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 antennas 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 1) the preprocessing based only on the spatial correlation and 2) 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

A part of this work was submitted to the 22nd European Signal Processing Conference 2014.
J. Park and B. Clerckx are with the Department of Electrical and Electronic Engineering, Imperial College London, South Kensington Campus London SW7 2AZ, United Kingdom (e-mail:{j.park, b.clerckx}@imperial.ac.uk).
This work is partially supported by Huawei Technologies Co., Ltd.
I. INTRODUCTION

By deploying a large number of antennas 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]–[8]. In [5], single-user beamforming in MISO with a large number of transmit antennas 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).

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 antennas) are concatenated, can be exploited to reduce the feedback overhead efficiently [8], [9]. 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 [9], [10]. 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 antennas. Another challenge of the massive MIMO system is the antenna space limitation. An increasing number of antennas 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 [11], [12]. The multi-polarized antennas can be one solution to alleviate the space constraint [13], [14]. The multi-polarized antenna systems have been investigated under various communication scenarios including the picocell/microcell [15], indoor/outdoor [16], and the line of sight (LOS)/non LOS (NLOS) [14] environments. However, despite the importance of polarized
In this paper, we first model the massive multi-user MIMO system, where BS is equipped with a large number of multi-polarized antennas and MSs are equipped with a single single-polarized antenna and present the dual structured precoding based on long-term/short-term CSIT. As done in the massive MIMO system with single-polarized antennas [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 [17]. However, 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 antennas 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].

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 the cross-polar discrimination (XPD) parameter. Here, XPD refers the long-term statistics of the antennas 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, but undergoes a performance degradation according to the XPD compared to that with spatial grouping only. 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. Based on the region, we propose to switch the precoding mode between two dual structured precoding strategies relying on 1) the spatial grouping only and 2) the subgrouping based on both the spatial correlation and polarization. Finally, motivated by 3D beamforming [8], [18], we extend the design to the 3D dual structured precoding in which the spatial correlation depends on both azimuth and elevation angles.

The rest of this paper is organized as follows. In Section [II] we introduce the massive MIMO system model with multiple multi-polarized antennas at BS and a (either vertically or horizontally) single
polarized antenna at 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 \((A)^T\), \((A)^H\), \([A]_i\), \([A]_j\), \(tr(A)\), and \(\det(A)\) denote the transpose, conjugate transpose, the \(i\)th row, the \(i\)th column, the trace, and the determinant of a matrix \(A\), respectively. In addition, \([A]_{i:j}\) denotes the submatrix from the \(i\)th column to the \(j\)th column of \(A\). The matrix norm \(\|A\|\) and the vector norm \(\|a\|\) denote the 2-norms of a matrix \(A\) and a vector \(a\), respectively. In addition, \(A \succeq 0\) means that a matrix \(A\) 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, \(I_M\), \(1_{M \times N}\), and \(0_{M \times N}\) 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 antennas and \(N\) active MSs, each with a single polarized antenna, 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 with either vertical or horizontal polarization[1]. Furthermore, since human activity

1Throughout the paper, we consider the dual-polarized antennas at the BS for ease of explanation, but our approach can be extended to the tri-polarized case without difficulty.
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 $y_g \in \mathbb{C}^{N_g}$ of the $g$th group with an assumption of flat-fading channel is given by

$$y_g = \begin{bmatrix} y^v_g \\ y^h_g \end{bmatrix} = H^H_g x + n_g,$$

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

$$x = \sum_{g=1}^{G} V_g d_g, \quad V_g \in \mathbb{C}^{M \times \bar{N}},$$

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

A. Channel model

By using the Karhunen-Loeve transform [17] and the polarized MIMO channel modeling with infinitesimally small antenna elements described in [9], [13], the downlink channel to the $g$th group, $H_g$, can be represented as

$$H_g = \left( \begin{array}{cc} 1 & r_{xp} \\ r_{xp} & 1 \end{array} \right) \otimes \left( \mathbf{U}_g \Lambda_g^{\frac{1}{2}} \right) \left( G_g \otimes (X \otimes \mathbf{1}_{r \times \frac{\bar{N}}{2}}) \right)$$

where $r_{xp}$ is the correlation coefficient between vertically and horizontally polarized antennas, $\Lambda_g$ 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, where the eigenvalues are assumed to be ordered in decreasing order of magnitude. Note that generally, $r_g \ll M$, and $U_g \in \mathbb{C}^{M \times r_g}$ has the associated eigenvectors of $\mathbf{R}_g^s$ as columns. The matrix $G_g$ is defined as

$$G_g = \begin{bmatrix} G_{vv}^g & G_{hv}^g \\ G_{vh}^g & G_{hh}^g \end{bmatrix},$$

