IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 6, 2019, accepted June 22, 2019, date of publication June 26, 2019, date of current version July 17, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2925321

Differential Downlink Transmission in Massive MU-MIMO Systems

FAHAD ALSIFIANY[1], AISSA IKHLEF[2], (Senior Member, IEEE), MAHMOUD ALAGELI[3], AND JONATHON CHAMBERS[4], (Fellow, IEEE)

[1]School of Engineering, Intelligent Sensing and Communications, Newcastle University, Newcastle upon Tyne NE1 7RU, U.K.
[2]Department of Engineering, Durham University, Durham DH1 3LE, U.K.
[3]Faculty of Engineering, Garaboulli, Elmergib University, Garaboulli, Libya
[4]Department of Engineering, University of Leicester, Leicester LE1 7RU, U.K.

Corresponding author: Fahad Alsifiany (f.a.n.alsifiany2@newcastle.ac.uk)

This work was supported by the Engineering and Physical Sciences Research Council (EPSPRC), U.K., under Grant EP/R006377/1.

ABSTRACT In this paper, a differential downlink transmission scheme is proposed for a massive multiple-input multiple-output (MIMO) system without explicit channel estimation. In particular, we use a downlink precoding technique combined with a different encoding scheme to simplify the overall system complexity. A novel precoder is proposed, which, with a large number of transmit antennas, can effectively precancel the multiple access interference (MAI) for each user, thus enhancing the system performance. Maximizing the worst case signal-to-interference-plus-noise ratio (SINR) is used to optimize the precoder for the users in which full power space profile (PSP) knowledge is available to the base station (BS). In addition, we provide two suboptimal solutions based on the matched and the orthogonality approach of the PSP to separate the data streams of multiple users. The decision feedback differential detection (DFDD) technique is employed to further improve the performance. The proposed schemes eliminate the MAI, enhance system performance, and achieve a simple low complexity transmission scheme. Moreover, transmission overheads are significantly reduced using the proposed scheme, since it avoids explicit channel estimation at both ends. The Monte Carlo simulation results demonstrate the effectiveness of the proposed schemes.

INDEX TERMS Massive (MIMO), differential modulation, multiple symbol differential detection, precoding.

I. INTRODUCTION

Multiple-input multiple-output (MIMO) technology helps in improving wireless multiple access and can be used to increase the spectral efficiency and improve the link reliability at low power operation [1], [2]. With multiple transmit antennas at the base station (BS), the system can spatially multiplex multiple data streams for multiple users at the same frequency and time. The spatial multiplexing property becomes more effective as the number of antennas becomes large where the system is referred to as massive MIMO [3]. Such properties render a massive MIMO architecture as an important part of many wireless communications standards, such as LTE and 5G networks.

Much of the research on MIMO downlink transmission designs assumes perfect channel state information (CSI) at the transmitter. The availability of CSI at both ends makes it possible for the system to eliminate the multiple access interference (MAI) between users. However, due to various reasons, such as pilot contamination from training sequence reuse in massive MIMO, perfect CSI estimation is unattainable [4]. In [5], the authors proposed a framework that uses the block diagonalisation method to cancel MAI between users. The proposed method provides a substantial gain in terms of spatial diversity with a low decoding complexity. However, for the decoding process, each user still needs to know the channel in order to decode the information signal coherently.

The authors in [6] proposed a downlink spreading scheme combined with differential detection (DD) to eliminate the need of estimating the CSI at the BS and users. The scheme provides both low complexity transceivers and good performance. However, for large number of users, the proposed scheme in [6] does not provide a comprehensive high rate differential scheme in a downlink scenario due to the long length of the spreading code. In [7], the authors proposed a

The associate editor coordinating the review of this manuscript and approving it for publication was Pietro Savazzi.

full rate downlink algebraic transmission scheme combined with a differential space-time scheme. The proposed scheme provides a full rate full diversity system and does not require any knowledge of the CSI to separate the data streams of multiple users. In this approach, however, the BS typically employs only a few antennas, and thus the corresponding improvement in spectral efficiency and system simplicity is still relatively modest.

In order to improve the spectral efficiency and to simplify the required signal processing, a massive MIMO downlink system is employed [8], [9], where the BS is equipped with a very large number of transmit antennas. In practice, the demodulation reference signals (DM-RS) are used to support channel estimation and data demodulation. In DM-RS, the estimation of the channel for coherent detection is often obtained by training and tracking, e.g. using reference signals (RS), or pilots. However, it is not always feasible to use training-based schemes, with systems that have a large number of antennas. As the number of transmit antennas grows large such as in the case of massive MIMO, the channel estimation process, system overheads, latency, and power consumption will grow proportionately [10]. Discussion of DM-RS improvements are ongoing in 3GPP release 15 standardization [11], hence, it is natural to adopt differential modulation with massive MIMO to reduce the overhead and latency of DM-RS.

A well-established method to enhance DD is multiple symbol differential detection (MSDD). The authors in [12] point out a 3dB performance improvement simply by demodulating the received symbols jointly as a block, instead of one at a time using the MSDD detection technique. The authors of [13], [14] developed MSDD detection for the uplink MIMO system in ultra-wideband (UWB) systems. Essentially, the authors in [14] adopted decision feedback differential detection (DFDD) for a massive MIMO system, as this approach improves the performance of MSDD. However, the multiuser transmission scheme in [14] suffers from severe MAI without a proper precoding design scheme. Furthermore, prior research on MSDD and MAI cancellation has mainly been focused on uplink transmission, where cancellation was implemented at the BS receiver, and therefore complexity was not a significant concern [13], [14]. For downlink transmission, however, interference cancellation at end users increases receiver complexity, and for this reason, we account for interference cancellation at the BS instead of receivers. In particular, in order to have low complexity receivers, it is assumed that the transmitted signals are precoded at the BS.

In this paper, we therefore propose a differential MIMO downlink transmission framework, in which a BS is equipped with a massive antenna array that precodes transmit signals to separate the data streams of multiple users. In particular, to achieve a low-complexity differential massive MIMO system, a novel downlink precoding design is proposed by employing knowledge of the power space profile (PSP) of users. It is assumed that the PSP for each user is estimated at the BS, since it can tolerate more complexity compared

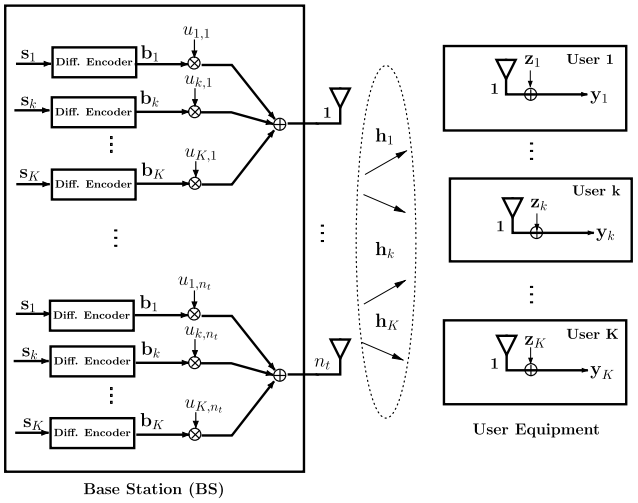

FIGURE 1. Differential massive MIMO downlink system model.

to receivers. Once the PSPs are estimated at the BS, the transmitter computes the precoder. More precisely, we provide an optimal solution for the precoder based on a max-min signal-to-interference-plus-noise ratio (SINR) problem formulation. The optimized precoder can effectively precancel the interference between users, thus enhancing overall system performance. In addition, we provide two suboptimal solutions suitable for the low interference system based on the matched and the orthogonality approach of PSP of each user. The proposed schemes facilitate precanceling MAI, enhance system performance, and provide simple transmitter and receiver schemes. Consequently, since the proposed scheme avoids channel estimation, the system overheads and latency will be reduced significantly.

The remainder of this paper is organized as follows. Section II introduces the system model of the differential massive MIMO system. Section III describes the downlink transmit precoding approach. Section IV presents differential detection for a massive MIMO system. In Section V, simulation results are shown. Finally, conclusions are drawn in Section VI.

Notations: Vectors are denoted by boldface lower case letters and matrices by boldface upper case letters. $\boldsymbol{I}_n$, $\mathbf{1}_n$, and $\mathbf{0}_{m\times n}$ denote an $n \times n$ identity matrix, $n \times 1$ identity vector and $m \times n$ zero matrix, respectively. The operators $(.)^T$, $(.)^*$ $(.)^H$, trace(.), log(·), $\log_2(\cdot)$, $|\cdot|$, $\|\cdot\|_F$, and diag(·) denote transpose, complex conjugate, conjugate transpose, trace of a matrix, natural logarithm, logarithm to base 2, absolute value of a scalar, Frobenius norm of a matrix, and diagonal of a matrix, respectively. $\Re(\cdot)$, $\Im(\cdot)$ denote real and imaginary part of a complex number, respectively. $\{\boldsymbol{x}_n\}$ denotes a set of all vectors indexed by n. $\mathbb{C}^{m\times n}$ and $\mathbb{R}^{m\times n}$ denote the set of all complex and the set of all real $m \times n$ matrices, respectively. $\mathbb{E}$ denotes the expected value of a discrete random variable. $\text{cov}(x, y)$ and $\text{var}(x)$ denote the covariance between the random variables x and y, and the variance of x, respectively.

II. SYSTEM MODEL

Consider a single-cell massive MIMO downlink broadcast channel. The BS has n_t transmit antennas, which simultaneously transmit multiple streams to K single-antenna users, as shown in Fig.1. The number of transmit antennas is assumed to be very large[1] ($n_t \gg 1$). We assumed that all users are equipped with a single-antenna for the decoding process which is a realistic assumption for the massive MIMO system, where the large number of transmit antennas at the BS provides a mutual orthogonality among the vector-valued channels to the users (so-called favorable propagation) [16]. For downlink massive MIMO transmission, multiple-antennas at each user increases receiver complexity and overhead. Instead, we would like to have a simple, inexpensive, and power efficient single-antenna receiver. Further, equivalent capacity can be achieved by serving K single-antenna users instead of one user having K-multiple-antennas users, thereby serving more users in the cell [16].

A. DIFFERENTIAL MASSIVE MIMO SYSTEM MODEL

For any kth user, $\boldsymbol{s}_k = [s_{k,1}, s_{k,2}, \cdots, s_{k,N}] \in \mathbb{C}^{1\times N}$, $1 \le k \le K$, is the information vector with elements drawn from an M-ary PSK constellation as:

$$\mathcal{M} = \left\{ e^{j2\pi i/M} \mid i = 0, 1, \cdots, M-1 \right\}, \tag{1}$$

where N denotes the block length of the coherence time intervals. In the context of differential massive MIMO system, a sequence of symbols of the kth user $s_{k,\tau}$, $1 \le \tau \le N$, is differentially encoded into the transmit symbol vector $\boldsymbol{b}_k \in \mathbb{C}^{1\times(N+1)}$ via the rule

$$b_{k,\tau} = s_{k,\tau} b_{k,\tau-1} = b_{k,0} \prod_{i=1}^{\tau} s_{k,i}. \tag{2}$$

The transmit information signal vector $\boldsymbol{b}_k$ comprises the initial reference symbol $b_{k,0} = 1$ and the following N differentially encoded symbols in the form of $\boldsymbol{b}_k = [b_{k,0}, b_{k,1}, \cdots, b_{k,N}]$.

