

# Multi-user Linear Precoding for Multi-polarized Massive MIMO System under Imperfect CSIT

Jaehyun Park, *Member, IEEE*, Bruno Clerckx, *Member, IEEE*,

**Abstract**—The space limitation and the channel acquisition prevent Massive MIMO from being easily deployed in a practical setup. Motivated by current deployments of LTE-Advanced, the use of multi-polarized antenna elements can be an efficient solution to address the space constraint. Furthermore, the dual-structured precoding, in which a preprocessing based on the spatial correlation and a subsequent linear precoding based on the short-term channel state information at the transmitter (CSIT) are concatenated, can reduce the feedback overhead efficiently. By grouping and preprocessing spatially correlated mobile stations (MSs), the dimension of the precoding signal space is reduced and the corresponding short-term CSIT dimension is reduced. In this paper, to reduce the feedback overhead further, we propose a dual-structured multi-user linear precoding, in which the subgrouping method based on co-polarization is additionally applied to the spatially grouped MSs in the preprocessing stage. Furthermore, under imperfect CSIT, the proposed scheme is asymptotically analyzed based on random matrix theory. By investigating the behavior of the asymptotic performance, we also propose a new dual-structured precoding in which the precoding mode is switched between two dual-structured precoding strategies with i) the preprocessing based only on the spatial correlation and ii) the preprocessing based on both the spatial correlation and polarization. Finally, we extend it to 3D dual-structured precoding.

**Index Terms**—Multi-polarized Massive MIMO, Dual structured precoding with long-term/short-term CSIT

## I. INTRODUCTION

By deploying a large number of antenna elements at the base station (BS), high spectral efficiencies can be achieved. Furthermore, because low-complexity linear precoding schemes can be efficiently exploited to serve multiple mobile stations (MSs) simultaneously, Massive MIMO plays a key role in beyond 4G cellular networks ([1]–[4] and references therein).

Assuming large scale arrays, several linear single-user and multi-user precoding schemes are asymptotically analyzed by using random matrix theory in [5]–[9]. In [5], single-user beamforming in MISO with a large number of transmit antenna elements has been analyzed under the per-antenna constant-envelope constraints and its extension to multi-user MIMO is treated in [6]. In [7], by utilizing a Stieltjes transform of random positive semi-definite matrices, the asymptotic

signal-to-interference-and-noise ratio (SINR) of linear precoding in a correlated Massive MISO broadcasting channel has been derived under imperfect channel state information at the transmitter (CSIT). In [9], by considering multi-cell downlink with massive transmit antenna elements, the asymptotic SINR is analyzed by using random matrix theory which gives an insight about the cooperative transmission strategy of BSs.

However, in FDD, where channel reciprocity is not exploitable, the multi-antenna channel acquisition at the transmitter prevents Massive MIMO from being easily deployed. To resolve the channel acquisition burden at BS, a dual structured precoding, in which a preprocessing based on the long-term CSIT (mainly, spatial correlation) and a subsequent linear precoding based on the short-term CSIT (that generally has lower dimension than the number of transmit antenna elements) are concatenated, can be exploited to reduce the feedback overhead efficiently [8], [10]. Note that the long-term CSI is slowly-varying and can be obtained accurately with a low feedback overhead. Because of the low feedback overhead and the attractive performance, the dual structured precoding has been also considered in the 4G and beyond 4G wireless standards [10], [11]. In [8], when MSs are clustered as several spatially correlated groups, joint spatial division and multiplexing scheme has been proposed in which the precoding matrix is composed of the *prebeamforming* matrix based on the spatial correlation and the *classical precoder* based on short-term CSIT. Furthermore, its performance is asymptotically analyzed for a large number of transmit antenna elements.

Another challenge of the Massive MIMO system is the antenna space limitation. An increasing number of antenna elements is difficult to be packed in a limited space and if it can be deployed, the high spatial correlation and the mutual coupling among the antennas elements may cause some system performance degradation, especially for a small numbers of active MSs [12], [13]. The multi-polarized antenna elements can be one solution to alleviate the space constraint [14], [15]. The multi-polarized antenna systems have been investigated under various communication scenarios including the picocell/microcell [16], indoor/outdoor [17], and the line of sight (LOS)/non LOS (NLOS) [15] environments. However, despite the importance of polarized antennas in practical deployments, the Massive MIMO system with multi-polarized antenna elements has not been addressed so far together with the multi-user linear precoding. Note that due to space constraints, closely spaced dual-polarized antennas is considered as the first priority deployment scenario for MIMO in LTE-A and is therefore likely to remain so as the number

A part of this work was published in the 22nd European Signal Processing Conference EUSIPCO 2014.

J. Park is with the Department of Electronic Engineering, Pukyong National University, Republic of Korea. B. Clerckx is with the Department of Electrical and Electronic Engineering, Imperial College London, United Kingdom and also with School of Electrical Engineering, Korea University (e-mail: jaehyun@pku.ac.kr, b.clerckx@imperial.ac.uk). B. Clerckx is the corresponding author.

This work is partially supported by Huawei Technologies Co., Ltd.

of antennas at the base station increases [10], [11], [18].

In this paper, we first model the Massive multi-user MIMO system, where the BS is equipped with a large number of multi-polarized antenna elements. For simplicity and practical issue of MS, we consider that MSs are equipped with a single single-polarized antenna element. We then present the dual structured precoding based on long-term/short-term CSIT. As done in the Massive MIMO system with a single-polarized antenna element [8], by grouping the spatially correlated MSs and multiplying the channel matrix of the grouped MSs with the same preprocessing matrix based on spatial correlation, the dimension of the precoding signal space is reduced and the corresponding short-term CSIT dimension can also be reduced from the Karhunen-Loeve transform [19]. Considering the multi-polarized Massive multi-user MIMO system for the first time, the contributions of this paper are listed below:

- To reduce the feedback overhead further, we first propose a dual structured linear precoding, in which the subgrouping method based on the polarization is additionally applied to the spatially grouped MSs in the preprocessing stage. That is, by subgrouping co-polarized MSs in each group, we let the MS report the CSI from the transmit antenna elements having the same polarization as its polarization. Then, the MS can further reduce the short-term CSI feedback overhead, compared to the case with the conventional preprocessing based only on the spatial correlation.
- Under the imperfect CSIT, two different dual structured precodings with preprocessing of i) grouping based only on the spatial correlation (i.e., spatial grouping) and ii) subgrouping based on both the spatial correlation and polarization are asymptotically analyzed based on random matrix theory with a large dimension [7]. Because, in this paper, the polarized Massive MU-MIMO channel is considered, the asymptotic inter/intra interferences are evaluated over the polarization domain as well as the spatial domain, which addresses a more general (polarized) channel environment compared to [7], [8]. Accordingly, the asymptotic performance can be further analyzed in terms of both the polarization and the spatial correlation and therefore, we can understand the performance behavior of dual structured precoding with respect to the long-term CSIT.
- We then propose a new dual precoding method to switch the precoding mode between two dual structured precoding strategies relying on i) the spatial grouping only and ii) the subgrouping based on both the spatial correlation and polarization.
- Motivated by 3D beamforming [8], [20], we extend the design to the 3D dual structured precoding in which the spatial correlation depends on both azimuth and elevation angles.
- Finally, we also discuss how we can modify the proposed precoding mode switching scheme when the polarization at the BS and MS is mismatched (due to e.g. random MS orientation).

Note that, from the asymptotic results, we can find that even

though the proposed dual precoding using the subgrouping can reduce the feedback overhead, its performance can be affected by the cross-polar discrimination (XPD) parameter. Here, XPD refers to the long-term statistics of the antenna elements and channel depolarization that measures the ability to distinguish the orthogonal polarization. That is, under the same feedback overhead, the dual precoding with subgrouping can utilize more accurate CSIT on half of the array, compared with precoding with spatial grouping, but exhibits performance more sensitive to the XPD. The precoding with spatial grouping can only utilize less accurate CSIT, but its performance is not affected by the XPD. Accordingly, we identify the region where the dual precoding with subgrouping outperforms that with spatial grouping. The region depends on the XPD, the spatial correlation, and the short-term CSIT quality, which motivated us to develop the new dual precoding method.

The rest of this paper is organized as follows. In Section II, we introduce the Massive MIMO system model with multiple multi-polarized antenna elements at the BS and a (either vertically or horizontally) single polarized antenna element at the multiple MSs. In Section III, we discuss the dual structured precoding based on the long-term/short-term CSIT. In Section IV, we investigate the asymptotic performance of the dual precoding schemes and their behavior over the XPD parameter. Based on the asymptotic results, in Section V, we propose a new dual structured precoding/feedback. In Section VI, we provide several discussion and simulation results, respectively, and in Section VII we give our conclusions.

Throughout the paper, matrices and vectors are represented by bold capital letters and bold lower-case letters, respectively. The notations  $(\mathbf{A})^T$ ,  $(\mathbf{A})^H$ ,  $(\mathbf{A})_i$ ,  $[\mathbf{A}]_i$ ,  $tr(\mathbf{A})$ , and  $\det(\mathbf{A})$  denote the transpose, conjugate transpose, the  $i$ th row, the  $i$ th column, the trace, and the determinant of a matrix  $\mathbf{A}$ , respectively. In addition,  $[\mathbf{A}]_{i:j}$  (resp.,  $(\mathbf{A})_{i:j}$ ) denotes the submatrix from the  $i$ th column (resp., row) to the  $j$ th column (resp., row) of  $\mathbf{A}$ . The matrix norm  $\|\mathbf{A}\|$  and the vector norm  $\|\mathbf{a}\|$  denote the 2-norms of a matrix  $\mathbf{A}$  and a vector  $\mathbf{a}$ , respectively. In addition,  $\mathbf{A} \succeq 0$  means that a matrix  $\mathbf{A}$  is positive semi-definite,  $\otimes$  denotes the Kronecker product, and  $\odot$  denotes the Hadamard product. The operation  $E_g[A_g]$  means the average of  $A_g$  over index  $g$ . Finally,  $\mathbf{I}_M$ ,  $\mathbf{1}_{M \times N}$ , and  $\mathbf{0}_{M \times N}$  denote the  $M \times M$  identity matrix, the  $M$  by  $N$  matrix with all 1 entries, and the  $M$  by  $N$  matrix with all 0 entries, respectively.

## II. SYSTEM MODEL

We consider a single-cell downlink system with one BS with  $M$  polarized antenna elements and  $N$  active MSs, each with a single polarized antenna element, where  $M$  and  $N$  are assumed to be even numbers. As in Fig 1, the BS has  $\frac{M}{2}$  pairs of co-located vertically/horizontally polarized antenna elements and the MSs have a single antenna element with either vertical or horizontal polarization<sup>1</sup>. Furthermore, since

<sup>1</sup>Throughout the paper, we consider the dual-polarized antenna elements at the BS for ease of explanation, but our approach can be extended to the tri-polarized case without difficulty. In addition, all the antenna elements are perfectly aligned with either vertical or horizontal polarization. In Section V-C, we discuss the polarization mismatch due to the random orientation of MSs.

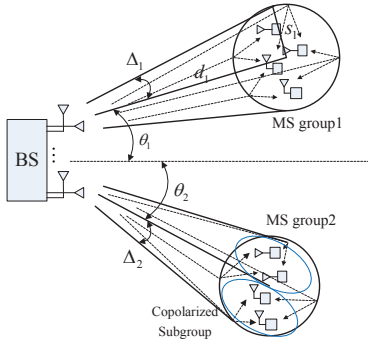


Fig. 1. Block diagram of a dual-polarized multi-user downlink system.

human activity is usually confined in small clustered regions such as buildings, locations of MSs tend to be spatially clustered, e.g.,  $G$  groups. Then, the received signal  $\mathbf{y}_g \in \mathbb{C}^{N_g}$  of the  $g$ th group with an assumption of flat-fading channel is given by

$$\mathbf{y}_g = \begin{bmatrix} \mathbf{y}_g^v \\ \mathbf{y}_g^h \end{bmatrix} = \mathbf{H}_g^H \mathbf{x} + \mathbf{n}_g, \quad (1)$$

where  $\mathbf{y}_g^v$  and  $\mathbf{y}_g^h$  are the received signal for MSs with vertical and horizontal polarization, respectively, and  $\mathbf{n}_g = \begin{bmatrix} \mathbf{n}_g^v \\ \mathbf{n}_g^h \end{bmatrix}$  is a zero-mean complex Gaussian noise vector having a covariance matrix  $\mathbf{I}_{N_g}$ , denoted as  $\mathbf{n}_g \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_g})$ . Here,  $N_g$  denotes the number of MSs in the  $g$ th group. For simplicity, it is assumed that  $N_1 = \dots = N_G = \bar{N}$ , where  $\bar{N}$  is even, and both  $\mathbf{y}_g^v$  and  $\mathbf{y}_g^h$  are  $\frac{\bar{N}}{2} \times 1$  vectors. The channel of the  $k$ th MS in the  $g$ th group is then given as  $\mathbf{h}_{gk} = [\mathbf{H}_g]_k$ . The  $M \times 1$  vector,  $\mathbf{x}$ , is the linearly precoded transmit signal expressed as

$$\mathbf{x} = \sum_{g=1}^G \mathbf{V}_g \mathbf{d}_g, \quad \mathbf{V}_g \in \mathbb{C}^{M \times \bar{N}}, \quad (2)$$

where  $\mathbf{V}_g$  and  $\mathbf{d}_g = \begin{bmatrix} \mathbf{d}_g^v \\ \mathbf{d}_g^h \end{bmatrix}$  are the linear precoding matrix and the data symbol vector for the MSs in the  $g$ th group, respectively. The precoded signal  $\mathbf{x}$  should satisfy the power constraint  $E[\|\mathbf{x}\|^2] \leq P$ .

### A. Channel model

By using the Karhunen-Loeve transform [19] and the polarized MIMO channel modeling with infinitesimally small antenna elements described in [10], [14], the downlink channel to the  $g$ th group,  $\mathbf{H}_g$ , can be represented as

$$\mathbf{H}_g = \left( \begin{bmatrix} 1 & r_{xp} \\ r_{xp} & 1 \end{bmatrix} \otimes (\mathbf{U}_g \mathbf{\Lambda}_g^{\frac{1}{2}}) \right) \left( \mathbf{G}_g \odot (\mathbf{X} \otimes \mathbf{1}_{r \times \frac{\bar{N}}{2}}) \right) \quad (3)$$

where  $r_{xp}$  is the correlation coefficient between vertically and horizontally polarized antenna elements,  $\mathbf{\Lambda}_g$  is an  $r_g \times r_g$  diagonal matrix with the non-zero eigenvalues of the spatial covariance matrix  $\mathbf{R}_g^s$  for the  $g$ th group<sup>2</sup>, where the eigenvalues are assumed to be ordered in decreasing order of

<sup>2</sup>For simplicity, it is assumed that the spatial covariance matrix is the same for both polarizations.

magnitude. Note that generally,  $r_g \ll M$ , and  $\mathbf{U}_g \in \mathbb{C}^{\frac{M}{2} \times r_g}$  is composed of the eigenvectors of  $\mathbf{R}_g^s$  as its columns. The matrix  $\mathbf{G}_g$  is defined as

$$\mathbf{G}_g = \begin{bmatrix} \mathbf{G}_g^{vv} & \mathbf{G}_g^{hv} \\ \mathbf{G}_g^{vh} & \mathbf{G}_g^{hh} \end{bmatrix}, \quad (4)$$

and the elements of  $\mathbf{G}_g^{pq} \in \mathbb{C}^{r_g \times \frac{\bar{N}}{2}}$ ,  $p, q \in \{h, v\}$  are i.i.d. complex Gaussian distributed with zero mean and unit variance. The matrix  $\mathbf{X}$  describes the power imbalance between the orthogonal polarizations and is given as

$$\mathbf{X} = \begin{bmatrix} 1 & \sqrt{\chi} \\ \sqrt{\chi} & 1 \end{bmatrix}, \quad (5)$$

where the parameter  $0 \leq \chi \leq 1$  is the inverse of the XPD, where  $1 \leq \text{XPD} \leq \infty$ . Note that, based on the reported measurement in [21], [22],  $r_{xp} \approx 0$ . Accordingly, (3) can be rewritten as

$$\begin{aligned} \mathbf{H}_g &= \left( \mathbf{I}_2 \otimes (\mathbf{U}_g \mathbf{\Lambda}_g^{\frac{1}{2}}) \right) \begin{bmatrix} \mathbf{G}_g^{vv} & \sqrt{\chi} \mathbf{G}_g^{hv} \\ \sqrt{\chi} \mathbf{G}_g^{vh} & \mathbf{G}_g^{hh} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{U}_g \mathbf{\Lambda}_g^{\frac{1}{2}} \mathbf{G}_g^{vv} & \sqrt{\chi} \mathbf{U}_g \mathbf{\Lambda}_g^{\frac{1}{2}} \mathbf{G}_g^{hv} \\ \sqrt{\chi} \mathbf{U}_g \mathbf{\Lambda}_g^{\frac{1}{2}} \mathbf{G}_g^{vh} & \mathbf{U}_g \mathbf{\Lambda}_g^{\frac{1}{2}} \mathbf{G}_g^{hh} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{H}_g^{vv} & \mathbf{H}_g^{hv} \\ \mathbf{H}_g^{vh} & \mathbf{H}_g^{hh} \end{bmatrix}, \end{aligned} \quad (6)$$

and its covariance matrix is given as

$$\begin{aligned} \mathbf{R}_g &= \begin{bmatrix} (1 + \chi) \mathbf{R}_g^s & \mathbf{0} \\ \mathbf{0} & (1 + \chi) \mathbf{R}_g^s \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{R}_g^s & \mathbf{0} \\ \mathbf{0} & \chi \mathbf{R}_g^s \end{bmatrix} + \begin{bmatrix} \chi \mathbf{R}_g^s & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_g^s \end{bmatrix} = \mathbf{R}_{gv} + \mathbf{R}_{gh}, \end{aligned} \quad (7)$$

where  $\mathbf{R}_{gv}$  and  $\mathbf{R}_{gh}$  are the covariance matrices of the vertically and the horizontally co-polarized MS subgroups, respectively.