$^2$For simplicity, it is assumed that the spatial covariance matrix is the same for both polarizations.
and the elements of $G^p q_g \in \mathbb{C}^{r_x \times \tau}$, $p, q \in \{ h, v \}$ are complex Gaussian distributed with zero mean and unit variance. The matrix $X$ describes the power imbalance between the orthogonal polarizations and is given as

$$X = \begin{bmatrix} 1 & \sqrt{\chi_0} \\ \sqrt{\chi_0} & 1 \end{bmatrix},$$

(5)

where the parameter $0 \leq \chi_0 \leq 1$ is the inverse of the XPD, where $1 \leq \text{XPD} \leq \infty$. Note that, based on the reported measurement in [19], [20], $r_{xp} \approx 0$. Accordingly, (3) can be rewritten as

$$H_g = \left( I_2 \otimes (U_g A_{g}^{\frac{1}{2}}) \right) \begin{bmatrix} G_{g}^{vv} & \sqrt{\chi_0} G_{g}^{bh} \\ \sqrt{\chi_0} G_{g}^{vh} & G_{g}^{hh} \end{bmatrix} = \begin{bmatrix} G_{g}^{vv} & \sqrt{\chi_0} G_{g}^{bh} \\ \sqrt{\chi_0} G_{g}^{vh} & G_{g}^{hh} \end{bmatrix},$$

(6)

and its covariance matrix is given as

$$R_g = \begin{bmatrix} (1 + \chi_0) R_g^s & 0 \\ 0 & (1 + \chi_0) R_g^s \end{bmatrix} = \begin{bmatrix} R_g^s & 0 \\ 0 & \chi_0 R_g^s \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} R_g^h = R_{gv} + R_{gh},$$

(7)

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

Note that the long-term parameters $R_g^s$ and $\chi_0$ are slowly-varying and assumed to be obtained accurately with a low feedback overhead. However, the short-term CSI parameter $G_g$ 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 $\hat{G}_g$ available at the transmitter is modeled as

$$\hat{G}_g = \sqrt{1 - \tau_g} G_g + \tau_g Z_g,$$

(8)

where the elements of $Z_g$ 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), $\hat{H}_g$ and $\hat{H}_g^{pp}$ can be defined as the imperfect CSI knowledge at the transmitter of $H_g$ and $H_g^{pp}$, respectively, by using $\hat{G}_g$. Throughout the paper, it is assumed that $\tau_1 = \ldots = \tau_G = \tau$, 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., $R_g^s$), the $k$th MS in the $g$th group quantizes $g_{gk} = [G_g]_k$ by using the random codebook [21] 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 \( g_{gk} \in \mathbb{C}^{2r_g \times 1} \) rather than \( h_{gk} \).

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]. That is, the precoding matrix for the \( g \)th group is given as

\[
V_g = B_g P_g,
\]

where \( B_g \in \mathbb{C}^{M \times \bar{B}} \) is the preprocessing matrix based on the long-term channel statistics with \( \bar{N} \leq \bar{B} \leq 2r_g \ll M \) and \( P_g \in \mathbb{C}^{\bar{B} \times \bar{N}} \) is the precoding matrix for the effective (instantaneous) channel \( H_g B_g \). 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

\[
y_g = H_g^H B_g P_g d_g + \sum_{l=1, l \neq g}^{G} H_g^H B_l P_l d_l + n_g.
\]

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 \( B_g \) based on the spatial correlation is designed as

\[
H_l^H B_g \approx 0, \quad \text{for } l \neq g.
\]

Then, \( P_g \) in (10) can be computed based on the decoupled system model \( y_g \approx H_g^H B_g P_g d_g + n_g \) where the inter-group interferences have been eliminated.