We consider that no prior information about the channel is available at the BS. The channel vector between the BS and user k, $\boldsymbol{h}_k = [h_{k,1}, h_{k,2}, \cdots, h_{k,n_t}]^T \in \mathbb{C}^{n_t\times 1}$, models independent fast fading and slow fading PSP attenuation, where the PSP is denoted as $g_{k,m}$, $1 \le m \le n_t$, we specify PSP later in Section II-B. It is assumed that the channel coefficients remain constant over the block length and vary independently from one block to another. The coefficient $h_{k,m}$ can be written as

$$h_{k,m} = \sqrt{g_{k,m}}\, \tilde{h}_{k,m}, \quad m = 1, 2, \cdots, n_t, \tag{3}$$

where $\tilde{h}_{k,m}$ is the fast-fading coefficient from the kth user to the mth transmit antenna of the BS, which is modeled as an independent over m and identically distributed (i.i.d.) complex Gaussian random variable with zero-mean and unit-variance, i.e., $\tilde{h}_{k,m} \sim \mathcal{CN}(0, 1)$. $g_{k,m}$ models the PSP attenuation between the mth antenna at the BS and user k, which is assumed to be independent over m and to be constant over many coherence time intervals N and known *a priori* to the BS. We consider that the value of $\tilde{h}_{k,m}$ remain stationary for a sufficiently long transmission time. Then, we have

$$\boldsymbol{h}_k = \boldsymbol{G}_k^{1/2}\, \tilde{\boldsymbol{h}}_k, \qquad k = 1, \cdots, K, \tag{4}$$

where $\tilde{\boldsymbol{h}}_k = [\tilde{h}_{k,1}, \tilde{h}_{k,2}, \cdots, \tilde{h}_{k,n_t}]^T \in \mathbb{C}^{n_t\times 1}$, and $\boldsymbol{G}_k = \mathrm{diag}(\boldsymbol{g}_k) = \mathrm{diag}(g_{k,1}, g_{k,2}, \cdots, g_{k,n_t}) \in \mathbb{R}^{n_t\times n_t}$. Therefore, the variance of $\{\boldsymbol{h}_k\}$ is determined by the user PSP, where the channel variance is equal to the power profile, i.e., $h_{k,m} \sim \mathcal{CN}(0, g_{k,m})$.

We assume that the multiuser system adopts a linear transmission and reception strategy. The BS performs transmit beamforming and communicates simultaneously with all users. The instantaneous transmitted signal matrix $\boldsymbol{B} \in \mathbb{C}^{n_t\times(N+1)}$ for the kth user can then be expressed as

$$\boldsymbol{B} = \sum_{k=1}^{K} \sqrt{p_k}\, \boldsymbol{u}_k \boldsymbol{b}_k, \tag{5}$$

where $\boldsymbol{u}_k = \left[u_{k,1}, u_{k,2}, \cdots, u_{k,n_t}\right]^T \in \mathbb{R}^{n_t\times 1}$ is the normalized differential transmit precoder (beamformer) of the kth user, where $\|\boldsymbol{u}_k\|^2 = 1$. p_k is the downlink average transmit power of the kth user. We consider a total power constraint at the BS is

$$\mathbb{E}\left\{ \mathrm{trace}(\boldsymbol{B}^H \boldsymbol{B}) \right\} = \bar{P}. \tag{6}$$

The received signal vector $\boldsymbol{y}_k \in \mathbb{C}^{1\times(N+1)}$ for the kth user is given by

$$\boldsymbol{y}_k = \sqrt{p_k}\boldsymbol{h}_k^H \boldsymbol{u}_k \boldsymbol{b}_k + \boldsymbol{h}_k^H \sum_{\substack{q=1 \\ q\neq k}}^{K} \sqrt{p_q}\boldsymbol{u}_q \boldsymbol{b}_q + \boldsymbol{z}_k \tag{7}$$

$$= \sqrt{p_k}\boldsymbol{h}_k^H \boldsymbol{u}_k \boldsymbol{b}_k + \boldsymbol{w}_k + \boldsymbol{z}_k, \tag{8}$$

where the term $\sqrt{p_k}\boldsymbol{h}_k^H \boldsymbol{u}_k \boldsymbol{b}_k$ represents the desired signal at the kth user, $\boldsymbol{w}_k = \boldsymbol{h}_k^H \sum_{\substack{q=1 \\ q\neq k}}^{K} \sqrt{p_q}\boldsymbol{u}_q \boldsymbol{b}_q \in \mathbb{C}^{1\times(N+1)}$ is the MAI component against the kth user, and $\boldsymbol{z}_k \in \mathbb{C}^{1\times(N+1)}$ is the noise vector modeled as zero-mean complex circularly symmetric Gaussian random variables, i.e., $\boldsymbol{z}_k \sim \mathcal{CN}(0, \sigma_{z_k}^2 \boldsymbol{I}_{n_t})$.

Assuming that the information transmitted symbols $\boldsymbol{b}_k$ are uncorrelated, the average SINR_k at the kth user can be expressed as follows

$$\mathrm{SINR}_k = \mathbb{E}\left[\frac{p_k \left|\boldsymbol{h}_k^H \boldsymbol{u}_k\right|^2}{\sum_{\substack{q=1 \\ q\neq k}}^{K} p_q \left|\boldsymbol{h}_k^H \boldsymbol{u}_q\right|^2 + \sigma_{z_k}^2} \right]. \tag{9}$$

[1]The assumption of $n_t \to \infty$ is valid and commonly used in the massive MIMO literature. However, the system's performance can be tested for the practical case of large but limited number of antennas [15].

B. CO-LOCATED ANTENNA SYSTEM WITH A GEOMETRICAL MODEL

Now, we summarize the PSP model construction following the approach in [14]. As shown in Fig. 2, the users are randomly distributed in front of a large uniform antenna array at the BS. We assume that the BS has full knowledge of any user's location information. The location of the users is determined by the following parameters: $r_{k,m}$ is the distance between the antenna index m and the kth user; l_k is the direct orthogonal distance between the kth user and the array; and, l_a is the antenna spacing.[2] Let m_k denote the antenna element closest to the kth user according to the Euclidean distance. From algebraic geometry, we have

$$r_{k,m} = \sqrt{l_k^2 + l_a^2|m - m_k|^2} \tag{10}$$

$$= l_k\sqrt{1 + |m - m_k|^2/l_{k,r}^2}, \quad m = 1, \cdots, n_t, \tag{11}$$

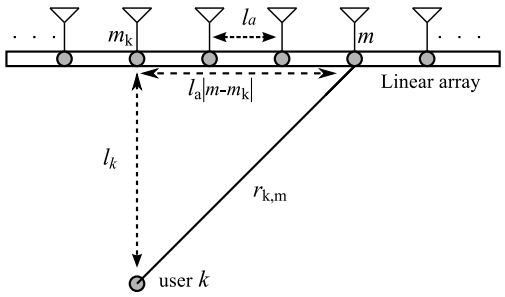

FIGURE 2. Geometric system model of user's location.

where $l_{k,r} \overset{\text{def}}{=} l_k/l_a$ denotes the normalized relative distance of the kth user to the array. We assume that the average transmit power obeys the path loss model with *path loss exponent* γ. Hence, the path loss for the kth user at antenna m is given by

$$g_{k,m} = r_{k,m}^{-\gamma}. \tag{12}$$

Using exponential and logarithmic properties, we have

$$g_{k,m} = \exp\left\{-\gamma \log\left(l_k\sqrt{1 + |m - m_k|^2/l_{k,r}^2}\right)\right\}$$

$$= \exp\left\{-\gamma \log(l_k) + \frac{-\gamma \log\left(1 + |m - m_k|^2/l_{k,r}^2\right)}{2}\right\}.$$

Since $\log(1 + x) \approx x$ for small x, we have

$$g_{k,m} = \exp\left\{-\gamma \log(l_k)\right\} \cdot \exp\left\{\frac{-\gamma|m - m_k|^2}{2l_{k,r}^2}\right\}$$

$$= \beta_k \cdot \exp\left\{\frac{-|m - m_k|^2}{2\zeta_k^2}\right\}, \tag{13}$$

where $\beta_k = \exp\left\{-\gamma \log(l_k)\right\}$ and $\zeta_k^2 \overset{\text{def}}{=} l_{k,r}^2/\gamma$. Therefore, the PSP is well approximated by a Gaussian function with mean m_k and channel variance ζ_k^2.

[2] we assume there is no correlation between transmit antennas at the BS as they are spaced at a minimum of 0.5λ.

Remark 1: In practical systems, the power space profile (PSP), $g_{k,m}$, of each user (which includes the path loss exponent) varies very slowly with time compared to the fast fading coefficients $\tilde{h}_{k,m}$. In this context, for massive MIMO systems, it is reasonable to assume that the PSPs of the users of the system are known at the BS [17]–[19]. In the uplink, we assume that each user transmits N (i.i.d.) symbols as $\boldsymbol{s}_k = [s_{k,1}, s_{k,2}, \cdots, s_{k,\tau}, \cdots, s_{k,N}]$. For each symbol the BS can calculate the PSP for each user accurately by ***averaging*** *the received uplink signal over different data slots indexed by τ as [17]*

$$\hat{\boldsymbol{G}}_k = diag(\hat{\boldsymbol{g}}_k) = \mathbb{E}\left[\boldsymbol{r}_{k,\tau}\boldsymbol{r}_{k,\tau}^H\right], \tag{14}$$

where $\boldsymbol{r}_{k,\tau}$ is the uplink received signal vector for the kth user at the antennas of the BS during the τth time slot, which is given as

$$\boldsymbol{r}_{k,\tau} = \boldsymbol{h}_k s_{k,\tau} + \boldsymbol{z}_{k,\tau}. \tag{15}$$

Hence, with the assumption of channel reciprocity, the PSP is calculated for each user during the uplink as in (14), which is assumed to be equivalent to the PSP in the downlink.

Note that PSP profile estimation is less challenging than estimating the actual channel state information, in which the PSP can remain constant over many coherence time intervals. However, the actual estimation process of $g_{k,m}$ is beyond the scope of this paper, thus we assumed the PSP profiles to be (perfectly) known in our system.