Note that the long-term parameters  $\mathbf{R}_g^s$  and  $\chi$  are slowly-varying and assumed to be obtained accurately with a low feedback overhead. However, the short-term CSI parameter  $\mathbf{G}_g$  is varying independently over the short-term coherence time. The feedback to the BS of the CSI is imperfect (due to e.g. quantization) and incurs a significant overhead. Accordingly, the imperfect CSIT  $\hat{\mathbf{G}}_g$  available at the transmitter is modeled as

$$\hat{\mathbf{G}}_g = \sqrt{1 - \tau_g} \mathbf{G}_g + \tau_g \mathbf{Z}_g, \quad (8)$$

where the elements of  $\mathbf{Z}_g$  are complex Gaussian distributed with zero mean and unit variance and  $\tau_g \in [0, 1]$  indicates the accuracy of available CSIT for the  $g$ th group. That is, the case of  $\tau_g = 0$  implies the perfect CSIT. From (6),  $\hat{\mathbf{H}}_g$  and  $\hat{\mathbf{H}}_g^{pp}$  can be defined as the imperfect CSI knowledge of  $\mathbf{H}_g$  and  $\mathbf{H}_g^{pp}$  at the transmitter, respectively, by using  $\hat{\mathbf{G}}_g$ . Throughout the paper, it is assumed that  $\tau_1 = \dots = \tau_G = \tau$ , but it can be easily extended to the scenario that  $\tau_i \neq \tau_j$  for  $i \neq j$ .

*Remark 1:* The imperfect channel model (8) comes from the scenario that when both BS and MS know the long-term statistics perfectly (i.e.,  $\mathbf{R}_g^s$ ), the  $k$ th MS in the  $g$ th group quantizes  $\mathbf{g}_{gk} = [\mathbf{G}_g]_k$  by using the random codebook [23] and feeds the codeword index back to the BS. Note that the

$k$ th MS in the  $g$ th group can have a much smaller feedback overhead by sending the essential channel information of  $\mathbf{g}_{gk} \in \mathbb{C}^{2r_g \times 1}$  rather than  $\mathbf{h}_{gk}$ .

### III. DUAL STRUCTURED PRECODING BASED ON LONG-TERM/SHORT-TERM CSIT

Thanks to the computational complexity reduction and the feedback overhead reduction (i.e., the dimension reduction using long-term statistics), the dual precoding scheme based on long-term/short-term CSIT has been widely utilized [8], [10], [11]. That is, the precoding matrix for the  $g$ th group is given as

$$\mathbf{V}_g = \mathbf{B}_g \mathbf{P}_g, \quad (9)$$

where  $\mathbf{B}_g \in \mathbb{C}^{M \times \bar{B}}$  is the preprocessing matrix based on the long-term channel statistics with  $\bar{N} \leq \bar{B} \leq 2r_g \ll M$  and  $\mathbf{P}_g \in \mathbb{C}^{\bar{B} \times \bar{N}}$  is the precoding matrix for the effective (instantaneous) channel  $\mathbf{H}_g^H \mathbf{B}_g$ . Here,  $\bar{B}$  is a design parameter that determines the dimension of the transformed channel using the long-term CSIT. The system equation (1) can then be rewritten as

$$\mathbf{y}_g = \mathbf{H}_g^H \mathbf{B}_g \mathbf{P}_g \mathbf{d}_g + \sum_{l=1, l \neq g}^G \mathbf{H}_g^H \mathbf{B}_l \mathbf{P}_l \mathbf{d}_l + \mathbf{n}_g. \quad (10)$$

In what follows, we introduce the conventional dual structured precoding scheme with the preprocessing using block diagonalization (BD) based on spatial correlation and the regularized ZF precoding for each decoupled group. Then, we propose the dual precoding scheme with BD and subgrouping (BDS) exploiting both the spatial correlation and the polarization (another long-term channel statistics parameter).

#### A. Preprocessing using block diagonalization based on spatial correlation

To null out the leakage to other groups, it is desirable that the preprocessing matrix  $\mathbf{B}_g$  based on the spatial correlation is designed as

$$\mathbf{H}_l^H \mathbf{B}_g \approx \mathbf{0}, \text{ for } l \neq g. \quad (11)$$

Then,  $\mathbf{P}_g$  in (10) can be computed based on the decoupled system model  $\mathbf{y}_g \approx \mathbf{H}_g^H \mathbf{B}_g \mathbf{P}_g \mathbf{d}_g + \mathbf{n}_g$  where the inter-group interferences have been eliminated.<sup>3</sup>

To obtain  $\mathbf{B}_g$  satisfying the condition (11), the BD can be utilized. That is, due to the block diagonal structure in (7), we first define

$$\mathbf{U}_{-g} = [\mathbf{U}_1^a, \dots, \mathbf{U}_{g-1}^a, \mathbf{U}_{g+1}^a, \dots, \mathbf{U}_G^a] \in \mathbb{C}^{\frac{M}{2} \times \sum_{l \neq g} r_l^a}, \quad (12)$$

where  $\mathbf{U}_g^a = [\mathbf{U}_g]_{1:r_g^a}$  and  $r_g^a (\leq r_g)$  is a design parameter reflecting the number of *dominant* eigenvalues of  $\mathbf{R}_g^s$ . That is,

<sup>3</sup>Note that if the BS has a large number of antenna elements and the number of antenna elements at the BS is larger than the number of MSs, we can find the ‘‘approximated’’ null-space satisfying (11), in general. Furthermore, from [8], [24] considering the one-ring channel model (see also Section VI), if the angle-of-departures (AoDs) of the multipaths from different groups are disjointed, the spatial covariance matrices of different groups become asymptotically orthogonal to each other as the number of antenna elements increases.

if we increase  $r_g^a$  close to  $r_g$ , the BD can find the subspace more orthogonal to the signal subspace spanned by other groups’ channel (the perfect orthogonality is guaranteed when  $r_g^a = r_g$ ), while the dimension of corresponding orthogonal subspace decreases as  $\frac{M}{2} - \sum_{l \neq g} r_l^a$ .

The matrix  $\mathbf{U}_{-g}$  in (12) then has a singular value decomposition (SVD) as

$$\mathbf{U}_{-g} = [\mathbf{E}_{-g}^{(1)}, \mathbf{E}_{-g}^{(0)}] \begin{bmatrix} \mathbf{\Lambda}_{-g}^{(1)} & \\ & \mathbf{\Lambda}_{-g}^{(0)} \end{bmatrix} \mathbf{V}_{-g}^H, \mathbf{E}_{-g}^{(0)} \in \mathbb{C}^{\frac{M}{2} \times \frac{M}{2} - \sum_{l \neq g} r_l^a}, \quad (13)$$

where  $\mathbf{E}_{-g}^{(1)}$  (respectively,  $\mathbf{E}_{-g}^{(0)}$ ) is the left singular vectors associated with the  $\sum_{l \neq g} r_l^a$  dominant (respectively,  $\frac{M}{2} - \sum_{l \neq g} r_l^a$  non-dominant) singular values  $\mathbf{\Lambda}_{-g}^{(1)}$  (respectively,  $\mathbf{\Lambda}_{-g}^{(0)}$ ). Then, because  $(\mathbf{E}_{-g}^{(0)})^H \mathbf{U}_{-g} = \mathbf{0}$ , by defining  $\tilde{\mathbf{H}}_g = (\mathbf{I}_2 \otimes \mathbf{E}_{-g}^{(0)})^H \mathbf{H}_g$ ,  $\tilde{\mathbf{H}}_g$  is orthogonal to the dominant eigen-space spanned by other groups’ channel. Note that the covariance matrix of  $\tilde{\mathbf{H}}_g$  is then given by

$$\tilde{\mathbf{R}}_g = (\mathbf{I}_2 \otimes \mathbf{E}_{-g}^{(0)})^H \mathbf{R}_g (\mathbf{I}_2 \otimes \mathbf{E}_{-g}^{(0)}), \quad (14)$$

and by defining  $\tilde{\mathbf{R}}_g^s = (\mathbf{E}_{-g}^{(0)})^H \mathbf{R}_g^s \mathbf{E}_{-g}^{(0)}$ , we have its eigenvalue decomposition (EVD) as

$$\tilde{\mathbf{R}}_g^s = \mathbf{F}_g \tilde{\mathbf{\Lambda}}_g \mathbf{F}_g^H, \quad (15)$$

where  $\mathbf{F}_g$  is the eigenvectors of  $\tilde{\mathbf{R}}_g^s$ . Then by letting  $\mathbf{F}_g^{(1)} = [\mathbf{F}_g]_{1:\frac{\bar{B}}{2}}$ , the preprocessing matrix can be given as

$$\mathbf{B}_g = \mathbf{I}_2 \otimes \mathbf{B}_g^s, \quad \mathbf{B}_g^s = \mathbf{E}_{-g}^{(0)} \mathbf{F}_g^{(1)}. \quad (16)$$

Accordingly, through the preprocessing matrix  $\mathbf{B}_g$ , we can transform the transmit signal for the  $g$ th group into the  $\bar{B}$  dimensional dominant eigen-space that is orthogonal to the subspace spanned by other groups’ channel. Note that, from (9) and (12),  $\bar{B}$  and  $r_g^a$  should be chosen properly to satisfy the conditions of  $\bar{N} \leq \bar{B} \leq 2(\frac{M}{2} - \sum_{l \neq g} r_l^a)$  and  $\bar{B} \leq 2r_g$ . Without loss of generality, we assume that  $r_1^a = \dots = r_G^a = r$  with a fixed  $r$  satisfying the above two constraints.

#### B. Multi-user precoding for each decoupled group

Because, from (10), the effective channel for the  $g$ th group is  $\bar{\mathbf{H}}_g = \mathbf{B}_g^H \mathbf{H}_g$ , the corresponding covariance matrix is given by

$$\begin{aligned} \bar{\mathbf{R}}_g &= \mathbf{B}_g^H \mathbf{R}_g \mathbf{B}_g = \mathbf{B}_g^H (\mathbf{R}_{g^v} + \mathbf{R}_{g^h}) \mathbf{B}_g = \bar{\mathbf{R}}_{g^v} + \bar{\mathbf{R}}_{g^h} \\ &= (1 + \chi) \begin{bmatrix} (\mathbf{B}_g^s)^H \mathbf{R}_g^s \mathbf{B}_g^s & \mathbf{0} \\ \mathbf{0} & (\mathbf{B}_g^s)^H \mathbf{R}_g^s \mathbf{B}_g^s \end{bmatrix}. \end{aligned} \quad (17)$$

Furthermore, due to the preprocessing, the interferences from other groups in (10) are almost nulled out. Accordingly, the precoding matrix  $\mathbf{P}_g$  is designed such that the intra-group interferences are nulled out based on the short-term CSIT of the  $g$ th group. That is, assuming the equal power allocation, the regularized ZF precoding matrix [8], [25] with imperfect CSIT can be computed as

$$\mathbf{P}_g = \xi_g \hat{\mathbf{K}}_g \hat{\mathbf{H}}_g, \quad (18)$$

where

$$\hat{\mathbf{K}}_g = \left( \hat{\mathbf{H}}_g \hat{\mathbf{H}}_g^H + \bar{B} \alpha \mathbf{I}_{\bar{B}} \right)^{-1}, \quad \hat{\mathbf{H}}_g = \mathbf{B}_g^H \hat{\mathbf{H}}_g. \quad (19)$$

Here,  $\hat{\mathbf{H}}_g$  is the effective channel estimate that is available at the BS and  $\alpha$  is a regularization parameter. Throughout the paper, it is set as  $\alpha = \frac{\bar{N}}{\bar{B}P}$ , which is equivalent with the MMSE linear filter [26]. The normalization factor  $\xi_g$  is then given as

$$\xi_g^2 = \frac{\bar{N} \frac{P}{N}}{\frac{P}{N} \text{tr}(\hat{\mathbf{H}}_g^H \hat{\mathbf{K}}_g^H \mathbf{B}_g^H \mathbf{B}_g \hat{\mathbf{K}}_g \hat{\mathbf{H}}_g)} = \frac{\bar{N}}{\text{tr}(\hat{\mathbf{H}}_g^H \hat{\mathbf{K}}_g^H \hat{\mathbf{K}}_g \hat{\mathbf{H}}_g)}, \quad (20)$$

where the second equality is due to the fact that  $\mathbf{B}_g^H \mathbf{B}_g = \mathbf{I}_{\bar{B}}$  from (16). Denoting  $\hat{\mathbf{h}}_{gk} = [\hat{\mathbf{H}}_g]_k$  as the effective channel estimate of the  $k$ th MS in the  $g$ th group, the SINR of the  $k$ th MS in the  $g$ th group with  $p$  polarization is then given as shown at the top of the next page. Accordingly, the sum rate is given by

$$R_{BD} = \sum_{g=1}^G \sum_{p \in \{v, h\}} \sum_{k=1}^{\bar{N}/2} \log_2(1 + \gamma_{gpk}^{BD}), \quad (22)$$

where the subscript and superscript  $BD$  indicate the dual precoding with Block Diagonalization based on spatial correlation.