To obtain \( B_g \) satisfying the condition (11), the BD can be utilized. That is, due to the block diagonal structure in (7), we first define

\[
U_{-g} = [U_1^a, \ldots, U_{g-1}^a, U_{g+1}^a, \ldots, U_G^a] \in \mathbb{C}^{\frac{M}{r_g} \times \sum_{i \neq g} r_i^a},
\]
where $U^a_g = [U_{g,1:r_g^a}]_{1:r_g^a}$ and $r_g^a(\leq r_g)$ is the number of dominant eigenvalues of $R^s_g$, a design parameter. That is, 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 $\frac{M_2}{2} - \sum_{l \neq g} r_l^a$.

The matrix $U_{-g}$ in (12) then has a singular value decomposition (SVD) as

$$U_{-g} = [E_{-g}^{(1)}, E_{-g}^{(0)}] \begin{bmatrix} \Lambda_{-g}^{(1)} & 0 \\ 0 & \Lambda_{-g}^{(0)} \end{bmatrix} \mathbf{V}_g^H,$$

where $E_{-g}^{(1)}$ (respectively, $E_{-g}^{(0)}$) is the left singular vectors associated with the $\sum_{l \neq g} r_l^a$ dominant (respectively, $M_2 - \sum_{l \neq g} r_l^a$ non-dominant) singular values $\Lambda_{-g}^{(1)}$ (respectively, $\Lambda_{-g}^{(0)}$). Then, because $(E_{-g}^{(0)})^H U_{-g} = 0$, by defining $\hat{H}_g = (I_2 \otimes E_{-g}^{(0)})^H H_g$, $\hat{H}_g$ is orthogonal to the dominant eigen-space spanned by other groups’ channel. Note that the covariance matrix of $\hat{H}_g$ is then given by

$$\hat{R}_g = (I_2 \otimes E_{-g}^{(0)})^H R_g (I_2 \otimes E_{-g}^{(0)}),$$

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

$$\hat{R}_g^s = F_g \hat{A}_g F_g^H,$$

where $F_g$ is the eigenvectors of $\hat{R}_g^s$. Then by letting $F_{g,1:r_g^a}^H = [F_g]_{1:r_g^a}$, the preprocessing matrix can be given as

$$B_g = I_2 \otimes B_g^s, \quad B_g^s = E_{-g}^{(0)} F_{g,1:r_g^a}^H.$$

Accordingly, through the preprocessing matrix $B_g$, we can transform the transmit signal for the $g$th group into the $\hat{B}$ dimensional dominant eigen-space that is orthogonal to the subspace spanned by other groups’ channel. Note that, from (9) and (12), $\hat{B}$ and $r_g^a$ should be chosen properly to satisfy the conditions of $\hat{N} \leq \hat{B} \leq 2(M_2 - \sum_{l \neq g} r_l^a)$ and $\hat{B} \leq 2r_g$. Without loss of generality, we assume that $r_1^a = \ldots = r_G^a = r$ 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 $H_g = B_g^H H_g$, the corresponding covariance matrix is given by

$$\hat{R}_g = B_g^H R_g B_g = B_g^H (R_{gv} + R_{gh}) B_g = \hat{R}_{gv} + \hat{R}_{gh} = (1 + \chi_o) \begin{bmatrix} (B_g^s)^H R_g^s B_g^s & 0 \\ 0 & (B_g^s)^H R_g^s B_g^s \end{bmatrix}.\ (17)$$

Furthermore, due to the preprocessing, the interferences from other groups in (10) are almost nulled out. Accordingly, the precoding matrix $P_g$ 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], [22] with imperfect CSIT can be computed as

$$ P_g = \xi_g \hat{K}_g \hat{H}_g, $$

where

$$ \hat{K}_g = \left( \hat{H}_g \hat{H}_g^H + B_0 I_B \right)^{-1}, \quad \hat{H}_g = B_g^H \hat{H}_g. $$