III. DOWNLINK TRANSMIT PRECODING

In massive MIMO, transmit precoding is used to cancel inter-user interference. Conventional transmit precoding design requires channel knowledge at the transmitter. However, in massive MIMO, the number of transmit antennas is very large, i.e., $n_t \gg 1$. Hence, the estimation of all channel coefficients $h_{k,m}$ quickly becomes unfeasible. Instead, differential transmit precoding schemes could be considered which avoid the need for explicit channel estimation. After estimating the PSP profile $g_{k,m}$ at the BS, we use this knowledge to design the transmit precoder for each user k to separate different users. Now, we present an asymptotic analysis of SINR and the proposed precoder design strategies for the differential massive MIMO framework.

A. ASYMPTOTIC ANALYSIS OF SINR

As a consequence of employing large number of antennas at the BS $n_t \to \infty$ (as our case of massive MIMO), the downlink channel vectors of independent users have a large degree of orthogonality, i.e.,

$$\frac{1}{n_t}\boldsymbol{h}_k^H\boldsymbol{h}_k \overset{N\to\infty}{\to} \frac{\boldsymbol{g}_k^T\boldsymbol{g}_k}{n_t}, \quad \frac{1}{n_t}\boldsymbol{h}_k^H\boldsymbol{h}_{q\neq k} \overset{N\to\infty}{\to} \frac{\boldsymbol{g}_k^T\boldsymbol{g}_{q\neq k}}{n_t}.$$

The orthogonality between different user's channels is determine by the orthogonality between the small fading vectors $\{\tilde{\boldsymbol{h}}_k\}$, and the orthogonality between the PSPs $\{\boldsymbol{g}_k\}$.

Theorem 1: From the law of large random numbers and under the most favorable propagation conditions, where the

column-vectors of the propagation vectors are asymptotically orthogonal, the expected value of SINR_k *can be calculated when* $n_t \to \infty$. *Since* $\{\boldsymbol{h}_k^H\}$ *has Gaussian distribution with zero-mean and covariance of* $\boldsymbol{G}_k = \mathrm{diag}(\boldsymbol{g}_k)$, *hence the desired signal* $\sqrt{p_k}\boldsymbol{h}_k^H\boldsymbol{u}_k$ *is also Gaussian distributed with zero mean and variance* $p_k\sum_{m=1}^{n_t} g_{k,m}u_{k,m}^2$, *where the sum of multiple Gaussian variables is also a Gaussian variable. Similarly, the interference component of* SINR_k *is also a Gaussian signal with variance* $\sum_{q=1,q\neq k}^{K} p_q\sum_{m=1}^{n_t} g_{k,m}u_{q,m}^2$. *The variance for the AWGN noise is* $\sigma_{z_k}^2$. *Therefore, the expected value of* SINR_k *as* $n_t \to \infty$ *is*

$$\mathrm{SINR}_k \overset{N\to\infty}{\to} \frac{\mathbb{E}\left[p_k\left|\boldsymbol{h}_k^H\boldsymbol{u}_k\right|^2\right]}{\mathbb{E}\left[\sum_{\substack{q=1\\q\neq k}}^{K} p_q\left|\boldsymbol{h}_k^H\boldsymbol{u}_q\right|^2+\sigma_{z_k}^2\right]} \tag{16}$$

$$= \frac{p_k\sum_{m=1}^{n_t} g_{k,m}u_{k,m}^2}{\sum_{\substack{q=1\\q\neq k}}^{K} p_q\sum_{m=1}^{n_t} g_{k,m}u_{q,m}^2+\sigma_{z_k}^2}. \tag{17}$$

Proof: See Appendix I □

B. SUBOPTIMAL PRECODERS

1) MATCHED PSP PRECODER

The first precoder design strategy is to match the beamformer vector to the PSP profile of the transmit antennas to separate different users, i.e., $u_{k,m}^2 \sim g_{k,m}$, which can be written as

$$u_{k,m} = \sqrt{\beta_k \cdot \exp\left\{\frac{-|m-m_k|^2}{2\zeta_k^2}\right\}}, \tag{18}$$

where we assume the BS has full knowledge of the channel parameter ζ_k, and the antenna index m_k which is the closest to the user k with maximum average power. For the power allocation in matched PSP scheme, we allocate the downlink transmit power equally between users, i.e., $p_k = \bar{P}/K$.

2) ORTHOGONAL PSP PRECODER

In the orthogonal PSP precoder, the beamformer for each user has to be distinguished and identified from other users. In the orthogonal precoder scheme, each user is assigned a unique orthogonal PSP to enhance data separation between users. The orthogonal PSP for each user is then multiplexed by its own power profile.

The orthogonal precoder for each user can be constructed using the Gram-Schmidt process (GSP). Let the vector $v_k = \left[v_{k,1}, v_{k,2}, \cdots, v_{k,n_t}\right]^T \in \mathbb{R}^{n_t\times 1}$ represent the user's PSP vector. The elements of vector v_k are computed by matching their value to the power profile of the transmit antennas, i.e., $v_{k,m}^2 \sim g_{k,m}$. The Gram-Schmidt process takes a finite, linearly independent set $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k, \cdots, \boldsymbol{v}_K\}$ for $K \leq n_t$ and generates an orthogonal set $\bar{S} = \{\bar{\boldsymbol{v}}_1, \bar{\boldsymbol{v}}_2, \cdots, \bar{\boldsymbol{v}}_k, \cdots, \bar{\boldsymbol{v}}_K\}$ which spans the same K-dimensional subspace of $\mathbb{R}^{n_t\times 1}$ as S. We define the projection operator as [20]

$$\mathrm{proj}_{\bar{\boldsymbol{v}}}(\boldsymbol{v}) = \frac{\langle\bar{\boldsymbol{v}},\boldsymbol{v}\rangle}{\langle\bar{\boldsymbol{v}},\bar{\boldsymbol{v}}\rangle}\bar{\boldsymbol{v}}, \tag{19}$$

where $\langle\bar{\boldsymbol{v}},\boldsymbol{v}\rangle$ denotes the inner product of the vectors $\bar{\boldsymbol{v}}$ and $\boldsymbol{v}$, i.e., $\langle\bar{\boldsymbol{v}},\boldsymbol{v}\rangle = \bar{\boldsymbol{v}}^T\boldsymbol{v}$ for vectors in $\mathbb{R}^{n_t\times 1}$. The Gram-Schmidt process then works as follows:

$$\begin{aligned}
\bar{\boldsymbol{v}}_1 &= \boldsymbol{v}_1,\\
\bar{\boldsymbol{v}}_2 &= \boldsymbol{v}_2 - \mathrm{proj}_{\bar{\boldsymbol{v}}_1}(\boldsymbol{v}_2),\\
\bar{\boldsymbol{v}}_3 &= \boldsymbol{v}_3 - \mathrm{proj}_{\bar{\boldsymbol{v}}_1}(\boldsymbol{v}_3) - \mathrm{proj}_{\bar{\boldsymbol{v}}_2}(\boldsymbol{v}_3),\\
&\vdots\\
\bar{\boldsymbol{v}}_k &= \boldsymbol{v}_k - \sum_{i=1}^{k-1}\mathrm{proj}_{\bar{\boldsymbol{v}}_i}(\boldsymbol{v}_k).
\end{aligned} \tag{20}$$

Note that, the Gram-Schmidt precoder for the first user is equal to the original PSP for the first user, i.e., $\bar{\boldsymbol{v}}_1 = \boldsymbol{v}_1$, hence the user's separation works only for the received signal of the first user. To enhance the separation for the received signal of other users, we multiply each element of the orthonormal vector $\bar{\boldsymbol{v}}_k$ by its own specific original power profile elements of $\boldsymbol{v}_k$ and then normalize them, which yields

$$\boldsymbol{u}_k = \frac{\boldsymbol{v}_k \circ \bar{\boldsymbol{v}}_k}{\|\boldsymbol{v}_k \circ \bar{\boldsymbol{v}}_k\|}, \tag{21}$$

where $\circ$ denotes the Hadamard product. For power allocation in orthogonal PSP precoder, we allocate the downlink transmit power equally between users, i.e., $p_k = \bar{P}/K$.

C. OPTIMAL PSP PRECODERS

In this precoder, we consider the joint optimization of power and downlink precoder for the PSP among all users simultaneously using the max-min formulation problem. A max-min formulation guarantees a fair quality of service among all users.

1) SINR OPTIMAL PSP PRECODER

In optimal PSP precoder, we maximize the worst case SINR jointly among all user. Starting from (17), the corresponding optimization problem can be written as

$$\underset{\substack{p_k,u_{k,m}\\k\in[1,K],m\in[1,n_t]}}{\text{maximise}} \ \min_{k\in[1,K]}\left(\frac{p_k\sum_{m=1}^{n_t} g_{k,m}u_{k,m}^2}{\sum_{\substack{q=1\\q\neq k}}^{K} p_q\sum_{m=1}^{n_t} g_{k,m}u_{q,m}^2+\sigma_{z_k}^2}\right), \tag{22a}$$

$$\text{subject to} \sum_{k=1}^{K}\sum_{m=1}^{n_t} p_k u_{k,m}^2 \leq \bar{P}, \tag{22b}$$

$$p_k u_{k,m}^2 \geq 0, \quad \forall k,m. \tag{22c}$$

Problem (22) can be recast as

$$\underset{\substack{p_k,\boldsymbol{c}_k\\k\in[1,K]}}{\text{maximise}} \ \min_{k\in[1,K]}\left(\frac{p_k\boldsymbol{f}_k^T\boldsymbol{c}_k}{\boldsymbol{f}_k^T\sum_{\substack{q=1\\q\neq k}}^{K} p_q\boldsymbol{c}_q+1}\right), \tag{23a}$$

$$\text{subject to} \sum_{k=1}^{K} p_k\mathbf{1}_{n_t}^T\boldsymbol{c}_k \leq \bar{P}, \tag{23b}$$

$$p_k\boldsymbol{c}_k \geq 0, \quad \forall k, \tag{23c}$$

where $\boldsymbol{f}_k$ and $\boldsymbol{c}_k$ are defined as

$$\boldsymbol{f}_k = \frac{1}{\sigma^2_{z_k}}\left[g_{k,1}, g_{k,2}, \cdots, g_{k,m}, \cdots, g_{k,n_t}\right]^T, \tag{24}$$

and

$$\boldsymbol{c}_k = \left[u^2_{k,1}, u^2_{k,2}, \cdots, u^2_{k,m} \cdots, u^2_{k,n_t}\right]^T. \tag{25}$$

It can be seen that the cost-function in (23a) is non-linear and non-convex over the optimization variables p_k, and $\boldsymbol{c}_k$ for $k \in [1, K]$. In the following, we provide optimal solutions for the design problems. The feasibility of problem (23) can be examined by solving it with the objective function replaced by constant values, i.e., finding a common domain which satisfies all problem constraints. Without loss of generality, we assume that our problem is feasible. Next, we solve our optimization problem optimally through recasting the non-convex constraints. Now, let's define a $Kn_t \times 1$ vector $\boldsymbol{v}$ as

$$\boldsymbol{v} = \left[p_1(\boldsymbol{c}_1)^T, p_2(\boldsymbol{c}_2)^T, \cdots, p_K(\boldsymbol{c}_K)^T\right]^T. \tag{26}$$