### C. Dual precoding using block diagonalization and subgrouping based on both spatial correlation and polarization

In Section III-A, the preprocessing matrix is computed based only on spatial correlation. However, when  $\chi$  becomes small (i.e., the antenna elements can favorably discriminate the orthogonally polarized signals), the interference signals through the cross-polarized channels can be naturally nulled out. This suggests that we can make the subgroups of co-polarized MSs in each group (see the second MS group in Fig. 1.) and let the BS precode the signal for the co-polarized subgroup by using the short-term CSIT of the transmit antenna elements having the same polarization with the associated subgroup. That is, from (1) and (10), the received signal for the co-polarized subgroup with  $p$  polarization, for  $p \in \{h, v\}$ , in the  $g$ th group can be written as

$$\begin{aligned} \mathbf{y}_g^p &= \mathbf{H}_{gp}^H \mathbf{B}_{gp} \mathbf{P}_{gp} \mathbf{d}_g^p + \sum_{\substack{q \in \{h, v\} \\ q \neq p}} \mathbf{H}_{gp}^H \mathbf{B}_{gq} \mathbf{P}_{gq} \mathbf{d}_g^q \\ &+ \sum_{l=1, l \neq g}^G \sum_{q \in \{h, v\}} \mathbf{H}_{gp}^H \mathbf{B}_{lq} \mathbf{P}_{lq} \mathbf{d}_l^q + \mathbf{n}_g^p, \end{aligned} \quad (23)$$

where  $\mathbf{H}_{gv} = \begin{bmatrix} \mathbf{H}_{gv}^{vv} \\ \mathbf{H}_{gv}^{vh} \end{bmatrix}$  and  $\mathbf{H}_{gh} = \begin{bmatrix} \mathbf{H}_{gh}^{hv} \\ \mathbf{H}_{gh}^{hh} \end{bmatrix}$  from (6). Here,  $\mathbf{B}_{gp}$  for  $p \in \{h, v\}$  are given as

$$\mathbf{B}_{gv} = \begin{bmatrix} \mathbf{B}_g^s \\ \mathbf{0} \end{bmatrix}, \quad \mathbf{B}_{gh} = \begin{bmatrix} \mathbf{0} \\ \mathbf{B}_g^s \end{bmatrix}, \quad (24)$$

where  $\mathbf{B}_g^s$  is given in (16). Note that, when  $\chi \approx 0$ , we can easily find that

$$\mathbf{H}_{lp}^H \mathbf{B}_{gq} \approx \mathbf{0}, \quad \text{for } p \neq q. \quad (25)$$

Furthermore, because  $\mathbf{H}_{gq}$ ,  $q \neq p$  has no influence on  $\mathbf{P}_{gp}$ , the MSs do not need to feed back the instantaneous CSI from cross polarized transmit antenna elements at BS. That is, the  $k$ th MS having vertical (horizontal) polarization in the  $g$ th group can quantize the first (last)  $r$  entries of  $\mathbf{g}_{gk}$  (see also Remark 1) and feed them back to the BS with the feedback amount reduced in half.

The precoding matrix  $\mathbf{P}_{gp}$  is then designed such that the intra-subgroup interferences are nulled out using the co-polarized short-term CSIT. That is, letting  $\hat{\mathbf{H}}_{gp}$  denote the imperfect CSI knowledge at the transmitter of  $\mathbf{H}_{gp}$ ,  $p \in \{h, v\}$ , the regularized ZF precoding matrix with imperfect CSIT can be computed as

$$\mathbf{P}_{gp} = \xi_{gp} \hat{\mathbf{K}}_{gp} \hat{\mathbf{H}}_{gp}, \quad (26)$$

where  $\hat{\mathbf{K}}_{gp} = \left( \hat{\mathbf{H}}_{gp} \hat{\mathbf{H}}_{gp}^H + \frac{\bar{B}}{2} \alpha \mathbf{I}_{\frac{\bar{B}}{2}} \right)^{-1}$  and  $\hat{\mathbf{H}}_{gp} = \mathbf{B}_{gp}^H \hat{\mathbf{H}}_{gp} = (\mathbf{B}_g^s)^H \hat{\mathbf{H}}_g^{pp}$ , the effective channel estimate that is available at the BS. The normalization factor  $\xi_{gp}$  is then given as

$$\xi_{gp}^2 = \frac{\bar{N}/2}{\text{tr}(\hat{\mathbf{H}}_{gp}^H \hat{\mathbf{K}}_{gp}^H \hat{\mathbf{K}}_{gp} \hat{\mathbf{H}}_{gp})}. \quad (27)$$

Assuming equal power allocation, the SINR of the  $k$ th MS in the subgroup with  $p$  polarization of the  $g$ th group is then given by

$$\gamma_{gpk}^{BDS} = \frac{\frac{P}{N} \xi_{gp}^2 |\mathbf{h}_{gpk}^H \mathbf{B}_{gp} \hat{\mathbf{K}}_{gp} \mathbf{B}_{gp}^H \hat{\mathbf{h}}_{gpk}|^2}{IN_{gpk}}, \quad (28)$$

where

$$\begin{aligned} IN_{gpk} &= \frac{P}{N} \sum_{j \neq k} \xi_{gp}^2 |\mathbf{h}_{gpk}^H \mathbf{B}_{gp} \hat{\mathbf{K}}_{gp} \mathbf{B}_{gp}^H \hat{\mathbf{h}}_{gpj}|^2 \\ &+ \frac{P}{N} \sum_{q \neq p} \sum_j \xi_{gq}^2 |\mathbf{h}_{gpk}^H \mathbf{B}_{gq} \hat{\mathbf{K}}_{gq} \mathbf{B}_{gq}^H \hat{\mathbf{h}}_{gqj}|^2 \\ &+ \frac{P}{N} \sum_{l \neq g} \sum_q \sum_j \xi_{lq}^2 |\mathbf{h}_{gpk}^H \mathbf{B}_{lq} \hat{\mathbf{K}}_{lq} \mathbf{B}_{lq}^H \hat{\mathbf{h}}_{lqj}|^2 + 1 \end{aligned} \quad (29)$$

and  $\mathbf{h}_{gpk} = [\mathbf{H}_{gp}]_k$  and  $\hat{\mathbf{h}}_{gpk} = [\hat{\mathbf{H}}_{gp}]_k$ , respectively. Accordingly, the sum rate is given by

$$R_{BDS} = \sum_{g=1}^G \sum_{p \in \{v, h\}} \sum_{k=1}^{\frac{\bar{N}}{2}} \log_2(1 + \gamma_{gpk}^{BDS}), \quad (30)$$

where the subscript and superscript  $BDS$  indicate the dual precoding with Block Diagonalization and Subgrouping based on both spatial correlation and polarization. Note that because, when  $\chi = 0$ , the interference from cross-polarized groups are perfectly nulled out, we can easily find that

$$\gamma_{gpk}^{BDS} = \gamma_{gpk}^{BD}. \quad (31)$$

## IV. ASYMPTOTIC PERFORMANCE ANALYSIS FOR DUAL PRECODING METHODS

In [7], when the number of transmit antenna elements ( $M$ ) is large, the asymptotic SINR of the regularized ZF precoding has been analyzed in spatially correlated MISO broadcasting systems with uni-polarized antennas under the imperfect CSIT and, in [8], the asymptotic SINR of the dual precoding with BD has been analyzed under the perfect CSIT and the uni-polarized antenna system. In this section,

$$\begin{aligned}\gamma_{gpk}^{BD} &= \frac{\frac{P}{N}\xi_g^2 |\mathbf{h}_{gk}^H \mathbf{B}_g \hat{\mathbf{K}}_g \hat{\mathbf{h}}_{gk}|^2}{\frac{P}{N} \sum_{j \neq k} \xi_g^2 |\mathbf{h}_{gk}^H \mathbf{B}_g \hat{\mathbf{K}}_g \hat{\mathbf{h}}_{gj}|^2 + \frac{P}{N} \sum_{l \neq g} \sum_j \xi_l^2 |\mathbf{h}_{gk}^H \mathbf{B}_l \hat{\mathbf{K}}_l \hat{\mathbf{h}}_{lj}|^2 + 1}}, \\ &= \frac{\frac{P}{N}\xi_g^2 |\mathbf{h}_{gk}^H \mathbf{B}_g \hat{\mathbf{K}}_g \mathbf{B}_g^H \hat{\mathbf{h}}_{gk}|^2}{\frac{P}{N} \sum_{j \neq k} \xi_g^2 |\mathbf{h}_{gk}^H \mathbf{B}_g \hat{\mathbf{K}}_g \mathbf{B}_g^H \hat{\mathbf{h}}_{gj}|^2 + \frac{P}{N} \sum_{l \neq g} \sum_j \xi_l^2 |\mathbf{h}_{gk}^H \mathbf{B}_l \hat{\mathbf{K}}_l \mathbf{B}_l^H \hat{\mathbf{h}}_{lj}|^2 + 1}}.\end{aligned}\quad (21)$$

based on random matrix theory results [4], [7], [27], we first derive the asymptotic SINR for two different dual precoding schemes – dual precoding with i) BD and ii) BDS under the imperfect CSIT and dual-polarized antenna system. Note that the asymptotic inter/intra interferences are evaluated over the polarization domain as well as the spatial domain, which can encounter a more generalized channel environment with a polarization. Based on the asymptotic results, we analyze the performance as a function of the XPD parameter  $\chi$ , and propose a new dual precoding/feedback scheme in the next section.

#### A. Dual precoding with block diagonalization based on spatial correlation

Before we proceed with the derivation of the asymptotic SINR for the dual precoding with BD, we introduce an important theorem about the asymptotic behavior of a random matrix with a large dimension developed in [7].

*Theorem 1:* ([7], Theorem 1) Let  $\mathbf{H}$  be the  $M \times N$  matrix, in which each column is a zero-mean complex Gaussian random vector having a covariance matrix  $\mathbf{R}_i$  for  $i = 1, \dots, N$ . In addition, let  $\mathbf{S}, \mathbf{Q} \in \mathbb{C}^{M \times M}$  be Hermitian nonnegative definite. Assume  $\limsup_{M \rightarrow \infty} \sup_{1 \leq i \leq N} \|\mathbf{R}_i\| < \infty$  and  $\mathbf{Q}$  has uniformly bounded spectrum norm. Then, for  $z < 0$ ,

$$\frac{1}{M} \text{tr}(\mathbf{Q}(\mathbf{H}\mathbf{H}^H + \mathbf{S} - z\mathbf{I}_M)^{-1}) - \frac{1}{M} \text{tr}(\mathbf{Q}\mathbf{T}(z)) \xrightarrow{M \rightarrow \infty} 0, \quad (32)$$

where

$$\mathbf{T}(z) = \left( \frac{1}{M} \sum_{j=1}^N \frac{\mathbf{R}_j}{1 + e_j(z)} + \mathbf{S} - z\mathbf{I}_M \right)^{-1}. \quad (33)$$

Here,  $e_i(z)$  for  $i = 1, \dots, N$  are the unique solution of

$$e_i(z) = \frac{1}{M} \text{tr} \left( \mathbf{R}_i \left( \frac{1}{M} \sum_{j=1}^N \frac{\mathbf{R}_j}{1 + e_j(z)} + \mathbf{S} - z\mathbf{I}_M \right)^{-1} \right), \quad (34)$$

which can be solved by the fixed-point algorithm and its convergence is also proved in [7].

Then, by using Theorem 1, the asymptotic SINR for the dual precoding with BD can be derived.

*Theorem 2:* When  $M, N, \bar{B}$  goes to infinity and  $\frac{N}{\bar{B}}$  is fixed, the SINR,  $\gamma_{gpk}^{BD}$  in (21) asymptotically converges as

$$\gamma_{gpk}^{BD} - \gamma_{gpk}^{BD,o} \xrightarrow{M \rightarrow \infty} 0, \quad (35)$$

where  $\gamma_{gpk}^{BD,o}$  is the asymptotic SINR, given as shown at the top of the page. Here,  $m_{gp}^o$ ,  $\Upsilon_{glp}^o$  and  $\Psi_g^o$  are the unique solutions

of

$$m_{gp}^o = \frac{1}{\bar{B}} \text{tr}(\bar{\mathbf{R}}_{gp} \mathbf{T}_g), \quad \mathbf{T}_g = \left( \frac{\bar{N}}{2\bar{B}} \sum_{q \in \{h,v\}} \frac{\bar{\mathbf{R}}_{gq}}{1+m_{gq}^o} + \alpha \mathbf{I}_{\bar{B}} \right)^{-1},$$

$$\Psi_g^o = \frac{1}{2\bar{B}} \frac{P}{G} \sum_{q \in \{h,v\}} \frac{m'_{gq}}{(1+m_{gq}^o)^2}, \quad (37)$$

$$\Upsilon_{gpp}^o = \frac{\bar{N}/2-1}{\bar{B}} \frac{P}{N} \frac{m'_{gpp}}{(1+m_{gp}^o)^2} + \frac{\bar{N}}{2\bar{B}} \frac{P}{N} \frac{m'_{gppq}}{(1+m_{gq}^o)^2},$$

$$\Upsilon_{glp}^o = \frac{P}{2N} \frac{\bar{N}}{\bar{B}} \sum_{q \in \{h,v\}} \frac{m'_{glpq}}{(1+m_{lq}^o)^2}. \quad (38)$$

In addition,  $\mathbf{m}'_g = [m'_{gv}, m'_{gh}]^T$ , and  $\mathbf{m}'_{gpp} = [m'_{gppv}, m'_{gpph}]^T$  given by

$$\mathbf{m}'_g = (\mathbf{I}_2 - \mathbf{J})^{-1} \mathbf{v}_g, \quad \mathbf{m}'_{gpp} = (\mathbf{I}_2 - \mathbf{J})^{-1} \mathbf{v}_{gpp}, \quad (39)$$

where

$$\mathbf{J} = \frac{\bar{N}}{2\bar{B}} \begin{bmatrix} \frac{\text{tr}(\bar{\mathbf{R}}_{gv} \mathbf{T}_g \bar{\mathbf{R}}_{gv} \mathbf{T}_g)}{B(1+m_{gv}^o)^2} & \frac{\text{tr}(\bar{\mathbf{R}}_{gv} \mathbf{T}_g \bar{\mathbf{R}}_{gh} \mathbf{T}_g)}{B(1+m_{gh}^o)^2} \\ \frac{\text{tr}(\bar{\mathbf{R}}_{gh} \mathbf{T}_g \bar{\mathbf{R}}_{gv} \mathbf{T}_g)}{B(1+m_{gv}^o)^2} & \frac{\text{tr}(\bar{\mathbf{R}}_{gh} \mathbf{T}_g \bar{\mathbf{R}}_{gh} \mathbf{T}_g)}{B(1+m_{gh}^o)^2} \end{bmatrix}, \quad (40)$$

$$\mathbf{v}_g = \frac{1}{\bar{B}} \begin{bmatrix} \text{tr}(\bar{\mathbf{R}}_{gv} \mathbf{T}_g^2) \\ \text{tr}(\bar{\mathbf{R}}_{gh} \mathbf{T}_g^2) \end{bmatrix}, \quad \mathbf{v}_{gpp} = \frac{1}{\bar{B}} \begin{bmatrix} \text{tr}(\bar{\mathbf{R}}_{gv} \mathbf{T}_g \bar{\mathbf{R}}_{gp} \mathbf{T}_g) \\ \text{tr}(\bar{\mathbf{R}}_{gh} \mathbf{T}_g \bar{\mathbf{R}}_{gp} \mathbf{T}_g) \end{bmatrix}$$

with  $\bar{\mathbf{R}}_{gp}$  defined in (17). In addition,  $\mathbf{m}'_{glp} = [m'_{glpv}, m'_{glph}]^T$  given by

$$\mathbf{m}'_{glp} = (\mathbf{I}_2 - \mathbf{J})^{-1} \mathbf{v}_{glp}, \quad (41)$$

where

$$\mathbf{J} = \frac{\bar{N}}{2\bar{B}} \begin{bmatrix} \frac{\text{tr}(\bar{\mathbf{R}}_{lv} \mathbf{T}_l \bar{\mathbf{R}}_{lv} \mathbf{T}_l)}{B(1+m_{lv}^o)^2} & \frac{\text{tr}(\bar{\mathbf{R}}_{lv} \mathbf{T}_l \bar{\mathbf{R}}_{lh} \mathbf{T}_l)}{B(1+m_{lh}^o)^2} \\ \frac{\text{tr}(\bar{\mathbf{R}}_{lh} \mathbf{T}_l \bar{\mathbf{R}}_{lv} \mathbf{T}_l)}{B(1+m_{lv}^o)^2} & \frac{\text{tr}(\bar{\mathbf{R}}_{lh} \mathbf{T}_l \bar{\mathbf{R}}_{lh} \mathbf{T}_l)}{B(1+m_{lh}^o)^2} \end{bmatrix},$$

$$\mathbf{v}_{glq} = \frac{1}{\bar{B}} \begin{bmatrix} \text{tr}(\bar{\mathbf{R}}_{lv} \mathbf{T}_l \mathbf{B}_l^H \mathbf{R}_{gp} \mathbf{B}_l \mathbf{T}_l) \\ \text{tr}(\bar{\mathbf{R}}_{lh} \mathbf{T}_l \mathbf{B}_l^H \mathbf{R}_{gp} \mathbf{B}_l \mathbf{T}_l) \end{bmatrix}. \quad (42)$$

*Proof:* See Appendix A. ■

Thanks to the structure of the dual polarized antenna elements, we can have a simple asymptotic SINR in Theorem 2 compared to that for the case of the general antenna covariance matrices [7]. This gives a useful insight into the behavior of the asymptotic SINR as a function of the polarization parameter in Section IV-C. Furthermore, when the spatial covariance matrix is the same for both polarizations, we can further simplify the asymptotic SINR in Theorem 2.

