^2 = 1$. The modulation formats are BPM and BPSK.

The ergodic capacity performance comparison between the proposed scheme and the signal-user (SU) scheme for the massive MIMO-OFDM uplink system is shown in
Figure 7 when \( M = 300 \) and \( \chi = 1 \). The phase noise variance of LO at the transmitter and receiver is the same, \( \sigma_r = \sigma_t \). Phase noise is the dominant effect leading to the estimated CSI of the pilot period and the CSI of the data transmission period which are different while SNR grows larger. Therefore, it can be seen that the ergodic capacity of the SU scheme is decreased in the presence of phase noise. In contrast, the proposed scheme can achieve the phase noise-free performance even according to (28), (32), and Proposition 3.

The TU scheme for the massive MIMO-OFDM uplink system is compared with the proposed scheme in terms of ergodic capacity performance as shown in Figure 8. The signal of these two schemes is the polarized signal which is transmitted under the same channel model in (7). We consider \( M = 300 \) and \( \chi = 1 \); the ergodic capacity performance comparison between the proposed scheme and the TU scheme for the massive MIMO-OFDM uplink system is illustrated in Figure 8. The TU scheme has two users which transmit two data information in uplink for phase noise-free case. Therefore, the ergodic capacity of the TU scheme is approximately twice that of the proposed scheme which has only one user. Similar to the SU scheme, phase noise also distorts the ergodic capacity performance of the TU scheme. As shown in Figure 8, phase noise with larger \( \sigma_i \) (\( \sigma_i = \sigma_t = 0.2 \)) causes more serious ergodic capacity performance degradation, especially. The reason is that larger phase noise leads to greater difference of the CSI between the pilot and data transmission periods. So, the proposed scheme has better performance under a larger phase noise scenario. The realization of affordable massive MIMO requires the use of a large number of inexpensive and potentially low-quality LO leading to larger phase noise. Our proposed scheme is suitable for deployment in the massive MIMO-OFDM system with larger phase noise.

According to (32), the variance of \( \mathbf{H}_{0,0} \) is modified from 1 to 4 as the range of \( \chi \) is \( 0 \leq \chi \leq 1 \). Therefore, the variance of the AWGN is changed and determined by the XPD effect which increases and alters the ergodic capacity as seen in (43). Figure 9 depicts the ergodic capacity comparison of the proposed scheme and the TU scheme under different \( \chi \) while considering \( M = 300 \) and \( \sigma_i = \sigma_t = 0.2 \). It can be seen that the ergodic capacity performance of these two schemes is better as \( \chi \) tending to 1 due to the variance of the AWGN is greater than that under \( \chi \) tending to 0.1.

The ergodic capacity performance comparison between the joint modulation scheme and the TU scheme for massive MIMO-OFDM uplink is illustrated in Figure 10. The ergodic capacity of the joint modulation scheme is nearly equal to the sum of the ergodic of SU-PSK scheme and the proposed scheme according to (48). And the ergodic capacity performance of the joint modulation scheme is significantly better than that of the TU scheme due to the polarization modulation unit which can cancel the effect caused by phase noise.

7. Conclusion

In this paper, the phase noise cancelation scheme based on polarization modulation was proposed for the massive MIMO-OFDM uplink system. With respect to the ergodic capacity, the proposed scheme achieves the phase noise-free performance, which is significantly better than the SU scheme. Especially, in the large phase noise scenario, the proposed scheme has the obvious performance.
increase compared with the TU scheme. Moreover, the joint modulation scheme is proposed to increase the ergodic capacity performance further. Therefore, the proposed scheme is suitable for massive MIMO-OFDM uplink systems with large phase noise, in which the number of BS antennas $M$ is large.

Appendix

A. Proof of (26) based on the law of large numbers

We follow some results of random vectors in [49] to justify Proposition 1. We have a pair of mutually independent vectors $a \triangleq [a_1 \ldots a_n]^T$ and $b \triangleq [b_1 \ldots b_n]^T$ of dimension $n \times 1$. And the elements of these vectors are i.i.d. zero-mean random variables with $E(a_i^2) = \sigma_a^2$ and $E(b_i^2) = \sigma_b^2$, $i = 1, \ldots, n$. As $n \to \infty$, by the law of large numbers, we have

$$\frac{1}{n} a^H a \xrightarrow{a.s.} \sigma_a^2,$$

and

$$\frac{1}{n} a^H b \xrightarrow{a.s.} 0.$$

In (26), the $\mathbf{H}_{0,0}$ is a $M \times 2$ matrix, which has two-column random vectors. These two columns are mutually independent vectors with i.i.d. random variables, $h_{mk} \sim \mathcal{CN}(0, 1)$. According to (49), we have

$$\frac{1}{M} \mathbf{H}_{0,0}^H \mathbf{H}_{0,0} = \frac{1}{M} \begin{pmatrix} \sum_{m=1}^{M} h_{m,1}^2 & \sum_{m=1}^{M} h_{m,1} h_{m,2} \\ \sum_{m=1}^{M} h_{m,1} h_{m,2} & \sum_{m=1}^{M} h_{m,2}^2 \end{pmatrix} \xrightarrow{a.s.} \begin{pmatrix} E(h_{m,1}^2) & E(h_{m,1} h_{m,2}) \\ E(h_{m,1} h_{m,2}) & E(h_{m,2}^2) \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

By modeling ICI as additive Gaussian noise independent of $\mathbf{H}_{0,0}$, (27) is obtained by (50).

B. Proof of (30) based on the central limit theorem

Also, from the assumption of the vectors $a$ and $b$ in Appendix A and based on the Lindeberg-Lévy central limit theorem, we have

$$\frac{1}{\sqrt{n}} a^H b \xrightarrow{d} \mathcal{CN}(0, \sigma_a^2 \sigma_b^2).$$

Each column of $\mathbf{H}_{0,0}$ is mutually independent and independent of $\mathbf{W}_{D,0}$. According to (52), we have
\[
\frac{1}{\sqrt{M}} \mathbf{R}_{m,0}^{\text{d}} W_{0,0} = \frac{1}{\sqrt{M}} \begin{pmatrix}
\sum_{n=1}^{M} h_{m,n} W_m \\
\sum_{n=1}^{M} h_{m,2,n} W_m \\
\var(h_{m,n}) W_m \\
\cov(h_{m,n}, h_{m,2,n}) W_m \\
\cov(h_{m,n}, h_{m,2,n}) W_m \\
\var(h_{m,n}) W_m
\end{pmatrix} \xrightarrow{d} \mathcal{CN} \left( \begin{pmatrix}
\mathbb{E}(h_{m,n} W_m) \\
\mathbb{E}(h_{m,2,n} W_m) \\
\cov(h_{m,n}, h_{m,2,n}) W_m \\
\cov(h_{m,n}, h_{m,2,n}) W_m \\
\cov(h_{m,n}, h_{m,2,n}) W_m \\
\var(h_{m,n} W_m)
\end{pmatrix} \right).
\]

(53)

Data Availability

The (simulation for phase noise cancelation) data used to support the findings of this study have been deposited in the Mendeley repository (10.17632/rkk63htc2s1).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported by the National Science Foundation of China Grant no. 61501050 and no. 61271177, Science Foundation of West Anhui University no. WXZR201630, and Major Program of the National Natural Science Foundation of Anhui Universities no. KJ2015ZD44.

References