In addition, let's define other variables $\boldsymbol{w}_k$ and $\bar{\boldsymbol{w}}_k$ of size $Kn_t \times 1$ as

$$\boldsymbol{w}_k = \left[\mathbf{0}^T_{(k-1)n_t \times 1}, \boldsymbol{f}^T_k, \mathbf{0}^T_{(K-k)n_t \times 1}\right]^T, \tag{27}$$

and

$$\bar{\boldsymbol{w}}_k = \left[\boldsymbol{f}^T_1, \cdots, \boldsymbol{f}^T_{k-1}, \mathbf{0}^T_{n_t \times 1}, \boldsymbol{f}^T_{k+1}, \cdots, \boldsymbol{f}^T_K\right]^T, \tag{28}$$

where $\mathbf{0}_{m \times 1}$ denotes an $m \times 1$ vector whose elements are zero. Next, the SINR_k optimization problem in (23) may be written in a more convenient form by using (26), (27), and (28), which yields

$$\underset{\boldsymbol{v}}{\text{maximise}} \ \min_{k \in [1,K]} \left(\frac{\boldsymbol{w}^T_k \boldsymbol{v}}{\bar{\boldsymbol{w}}^T_k \boldsymbol{v} + 1}\right), \tag{29a}$$

$$\text{subject to } \mathbf{1}^T_{n_t} \boldsymbol{v} \leq \bar{P}, \tag{29b}$$

$$\boldsymbol{v} \geq 0. \tag{29c}$$

To convexify the cost-function (29a), which comprises a product of fractional terms, we substitute the numerators and denominators of the fractions by exponential variables as follows [21]

$$e^{\alpha_k} = \boldsymbol{w}^T_k \boldsymbol{v}, \quad \forall k, \tag{30a}$$

$$e^{\tilde{\alpha}_k} = \bar{\boldsymbol{w}}^T_k \boldsymbol{v} + 1, \quad \forall k. \tag{31b}$$

Then, by using the properties of the exponential and according to (30a) and (30a), the problem in (29) can be formalized as

$$\underset{\substack{\boldsymbol{v}, \alpha_k, \tilde{\alpha}_k \\ k \in [1,K]}}{\text{maximise}} \ \min_{k \in [1,K]} \left(e^{(\alpha_k - \tilde{\alpha}_k)}\right), \tag{32a}$$

$$\text{subject to } \mathbf{1}^T_{n_t} \boldsymbol{v} \leq \bar{P}, \tag{32b}$$

$$\boldsymbol{v} \geq 0, \tag{32c}$$

$$e^{\alpha_k} \leq \boldsymbol{w}^T_k \boldsymbol{v}, \quad \forall k, \tag{32d}$$

$$e^{\tilde{\alpha}_k} \geq \bar{\boldsymbol{w}}^T_k \boldsymbol{v} + 1, \quad \forall k. \tag{32e}$$

It can be seen that the exponential parameters e^{α_k} and $e^{\tilde{\alpha}_k}$ in (32d) and (32e) are constrained by the expressions on the right hand sides of (30a) and (30a), respectively. The objective function in (32a) consists of an exponential function which is non-convex, and thus we can linearize it using the monotonicity property of the exponential function. Hence, the objective function in (32a) can be defined as follows

$$e^{(\alpha_k - \tilde{\alpha}_k)} \overset{\text{def}}{=} \alpha_k - \tilde{\alpha}_k, \quad \forall k. \tag{33}$$

Next, to deal with the non-convex constraint (32e), we linearize the exponential term $e^{\tilde{\alpha}_k}$ using the first order Taylor approximation as follows [22]

$$e^{\tilde{\alpha}_k} = e^{\ddot{\alpha}_k}\left(1 + \tilde{\alpha}_k - \ddot{\alpha}_k\right), \quad \forall k, \tag{34}$$

where $\ddot{\alpha}_k$ is the point where the linear approximation is made. Therefore, from (33) and (34), problem (32) can be reformulated as

$$\underset{\substack{\boldsymbol{v}, \alpha_k, \tilde{\alpha}_k \\ k \in [1,K]}}{\text{maximise}} \ \min_{k \in [1,K]} \alpha_k - \tilde{\alpha}_k, \tag{35a}$$

$$\text{subject to } \mathbf{1}^T_{n_t} \boldsymbol{v} \leq \bar{P}, \tag{35b}$$

$$\boldsymbol{v} \geq 0, \tag{35c}$$

$$e^{\alpha_k} \leq \boldsymbol{w}^T_k \boldsymbol{v}, \quad \forall k, \tag{35d}$$

$$e^{\ddot{\alpha}_k}\left(1 + \tilde{\alpha}_k - \ddot{\alpha}_k\right) \geq \bar{\boldsymbol{w}}^T_k \boldsymbol{v} + 1, \quad \forall k. \tag{35e}$$

Now the above problem (35) is convex and can be solved iteratively using CVX optimization software [23]. The initial value of $\ddot{\alpha}_k$ is updated by the optimized value of $\tilde{\alpha}_k$, $\forall k$, obtained in the previous iteration. The iterations continue until the error, $\sum_{k=1}^{K} |\ddot{\alpha}_k - \tilde{\alpha}_k|$, converges to a certain threshold. Algorithm 1 is provided to solve the above optimization function. Here $\boldsymbol{\alpha} = [\alpha_1 \cdots \alpha_K]^T$, $\tilde{\boldsymbol{\alpha}} = \left[\tilde{\alpha}_1 \cdots \tilde{\alpha}_K\right]^T$, and $\ddot{\boldsymbol{\alpha}} = [\ddot{\alpha}_1 \cdots \ddot{\alpha}_K]^T$.

Algorithm 1 Algorithm for Solving Problem (35)

1: Set threshold = ϵ
2: Initialize $\ddot{\boldsymbol{\alpha}}^{[i]}$, $\tilde{\boldsymbol{\alpha}}^{[i]}$, $i = 0$
3: **while** $\mathbf{1}^T_{n_t}\left(|\ddot{\boldsymbol{\alpha}} - \tilde{\boldsymbol{\alpha}}|\right) > \epsilon$ or $i = 0$ **do**
4: increment $i = i + 1$.
5: update the initial values $\ddot{\boldsymbol{\alpha}}^{[i]} = \tilde{\boldsymbol{\alpha}}^{[i-1]}$.
6: solve problem (35) using CVX and calculate $\boldsymbol{v}^{[i]}$, $\boldsymbol{\alpha}^{[i]}$, $\tilde{\boldsymbol{\alpha}}^{[i]}$.
7: **until** Convergence.
8: **end while**
9: Find $\boldsymbol{u}_k$ and p_k of each user from $\boldsymbol{v}$ as in (25) and (26).

Remark 2: There is an alternative approach for designing the transmit precoder based on maximizing the worst case of signal-to-leakage-noise ratio (SLNR). The SLNR is defined as the ratio of received signal power at the desired user to received signal power at the other users (the leakage) [24].

The average SLNR_k *at the kth user can be expressed as*

$$\mathrm{SLNR}_k \overset{N\to\infty}{\to} \frac{\mathbb{E}\left[p_k \left|\boldsymbol{h}_k^H \boldsymbol{u}_k\right|^2\right]}{\mathbb{E}\left[p_k \sum_{\substack{q=1\\ q\neq k}}^{K} \left|\boldsymbol{h}_q^H \boldsymbol{u}_k\right|^2 + \sigma_{z_k}^2\right]} \tag{36}$$

$$= \frac{p_k \sum_{m=1}^{n_t} g_{k,m} u_{k,m}^2}{p_k \sum_{\substack{q=1\\ q\neq k}}^{K} \sum_{m=1}^{n_t} g_{q,m} u_{k,m}^2 + \sigma_{z_k}^2}. \tag{37}$$

The proof of (37) is similar to the proof of (17) in Theorem 1. The optimization solution for maximizing the worst case SLNR of (37), (max-min SLNR), jointly among all users provides the same performance as in the proposed optimal PSP SINR precoder, (max-min SINR).

D. COMPUTATIONAL COMPLEXITY ANALYSIS FOR THE PSP PRECODERS

In this section, we quantify the computational complexity for the proposed PSP precoders for the optimal and the suboptimal solutions. The computational process is done based on the size of input data, the floating point operation (FLOPs), the type of the optimization problems, the number of the required iterations, and the methods used in finding the solution.

1) COMPLEXITY OF SUBOPTIMAL SOLUTIONS

The notion of FLOPs is introduced. We use the total number of FLOPs to measure the computational complexity of matrix operations. We summarize the total FLOPs needed for some matrix operations below [25]:

- Multiplication of $m \times n$ and $n \times p$ complex matrices: $\mathcal{O}(8mnp - 2mp)$;
- Inversion of an $m \times m$ real matrix using Gauss-Jordan elimination: $\mathcal{O}(4m^3/3)$.
- GSP to an $m \times n$ $(m \geq n)$ complex matrix: $\mathcal{O}(8n^2(m - n/3))$.
- Hadamard product for two $m \times m$ matrices: $\mathcal{O}(m)$.
- L^2-norm of an $m \times 1$ real vector: $\mathcal{O}(3m)$.

According to the aforementioned summary of FLOPs operations, the computational complexity of the suboptimal PSP precoder is

$$\mathcal{O}\left(K\left(13n_t - \frac{8}{3}\right)\right). \tag{38}$$

2) COMPLEXITY OF OPTIMAL SOLUTION

Now, we calculate the complexity of optimizing the downlink PSP precoder which is formulated as a linear programming (LP) problem in (35). The computational complexity of such LP problems has been studied in Chapter 6 in [26] where the complexity is calculated in terms of the number of optimization variables n, number of constraints m and the size of input data $\dim(\boldsymbol{p})$, where $\boldsymbol{p}$ is the vector of input data. To apply the complexity evaluation steps given in chapter 6 in [26], problem (35) is recast into its standard LP form. This can be achieved by replacing the min operator in the objective function by new slack variable π and K scalar constraints (see (39d)). Therefore, (39) is an equivalent and standard LP recast of the original problem (35). Note that the constraint (35d) is linearized in a similar way as used for (35e) since the used CVX's solvers such as SDPT3 and SeDuMi do not support the exponential function.

$$\underset{\substack{\boldsymbol{v},\,\alpha_k,\,\tilde{\alpha}_k,\,\pi\\ k\in[1,K]}}{\text{maximise}} \quad \pi \tag{39a}$$

$$\text{subject to } \mathbf{1}_{n_t}^T \boldsymbol{v} \leq \bar{P}, \tag{39b}$$

$$\boldsymbol{v} \geq 0, \tag{39c}$$

$$\alpha_k - \tilde{\alpha}_k \geq \pi, \quad \forall k, \tag{39d}$$

$$e^{\hat{\alpha}_k}\left(1 + \alpha_k - \hat{\alpha}_k\right) \leq \boldsymbol{w}_k^T \boldsymbol{v}, \quad \forall k, \tag{39e}$$

$$e^{\ddot{\alpha}_k}\left(1 + \tilde{\alpha}_k - \ddot{\alpha}_k\right) \geq \bar{\boldsymbol{w}}_k^T \boldsymbol{v} + 1, \quad \forall k. \tag{39f}$$