*Corollary 1:* When the spatial covariance matrix is the same for both polarization, i.e., infinitesimally small dual-polarized antenna elements are co-located (see footnote 2), the asymptotic SINR  $\gamma_{gpk}^{BD,o}$  in (35) can be written as in a simpler form as shown at the top of the page. Here,  $m_g^o$ ,  $\Upsilon_{gl}^o$ ,

$$\gamma_{gpk}^{BD,o} = \frac{\frac{P}{N}(\xi_g^o)^2(1-\tau^2)(m_{gp}^o)^2}{(\xi_g^o)^2\Upsilon_{gpp}^o(1-\tau^2(1-(1+m_{gp}^o)^2)) + (1 + \sum_{l \neq g} (\xi_l^o)^2 \Upsilon_{glp}^o)(1+m_{gp}^o)^2}, \quad (36)$$

$$\text{with } (\xi_g^o)^2 = \frac{P}{G\Psi_g^o}.$$

$$\gamma_{gk}^{\prime BD,o} = \frac{\frac{P}{N}(\xi_g^o)^2(1-\tau^2)(m_g^o)^2}{(\xi_g^o)^2\Upsilon_{gg}^o(1-\tau^2(1-(1+m_g^o)^2)) + (1 + \sum_{l \neq g} (\xi_l^o)^2 \Upsilon_{gl}^o)(1+m_g^o)^2}, \quad (43)$$

$$\text{with } (\xi_g^o)^2 = \frac{P}{G\Psi_g^o}.$$

and  $\Psi_g^o$  are the unique solutions of

$$m_g^o = \frac{1}{B} \text{tr}(\bar{\mathbf{R}}_g' \mathbf{T}_g), \quad \mathbf{T}_g = \left( \frac{\bar{N}}{B} \frac{\bar{\mathbf{R}}_g'}{1+m_g^o} + \alpha \mathbf{I}_{\bar{B}} \right)^{-1}, \quad (44)$$

$$\Psi_g^o = \frac{1}{B} \frac{P}{G} \frac{m_g'}{(1+m_g^o)^2}, \quad m_g' = \frac{\frac{1}{B} \text{tr}(\bar{\mathbf{R}}_g' \mathbf{T}_g^2)}{1 - \frac{\frac{N}{B} \text{tr}(\bar{\mathbf{R}}_g' \mathbf{T}_g \bar{\mathbf{R}}_g' \mathbf{T}_g)}{B(1+m_g^o)^2}}, \quad (45)$$

$$\Upsilon_{gg}^o = \frac{\bar{N}-1}{B} \frac{P}{N} \frac{m_{gg}'}{(1+m_g^o)^2}, \quad \Upsilon_{gl}^o = \frac{P}{N} \frac{\bar{N}}{B} \frac{m_{gl}'}{(1+m_l^o)^2}, \quad (46)$$

$$m_{gg}' = \frac{\frac{1}{B} \text{tr}(\bar{\mathbf{R}}_g' \mathbf{T}_g \bar{\mathbf{R}}_g' \mathbf{T}_g)}{1 - \frac{\frac{N}{B} \text{tr}(\bar{\mathbf{R}}_g' \mathbf{T}_g \bar{\mathbf{R}}_g' \mathbf{T}_g)}{B(1+m_g^o)^2}}, \quad m_{gl}' = \frac{\frac{1}{B} \text{tr}(\bar{\mathbf{R}}_l' \mathbf{T}_l \mathbf{B}_l^H \bar{\mathbf{R}}_g' \mathbf{B}_l \mathbf{T}_l)}{1 - \frac{\frac{N}{B} \text{tr}(\bar{\mathbf{R}}_l' \mathbf{T}_l \bar{\mathbf{R}}_l' \mathbf{T}_l)}{B(1+m_l^o)^2}}, \quad (47)$$

where  $\mathbf{R}_g' = \frac{1}{2} \mathbf{R}_g$  and  $\bar{\mathbf{R}}_g' = \frac{1}{2} \bar{\mathbf{R}}_g$ . Here,  $\mathbf{R}_g$  and  $\bar{\mathbf{R}}_g$  are defined in (7) and (17).

*Proof:* From (7) and (17), we can see that  $m_{gh}^o = m_{gv}^o = \frac{1}{B} \text{tr}(\frac{1}{2} \bar{\mathbf{R}}_g \mathbf{T}_g)$  in (37). Accordingly, by letting  $m_g^o = \frac{1}{B} \text{tr}(\frac{1}{2} \bar{\mathbf{R}}_g \mathbf{T}_g)$ ,  $\mathbf{T}_g$  in (37) can be rewritten as that in (44). Furthermore, because, in (40),

$$\text{tr}(\bar{\mathbf{R}}_{gv} \mathbf{T}_g^2) = \text{tr}(\bar{\mathbf{R}}_{gh} \mathbf{T}_g^2) = \frac{1}{2} \text{tr}(\bar{\mathbf{R}}_g \mathbf{T}_g^2), \quad (48)$$

$$\text{tr}(\bar{\mathbf{R}}_{gv} \mathbf{T}_g \bar{\mathbf{R}}_{gv} \mathbf{T}_g) + \text{tr}(\bar{\mathbf{R}}_{gh} \mathbf{T}_g \bar{\mathbf{R}}_{gh} \mathbf{T}_g) = \frac{1}{4} \text{tr}(\bar{\mathbf{R}}_g \mathbf{T}_g \bar{\mathbf{R}}_g \mathbf{T}_g),$$

$m_{gv}' = m_{gh}' = m_g'$  as in (45). Similarly, we can prove that the parameters  $m_{glpq}'$  are given as (47). By substituting  $m_g'$ ,  $m_{gg}'$ ,  $m_{gl}'$  into  $\Psi_g^o$ ,  $\Upsilon_{gpp}^o$ , and  $\Upsilon_{glp}^o$  of (37) and (38), we can prove that  $\Psi_g^o$ ,  $\Upsilon_{gpp}^o$ , and  $\Upsilon_{glp}^o$  can be written as in (45) and (46). ■

From Corollary 1, the asymptotic SINR is independent of MS index  $k$  and polarization index  $p$ . Accordingly, by letting  $\gamma_{gk}^{\prime BD,o} \triangleq \gamma_g^{\prime BD,o}$  for  $k = 1, \dots, \bar{N}$ , the asymptotic sum rate can be approximated as

$$\begin{aligned} R_{BD}^o &\approx \sum_{g=1}^G \sum_{p \in \{v,h\}} \sum_{k=1}^{\bar{N}/2} \log_2(1 + \gamma_{gk}^{\prime BD,o}) \\ &= \sum_{g=1}^G \bar{N} \log_2(1 + \gamma_g^{\prime BD,o}). \end{aligned} \quad (49)$$

*Remark 2:* We note that, when  $\tau = 0$ ,  $\gamma_{gk}^{\prime BD,o}$  in Corollary 1 is analogous to the asymptotic SINR of the dual precoding with BD derived in [8] under the perfect CSIT and uni-polarized system. That is, when the infinitesimally small dual-polarized antenna elements are co-located, the asymptotic SINRs of the MSs in the same group are the same irrespective of their antenna deployment, i.e., vertical or horizontal polarization. In addition, the effective covariance matrix is given

by  $\mathbf{R}_g' = \frac{1}{2} \mathbf{R}_g$ . That is, it can be described as if the BS and MSs are co-polarized and the correlation matrices for the MSs in the  $g$ th group are the same as  $\mathbf{I}_2 \otimes \mathbf{R}_g^s$  and the effective transmit power of BS is reduced from  $P$  to  $\frac{1+\chi}{2} P$ . Note that this is valid only when the spatial covariance matrix is the same for both polarizations. That is, Theorem 2 is extended to more general covariance matrices addressing the polarization of antenna elements.

### B. Dual precoding with block diagonalization and subgrouping based on both spatial correlation and polarization

By using Theorem 1 and an approach similar as that used for the dual precoding with BD, the asymptotic SINR for the dual precoding with BDS can be derived.

*Theorem 3:* When  $M$ ,  $N$ ,  $\bar{B}$  go to infinity and  $\frac{N}{\bar{B}}$  is fixed, the SINR,  $\gamma_{gpk}^{BDS}$  in (28) asymptotically converges as

$$\gamma_{gpk}^{BDS} - \gamma_{gpk}^{BDS,o} \xrightarrow{M \rightarrow \infty} 0, \quad (50)$$

where  $\gamma_{gpk}^{BDS,o}$  is given by

$$\gamma_{gpk}^{BDS,o} = \frac{\frac{P}{N}(\xi_{gp}^o)^2(1-\tau^2)(m_{gp}^o)^2}{IN_{gpk}^o}, \quad (51)$$

where

$$\begin{aligned} IN_{gpk}^o &= (\xi_{gp}^o)^2 \Upsilon_{gpp}^o (1 - \tau^2 (1 - (1 + m_{gp}^o)^2)) + \\ &(1 + \sum_{q \neq p} (\xi_{gq}^o)^2 \Upsilon_{gppq}^o + \sum_{l \neq g} \sum_q (\xi_{lq}^o)^2 \Upsilon_{glpq}^o) (1 + m_{gp}^o)^2, \end{aligned} \quad (52)$$

with  $(\xi_{gp}^o)^2 = \frac{P}{G\Psi_{gp}^o}$ , where  $m_{gp}^o$ ,  $\Upsilon_{glpq}^o$ , and  $\Psi_{gp}^o$  are the unique solutions of

$$\begin{aligned} m_{gp}^o &= \frac{2}{B} \text{tr}(\bar{\mathbf{R}}_{gp} \mathbf{T}_{gp}), \quad \mathbf{T}_{gp} = \left( \frac{\bar{N}}{B} \frac{\bar{\mathbf{R}}_{gp}}{1+m_{gp}^o} + \alpha \mathbf{I}_{\bar{B}/2} \right)^{-1}, \\ \Psi_{gp}^o &= \frac{1}{B} \frac{P}{G} \frac{m_{gp}'}{(1+m_{gp}^o)^2}, \end{aligned} \quad (53)$$

$$\Upsilon_{gpp}^o = \frac{\bar{N}/2-1}{B/2} \frac{P}{N} \frac{m_{gppp}'}{(1+m_{gp}^o)^2}, \quad \Upsilon_{glpq}^o = \frac{P}{N} \frac{\bar{N}}{B} \frac{m_{glpq}'}{(1+m_{lq}^o)^2}, \quad (54)$$

$$\text{with } m_{gp}' = \frac{\frac{2}{B} \text{tr}(\bar{\mathbf{R}}_{gp} \mathbf{T}_{gp}^2)}{1 - \frac{\frac{N}{B} \text{tr}(\bar{\mathbf{R}}_{gp} \mathbf{T}_{gp} \bar{\mathbf{R}}_{gp} \mathbf{T}_{gp})}{B/2(1+m_{gp}^o)^2}},$$

$$m_{gppp}' = \frac{\frac{2}{B} \text{tr}(\bar{\mathbf{R}}_{gp} \mathbf{T}_{gp} \bar{\mathbf{R}}_{gp} \mathbf{T}_{gp})}{1 - \frac{\frac{N}{B} \text{tr}(\bar{\mathbf{R}}_{gp} \mathbf{T}_{gp} \bar{\mathbf{R}}_{gp} \mathbf{T}_{gp})}{B/2(1+m_{gp}^o)^2}}, \quad (55)$$

$$m_{glpq}' = \frac{\frac{2}{B} \text{tr}(\bar{\mathbf{R}}_{lq} \mathbf{T}_{lq} \mathbf{B}_{lq}^H \bar{\mathbf{R}}_{gp} \mathbf{B}_{lq} \mathbf{T}_{lq})}{1 - \frac{\frac{N}{B} \text{tr}(\bar{\mathbf{R}}_{lq} \mathbf{T}_{lq} \bar{\mathbf{R}}_{lq} \mathbf{T}_{lq})}{B/2(1+m_{lq}^o)^2}}, \quad (56)$$

where  $\bar{\mathbf{R}}_{gp}$  is defined in (17).

*Proof:* Because the proof is similar to that of Theorem 2, it is omitted. ■

From Theorem 3, the asymptotic SINR is independent of MS index  $k$ . Furthermore, because  $\mathbf{X}$  in (5) is symmetric,  $\gamma_{gvk}^{BDS,o} = \gamma_{ghk}^{BDS,o}$ . Accordingly, by letting  $\gamma_{gvk}^{BDS,o} = \gamma_{ghk}^{BDS,o} \triangleq \gamma_g^{BDS,o}$ , the asymptotic sum rate can be approximated as

$$\begin{aligned} R_{BDS}^o &\approx \sum_{g=1}^G \sum_{p \in \{v,h\}} \frac{\bar{N}}{2} \log_2(1 + \gamma_{gpk}^{BDS,o}) \\ &= \sum_{g=1}^G \bar{N} \log_2(1 + \gamma_g^{BDS,o}). \end{aligned} \quad (57)$$

### C. Asymptotic performance analysis as a function of the XPD parameter $\chi$

In this section, we investigate the effect of the XPD parameter  $\chi$  on the asymptotic SINRs of the dual precoding schemes. Based on the asymptotic results in Theorems 2 and 3 and Corollary 1, we can have the following propositions.

*Proposition 1:* For a large  $M$ , the asymptotic SINR of the dual precoding with BD in (36) is approximately independent of the XPD parameter  $\chi$ . That is, if we define the asymptotic SINR as the function  $\gamma_{gpk}^{BD,o}(\chi)$  of  $\chi$ , then  $\gamma_{gpk}^{BD,o}(\chi) \approx \gamma_{gpk}^{BD,o}(0)$ .

*Proof:* See Appendix B. ■

*Proposition 2:* For a large  $M$ , the asymptotic SINR of the dual precoding with BDS in (51) can be approximately written as the function  $\gamma_{gpk}^{BDS,o}(\chi)$  of  $\chi$  given by

$$\gamma_{gpk}^{BDS,o}(\chi) \approx \frac{\gamma_{gpk}^{BDS,o}(0)}{1 + c_0\chi}, \quad (58)$$

where

$$c_0 = E_{g,p} \left[ \frac{(\xi_{gp}^o(0))^2 \Upsilon_{ggpp}^o(0)}{\frac{(\xi_{gp}^o(0))^2 \Upsilon_{ggpp}^o(0)}{(1+m_{gp}^o(0))^2} (\tau^2((1+m_{gp}^o(0))^2 - 1) + 1) + 1} \right]. \quad (59)$$

*Proof:* See Appendix C. ■

*Remark 3:* Note that, from Proposition 2, the asymptotic SINR of the dual precoding with BDS decreases when  $\chi$  increases. This is because the subgroups are formed with the assumption that the interferences through the cross-polarized channels are perfectly nulled out in Section III-C and  $\mathbf{P}_{gp}$  in (26) is determined based only on co-polarized CSIT. Therefore, the interference power increases proportionally to  $\chi$ . In contrast, because the dual precoding with BD nulls out the intra-group interferences based on both co/cross polarized CSIT, it exhibits performances somehow robust to the variation of the polarization parameter  $\chi$ . In addition, from Theorem 2, 3, and Corollary 1, we can also see that  $\gamma_{gpk}^{BDS,o}(0) = \gamma_{gpk}^{BD,o}(0)$ .

## V. DISCUSSION

### A. A new dual structured precoding/feedback

Even though the sum-rate performance of the dual precoding with BDS decreases as  $\chi$  increases, it can utilize more

accurate short-term CSIT compared to the dual precoding with BD under the same number of feedback bits as stated in Section III-C. Assuming the CSI is perfectly estimated at MSs, when random vector quantization (RVQ) with  $N_B$  bits is utilized [23], the quantization error for the short-term CSIT in the dual precoding with BD (i.e., the columns of  $\mathbf{G}_g$  in (4)) is upper bounded as<sup>4</sup>

$$\tau_{BD}^2 < 2^{-\frac{N_B}{2r-1}}. \quad (60)$$

For the dual precoding with BDS, the quantization error is upper bounded as

$$\tau_{BDS}^2 < 2^{-\frac{N_B}{r-1}}. \quad (61)$$

Because the bound is tight for a large  $N_B$  [23], by assuming  $\tau_{BD}^2 = 2^{-\frac{N_B}{2r-1}} (\approx \tau_{BDS})$ , we have the following proposition.