Here, $\hat{H}_g$ is the effective channel estimate that is available at the BS and $\alpha$ is a regularization parameter. Throughout the paper, it is set as $\alpha = \bar{N}/B_T$, which is equivalent with the MMSE linear filter [23]. The normalization factor $\xi_g$ is then given as

$$ \xi_g^2 = \frac{\bar{N} P}{\mathcal{N}_T tr(\hat{H}_g^H \hat{K}_g \hat{H}_g)} = \frac{\bar{N}}{tr(\hat{H}_g^H \hat{K}_g \hat{H}_g)}, $$

where the second equality is due to the fact that $B_g^H B_g = I_B$ from (16). Denoting $\hat{h}_{gk} = [\hat{H}_g]_k$ 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 by

$$ \gamma_{gpk}^{BD} = \frac{\mathcal{P}_N \xi_g^2 |h_{gk}^H B_{gk} \hat{K}_g \hat{h}_{gk}|^2}{\mathcal{N}_T \sum_{j \neq k} \xi_j^2 |h_{gk}^H B_{gj} \hat{K}_g \hat{h}_{gj}|^2 + \mathcal{P}_N \sum_{l \neq g} \sum_j \xi_j^2 |h_{gk}^H B_{lk} \hat{K}_l \hat{h}_{lj}|^2 + 1}, $$

and accordingly, the sum rate is given by

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

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 $\chi_0$ becomes small (i.e., the antennas 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 antennas 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

$$y_g^p = \mathbf{H}_g^H \mathbf{B}_{gp} \mathbf{P}_{gp} d_g^p + \sum_{q \in \{h,v\}, q \neq p} \mathbf{H}_g^H \mathbf{B}_{gq} \mathbf{P}_{gp} d_g^q + \sum_{l=1, l \neq g}^G \sum_{q \in \{h,v\}} \mathbf{H}_g^H \mathbf{B}_{lq} \mathbf{P}_{lq} d_l^q + n_g^p,$$  

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

$$\mathbf{B}_{gv} = \begin{bmatrix} \mathbf{B}_g^v \\ 0 \end{bmatrix}, \quad \mathbf{B}_{gh} = \begin{bmatrix} 0 \\ \mathbf{B}_g^h \end{bmatrix},$$

where $\mathbf{B}_g^s$ is given in (16). Note that, when $\chi_0 \approx 0$, we can easily find that

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

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 antennas 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

$$\mathbf{P}_{gp} = \xi_{gp} \hat{\mathbf{K}}_{gp} \hat{\mathbf{H}}_{gp}, \tag{26}$$

where $\hat{\mathbf{K}}_{gp} = \left( \hat{\mathbf{H}}_{gp}^H \hat{\mathbf{H}}_{gp} + \frac{\bar{\epsilon}^2}{\bar{N}} \mathbf{I}_2 \right)^{-1}$ and $\hat{\mathbf{H}}_{gp} = \mathbf{B}_{gp}^H \hat{\mathbf{H}}_{gp} = (\mathbf{B}_g^s)^H \hat{\mathbf{H}}_{gp}^p$, the effective channel estimate that is available at the BS. The normalization factor $\xi_{gp}$ is then given as

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

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

$$\gamma_{BDS, gpk} = \frac{P \xi_{gp}^2 |\mathbf{h}_{gpk}^H \mathbf{B}_{gp} \hat{\mathbf{K}}_{gp} \hat{\mathbf{B}}_{gp}^H \hat{\mathbf{h}}_{gpk}|^2}{IN_{gpk}}, \tag{28}$$
where
\[
IN_{pgk} = \frac{P}{N} \sum_{j \neq k} \xi_{gp}^2 |h_{gp}^H B_{gp} \hat{K}_{gp} B_{gp}^H \hat{h}_{gp}|^2 + \frac{P}{N} \sum_{q \neq p} \sum_{j} \xi_{gq}^2 |h_{gq}^H B_{gq} \hat{K}_{gq} B_{gq}^H \hat{h}_{gq}|^2 \\
+ \frac{P}{N} \sum_{l \neq g} \sum_{q} \sum_{j} \xi_{lq}^2 |h_{lp}^H B_{lp} \hat{K}_{lp} B_{lp}^H \hat{h}_{lp}|^2 + 1 \tag{29}
\]

and \( h_{gpk} = [H_{gp}]_k \) and \( \hat{h}_{gpk} = [\hat{H}_{gp}]_k \), respectively. Accordingly, the sum rate is given by
\[
R_{BDS} = \sum_{g=1}^{G} \sum_{p \in \{v,h\}} \sum_{k=1}^{N/2} \log_2 (1 + \gamma_{gpk}^{BDS}), \tag{30}
\]
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
\[
\gamma_{gpk}^{BDS} = \gamma_{gpk}^{BD}. \tag{31}
\]