Problem (39) contains $n = (n_t+2)K+1$ scalar variables, $m = (n_t+3)K$ scalar constraints, and require the input data vector $\boldsymbol{p} = [n,\ m,\ \boldsymbol{w}_1^T,\ \ldots,\ \boldsymbol{w}_K^T,\ \bar{\boldsymbol{w}}_1^T,\ \ldots,\ \bar{\boldsymbol{w}}_K^T,\ \hat{\alpha}_1,\ \ldots,\ \hat{\alpha}_K,\ \ddot{\alpha}_1,\ \ldots, \ddot{\alpha}_K]$. According to these problem parameters, the complexity of achieving a per-iteration solution within the an accuracy ϵ is [26]

$$\mathcal{O}(1)\sqrt{m+n}\ \ln\left(\frac{\dim(\boldsymbol{p}) + \|\boldsymbol{p}\|_1 + \epsilon^2}{\epsilon}\right), \tag{40}$$

where $\mathcal{O}(1)$ is the complexity of a real operation. According to (40) and the aforementioned problem parameters, the per-iteration complexity asymptotically (as $n_t,\ K \to \infty$ and $n_t \gg K$) converges to

$$\mathcal{O}\left(Kn_t\left[\ln(2K^2 n_t) + \ln\left(\frac{1}{\epsilon}\right)\right]\right). \tag{41}$$

Obviously, from (38) and (41), the optimal PSP precoder has lower computational complexity than the suboptimal PSP precoders, where the main parameters are the total number of users K and the total number of transmit antennas n_t. We will explore more on the comparison between (38) and (41) in Section V.

IV. DIFFERENTIAL DETECTION FOR MASSIVE MIMO WITH DOWNLINK TRANSMISSION

In this section, the differential encoding and decoding process for the downlink transmission in a massive MIMO system is discussed. Here, we assume that neither the transmitter nor the receiver has prior knowledge of the CSI.

A. MULTIPLE SYMBOLS DIFFERENTIAL DETECTION

The simpler suboptimal method of implementing DD with massive MIMO is to encode the transmitted data differentially and to decode only the last two consecutive received symbols, e.g. $N = 2$, without any knowledge of the CSI. In contrast, the optimal method is to decode a block of N consecutive information symbols jointly without any knowledge of the CSI by performing MSDD, e.g. $N \gg 2$, which results in a

3dB performance improvement compared to DD[3] [12], [13]. In MSDD for the downlink system, the differential transmissions are implemented in blocks, in which each user k receives the sum of all the transmit waveforms of other users; then, the received signal blocks for each user must be detected independently. The measurements at the receiver are collected by spatial autocorrelation, and then we resort to the generalized likelihood ratio test (GLRT) optimization criterion whereby the maximization of the likelihood function is performed not only over the unknown symbols but also over unknown channels [27].

When using the M-ary PSK constellation, the MSDD detection problem can be simplified as [13]

$$\hat{\boldsymbol{b}}_k = \underset{\tilde{\boldsymbol{b}}_k \in \mathcal{M}^{N+1}, \tilde{b}_o = 1}{\arg\max} \; \tilde{\boldsymbol{b}}_k \boldsymbol{Y}_k \tilde{\boldsymbol{b}}_k^H \tag{42}$$

$$= \underset{\tilde{\boldsymbol{b}}_k \in \mathcal{M}^{N+1}, \tilde{b}_o = 1}{\arg\max} \; \Re\left\{ \sum_{\tau=1}^{N} \tilde{b}_{k,\tau}^* \sum_{l=0}^{\tau-1} \tilde{b}_{k,l} \cdot y_{k,l,\tau} \right\}, \tag{43}$$

where $\boldsymbol{Y}_k = \boldsymbol{y}_k^H \boldsymbol{y}_k \in \mathbb{C}^{(N+1)\times(N+1)}$ is the autocorrelation matrix of the received signal comprised of the correlation coefficients $y_{k,l,\tau}$, $\tau = 1, \ldots, N$, $l = 0, \ldots, \tau - 1$, between the lth and the τth received differential signals. As we are interested in the information symbols $\boldsymbol{s}_k$, it can be seen that $\boldsymbol{s}_k$ is directly obtained as

$$s_{k,\tau} = b_{k,\tau} \cdot b_{k,\tau-1}^*. \tag{44}$$

In (42), the differential decoder uses one side of the complex-conjugate symmetry of the correlation coefficients, thus $y_{k,l,\tau} = y_{k,\tau,l}^*$. Further, the diagonal elements of $\boldsymbol{Y}_k$ can be neglected as they do not influence the decision metrics, i.e., $y_{k,l,l} = y_{k,\tau,\tau} = 0$.

B. DECISION FEEDBACK DIFFERENTIAL DETECTION

In order to improve the performance further, DFDD is adopted in this paper. This approach leads to better performance compared to MSDD. Different from [14], we construct the DFDD for the downlink transmission instead of the uplink. In DFDD, the decisions are made successively, adding all previous decisions in the decision of the current symbol. In this decoding algorithm, the decoder detects symbols one by one. After finding the best candidate for the first symbol, the effects of this symbol in all of the receiver equations are added and considered. Then, the second symbol is detected from the new sets of equations. The effects of the first and second detected symbols are added and then considered to derive a new set of equations. The process continues until all symbols are detected. Of course, the order in which the symbols are detected will impact the end solution. The algorithm includes three steps, i.e. decision, process, and ordering.

[3]In MSDD, there is a 3dB gain when using large values of N, yet the cardinality of the search set grows exponentially with N, i.e., $|\mathcal{M}| = \mathcal{M}^{N+1}$. However, to achieve low complexity design, one can resort to using an edge computing platform or the conventional DD.

1) DECISION PROCESS

From the description given above and starting with $b_{k,0} = 1$, the decision process means that the information symbols in (43) are detected one by one as

$$\hat{b}_{k,\tau} = \underset{\tilde{b}_{k,\tau} \in \mathcal{M}^{N+1}}{\arg\max} \; \Re\left\{ \tilde{b}_{k,\tau}^* \sum_{l=0}^{\tau-1} \tilde{b}_{k,l} \cdot y_{k,l,\tau} \right\} \tag{45}$$

$$= \exp\left(j \cdot \theta_{\mathrm{PSK}}\left\{ \sum_{l=0}^{\tau-1} \tilde{b}_{k,l} \cdot y_{k,l,\tau} \right\} \right), \tag{46}$$

where

$$\theta_{\mathrm{PSK}}\{x\} \overset{\mathrm{def}}{=} \frac{2\pi}{M} \cdot \left\lfloor \frac{M}{2\pi} \cdot \arg(x) \right\rceil, \tag{47}$$

and

$$\Delta\theta_{\mathrm{PSK}}\{x\} \overset{\mathrm{def}}{=} \left| \arg\left\{ \exp\left(j \cdot \left(\arg(x) - \theta_{\mathrm{PSK}}(x) \right) \right) \right\} \right| \tag{48}$$

quantizes the phase of a complex number $x \in \mathbb{C}$ to the M phase values of M-ary PSK, and computes the quantization error, respectively. The operation $\lfloor x \rceil$ in (47) takes as input a real number x and gives as output a reduction into the interval $[\pi, 2\pi]$. The purpose of this step is to decide which transmitted symbol to detect at each stage of the decoding.

2) OPTIMUM DECISION ORDERING

It is well known from decision feedback equalization in MIMO systems, also known as BLAST [28], that sorting the decisions in an optimized order improves performance. The symbol with lowest quantization error in (48) is the best in this step. The decision order can be achieved by reordering the columns and rows of the $\boldsymbol{Y}_k$ matrix. That is, we first denote the index for the best transmitted symbols in the $\boldsymbol{Y}_k$ matrix by $(\hat{\tau}_0, \hat{\tau}_1, \cdots, \hat{\tau}_N)$, where $\hat{\tau}_i, \hat{\tau}_l \in \{0, \cdots, N\}$, $\hat{\tau}_i \neq \hat{\tau}_l$ for $i \neq l$. Then, we define the symbols transmitted in the $\hat{\tau}_i$th index by $(b_{k,\hat{\tau}_0}, b_{k,\hat{\tau}_1}, \cdots, b_{k,\hat{\tau}_i}, \cdots, b_{k,\hat{\tau}_N})$.

Now, we start by setting the initial transmitted symbol to identity, i.e., $b_{k,\hat{\tau}_0} = 1$. Then, the first decided symbol should be the $\hat{\tau}_1$th symbol, where

$$[\hat{\tau}_0, \hat{\tau}_1] = \underset{\substack{l \in \{0,\cdots,N\}, \tau \in \{1,\cdots,N\} \\ l \leq \tau}}{arg\,min} \left| \Delta\theta_{\mathrm{PSK}}\{y_{k,l,\tau}\} \right|, \tag{49}$$

and the estimate for the $b_{k,\hat{\tau}_1}$ symbol is obtained from

$$b_{k,\hat{\tau}_1} = \exp\left(j \cdot \theta_{\mathrm{PSK}}\left\{ y_{k,\hat{\tau}_0,\hat{\tau}_1} \right\} \right). \tag{50}$$

Taking the previous decision into account, the symbol that is decided next can be obtained successively from

$$\hat{\tau}_i = \underset{\substack{\tau \in \{1,\cdots,N\} \\ /\{\hat{\tau}_0,\cdots,\hat{\tau}_{i-1}\}}}{arg\,min} \left| \Delta\theta_{\mathrm{PSK}}\left\{ \sum_{l=0}^{i-1} b_{k,\hat{\tau}_l} \cdot y_{k,\hat{\tau}_l,\tau} \right\} \right|, \tag{51}$$

and its value can be obtained from

$$b_{k,\hat{\tau}_i} = \exp\left(j \cdot \theta_{\mathrm{PSK}}\left\{ \sum_{l=0}^{i-1} b_{k,\hat{\tau}_l} \cdot y_{k,\hat{\tau}_l,\hat{\tau}_i} \right\} \right). \tag{52}$$

TABLE 1. The values of PSP parameters to be used in (13).

PSP Parameters List					
Path loss exponent γ	Antenna spacing l_a	Orthogonal distance l_k	Relative distance $l_{k,r}$	Peak amplitude β_k	Channel variance ζ_k
3.6	0.5m	3m	6m	0.2	3
3.6	0.5m	5m	9m	0.1	5
3.6	0.5m	9m	19m	0.03	10
3.6	0.5m	14m	28m	0.02	15
3.6	0.5m	19m	38m	0.01	20

This ordering scheme has attempted to provide reliable decisions for the first decided symbols, which will impact the decision for subsequent symbols, and thus improve performance. Further, it must be noted that the actual realizations of the channel vectors $\{\boldsymbol{h}_k\}$ are not needed to decode the information signals.