*Proposition 3:* For a given  $\chi$  and a large  $M$ , when

$$N_B \lesssim (2r-1) \left( \log_2 \left( 1 + E_{g,p} \left( \frac{(1+m_{gp}^o(0))^2 - 1}{(\xi_{gp}^o(0))^2 \Upsilon_{ggpp}^o(0) (1+m_{gp}^o(0))^2} \right) - \log_2 \chi \right) \right), \quad (62)$$

the dual precoding with BDS outperforms that with BD.

*Proof:* From (51) and Proposition 2, the SINR of the dual precoding with BDS can be written as

$$\gamma_{gpk}^{BDS,o}(\chi) = \frac{A_0(1 - \tau_{BDS}^2)}{(B_0(1 + D_0\tau_{BDS}^2) + (1 + E_0)(D_0 + 1))(1 + c_0\chi)} \quad (63)$$

where  $A_0 = \frac{P}{N} (\xi_{gp}^o(0))^2 (m_{gp}^o(0))^2$ ,  $B_0 = (\xi_{gp}^o(0))^2 \Upsilon_{ggpp}^o(0)$ ,  $D_0 = (1 + m_{gp}^o(0))^2 - 1$ , and  $E_0 = \sum_{l \neq g} (\xi_{lp}^o(0))^2 \Upsilon_{glpp}^o(0)$ . Note that  $\gamma_{gpk}^{BD,o}(0) = \gamma_{gpk}^{BDS,o}(0)$  in Remark 3, assuming the same channel accuracy (i.e.,  $\tau_{BD}^2 = \tau_{BDS}^2$ ). Therefore, by setting  $\chi = 0$  in (63), from Proposition 1, the SINR of the dual precoding with BD can then be given as

$$\gamma_{gpk}^{BD,o}(\chi) = \frac{A_0(1 - \tau_{BD}^2)}{B_0(1 + D_0\tau_{BD}^2) + (1 + E_0)(D_0 + 1)}. \quad (64)$$

Because the dual precoding with BDS outperforms that with BD when  $\gamma_{gpk}^{BDS,o}(\chi) \geq \gamma_{gpk}^{BD,o}(\chi)$ , by letting  $\tau_{BD}^2 \triangleq \tau^2 = 2^{-\frac{N_B}{2r-1}} (\approx \tau_{BDS})$ , we have

$$\begin{aligned} &\frac{(1 - \tau^4)}{(B_0(1 + D_0\tau^4) + (1 + E_0)(D_0 + 1))(1 + c_0\chi)} \\ &\geq \frac{(1 - \tau^2)}{B_0(1 + D_0\tau^2) + (1 + E_0)(D_0 + 1)}. \end{aligned} \quad (65)$$

After a simple calculation, we have

$$\chi \leq \frac{(B_0D_0 + B_0 + (1 + E_0)(D_0 + 1))\tau^2}{c_0(B_0D_0\tau^4 + B_0 + (1 + E_0)(D_0 + 1))}. \quad (66)$$

<sup>4</sup>Here, the codebook is fixed given  $r$  and there is no adaptive codebook that would adapt as a function of  $r$  and XPD. Of course, if  $r$  changes, the codebook can be adaptively designed with respect to  $r$ , but it is out of scope of this paper.



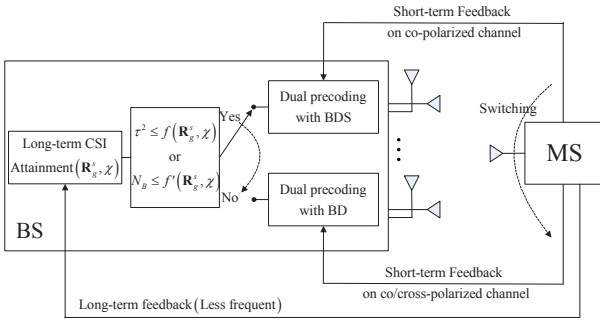


Fig. 2. Block diagram of a new dual precoding/feedback scheme.

From (59) and the fact that  $E_0 \ll 1$  due to the BD,  $c_0 \approx \frac{B_0(D_0+1)}{B_0D_0\tau^4+B_0+D_0+1}$  and we have

$$\begin{aligned} \chi &\leq \left(1 + \frac{D_0}{B_0(D_0+1)}\right)\tau^2 \\ &= \left(1 + \frac{(1+m_{gp}^o(0))^2-1}{(\xi_{gp}^o(0))^2\Upsilon_{gpp}^o(0)(1+m_{gp}^o(0))^2}\right)\tau^2, \end{aligned} \quad (67)$$

which induces (62) by taking the expectation over  $g$  and  $p$  in (67). ■

*Remark 4:* From Proposition 3, when the feedback bits are not enough to describe the short-term CSIT accurately, the dual precoding with BDS exhibits a better performance than that with BD. That is, it is preferable that by forming the co-polarized subgroup, each MS feeds back the short-term CSI from the co-polarized transmit antenna elements. In addition, from (62), when  $\chi \rightarrow 0$ , the dual precoding with BDS always exhibits better performance than that with BD. Note that for high SNR (i.e.,  $m_{gp}^o(0) \gg 1$ ), the expectation term in (62) is approximated as  $E_{g,p} \left( \frac{1}{(\xi_{gp}^o(0))^2\Upsilon_{gpp}^o(0)} \right)$ , which is inversely proportional to the intra-subgroup interference power from (52). Hence, a smaller intra-subgroup interference widens the region where BDS outperforms BD. That is, if the transmit signals to the co-polarized MSs can be asymptotically well separated by the linear precoding (less intra-subgroup interference), the feedback of the short-term CSI of the co-polarized channel with a higher accuracy is preferable.

Therefore, motivated by Proposition 3, a new dual precoding/feedback scheme can be described in Fig 2. Note that, depending on the long-term CSI (spatial correlation, polarization) and the number of feedback bits (or, the short-term CSIT accuracy  $\tau$ ), the dual precoding is switched between BD and BDS. In other words, by analyzing the asymptotic performance of BD and BDS, we can propose a new dual structured precoding which outperforms both BD and BDS by balancing the weakness of BD (low CSIT accuracy across the whole array) and BDS (performance degradation due to large polarization parameter  $\chi$ ). That is, given the same feedback overhead, for small  $\chi$ , the dual precoding with BDS will exhibit better performance, while, for large  $\chi$ , that with BD will show better performance. However, because the new dual precoding is switched between BD and BDS based on Proposition 3, the new dual precoding will show better performance than other two schemes.

### B. 3D dual structured precoding

Motivated by 3D beamforming [8], [20], the proposed scheme can also be extended to the scenario of 3D dual structured precoding. Assuming that the  $M_E \times M_A$  uniform planar array with dual-polarized antenna elements is exploited at BS and there are  $L$  elevation regions. For simplicity, each elevation region has the same number of groups,  $G$ , as in Fig 3. Note that the elevation angular spread depends on the distance  $d_{gl}$  between the BS and the  $g$ th group of the  $l$ th region and the radius of the ring of scatterer  $s_{gl}$ . By letting  $\mathbf{h}_{gkl}$  be the  $2M_A M_E \times 1$  vectorized channel of the  $k$ th MS in the  $g$ th group of the  $l$ th region, it can be written as

$$\mathbf{h}_{gkl} = ((\mathbf{I}_2 \otimes \mathbf{U}_{glA}) \otimes \mathbf{U}_{lE}) (\mathbf{\Lambda}_{glA}^{\frac{1}{2}} \otimes \mathbf{\Lambda}_{lE}^{\frac{1}{2}}) \mathbf{g}_{gkl}, \quad (68)$$

where  $\mathbf{\Lambda}_{glA}$  and  $\mathbf{\Lambda}_{lE}$  are the  $r_{glA} \times r_{glA}$  and  $r_{lE} \times r_{lE}$  diagonal matrices with non-zero eigenvalues of the spatial correlation matrices  $\mathbf{R}_{glA}^s$  and  $\mathbf{R}_{lE}^s$  over the azimuth and elevation directions, respectively, and  $\mathbf{U}_{glA}$  and  $\mathbf{U}_{lE}$  are the matrices of the associated eigenvectors. Here,  $\mathbf{g}_{gkl} = [\mathbf{g}_{gklv}^T, \chi \mathbf{g}_{gklh}^T]^T$  (resp.  $\mathbf{g}_{gkl} = [\chi \mathbf{g}_{gklv}^T, \mathbf{g}_{gklh}^T]^T$ ) for vertically (resp. horizontally) polarized MSs and  $\mathbf{g}_{gklp}$  is a  $r_{glA} r_{lE} \times 1$  vector whose elements are complex Gaussian distributed with zero mean and unit variance. Then, the 3D dual structured precoding signal can be given as

$$\mathbf{x} = \sum_{l=1}^L \left( \sum_{g=1}^G \mathbf{V}_{gl} \mathbf{d}_{gl} \right) \otimes \mathbf{q}_l, \quad (69)$$

where  $\mathbf{q}_l$  is the preprocessing vector based on  $\mathbf{R}_{lE}^s$  that nulls out the interferences from the other elevation regions. Similarly to Section III-A,  $\mathbf{q}_l$  can be computed such that  $\mathbf{q}_l^H \mathbf{U}_{-lE} = 0$  with  $\mathbf{U}_{-lE} = [\mathbf{U}_{1E}, \dots, \mathbf{U}_{l-1E}, \mathbf{U}_{l+1E}, \dots, \mathbf{U}_{LE}]$ . Then, after a simple manipulation, the received signal  $\mathbf{y}_{gl}$  of the  $g$ th group in the  $l$ th region is given as

$$\begin{aligned} \mathbf{y}_{gl} &= \sum_{g'=1}^G \mathbf{G}_{gl}^H (\mathbf{\Lambda}_{glA}^{\frac{1}{2}} \otimes \mathbf{\Lambda}_{lE}^{\frac{1}{2}}) ((\mathbf{I}_2 \otimes \mathbf{U}_{glA})^H \mathbf{V}_{g'l} \mathbf{d}_{g'l}) \otimes (\mathbf{U}_{lE}^H \mathbf{q}_l) + \mathbf{n}_{gl} \\ &= \sum_{g'=1}^G \sqrt{\tilde{\lambda}_l} \mathbf{G}_{gl}^H \mathbf{\Lambda}_{glA}^{\frac{1}{2}} (\mathbf{I}_2 \otimes \mathbf{U}_{glA})^H \mathbf{V}_{g'l} \mathbf{d}_{g'l} + \mathbf{n}_{gl}, \end{aligned} \quad (70)$$

where  $\tilde{\lambda}_l = \mathbf{q}_l^H \mathbf{R}_{lE}^s \mathbf{q}_l$ . Note that after the vertical preprocessing based on long-term CSIT (elevation), (70) is the (2-D spatial domain) equivalent system in Section II and the new dual structured precoding in Section V-A can be applied to (70) with a channel scaling constant  $\tilde{\lambda}_l$  for the  $l$ th elevation region.

### C. Polarization mismatch

When the antenna polarization between the transmitter and the receiver is not perfectly aligned (i.e., polarization mismatch), the performance can be degraded for both single polarized antenna system and dual polarized antenna system [21], [28]. In [28], the polarization mismatch can be expressed by the rotation matrix with a mismatched angle  $\theta^{ms}$  (i.e.  $\begin{bmatrix} \cos \theta^{ms} & -\sin \theta^{ms} \\ \sin \theta^{ms} & \cos \theta^{ms} \end{bmatrix}$ ). Accordingly, we let  $\theta_{gk}^{ms} \sim$

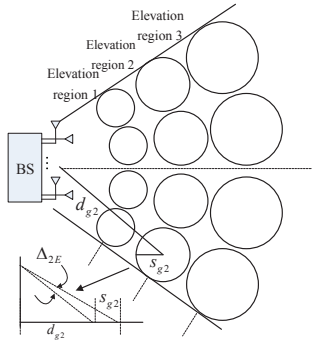


Fig. 3. Block diagram of a 3D dual polarized multi-user downlink system.

$U[-\theta_{\max}^{ms}, \theta_{\max}^{ms}]$  denote the mismatched polarization angle for the  $k$ th user in the  $g$ th group, where  $U[a, b]$  indicates the uniform distribution between  $a$  and  $b$  and  $\theta_{\max}^{ms}$  is the maximum value of the mismatched angle. Note that the case of  $\theta_{\max}^{ms} = 0$  indicates the MSs are perfectly aligned with either vertically or horizontally polarized antenna elements of the BS. Then, from (6) and the above observation, the polarization mismatched channel  $\mathbf{h}_{gk}^{ms}$  for the  $k$ th MS in the  $g$ th group can be given as

$$\mathbf{h}_{gk}^{ms} = \begin{bmatrix} \mathbf{U}_g \Lambda_g^{\frac{1}{2}} (\cos \theta_{gk}^{ms} - \sqrt{\chi} \sin \theta_{gk}^{ms}) (\mathbf{g}_{gk})_{1:r_g} \\ \mathbf{U}_g \Lambda_g^{\frac{1}{2}} (\sin \theta_{gk}^{ms} + \sqrt{\chi} \cos \theta_{gk}^{ms}) (\mathbf{g}_{gk})_{r_g+1:2r_g} \end{bmatrix}, \quad (71)$$

for vertical polarized MSs, and

$$\mathbf{h}_{gk}^{ms} = \begin{bmatrix} \mathbf{U}_g \Lambda_g^{\frac{1}{2}} (\sqrt{\chi} \cos \theta_{gk}^{ms} - \sin \theta_{gk}^{ms}) (\mathbf{g}_{gk})_{1:r_g} \\ \mathbf{U}_g \Lambda_g^{\frac{1}{2}} (\sqrt{\chi} \sin \theta_{gk}^{ms} + \cos \theta_{gk}^{ms}) (\mathbf{g}_{gk})_{r_g+1:2r_g} \end{bmatrix}, \quad (72)$$

for horizontal polarized MSs. From (71), the covariance matrix of the vertically co-polarized MS subgroup  $\mathbf{R}_{gv}^{ms}$  can be given as

$$\mathbf{R}_{gv}^{ms} = \begin{bmatrix} \mathbf{U}_g \Lambda_g^{\frac{1}{2}} \mathbf{D}_{vv}^{ms} \Lambda_g^{\frac{H}{2}} \mathbf{U}_g^H & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_g \Lambda_g^{\frac{1}{2}} \mathbf{D}_{hv}^{ms} \Lambda_g^{\frac{H}{2}} \mathbf{U}_g^H \end{bmatrix}, \quad (73)$$

where

$$\begin{aligned} \mathbf{D}_{vv}^{ms} &= E \left[ \frac{2}{N} \sum_{k=1}^{\frac{N}{2}} (\cos \theta_{gk}^{ms} - \sqrt{\chi} \sin \theta_{gk}^{ms})^2 \right] \mathbf{I}_{r_g} \\ &= \left( \frac{1}{2} + \frac{\sin 2\theta_{\max}^{ms}}{4\theta_{\max}^{ms}} + \chi \left( \frac{1}{2} - \frac{\sin 2\theta_{\max}^{ms}}{4\theta_{\max}^{ms}} \right) \right) \mathbf{I}_{r_g}, \\ \mathbf{D}_{hv}^{ms} &= E \left[ \frac{2}{N} \sum_{k=1}^{\frac{N}{2}} (\sin \theta_{gk}^{ms} + \sqrt{\chi} \cos \theta_{gk}^{ms})^2 \right] \mathbf{I}_{r_g} \\ &= \left( \frac{1}{2} - \frac{\sin 2\theta_{\max}^{ms}}{4\theta_{\max}^{ms}} + \chi \left( \frac{1}{2} + \frac{\sin 2\theta_{\max}^{ms}}{4\theta_{\max}^{ms}} \right) \right) \mathbf{I}_{r_g}. \end{aligned}$$

Therefore, we can have

$$\mathbf{R}_{gv}^{ms} = c_{eff} \begin{bmatrix} \mathbf{R}_g^s & \mathbf{0} \\ \mathbf{0} & \chi_{eff} \mathbf{R}_g^s \end{bmatrix}, \quad \mathbf{R}_{gh}^{ms} = c_{eff} \begin{bmatrix} \chi_{eff} \mathbf{R}_g^s & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_g^s \end{bmatrix} \quad (74)$$

where  $c_{eff} = \left( \frac{1}{2} + \frac{\sin 2\theta_{\max}^{ms}}{4\theta_{\max}^{ms}} + \chi \left( \frac{1}{2} - \frac{\sin 2\theta_{\max}^{ms}}{4\theta_{\max}^{ms}} \right) \right)$  and  $\chi_{eff} = \frac{1}{c_{eff}} \left( \frac{1}{2} - \frac{\sin 2\theta_{\max}^{ms}}{4\theta_{\max}^{ms}} + \chi \left( \frac{1}{2} + \frac{\sin 2\theta_{\max}^{ms}}{4\theta_{\max}^{ms}} \right) \right)$ . Note that  $c_{eff} \leq 1$  and the equality is satisfied when  $\theta_{\max}^{ms} = 0$ , i.e.,

polarization is perfectly aligned. Interestingly, it is known that the effect of the polarization mismatch can be captured by the multiplication of the long-term channel with a scalar ( $\leq 1$ ) [21], which is also consistent with  $c_{eff}$  in (74). Furthermore, considering the polarization mismatch, we can use  $\chi_{eff}$  instead of  $\chi$  in the dual precoding with the mode switching described in Remark 4 and Fig. 2.

## VI. SIMULATION RESULTS

Throughout the simulations, we consider the one-ring model for the spatially correlated channel [29], [30]. The correlation between the channel coefficients of antenna elements  $1 \leq m, n \leq M$  is given by

$$[\mathbf{R}_g^s]_{m,n} = \frac{1}{2\Delta_g} \int_{-\Delta_g}^{\Delta_g} e^{-j\pi\lambda_0^{-1}\Omega(\alpha+\theta_g)(\mathbf{r}_m-\mathbf{r}_n)} d\alpha, \quad (75)$$

where  $\theta_g$  and  $\Delta_g$  are, respectively, the azimuth angle at which the  $g$ th group is located and the angular spread of the departure waves to the  $g$ th group which is determined as  $\Delta_g \approx \tan^{-1}(s_g/d_g)$ . Here,  $s_g$  and  $d_g$  are, respectively, the radius of the ring of scatterers for the  $g$ th group and the distance between the BS and the  $g$ th group. (see Fig. 1). The parameter  $\lambda_0$  is the wavelength of signal,  $\mathbf{r}_m = [x_m, y_m]^T$  is the position vector of the  $m$ th antenna element, and  $\Omega(\alpha)$  indicates the wave vector with the angle-of-departure (AoD),  $\alpha$ , given by  $\Omega(\alpha) = (\cos(\alpha), \sin(\alpha))$ .

1) *Suitability of multi-polarized antenna elements in multi-user Massive MIMO system:* To see the suitability of the dual polarized antenna elements in multi-user Massive MIMO system, we have compared the performance of the linear array antennas with dual polarized antenna elements and single polarized antenna elements, when the dual structured precoding scheme with BD is applied. The number of antenna elements in both single polarized and dual polarized linear arrays are set as  $M = 120$ . The dual polarized linear array is then composed of 60 vertically and 60 horizontally polarized antenna elements. Here, the XPD parameter is set as  $\chi = 0.1$ . Total 32 MSs with a single antenna element are clustered into 4 groups having the same number of MSs per group ( $N = 32$ ,  $G = 4$ ,  $\bar{N} = 8$ ). Here, we assumed that when the single polarized linear array is deployed at the BS, the antenna elements of all MSs are co-polarized with those of BS (which is the optimal scenario for the single polarized linear array). For the dual polarized array case, in each group,  $\frac{\bar{N}}{2}$  vertically polarized and  $\frac{\bar{N}}{2}$  horizontally polarized MSs coexist.<sup>5</sup> For preprocessing, we set  $\bar{B} = 14$  for both the single/dual polarized cases such that  $\bar{N} \leq \bar{B} \leq (M - (G - 1)r')$  and  $\bar{B} \leq r'$ , where  $r'$  is the minimum among the rank of  $\mathbf{R}_g$ ,  $g = 1, \dots, G$ . In addition,  $\Delta_1 = \dots = \Delta_4 = \Delta = \frac{8}{180}\pi$  and  $\theta_g = -\frac{\pi}{4} + \frac{\pi}{6}(g - 1)$  for  $g = 1, \dots, 4$ . Fig. 4 shows the sum-rate curves for the arrays with dual polarized antenna elements,  $d_s = \frac{\lambda_0}{2}$  and with single polarized antenna elements,  $d_s = \{\frac{\lambda_0}{2}, \frac{\lambda_0}{4}\}$ , where  $d_s$  is the inter-antenna distance. We can see that the dual polarized array with the array size of  $30\lambda_0$  exhibits better

<sup>5</sup>Note that this environment can be made by choosing proper  $\frac{\bar{N}}{2}$  vertically polarized and  $\frac{\bar{N}}{2}$  horizontally polarized MSs when we have enough MSs in a cell. user selection (multi-user diversity) in this paper.

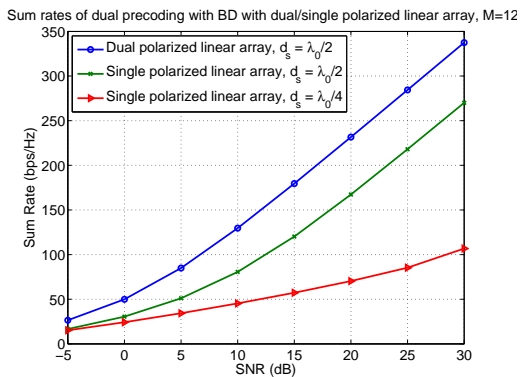


Fig. 4. Sum rates of the dual precoding with BD with dual/single polarized linear arrays.

performance than the single polarized one with the size of  $60\lambda_0$  (corresponding to  $d_s = \frac{\lambda_0}{2}$ ). Furthermore, we can see that, if we let the single polarized array have the same size as the dual polarized one by reducing its inter-antenna space, its performance worsens significantly. Accordingly, the multi-polarized antenna can be one possible solution that partially alleviates the space limitation of Massive MIMO system.

2) *Performance comparison of dual precodings with BD and BDS*: To evaluate the performance of the dual precoding schemes with BD and BDS, we have also run Monte-Carlo (MC) simulations. Here, it is assumed that BS has a dual polarized linear array antenna with  $M = 120$ . It is also assumed that  $N = 32$ ,  $G = 4$ ,  $\bar{N} = 8$ . For preprocessing, we set as  $\bar{B} = \min(2\bar{N}, 2r)$ , where  $r$  is the minimum among the rank of  $\mathbf{R}_g^s$ ,  $g = 1, \dots, G$ . In addition,  $\mathbf{R}_g^s$  is generated by (75) with  $d_s = \frac{\lambda_0}{2}$ ,  $\Delta = \frac{\pi}{12}$ , and  $\theta_g = -\frac{\pi}{4} + \frac{\pi}{6}(g-1)$  for  $g = 1, \dots, 4$ .

Fig. 5 shows the sum rate of the dual precoding with BD and BDS when the perfect CSIT is assumed (i.e.,  $\tau^2 = 0$ ). Note that when  $\chi = 0$ , the dual precoding with BD and BDS exhibit the same sum rate performance, as mentioned in Remark 3 and (31). However, when  $\chi = 0.1$ , the performance of the dual precoding with BDS is degraded, while that of the dual precoding with BD does not change significantly compared to the case of  $\chi = 0$ . Fig. 6 shows the sum rates when  $\chi = 0$  for the imperfect CSIT with  $\tau^2 = 0.1$  (i.e.,  $\tau_{BD}^2 = \tau^2$  and  $\tau_{BDS}^2 = \tau$  from (60) and (61)). We can see that the BDS exhibits better performance than the BD because the BDS can exploit more accurate short-term CSIT due to the feedback of the smaller dimensional channel instance. Interestingly, the gap between the analytic results and MC simulation results becomes larger as SNR increases. This is because the analytical derivation is based on Theorem 1 with  $z = -\alpha$  and the approximation error bound is proportional to  $\frac{1}{\alpha} = \frac{\bar{B}P}{N}$ , which is also addressed in [Proposition 12, 7]. In Fig. 7, we provide the sum rate curves of dual precoding schemes as a function of  $\chi$  for  $\tau^2 = \{0, 0.5, 1.0, 1.5\}$  when  $SNR = 15dB$ . Here, the approximated sum-rate is obtained from the approximated SINR in Proposition 1 and 2. Note that the performance of the dual precoding with BD is not affected by the variation of  $\chi$ , while that with BDS decreases

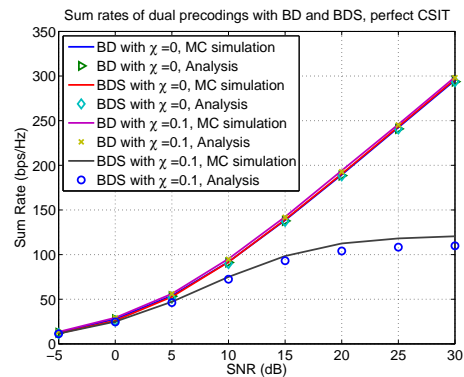


Fig. 5. Sum rates of the dual precoding with BD and BDS under the perfect CSIT.

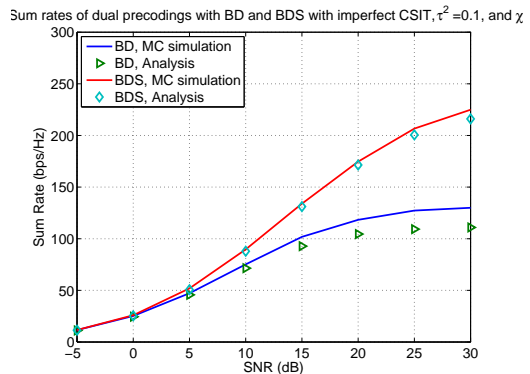


Fig. 6. Sum rates of the dual precoding with BD and BDS under the imperfect CSIT,  $\tau^2 = 0.1$ .

as  $\chi$  increases. However, the dual precoding with BD is largely affected by  $\tau^2$ . As  $\tau^2$  increases, the region that BDS outperforms BD becomes wider. Note that the crossing point in Fig. 7 is located where  $\chi \approx \tau^2$ , which agrees with (67) when  $B_0 \gg 1$ .

In Fig. 8, we also compare the sum rates of BD and BDS versus the number of feedback bits per user when  $\chi = \{0.1, 0.2\}$ ,  $SNR = 25$ . Per Remark 4, when the feedback bits are not enough to describe the short-term CSIT accurately, the BDS exhibits a better performance than the BD. Here, the cross point corresponds to  $N_B \approx (2r - 1) \left( \log_2 \left( 1 + E_{g,p} \left( \frac{(1+m_{gp}^o(0))^2 - 1}{(\xi_{gp}^o(0))^2 \tau_{gpp}^o(0)(1+m_{gp}^o(0))^2} \right) \right) - \log_2 \chi \right)$  in (62).

3) *Performance comparison of the proposed dual precoding*: To verify the performance of the proposed dual structured precoding in Section V, we have compared its sum rates with those of dual precodings with BD and BDS. In Fig. 9, we set  $N_B = \{50, 65\}$  and  $\chi$  is uniformly distributed on  $[0, 0.5]$ . We can find that the sum rates for  $N_B = 65$  is higher than that for  $N_B = 50$  and the sum rates of all schemes are saturated at high SNR due to the imperfect CSIT. Note that the proposed scheme exhibits higher performance than the other two schemes, as mentioned in Section V.A.

In Fig. 10, we have evaluated the 3D dual structured precoding in Section V-B when  $10 \times 50$  uniform planar array is deployed at BS with a height of  $60m$  and there are

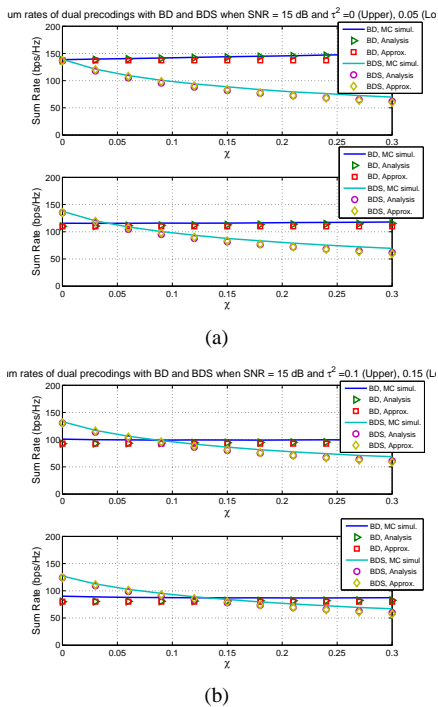


Fig. 7. Sum rates of the dual precoding with BD and BDS over  $\chi$  for (a)  $\tau^2 = \{0, 0.5\}$  (b)  $\tau^2 = \{1.0, 1.5\}$  when  $SNR = 15dB$ .

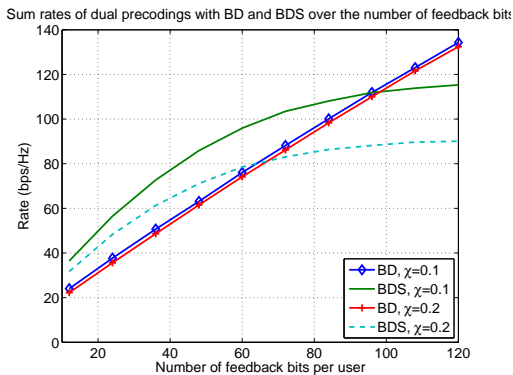


Fig. 8. Sum rates of the dual precoding with BD and BDS versus the number of feedback bits for  $\chi = \{0.1, 0.2\}$ .

three elevation regions with  $d_{gl} \in \{30, 60, 100\}m$ , each with 4 groups. We assume the equal power allocation over the elevation region,  $SNR = 25dB$ , and that  $\tau^2$  is uniformly distributed on  $[0, 1]$ . Each group has 8 MSs and all groups have the same angular spread  $\Delta = \frac{\pi}{12}$ . The radius of ring of scatterer is then given as  $s_{gl} = d_{gl}\tan(\Delta)$  and the path loss is modeled as  $P_{loss,l} = \frac{1}{1 + (\frac{d_{gl}}{60})^3}$ . Then, the elevation angle

spread can be computed as  $\tan^{-1}\left(\frac{60}{d_{gl}-s_{gl}}\right) - \tan^{-1}\left(\frac{60}{d_{gl}}\right)$  (see Fig. 3). In addition, the mismatch angle for each user is randomly generated as  $\theta_{gk}^{ms} \sim U[-\theta_{\max}^{ms}, \theta_{\max}^{ms}]$  with (a)  $\theta_{\max}^{ms} = 0$  and (b)  $\theta_{\max}^{ms} = 0.22\pi$ . We can see that the overall performances for  $\theta_{\max}^{ms} = 0.22\pi$  is outperformed by those for  $\theta_{\max}^{ms} = 0$  (i.e., the polarization is perfectly aligned). In addition, the performance of dual precoding with BDS is more sensitively degraded than that with the BD. For small  $\chi$  the

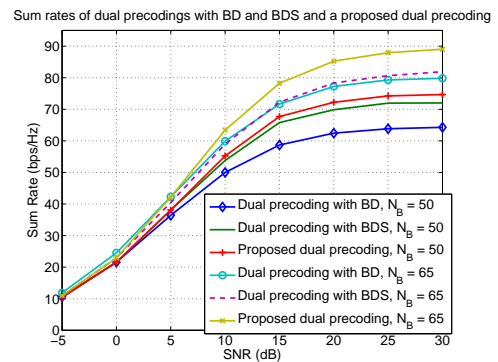


Fig. 9. Sum rates of dual precodings with BD and BDS and a proposed dual precoding with  $N_B = \{50, 65\}$ .