V. SIMULATION RESULTS AND DISCUSSION

In this section, the performance of the differential massive MIMO downlink transmission is examined. We assume the channel is modeled as quasi-static, where the block fading channel between the transmitter and receiver is constant (but unknown) during N successive channel uses, i.e., the block length of the coherence time intervals. The fast fading coefficients for each user $\tilde{\boldsymbol{h}}_k = [\tilde{h}_{k,1}, \cdots, \tilde{h}_{k,n_t}]^T$ are mutually independent and modeled as independent and identically distributed (i.i.d.) complex Gaussian random variable with zero-mean and unit-variance, i.e., $\tilde{h}_{k,m} \sim \mathcal{CN}(0, 1)$.

Throughout this section, we assume the following; urban area cellular radio model for γ, one receive antenna per user, the noise power $\sigma_{z_k}^2 = 0$ dB, the constellation size is 4-PSK, the length of the transmission block is set to $N = 200$, and we use the DFDD detection technique for differential detection. Table 1 shows the values of PSP parameters to be used in (13) whenever needed throughout the simulation section. Note that using ζ without the superscript k means that the values of ζ are equal for all users, i.e., $\zeta_1 = \cdots = \zeta_K = \zeta$. The Monte Carlo simulation is used to evaluate the performance in terms of bit error rate (BER).

A. SINGLE-USER SCENARIO

The BER performance curve is first simulated and plotted for only one user. We assume that the user's location is in front of the center of the antenna array, i.e., $m_1 = 50$. The BS has $n_t = 100$ transmit antennas. We examined this case using the three proposed precoders, e.g., matched PSP precoder, orthogonal PSP precoder, and optimal PSP precoder. In addition to this we compared them against the unity precoder (equal power allocation), where the precoder vector elements are all set to one, i.e., $\{\boldsymbol{u}_k\} = \mathbf{1}_{n_t}$ and then normalized. The channel parameter is set to $\zeta = 10$. When there is no interference, Fig. 3 shows that the performance of

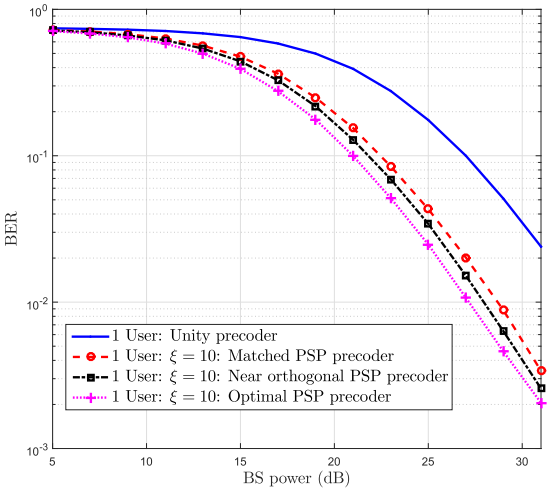

FIGURE 3. BER performance of the proposed differential MIMO downlink transmission with single user. $n_t = 100$, $\zeta = 10$.

the proposed precoders schemes, e.g., matched, orthogonal, and optimal precoders, outperforms the one that does not perform any kind of optimization for the precoding vector, e.g., the unity precoder. Clearly, in the interference-free system, the performance of the optimal PSP precoder is slightly better than the other two precoders but the difference is very small. It should be noted that in a coherent system, it is well known that the matched (to the channel) filter maximizes the SNR for the single user case. This is valid for both conventional and massive MIMO systems. However, in a noncoherent system, the matched PSP precoder is matched only to the PSP and not to the channel itself. Therefore, the matched PSP precoder does not necessarily maximize the SNR. In the optimal PSP precoder design, the optimizer tends to allocate the power to the channels that have significant gains. In other words, as the PSP coefficients are positive, the optimized precoder (that maximizes the SNR and improves the BER) will have only coefficients corresponding to the largest coefficients of the PSP greater than zero and the rest are equal to zero.

B. MULTIPLE-USER SCENARIO

Fig. 4 shows the coefficients of the proposed precoders, i.e., matched, orthogonal, and optimal PSP in the case of $K = 3$ users and $n_t = 100$. The users are placed in front of the uniform array at equal distance l_k from the BS but with different positions (angles) $m_1 = 20$, $m_2 = 50$, and $m_3 = 80$. Since we assume l_k is equal for all users, then $\zeta_1 = \zeta_2 = \zeta_3 = \zeta = 15$. In Fig 4-(a), the BS first uses (13) to generate the PSP, $\{\boldsymbol{g}_k\}$, for each user, in which the BS uses them as an input to the three designed precoders. Fig 4-(a) shows the generated PSP for the three users (blue: user 1; red: user 2; black: user 3). We observe that in the matched PSP precoder in Fig 4-(b), the precoder coefficients for the three users overlap significantly. For the orthogonal

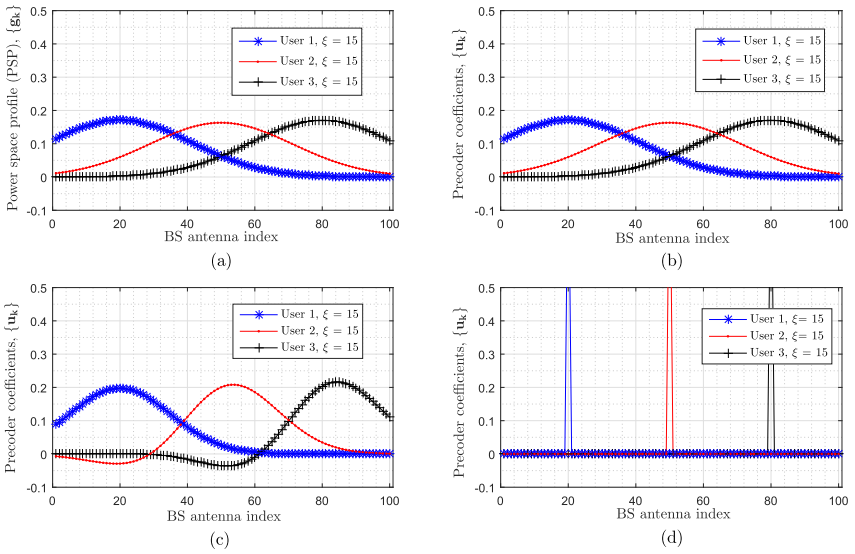

FIGURE 4. The proposed coefficients for the precoders of different users for $K = 3$, $n_t = 100$, and $\zeta = 15$. (a) normalized PSP of different users, (b) normalized matched PSP precoder's coefficients, (c) normalized orthogonal PSP precoder's coefficients, (d) normalized optimal PSP precoder's coefficients.

precoder in Fig 4-(c), the overlap between the precoder coefficients is reduced by using the Gram-Schmidt process. In the optimal PSP precoder in Fig 4-(d), the overlap between the precoder coefficients is minimized and the user is mostly separated. It is worth mentioning that if the following three conditions are satisfied, namely n_t is very large, ζ_k is small, and l_k is small, we have $\boldsymbol{g}_k^T \boldsymbol{u}_q \approx 0$ for $k \neq q$. The value of ζ_k is affected by the user's distance l_k from the BS, the shorter the user's distance to the BS the smaller the value of ζ_k, which minimizes the interference between users.

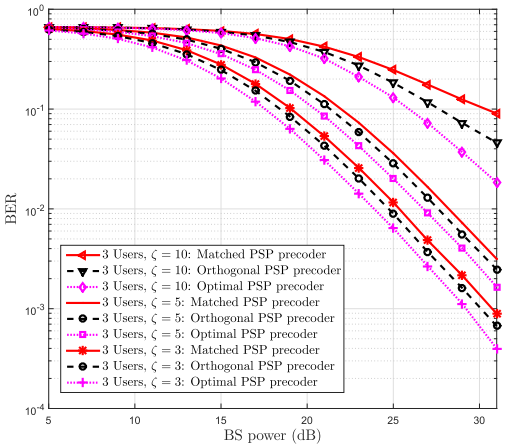

FIGURE 5. BER performance of the proposed differential MIMO downlink transmission with $K = 3$, $n_t = 100$. The values of PSP parameter are $\zeta = 3$, $\zeta = 5$ and $\zeta = 10$.

In Fig. 5, we compare the performance of the proposed PSP precoders in terms of BER. We assume $K = 3$, $m_1 = 20$, $m_2 = 50$, $m_3 = 80$, $n_t = 100$. In Fig. 5, for any value of ζ, the performance of the optimal PSP precoder outperforms the other precoders. The matched PSP precoder is not robust against interference at high BS power and thus has the worst performance. In the case of $\zeta = 3$ for all users, the performance of the precoders is almost the same and this is because of using small value of ζ in which the users do not overlap and hence are separated very well.

In Fig. 5 also, in the presence of interference between users, the value of the power profile parameters such as ζ can impact the precoders' performance. In Fig. 5, we show the effect of adjusting ζ on the performance of the matched, orthogonal, and optimal precoders. Note that when we increase the value of channel variance for all users from $\zeta = 3$ to $\zeta = 5$ and then $\zeta = 10$, the power profile significantly overlaps between users hence causing a degradation in the system performance. Hence, for large orthogonality between users' channels (small value of channel parameter ζ), the performance of matched precoder design is close to the optimal design performance. The larger the orthogonality the closer the performance.

In Fig. 6, we investigate the impact of increasing the number of users on the system performance in terms of BER using the three proposed precoders. We considered: $n_t = 100$, $\zeta = 5$, and $K = 2$, $K = 4$, and $K = 5$. For $K = 2$, the positions are set to [25 75], for $K = 4$, the positions are set to [20 40 60 80], and for $K = 5$, the positions are set to [20 35 50 65 80]. It is shown that differential massive MIMO systems with fewer users outperform those with a large number of users. However, using an optimal PSP precoder with the most appropriate number of n_t and/or value

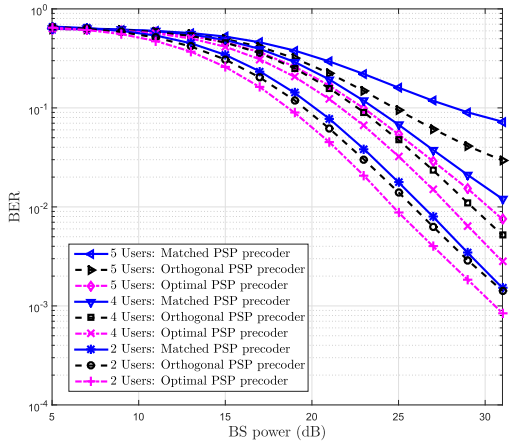

FIGURE 6. BER performance of the proposed differential MIMO downlink transmission with $n_t = 100$, $\zeta = 5$. Users cases are $K = 2$, $K = 4$, and $K = 5$.

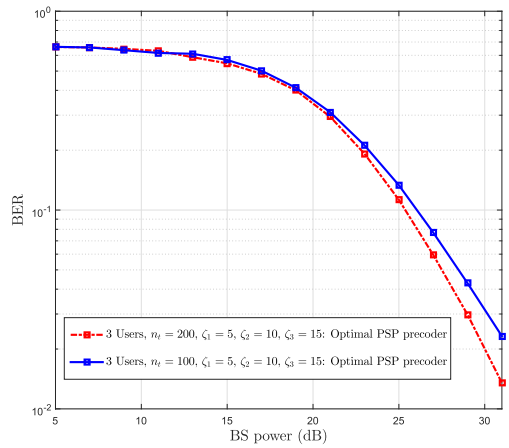

FIGURE 7. BER performance of the proposed differential MIMO downlink transmission with $K = 3$, $n_t = 100$, and $n_t = 200$, using different values of channel variance between users; $\zeta_1 = 5$, $\zeta_2 = 10$, and $\zeta_3 = 15$.

of ζ can minimize the overlap between users and thereby reduce loss of performance.

In Fig. 7, we examine the influence of increasing the number of transmit antennas, e.g., $n_t = 100$ to $n_t = 200$, on the system performance using the optimal PSP precoder. Three users, $K = 3$, are placed in front of the uniform array at different positions [20 50 80] and different distance l_k from the BS, which yields $\zeta_1 = 5$, $\zeta_2 = 10$, and $\zeta_3 = 15$. From Fig. 7, it can be seen that differential massive MIMO systems with higher number of transmit antennas outperform those with lower number of antennas. Therefore, as $n_t \to \infty$ the degree of orthogonality between users becomes large which can minimize the interference between users and improve the overall performance of the system. The larger the number of transmit antennas the better the performance.

In Fig. 8 and Fig. 9, we show the computational complexity of the system. In Fig. 8, we first set the number of users to $K = 6$ and increase the number of transmit antenna n_t.

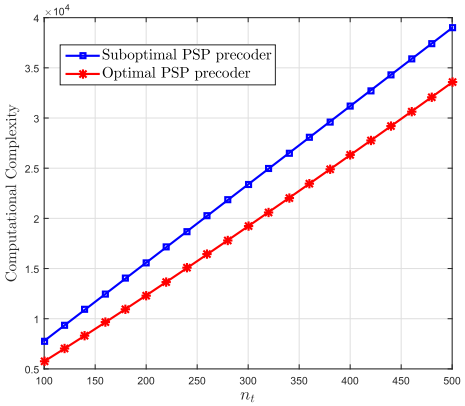

FIGURE 8. Comparison of the computational complexity for suboptimal PSP precoders and optimal PSP precoder with $K = 6$ and $\epsilon = 0.5$.

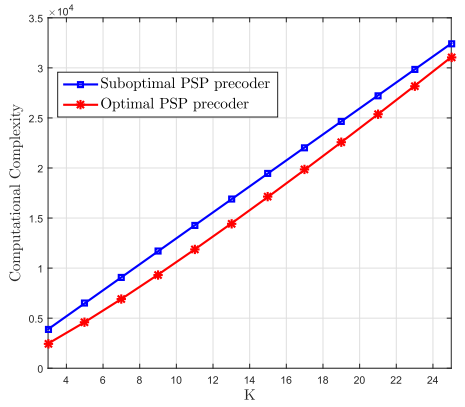

FIGURE 9. Comparison of the computational complexity for suboptimal PSP precoders and optimal PSP precoder with $n_t = 100$ and $\epsilon = 0.5$.

Similarly, in Fig. 9, the number of transmit antennas is fixed to be $n_t = 100$ while the number of users in the system increases gradually. From both figures, the computational complexity of the suboptimal PSP precoders are higher than the optimal PSP precoder. We also observe that varying the number of transmit antennas at the BS has higher impact on the complexity than varying the number of users. Therefore, the optimal PSP precoder yields a low complexity scheme while providing good performance.

VI. CONCLUSION

This paper proposed three precoding schemes, namely the matched, orthogonal and optimal PSP precoders, for downlink transmission in massive MIMO systems with differential encoding and detection. With a large number of transmit antennas at the BS and full knowledge of the PSP, the proposed low-complexity downlink precoding techniques allow MAI between users to be eliminated. In a multiuser scenario, the optimal PSP precoder can effectively separate the data streams of different users, thus enhancing the system performance. In the detection scenario, the DFDD technique is used to detect the differential information signals.

Simulations show that the proposed schemes are effective precoding techniques for a massive MIMO system in a scenario where the channel is unknown at both the transmitter and receiver.

For future work, it is of interest to investigate the following; the estimation of $g_{k,m}$ in real wireless communication systems; the capacity of noncoherent massive MIMO systems and compare it with that of coherent massive MIMO systems; a Rician fading channel could be tested as it is more suitable for the case of small cells; and finally the case of correlated transmitted symbols.

APPENDIX I. PROOFS OF THEOREM 1

The expected value of SINR_k at the kth user can be expressed as follows

$$\text{SINR}_k = \mathbb{E}\left\{\frac{p_k|\boldsymbol{h}_k^H\boldsymbol{u}_k|^2}{\sum_{\substack{q=1\\q\neq k}}^{K} p_q|\boldsymbol{h}_k^H\boldsymbol{u}_q|^2+\sigma_{z_k}^2}\right\} \tag{53}$$

$$= \left(p_k\,\mathbb{E}\left\{|\boldsymbol{h}_k^H\boldsymbol{u}_k|^2\right\}\right)\mathbb{E}\left\{\frac{1}{\sum_{\substack{q=1\\q\neq k}}^{K} p_q|\boldsymbol{h}_k^H\boldsymbol{u}_q|^2+\sigma_{z_k}^2}\right\}. \tag{54}$$

To calculate the expected value of the norm $|\boldsymbol{h}_k^H\boldsymbol{u}_k|^2$, we first expand it as follows

$$|\boldsymbol{h}_k^H\boldsymbol{u}_k|^2 = \sum_{m=1}^{n_t}|\tilde{h}_{k,m}|^2\,g_{k,m}u_{k,m}^2 + \sum_{\mathbb{I}}\tilde{h}_{k,i}^*\tilde{h}_{k,j}\sqrt{g_{k,i}g_{k,j}}u_{k,i}u_{k,j}, \tag{55}$$

where $\mathbb{I} = \big\{\{k,i\}_i \times \{k,j\}_j \mid \{k,i\} \neq \{k,j\}\big\}$. Since $\mathbb{E}\{|\tilde{h}_{k,m}|^2\} = 1$ and $\mathbb{E}\{\tilde{h}_{k,i}^*\tilde{h}_{k,j}\} = 0$ are always true,[4] then we have

$$\mathbb{E}\left\{|\boldsymbol{h}_k^H\boldsymbol{u}_k|^2\right\} = \sum_{m=1}^{n_t} g_{k,m}u_{k,m}^2. \tag{56}$$

By using the result of (56) in (54), we have

$$\text{SINR}_k = \left(p_k\sum_{m=1}^{n_t} g_{k,m}u_{k,m}^2\right) \times \mathbb{E}\left\{\frac{1}{\sum_{\substack{q=1\\q\neq k}}^{K} p_q|\boldsymbol{h}_k^H\boldsymbol{u}_q|^2+\sigma_{z_k}^2}\right\}. \tag{57}$$

Now, we consider the expectation $\mathbb{E}\left\{\frac{1}{\sum_{\substack{q=1\\q\neq k}}^{K} p_q|\boldsymbol{h}_k^H\boldsymbol{u}_q|^2+\sigma_{z_k}^2}\right\}$. By using the Taylor series expansion, we can write this expectation as [29]

$$\mathbb{E}\left\{\frac{1}{\sum_{\substack{q=1\\q\neq k}}^{K} p_q|\boldsymbol{h}_k^H\boldsymbol{u}_q|^2+\sigma_{z_k}^2}\right\} = \mathbb{E}\left\{\frac{X_k}{Y_k}\right\} = \frac{\mathbb{E}\{X_k\}}{\mathbb{E}\{Y_k\}} - \frac{\text{cov}\,(X_k,Y_k)}{(\mathbb{E}\{Y_k\})^2} + \frac{\text{var}\,(Y_k)\,\mathbb{E}\{X_k\}}{(\mathbb{E}\{Y_k\})^2\,\mathbb{E}\{Y_k\}}, \tag{58}$$

where $X_k = 1$ and $Y_k = \sum_{\substack{q=1\\q\neq k}}^{K} p_q|\boldsymbol{h}_k^H\boldsymbol{u}_q|^2 + \sigma_{z_k}^2$. Now, we calculate the values of $\mathbb{E}\{Y_k\}$ and $\text{var}(Y_k)$. Following similar calculation used to obtain (55) and (56), we have $\mathbb{E}\{|\boldsymbol{h}_k^H\boldsymbol{u}_q|^2\} = \sum_{m=1}^{n_t} g_{k,m}u_{q,m}^2$. Therefore

$$\mathbb{E}\{Y_k\} = \sum_{\substack{q=1\\q\neq k}}^{K} p_q\sum_{m=1}^{n_t} g_{k,m}u_{q,m}^2 + \sigma_{z_k}^2. \tag{59}$$

On the other hand, we have

$$\text{var}\{Y_k\} = \mathbb{E}\left\{\left|Y_k - \mathbb{E}\{Y_k\}\right|^2\right\} = \sum_{\substack{q=1\\q\neq k}}^{K} p_q^2\,\mathbb{E}\left\{\left|\sum_{\mathbb{I}}\tilde{h}_{k,i}^*\tilde{h}_{k,j}\sqrt{g_{k,i}g_{k,j}}u_{q,i}u_{q,j}\right|^2\right\} = \sum_{\substack{q=1\\q\neq k}}^{K} p_q^2\sum_{\mathbb{I}} g_{k,i}g_{k,j}u_{q,i}^2u_{q,j}^2. \tag{60}$$

Based on (59), (60) and the orthogonality between $\boldsymbol{g}_k$ and $\boldsymbol{u}_q$ as $n_t \to \infty$ for $k \neq q$, the following inequality $(\mathbb{E}\{Y_k\})^2 \gg \text{var}(Y_k)$ is true. Further, $\text{cov}(X_k, Y_k) = 0$. Applying these results to the series expansion in (58) we get

$$\mathbb{E}\left\{\frac{1}{\sum_{\substack{q=1\\q\neq k}}^{K} p_q|\boldsymbol{h}_k^H\boldsymbol{u}_q|^2+\sigma_{z_k}^2}\right\} \approx \frac{\mathbb{E}\{X_k\}}{\mathbb{E}\{Y_k\}} = \frac{1}{\sum_{\substack{q=1\\q\neq k}}^{K} p_q\sum_{m=1}^{n_t} g_{k,m}u_{q,m}^2+\sigma_{z_k}^2}. \tag{61}$$

Substituting (61) into (57), we obtain

$$\text{SINR}_k = \frac{p_k\sum_{m=1}^{n_t} g_{k,m}u_{k,m}^2}{\sum_{\substack{q=1\\q\neq k}}^{K} p_q\sum_{m=1}^{n_t} g_{k,m}u_{q,m}^2+\sigma_{z_k}^2}. \tag{62}$$

This concludes the proof.