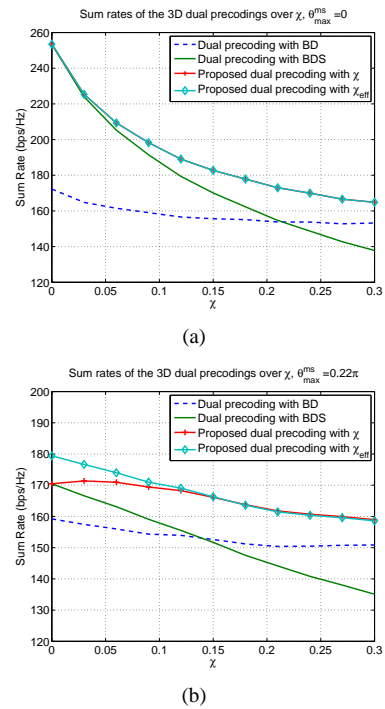


Fig. 10. Sum rates of the 3D dual precodings over  $\chi$  for (a)  $\theta_{\max}^{ms} = 0$  (b)  $\theta_{\max}^{ms} = 0.22\pi$ .

dual precoding with BDS exhibits better performance than that with BD regardless of polarization mismatch. Furthermore, the proposed dual precoding in Section V also outperforms the two other schemes. That is, by switching between BD and BDS based on the long-term CSIT parameters (spatial correlation and XPD) jointly with the number of short-term feedback bits (or, short-term CSIT quality), the performance of the dual structured precoding can be improved, regardless of the polarization mismatch. In addition, using the effective XPD parameter  $\chi_{eff}$  shows better performance than the actual XPD parameter  $\chi$ .

## VII. CONCLUSION

In this paper, we have investigated the dual structured linear precoding in the multi-polarized MU Massive MIMO system. In the dual precoding with BD, MSs are grouped

based only on the spatial correlation. However, in that with BDS, by subgrouping the co-polarized MSs in the spatially separated groups, we can further reduce the short-term CSI feedback overhead. Based on the random matrix theory, the system performances of dual structured precoding schemes are asymptotically analyzed. From the asymptotic results, we have found that the performance of the dual precoding with BD is insensitive to the XPD, while that of BDS is affected by the XPD parameter. Because the dual precoding with BDS can have more accurate CSIT (across half of the array) than that with BD under the same number of feedback bits, the region of the number of feedback bits where the BDS exhibits better performance than the BD is analytically derived. That is, assuming the CSI is perfectly estimated at MSs, when the number of feedback bits is not large enough to describe the short-term CSIT accurately and the XPD parameter ( $\chi$ ) is small, the dual precoding with BDS exhibits a better performance. Finally, based on that observation, we have proposed a new dual structured precoding/feedback in which the precoding mode is switched between BD and BDS depending on the XPD, spatial correlation, and the number of short-term feedback bits (short-term CSIT quality) and extended it to 3D dual structured precoding.

#### APPENDIX A PROOF OF THEOREM 2

We note that our derivation is based on the derivation of the asymptotic SINR of regularized ZF precoding in spatially correlated MISO broadcasting under the imperfect CSIT [7]. The main difference is that the asymptotic inter/intra interferences are evaluated over the polarization domain as well as the spatial domain. We first consider the normalization factor  $\xi_g^2$  in (20). By letting  $\Psi \triangleq \frac{P}{N} \text{tr}(\hat{\mathbf{H}}_g^H \hat{\mathbf{K}}_g^H \hat{\mathbf{K}}_g \hat{\mathbf{H}}_g)$ ,  $\xi_g^2$  can be rewritten as

$$\xi_g^2 = \frac{P}{G} \frac{1}{\Psi}. \quad (76)$$

By substituting  $\hat{\mathbf{K}}_g$  in (19), due to the matrix inversion lemma,  $\Psi$  can be written as shown at the top of the page. From Lemma 4 in [7], we have

$$\Psi - \frac{P}{N\bar{B}} \sum_{p \in \{v,h\}} \sum_{k=1}^{\bar{N}/2} \frac{\frac{1}{B} \text{tr}(\mathbf{B}_g^H \mathbf{R}_{gp} \mathbf{B}_g (\mathbf{A}_g^{-k} + \alpha \mathbf{I}_{\bar{B}})^{-2})}{(1 + \frac{1}{B} \text{tr}(\mathbf{B}_g^H \mathbf{R}_{gp} \mathbf{B}_g (\mathbf{A}_g^{-k} + \alpha \mathbf{I}_{\bar{B}})^{-1}))^2} \xrightarrow{M \rightarrow \infty} 0. \quad (78)$$

Because the covariance matrix of users in the same co-polarized subgroup is equal, from Lemma 6 of [7], we have

$$\Psi - \frac{\bar{N}P}{2N\bar{B}} \sum_{p \in \{v,h\}} \frac{\frac{1}{B} \text{tr}(\mathbf{B}_g^H \mathbf{R}_{gp} \mathbf{B}_g (\mathbf{A}_g + \alpha \mathbf{I}_{\bar{B}})^{-2})}{(1 + \frac{1}{B} \text{tr}(\mathbf{B}_g^H \mathbf{R}_{gp} \mathbf{B}_g (\mathbf{A}_g + \alpha \mathbf{I}_{\bar{B}})^{-1}))^2} \xrightarrow{M \rightarrow \infty} 0, \quad (79)$$

where  $\mathbf{A}_g = \frac{1}{B} \hat{\mathbf{H}}_g \hat{\mathbf{H}}_g^H$ . Then, we define  $m_{gp}(z) \triangleq \frac{1}{B} \text{tr}(\mathbf{B}_g^H \mathbf{R}_{gp} \mathbf{B}_g (\mathbf{A}_g - z \mathbf{I}_{\bar{B}})^{-1})$ , which is analogous to (32) in Theorem 1 by setting  $\mathbf{S} = \mathbf{0}$  and  $\mathbf{Q} = \mathbf{B}_g^H \mathbf{R}_{gp} \mathbf{B}_g (= \bar{\mathbf{R}}_{gp})$

from (17)). The trace in the denominator of (79) is then equal to  $m_{gp}(-\alpha)$  and, from Theorem 1,

$$\frac{1}{B} \text{tr}(\mathbf{B}_g^H \mathbf{R}_{gp} \mathbf{B}_g (\mathbf{A}_g + \alpha \mathbf{I}_{\bar{B}})^{-1}) - \frac{1}{B} \text{tr}(\bar{\mathbf{R}}_{gp} \mathbf{T}_g) \xrightarrow{M \rightarrow \infty} 0, \quad (80)$$

where  $\mathbf{T}_g$  is given by (37) from (33) and (34). Furthermore, the trace in the numerator of (79) is the derivative of  $m_{gp}(z)$  at  $z = -\alpha$ , i.e.,  $m'_{gp}(-\alpha)$ . Therefore,

$$\frac{1}{B} \text{tr}(\mathbf{B}_g^H \mathbf{R}_{gp} \mathbf{B}_g (\mathbf{A}_g + \alpha \mathbf{I}_{\bar{B}})^{-2}) - \frac{1}{B} \text{tr}(\bar{\mathbf{R}}_{gp} \mathbf{T}'_g) \xrightarrow{M \rightarrow \infty} 0, \quad (81)$$

where

$$\mathbf{T}'_g = \mathbf{T}_g \left( \frac{\bar{N}}{2\bar{B}} \sum_{p \in \{v,h\}} \frac{\bar{\mathbf{R}}_{gp} m'_{gp}}{(1 + m_{gp}(-\alpha))^2} + \mathbf{I}_{\bar{B}} \right) \mathbf{T}_g, \quad (82)$$

$$m'_{gp} = \frac{1}{B} \text{tr}(\bar{\mathbf{R}}_{gp} \mathbf{T}'_g). \quad (83)$$

Then, by putting  $\mathbf{T}'_g$  in (82) into (83),  $\mathbf{m}'_g$  can be obtained as in (39) and  $\Psi$  is converged as

$$\Psi - \frac{P\bar{N}}{2N\bar{B}} \sum_{p \in \{v,h\}} \frac{m'_{gp}}{(1 + m_{gp}^o)^2} \xrightarrow{M \rightarrow \infty} 0. \quad (84)$$

For the signal power component  $\frac{P}{N} |\mathbf{h}_{gk}^H \mathbf{B}_g \hat{\mathbf{K}}_g \hat{\mathbf{h}}_{gk}|^2$  and the intra-group interference component  $\frac{P}{N} \sum_{j \neq k} |\mathbf{h}_{gk}^H \mathbf{B}_g \hat{\mathbf{K}}_g \hat{\mathbf{h}}_{gj}|^2$  in (21), because  $\mathbf{G}_g$  and  $\mathbf{Z}_g$  in (8) are independent, by taking a similar approach described in (77) and (78) (See also Appendix II.B and C in [7]), we can have the following relations:

$$|\mathbf{h}_{gk}^H \mathbf{B}_g \hat{\mathbf{K}}_g \hat{\mathbf{h}}_{gk}|^2 - \frac{(1 - \tau^2)(m_{gp}^o)^2}{(1 + m_{gp}^o)^2} \xrightarrow{M \rightarrow \infty} 0, \quad (85)$$

$$\sum_{j \neq k} |\mathbf{h}_{gk}^H \mathbf{B}_g \hat{\mathbf{K}}_g \hat{\mathbf{h}}_{gj}|^2 - \frac{\Upsilon_{gkp}^k (1 - \tau^2 (1 - (1 + m_{gp}^o)^2))}{(1 + m_{gp}^o)^2} \xrightarrow{M \rightarrow \infty} 0, \quad (86)$$

where

$$\Upsilon_{gkp}^k = \frac{1}{B} \sum_{j \neq k} \hat{\mathbf{g}}_{gj}^H \mathbf{R}_{gp}^{1/2} \mathbf{B}_g \hat{\mathbf{K}}_g \mathbf{B}_g^H \mathbf{R}_{gp} \mathbf{B}_g \hat{\mathbf{K}}_g \mathbf{B}_g^H \mathbf{R}_{gp}^{1/2} \hat{\mathbf{g}}_{gj} + \frac{1}{B} \sum_{j=1}^{\bar{N}/2} \hat{\mathbf{g}}_{gj}^H \mathbf{R}_{gp}^{1/2} \mathbf{B}_g \hat{\mathbf{K}}_g \mathbf{B}_g^H \mathbf{R}_{gp} \mathbf{B}_g \hat{\mathbf{K}}_g \mathbf{B}_g^H \mathbf{R}_{gp}^{1/2} \hat{\mathbf{g}}_{gj}. \quad (87)$$

Note that by taking a similar approach described in (77) and (78),

$$\Upsilon_{gkp}^k - \frac{\bar{N}/2 - 1}{B} \frac{\text{tr}(\bar{\mathbf{R}}_{gp} (\mathbf{A}_g + \alpha \mathbf{I}_{\bar{B}})^{-1} \bar{\mathbf{R}}_{gp} (\mathbf{A}_g + \alpha \mathbf{I}_{\bar{B}})^{-1})}{(1 + \frac{1}{B} \text{tr}(\bar{\mathbf{R}}_{gp} (\mathbf{A}_g + \alpha \mathbf{I}_{\bar{B}})^{-1}))^2} - \frac{\bar{N}}{2B} \frac{\text{tr}(\bar{\mathbf{R}}_{gp} (\mathbf{A}_g + \alpha \mathbf{I}_{\bar{B}})^{-1} \bar{\mathbf{R}}_{gp} (\mathbf{A}_g + \alpha \mathbf{I}_{\bar{B}})^{-1})}{(1 + \frac{1}{B} \text{tr}(\bar{\mathbf{R}}_{gp} (\mathbf{A}_g + \alpha \mathbf{I}_{\bar{B}})^{-1}))^2} \xrightarrow{M \rightarrow \infty} 0. \quad (88)$$

By defining  $m_{gppq}(z) \triangleq \frac{1}{B} \text{tr}(\bar{\mathbf{R}}_{gp} (\mathbf{A}_g + \alpha \mathbf{I}_{\bar{B}} - z \bar{\mathbf{R}}_{gp})^{-1})$ , the trace in the numerator of (88) is the derivative of  $m_{gppq}(z)$  at  $z = 0$ . Furthermore, by using Theorem 1,

$$m_{gppq}(z) - \frac{1}{B} \text{tr}(\bar{\mathbf{R}}_{gp} \mathbf{T}_{gp}(z)) \xrightarrow{M \rightarrow \infty} 0, \quad (89)$$

$$\begin{aligned}\Psi &= \frac{P}{N} \sum_{k=1}^{\bar{N}} \hat{\mathbf{h}}_{gk}^H \mathbf{B}_g (\hat{\mathbf{H}}_g \hat{\mathbf{H}}_g^H + \bar{B} \alpha \mathbf{I}_{\bar{B}})^{-2} \mathbf{B}_g^H \hat{\mathbf{h}}_{gk} = \frac{P}{N\bar{B}} \sum_{k=1}^{\bar{N}} \frac{\hat{\mathbf{h}}_{gk}^H \mathbf{B}_g (\mathbf{A}_g^{-k} + \alpha \mathbf{I}_{\bar{B}})^{-2} \mathbf{B}_g^H \hat{\mathbf{h}}_{gk}}{(1 + \hat{\mathbf{h}}_{gk}^H \mathbf{B}_g (\mathbf{A}_g^{-k} + \alpha \mathbf{I}_{\bar{B}})^{-1} \mathbf{B}_g^H \hat{\mathbf{h}}_{gk})^2} \\ &= \frac{P}{N\bar{B}} \sum_{p \in \{v, h\}} \sum_{k=1}^{\bar{N}/2} \frac{\hat{\mathbf{g}}_{gk}^H \mathbf{R}_{gp}^{1/2} \mathbf{B}_g (\mathbf{A}_g^{-k} + \bar{B} \alpha \mathbf{I}_{\bar{B}})^{-2} \mathbf{B}_g^H \mathbf{R}_{gp}^{1/2} \hat{\mathbf{g}}_{gk}}{(1 + \hat{\mathbf{g}}_{gk}^H \mathbf{R}_{gp}^{1/2} \mathbf{B}_g (\mathbf{A}_g^{-k} + \alpha \mathbf{I}_{\bar{B}})^{-1} \mathbf{B}_g^H \mathbf{R}_{gp}^{1/2} \hat{\mathbf{g}}_{gk})^2},\end{aligned}\quad (77)$$

where  $\mathbf{A}_g^{-k} = \frac{1}{\bar{B}} \hat{\mathbf{H}}_g^{-k} (\hat{\mathbf{H}}_g^{-k})^H$  and  $\hat{\mathbf{H}}_g^{-k} = [\hat{\mathbf{h}}_{g1}, \dots, \hat{\mathbf{h}}_{gk-1}, \hat{\mathbf{h}}_{gk+1}, \dots, \hat{\mathbf{h}}_{g\bar{N}}]$ .

where  $\mathbf{T}_{gp}(z) = (\frac{\bar{N}}{2\bar{B}} \sum_{q \in \{v, h\}} \frac{\bar{\mathbf{R}}_{gq}}{1 + m_{ggqp}(z)} + \alpha \mathbf{I}_{\bar{B}} - z \bar{\mathbf{R}}_{gp})^{-1}$ . First, let us consider the high SNR regime (i.e., small  $\alpha$ ). Accordingly, we have Because

$$\begin{aligned}\mathbf{T}'_{gp}(z) &= \mathbf{T}_{gp}(z) \left( \frac{\bar{N}}{2\bar{B}} \sum_{q \in \{v, h\}} \frac{\bar{\mathbf{R}}_{gq} m'_{ggqp}(z)}{(1 + m_{ggqp}(z))^2} + \bar{\mathbf{R}}_{gp} \right) \mathbf{T}_{gp}(z). \quad (90) \quad \text{tr}(\mathbf{R}(\beta \mathbf{R} + \alpha \mathbf{I})^{-1}) = \frac{1}{\beta} \text{tr}(\mathbf{R}(\mathbf{R} + \alpha/\beta \mathbf{I})^{-1}) \\ &\approx \frac{1}{\beta} \text{tr}(\mathbf{R}(\mathbf{R} + \alpha \mathbf{I})^{-1}),\end{aligned}\quad (97)$$

Note that  $\mathbf{T}_{gp}(0) = \mathbf{T}_g$  in (37) and from (89),  $m_{ggpq}(0) - m_{gq}^o \xrightarrow{M \rightarrow \infty} 0$ . Together with  $m'_{ggqp}(z) - \frac{1}{\bar{B}} \text{tr}(\bar{\mathbf{R}}_{gq} \mathbf{T}'_{gp}(z)) \xrightarrow{M \rightarrow \infty} 0$ ,  $m'_{gp}$  can be obtained as (39). Accordingly,

$$\Upsilon_{ggp}^k - \Upsilon_{ggp}^o \xrightarrow{M \rightarrow \infty} 0, \quad (91)$$

where  $\Upsilon_{ggp}^o$  is defined in (38).

Now let us consider the inter-group interference component  $\Upsilon_{glp}^k \triangleq \sum_j |\mathbf{h}_{gk}^H \mathbf{B}_l \hat{\mathbf{K}}_l \mathbf{B}_l^H \hat{\mathbf{h}}_{lj}|^2$  in the denominator of (21) for  $l \neq g$ . Then,  $\Upsilon_{glp}^k$  can be rewritten as

$$\Upsilon_{glp}^k = \frac{1}{\bar{B}} \sum_{j=1}^{\bar{N}/2} \sum_{q \in \{v, h\}} \hat{\mathbf{g}}_{lj}^H \mathbf{R}_{lq}^{1/2} \mathbf{B}_l \hat{\mathbf{K}}_l \mathbf{B}_l^H \mathbf{R}_{gp} \mathbf{B}_l \hat{\mathbf{K}}_l \mathbf{B}_l^H \mathbf{R}_{lq}^{1/2} \hat{\mathbf{g}}_{lj}. \quad (92)$$

Note that (92) has a similar form with (87). Therefore, similarly done in (88), we have

$$\Upsilon_{glp}^k - \frac{\bar{N}}{2\bar{B}} \sum_{q \in \{v, h\}} \frac{\frac{1}{\bar{B}} \text{tr}(\bar{\mathbf{R}}_{lq} (\mathbf{A}_l + \alpha \mathbf{I}_{\bar{B}})^{-1} \mathbf{B}_l^H \mathbf{R}_{gp} \mathbf{B}_l (\mathbf{A}_l + \alpha \mathbf{I}_{\bar{B}})^{-1})}{(1v + \frac{1}{\bar{B}} \text{tr}(\bar{\mathbf{R}}_{lq} (\mathbf{A}_l + \alpha \mathbf{I}_{\bar{B}})^{-1}))^2} \xrightarrow{M \rightarrow \infty} 0.$$

Furthermore, by following the steps similarly done in (89)-(90), we can obtain

$$\Upsilon_{gl}^k - \Upsilon_{gl}^o \xrightarrow{M \rightarrow \infty} 0, \quad (94)$$

where  $\Upsilon_{glp}^o$  is defined in (38). By substituting (84), (85), (86), (91), and (94) into (21), we can have (35).

#### APPENDIX B PROOF OF PROPOSITION 1

Let  $\mathbf{R}_g(\chi) = (1 + \chi) \begin{bmatrix} \mathbf{R}_g^s & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_g^s \end{bmatrix}$  from (7). Then, we have  $\mathbf{R}_g(\chi) = (1 + \chi) \mathbf{R}_g(0)$  (respectively,  $\bar{\mathbf{R}}_g(\chi) = (1 + \chi) \bar{\mathbf{R}}_g(0)$ ) and, from Corollary 1,  $\gamma_{gpk}^{BD,o}(\chi)$  can be evaluated by using (43) with  $\frac{1}{2} \bar{\mathbf{R}}_g(\chi)$ . From (44), we define

$$\mathbf{T}_g(\chi) = \left( \frac{\bar{N}}{2\bar{B}} \frac{\bar{\mathbf{R}}_g(\chi)}{1 + m_g^o(\chi)} + \alpha \mathbf{I}_{\bar{B}} \right)^{-1}, \quad (95)$$

$$m_g^o(\chi) = \frac{1}{2\bar{B}} \text{tr}(\bar{\mathbf{R}}_g(\chi) \mathbf{T}_g(\chi)). \quad (96)$$

for any Hermitian nonnegative definite matrix  $\mathbf{R}$  and small  $\alpha$ , by substituting (95) into (96), we have

$$\begin{aligned}m_g^o(\chi) &= \frac{1}{2\bar{B}} \text{tr}((1 + \chi) \bar{\mathbf{R}}_g(0) \left( \frac{\bar{N}}{2\bar{B}} \frac{(1 + \chi) \bar{\mathbf{R}}_g(0)}{1 + m_g^o(\chi)} + \alpha \mathbf{I}_{\bar{B}} \right)^{-1}), \\ &\approx \frac{1}{2\bar{B}} \text{tr}(\bar{\mathbf{R}}_g(0) \left( \frac{\bar{N}}{2\bar{B}} \frac{\bar{\mathbf{R}}_g(0)}{1 + m_g^o(\chi)} + \alpha \mathbf{I}_{\bar{B}} \right)^{-1}),\end{aligned}\quad (98)$$

which implies that

$$m_g^o(\chi) = m_g^o(0). \quad (99)$$

Furthermore, by substituting (99) into (95),

$$\mathbf{T}_g(\chi) = \frac{1}{1 + \chi} \left( \frac{\bar{N}}{2\bar{B}} \frac{\bar{\mathbf{R}}_g(0)}{1 + m_g^o(0)} + \frac{\alpha}{1 + \chi} \mathbf{I}_{\bar{B}} \right)^{-1} \quad (100)$$

By using (97), (99) and (100), we can also derive

$$(93)'_g(\chi) \approx \frac{1}{1 + \chi} m'_g(0), \quad m'_{gg}(\chi) \approx m'_{gg}(0), \quad m'_{gl}(\chi) \approx m'_{gl}(0) \quad (101)$$

which implies that

$$\begin{aligned}\Psi_g^o(\chi) &\approx \frac{1}{1 + \chi} \Psi_g^o(0), \quad \Upsilon_{gg}^o(\chi) \approx \Upsilon_{gg}^o(0), \\ \Upsilon_{gl}^o(\chi) &\approx \Upsilon_{gl}^o(0), \quad \xi_g^o(\chi) \approx (1 + \chi) \xi_g^o(0).\end{aligned}\quad (102)$$

Then, based on (99) and (102) together with (43),  $\gamma_{gpk}^{BD,o}(\chi) \approx \gamma_{gpk}^{BD,o}(0)$ .

For low SNR regime (i.e., large  $\alpha$ ), in (95),  $\mathbf{T}_g(\chi) \approx \frac{1}{\alpha} \mathbf{I}_{\bar{B}}$  and accordingly,  $m_g^o(\chi) \approx (1 + \chi) m_g^o(0) \ll 1$ . Furthermore,

$$\begin{aligned}m'_g(\chi) &\approx (1 + \chi) m'_g(0), \quad m'_{gg}(\chi) \approx (1 + \chi)^2 m'_{gg}(0), \\ m'_{gl}(\chi) &\approx (1 + \chi)^2 m'_{gl}(0),\end{aligned}\quad (103)$$

and  $\Psi_g^o(\chi) \approx (1 + \chi) \Psi_g^o(0)$ ,  $\Upsilon_{gg}^o(\chi) \approx (1 + \chi)^2 \Upsilon_{gg}^o(0)$ ,  $\Upsilon_{gl}^o(\chi) \approx (1 + \chi)^2 \Upsilon_{gl}^o(0)$ , and  $\xi_g^o(\chi) \approx \frac{1}{1 + \chi} \xi_g^o(0)$ . Accordingly, in the low SNR regime, we can have that  $\gamma_{gpk}^{BD,o}(\chi) \approx \gamma_{gpk}^{BD,o}(0)$ . Accordingly, in the low SNR regime, we can have that  $\gamma_{gpk}^{BD,o}(\chi) \approx \gamma_{gpk}^{BD,o}(0)$ .

APPENDIX C  
PROOF OF PROPOSITION 2

From (7) and (24), we have  $\bar{\mathbf{R}}_{gv}(\chi) = \bar{\mathbf{R}}_{gh}(\chi) = (\mathbf{B}_g^s)^H \mathbf{R}_g^s \mathbf{B}_g^s$ . That is,  $\bar{\mathbf{R}}_{gv}(\chi)$  and  $\bar{\mathbf{R}}_{gh}(\chi)$  are independent of  $\chi$  and expressed as  $\bar{\mathbf{R}}_{gp}(\chi) = \bar{\mathbf{R}}_{gp}(0)$ . Accordingly, from (53) and (55),

$$\begin{aligned} m_{gp}^o(\chi) &= m_{gp}^o(0), & \mathbf{T}_{gp}(\chi) &= \mathbf{T}_{gp}(0), \\ m'_{gp}(\chi) &= m'_{gp}(0), & m'_{gpp}(\chi) &= m'_{gpp}(0). \end{aligned} \quad (104)$$

Because  $\mathbf{B}_{lq}^H \mathbf{R}_{gp}(\chi) \mathbf{B}_{lq} = \begin{cases} \mathbf{B}_{lq}^H \mathbf{R}_{gp}(0) \mathbf{B}_{lq} & \text{for } q = p \\ \chi \mathbf{B}_{lq}^H \mathbf{R}_{gq}(0) \mathbf{B}_{lq} & \text{for } q \neq p \end{cases}$ ,  
from (53),  $m'_{glpq}(\chi) = \begin{cases} m'_{glpq}(0) & \text{for } q = p \\ \chi m'_{glpq}(0) & \text{for } q \neq p \end{cases}$ . Therefore,  
from (51),

$$\gamma_{gpk}^{BDS,o}(\chi) = \frac{\frac{P}{N}(\xi_{gp}^o)^2(1-\tau^2)(m_{gp}^o)^2}{IN_{gpk}^o(\chi)}, \quad (105)$$

where  $IN_{gpk}^o(\chi)$  is given as shown at the top of the page.

Because of the preprocessing with BD,  $\Upsilon_{gpp}^o(0) \gg \Upsilon_{glpp}^o(0)$  for  $g \neq l$  and  $IN_{gpk}^o(\chi) \approx IN_{gpk}^o(0)(1 + c_{0,gpk}\chi)$ , where  $c_{0,gpk} = \frac{(\xi_{gp}^o(0))^2 \Upsilon_{gpp}^o(0)(1+m_{gp}^o(0))^2}{(\xi_{gp}^o(0))^2 \Upsilon_{gpp}^o(0)(\tau^2((1+m_{gp}^o(0))^2-1)+1)+(1+m_{gp}^o(0))^2}$ . By averaging  $c_{0,gpk}$  over  $g, p$ , we can have (59).

REFERENCES

- [1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
- [2] 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 Processing Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
- [3] H. Huh, G. Caire, H. C. Papadopoulos, and S. A. Ramprasad, "Achieving "massive MIMO" spectral efficiency with a not-so-large number of antennas," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3226–3239, Sept. 2012.
- [4] J. Hoydis, S. Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Select. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
- [5] S. K. Mohammed and E. G. Larsson, "Single-user beamforming in large-scale MISO systems with per-antenna constant-envelope constraints: the doughnut channel," *IEEE Trans. Wireless Commun.*, vol. 11, no. 11, pp. 3992–4005, Nov. 2012.
- [6] —, "Per-antenna constant envelope precoding for large multi-user MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 1059–1071, Mar. 2013.
- [7] S. Wagner, R. Couillet, M. Debbah, and D. T. M. Slock, "Large system analysis of linear precoding in correlated MISO broadcast channels under limited feedback," *IEEE Trans. Inform. Theory*, vol. 58, no. 7, pp. 4509–4537, July 2012.
- [8] A. Adhikary, J. Nam, J. Ahn, and G. Caire, "Joint spatial division and multiplexing—the large-scale array regime," *IEEE Trans. Inform. Theory*, vol. 59, no. 10, pp. 6441–6463, Oct. 2013.
- [9] J. Zhang, C. Wen, S. Jin, X. Gao, and K. Wong, "Large system analysis of cooperative multi-cell downlink transmission via regularized channel inversion with imperfect CSIT," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 4801–4813, Oct. 2013.
- [10] B. Clerckx and C. Oestges, *MIMO Wireless Networks: Channels, Techniques and Standards for Multi-Antenna, Multi-User and Multi-Cell Systems*. Oxford, UK: Academic Press (Elsevier), 2013.
- [11] C. Lim, T. Yoo, B. Clerckx, B. Lee, and B. Shim, "Recent trend of multiuser MIMO in LTE Advanced," *IEEE Commun. Mag.*, vol. 51, no. 3, pp. 127–135, Mar. 2013.
- [12] B. Clerckx, G. Kim, and S. Kim, "Correlated fading in broadcast MIMO channels: curse or blessing?" in *Proc. IEEE Global Telecommunications Conference, 2008*, Nov. 2008.
- [13] B. Clerckx, C. Craeye, D. Vanhoenacker-Janvier, and C. Oestges, "Impact of antenna coupling on 2x2 MIMO communications," *IEEE Trans. Veh. Technol.*, vol. 56, no. 3, pp. 1009–1018, May 2007.
- [14] T. Kim, B. Clerckx, D. J. Love, and S. J. Kim, "Limited feedback beamforming systems for dual-polarized MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3425–3439, Nov. 2010.
- [15] C. Oestges, B. Clerckx, M. Guillaud, and M. Debbah, "Dual-polarized wireless communications: from propagation models to system performance evaluation," *IEEE Trans. Wireless Commun.*, vol. 7, no. 10, pp. 4019–4031, Oct. 2008.
- [16] K. Sulonen, P. Suvikunnas, L. Vuokko, J. Kivinen, and P. Vainikainen, "Comparison of MIMO antenna configurations in picocell and microcell environments," *IEEE J. Select. Areas Commun.*, vol. 21, no. 5, pp. 703–712, June 2003.
- [17] L. Dong, H. Choo, J. R. W. Heath, and H. Ling, "Simulation of MIMO channel capacity with antenna polarization diversity," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1869–1873, July 2005.
- [18] F. Boccardi, B. Clerckx, A. Ghosh, E. Hardouin, G. Jöngren, K. Kusume, E. Onggosanusi, and Y. Tang, "Multiple antenna techniques in LTE-advanced," *IEEE Commun. Mag.*, vol. 50, no. 2, pp. 114–121, Mar. 2012.
- [19] K. R. Rao and P. C. Yip, *The transform and data compression handbook*. Florida, USA: CRC Press LLC, 2001.
- [20] Y. Li, X. Ji, D. Liang, and Y. Li, "Dynamic beamforming for three-dimensional MIMO technique in LTE-advanced networks," to be appeared in *International Journal of Antennas and Propagation*, 2013.
- [21] M. Coldrey, "Modeling and capacity of polarized MIMO channels," in *Proc. IEEE Vehicular Technology Conference, 2008*, May 2008, pp. 440–444.
- [22] H. Asplund, J. Berg, F. Harrysson, J. Medbo, and M. Riback, "Propagation characteristics of polarized radio waves in cellular communications," in *Proc. IEEE Vehicular Technology Conference, 2007*, Sept. 2007, pp. 839–843.
- [23] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inform. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.
- [24] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Select. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.
- [25] D. Hwang, B. Clerckx, and G. Kim, "Regularized channel inversion with quantized feedback in downlink multiuser channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 5785–5789, Dec. 2009.
- [26] D. Tse and P. Viswanath, *Fundamental of Wireless Communication*, 1st ed. Cambridge: Cambridge Univ. Press, 2005.
- [27] W. Hachem, P. Loubaton, and J. Najim, "Deterministic equivalents for certain functionals of large random matrices," *Annals of Applied Probability*, vol. 17, no. 3, pp. 875–930, June 2007.
- [28] H. Joung, H. Jo, C. Mun, and J. Yook, "Capacity loss due to polarization-mismatch and space-correlation on MISO channel," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 2124–2136, Apr. 2014.
- [29] D. S. Shiu, *Wireless communication using dual antenna array*. Massachusetts, USA: Kluwer Academic Publishers, 2000.
- [30] D. S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multi-element antenna systems," in *Proc. IEEE International Conference on Universal Personal Communications, 1998*, Oct. 1998, pp. 429–433.

$$\begin{aligned}
IN_{gpk}^o(\chi) &= (\xi_{gp}^o(0))^2 \Upsilon_{ggpp}^o(0) (1 - \tau^2 (1 - (1 + m_{gp}^o(0))^2)) \\
&+ (1 + \chi (\xi_{gp}^o(0))^2 \Upsilon_{ggpp}^o(0) + \sum_{l \neq g} (1 + \chi) (\xi_{lp}^o(0))^2 \Upsilon_{glpp}^o(0)) (1 + m_{gp}^o(0))^2, \\
&= IN_{gpk}^o(0) + \chi ((\xi_{gp}^o(0))^2 \Upsilon_{ggpp}^o(0) + \sum_{l \neq g} (\xi_{lp}^o(0))^2 \Upsilon_{glpp}^o(0)) (1 + m_{gp}^o(0))^2.
\end{aligned} \tag{106}$$


---