[4] Please note that the expectation is over the fast-fading randomness.

REFERENCES

[1] C. Xu, S. Sugiura, S. X. Ng, P. Zhang, L. Wang, and L. Hanzo, "Two decades of MIMO design tradeoffs and reduced-complexity MIMO detection in near-capacity systems," *IEEE Access*, vol. 5, pp. 18564–18632, May 2017.

[2] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.

[3] K. N. R. S. V. Prasad, E. Hossain, and V. K. Bhargava, "Energy efficiency in massive MIMO-based 5G networks: Opportunities and challenges," *IEEE Wireless Commun.*, vol. 24, no. 3, pp. 86–94, Jun. 2017.

[4] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 7, pp. 1169–1174, Jul. 2000.

[5] R. Chen, J. G. Andews, and R. W. Health, "Multiuser space-time block coded MIMO with downlink precoding," in *Proc. IEEE Int. Conf. Commun.*, vol. 5, Jun. 2004, pp. 2689–2693.
[6] F. Alsifiany, A. Ikhlef, and J. Chambers, "On differential modulation in downlink multiuser MIMO systems," in *Proc. 25th Eur. Signal Process. Conf. (EUSIPCO)*, Aug./Sep. 2017, pp. 558–562.
[7] F. Alsifiany, A. Ikhlef, and J. Chambers, "Exploiting high rate differential algebraic space-time block code in downlink multiuser MIMO systems," *IET Commun.*, vol. 12, no. 17, pp. 2188–2197, Oct. 2018.
[8] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
[9] H. Huh, G. Caire, H. C. Papadopoulos, and S. A. Ramprashad, "Achieving 'massive MIMO' spectral efficiency with a not-so-large number of antennas," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3226–3239, Sep. 2012.
[10] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.
[11] 3GPP. (2015). *TDocs (Written Contributions) at Meeting*. Accessed: Apr. 8, 2019. [Online]. Available: https://www.3gpp.org/DynaReport/TDocExMtg–R1-81–31256.htm
[12] M. K. Simon and M. S. Alouini, "Multiple symbol differential detection with diversity reception," *IEEE Trans. Commun.*, vol. 49, no. 8, pp. 1312–1319, Aug. 2001.
[13] V. Lottici and Z. Tian, "Multiple symbol differential detection for UWB communications," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1656–1666, May 2008.
[14] A. Schenk and R. F. H. Fischer, "Noncoherent detection in massive MIMO systems," in *Proc. 17th Int. ITG Workshop Smart Antennas (WSA)*, Mar. 2013, pp. 1–8.
[15] X. Wang and C. Zhai, "Simultaneous wireless information and power transfer for downlink multi-user massive antenna-array systems," *IEEE Trans. Commun.*, vol. 65, no. 9, pp. 4039–4048, Sep. 2017.
[16] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.
[17] Y. Wu, R. Schober, D. W. K. Ng, C. Xiao, and G. Caire, "Secure massive MIMO transmission with an active eavesdropper," *IEEE Trans. Inf. Theory*, vol. 62, no. 7, pp. 3880–3900, Jul. 2016.
[18] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Mar. 2013.
[19] Y. Wu, S. Jin, X. Gao, M. R. McKay, and C. Xiao, "Transmit designs for the MIMO broadcast channel with statistical CSI," *IEEE Trans. Signal Process.*, vol. 62, no. 17, pp. 4451–4466, Sep. 2014.
[20] T. S. Rappaport, *Wireless Communications: Principles and Practice*, vol. 2. Upper Saddle River, NJ, USA: Prentice-Hall, Jan. 1996.
[21] M. Alageli, A. Ikhlef, and J. Chambers, "Optimization for maximizing sum secrecy rate in MU-MISO SWIPT systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 1, pp. 537–553, Jan. 2018.
[22] M. Alageli, A. Ikhlef, and J. Chambers, "Optimal transmit power minimization in secure MU-MISO SWIPT systems," in *Proc. IEEE 17th Int. Conf. Ubiquitous Wireless Broadband (ICUWB)*, Sep. 2017, pp. 1–7.
[23] M. Grant and S. Boyd. (Apr. 2011). *CVX: MATLAB Software for Disciplined Convex Programming, Version 1.21*. [Online]. Available: cvxr.com/cvx
[24] A. Tarighat, M. Sadek, and A. H. Sayed, "A multi user beamforming scheme for downlink MIMO channels based on maximizing signal-to-leakage ratios," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 3, Mar. 2005, pp. iii-1129–iii-1132.
[25] G. H. Golub and C. F. Van Loan, *Matrix Computations*, vol. 3. Baltimore, MA, USA: Johns Hopkins Univ. Press, 1996.
[26] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. Philadelphia, PA, USA: SIAM, Jan. 2001.
[27] D. Warrier and U. Madhow, "Spectrally efficient noncoherent communication," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 651–668, Mar. 2002.
[28] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. URSI Int. Symp. Signals, Syst., Electron.*, Oct. 1998, pp. 295–300.
[29] M. Alageli, A. Ikhlef, and J. Chambers, "SWIPT massive MIMO systems with active eavesdropping," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 1, pp. 233–247, Jan. 2019.

FAHAD ALSIFIANY received the B.Sc. degree in electrical and electronic engineering from King Abdulaziz University, Jeddah, Saudi Arabia, in 2000, and the M.Sc. degree in telecommunications and networking from the University of Pittsburgh, Pittsburgh, USA, in 2011. He is currently pursuing the Ph.D. degree with the Intelligent Sensing and Communications Group, Newcastle University, Newcastle upon Tyne, U.K. From 2001 to 2015, he was a Lecturer with the Faculty of the King Fahad Security College, Riyadh, Saudi Arabia. His current research interests include noncoherent systems, physical layer security, and massive multiple-input multiple-output (MIMO).

AISSA IKHLEF (M'09–SM'17) was born in Constantine, Algeria. He received the B.S. degree in electrical engineering from the University of Constantine, Constantine, in 2001, the M.Sc. and Ph.D. degrees in electrical engineering from the University of Rennes 1, Rennes, France, in 2004 and 2008, respectively, and the Ph.D. degree from Supélec, France. From 2004 to 2008, he was with Supélec. From 2007 to 2008, he was a Lecturer with the University of Rennes 1. From 2008 to 2010, he was a Postdoctoral Fellow with the Communication and Remote Sensing Laboratory, Catholic University of Louvain, Louvain-La-Neuve, Belgium. In 2009, he was a Visiting Postdoctoral Fellow with The University of British Columbia, Vancouver, BC, Canada, where he was a Postdoctoral Fellow with the Data Communications Group, from 2010 to 2013. From 2013 to 2014, he was a Senior Research Engineer with Toshiba Research Europe Ltd, Bristol, U.K. From 2014 to 2016, he was a Lecturer (Assistant Professor) with the School of Electrical and Electronic Engineering, Newcastle University, Newcastle upon Tyne, U.K. Since 2016, he has been an Assistant Professor with the Department of Engineering, Durham University, Durham, U.K. His current research interests include machine learning, energy harvesting communications, physical layer security, and massive multiple-input multiple-output (MIMO). He has served as an Editor for the IEEE COMMUNICATIONS LETTERS, from 2014 to 2016.

MAHMOUD ALAGELI received the B.Sc. degree (Hons.) in electrical and electronic engineering from the Engineering Academy Tajoura, Tripoli, Libya, in 1999, the M.Eng. degree in communication and computer from the National University of Malaysia, Malaysia, in 2006, and the Ph.D. degree in electrical engineering from Newcastle University, Newcastle upon Tyne, U.K., in 2019. From 2007 to 2011, he was an Assistant Lecturer with the Engineering Academy Tajoura. From 2012 to 2013, he was a Lecturer with the Faculty of Engineering, Garaboulli, Elmergib University, Libya. His current research interests include energy harvesting communications, physical layer security, and massive multiple-input multiple-output (MIMO). He was a recipient of The Bright Bestowal 2006 organized by the Center of Graduate Studies, National University of Malaysia.

JONATHON CHAMBERS (S'83–M'90–SM'98–F'11) received the Ph.D. and D.Sc. degrees in signal processing from the Imperial College of Science, Technology and Medicine (Imperial College London), London, U.K., in 1990 and 2014, respectively.

From 1991 to 1994, he was a Research Scientist with the Schlumberger Cambridge Research Center, Cambridge, U.K. In 1994, he returned to Imperial College London as a Lecturer in signal processing and was promoted to Reader (Associate Professor), in 1998. From 2001 to 2004, he was the Director of the Center for Digital Signal Processing and a Professor of signal processing with the Division of Engineering, King's College London. From 2004 to 2007, he was a Cardiff Professorial Research Fellow with the School of Engineering, Cardiff University, Cardiff, U.K. From 2007 to 2014, he led the Advanced Signal Processing Group, School of Electronic, Electrical and Systems Engineering, Loughborough University, where he is currently a Visiting Professor. In 2015, he joined the School of Electrical and Electronic Engineering, Newcastle University, where he was a Professor of signal and information processing and led the ComS2IP Group and is also a Visiting Professor. In 2017, he became the Head of the Department of Engineering, University of Leicester. He is also an International Honorary Dean and a Guest Professor with Harbin Engineering University, China, with support from the 1000 Talents Scheme. He is a coauthor of the books: *Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability* (New York, NY, USA: Wiley, 2001) and *EEG Signal Processing* (New York, NY, USA: Wiley, 2007). He has advised more than 80 researchers through to Ph.D. graduation and published more than 500 conference papers and journal articles, many of them are in the IEEE journals. His research interests include adaptive signal processing, and machine learning and their applications. He is also a Fellow of the Royal Academy of Engineering, U.K., and the Institution of Electrical Engineers. In 2007, he received the first QinetiQ Visiting Fellowship for his outstanding contributions to adaptive signal processing and his contributions to QinetiQ, as a result of his successful industrial collaboration with the international defense systems company QinetiQ. He was the Technical Program Chair of the 15th International Conference on Digital Signal Processing and the 2009 IEEE Workshop on Statistical Signal Processing, both held at Cardiff, U.K., and a Technical Program Co-Chair of the 36th IEEE International Conference on Acoustics, Speech, and Signal Processing, Prague, Czech Republic. He has served on the IEEE Signal Processing Theory and Methods Technical Committee for six years and the IEEE Signal Processing Society Awards Board for three years, together with the Jack Kilby Award Committee. He has served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING for two terms over the periods 1997–1999 and 2004–2007 and as a Senior Area Editor, from 2011 to 2014.