IV. ASYMPTOTIC PERFORMANCE ANALYSIS FOR DUAL PRECODING METHODS

In [7], when the number of transmit antennas \( M \) is large, the asymptotic SINR of the regularized ZF precoding has been analyzed in spatially correlated MISO broadcasting 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, based on analytical results using random matrix theory [7], [24], [25], 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. Then, we analyze the performance as a function of the XPD parameter \( \chi_0 \), 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].

\textbf{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 \( i = 1, ..., N \). In addition, let \( \mathbf{S}, \mathbf{Q} \in \mathbb{C}^{M \times M} \) be Hermitian nonnegative definite. Assume \( \lim \sup_{M \to \infty} \sup_{1 \leq i \leq N} \| \mathbf{R}_i \| < \infty \) and \( \mathbf{Q} \) has uniformly bounded spectrum norm. Then, for \( z < 0 \),
\[
\frac{1}{M} tr(\mathbf{Q}(\mathbf{H}\mathbf{H}^H + \mathbf{S} - z \mathbf{I}_M)^{-1}) - \frac{1}{M} tr(\mathbf{Q}(z)^{(M \to \infty}) \to 0, \tag{32}
\]
Then, by using Theorem 1, the asymptotic SINR for the dual precoding with BD can be derived.

Here,
\[
\varepsilon_i(z) = \frac{1}{M} \text{tr} \left( \mathbf{R}_i \left( \frac{1}{M} \sum_{j=1}^{N} \frac{\mathbf{R}_j}{1 + e_j(z)} + \mathbf{S} - z \mathbf{I}_M \right)^{-1} \right),
\]
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 \(M, N, B\) goes to infinity and \(N^2 \to B\) is fixed, the SINR, \(\gamma_{gpk}^{BD}\), in (31) asymptotically converges as

\[
\gamma_{gpk}^{BD} = \gamma_{gpk}^{BD, o} \xrightarrow{M \to \infty} 0,
\]

where \(\gamma_{gpk}^{BD, o}\) is the asymptotic SINR, given by

\[
\gamma_{gpk}^{BD} = \frac{P}{\bar{\gamma}_{ggp}(1 - \tau^2)(m_{gpp}^o)^2} \left( (e_1^o)^2 \bar{\gamma}_{ggp}(1 - \tau^2(1 + m_{gpp}^o)^2) + (1 + \sum_{l \neq g} (e_l^o)^2 \bar{\gamma}_{gpl}^o)(1 + m_{gpl}^o)^2 \right)^{-1}
\]

with \((e_1^o)^2 = \frac{P}{C_{\text{g,\text{v}}}^g}\). Here, \(m_{gpp}^o, \bar{\gamma}_{gpl}^o, \) and \(\Psi_g\) are the unique solutions of

\[
m_{gpp}^o = \frac{1}{B} \text{tr} (\bar{\mathbf{R}}_{gpp} \mathbf{T}_g), \quad \mathbf{T}_g = \left( \frac{\bar{\mathbf{N}}}{2B} \sum_{q \in \{h,v\}} \frac{\bar{\mathbf{R}}_{gqq}}{1 + m_{gqq}^o} + \alpha \mathbf{I}_B \right)^{-1}, \quad \Psi_g = \frac{1}{2B \bar{\mathbf{G}}} \sum_{q \in \{h,v\}} \frac{m_{gqq}^o}{(1 + m_{gqq}^o)^2}, \quad \bar{\gamma}_{gpp}^o = \frac{\bar{\mathbf{N}}/2 - 1}{B} \frac{m_{gpp}^o}{\bar{\mathbf{N}}(1 + m_{gpp}^o)^2} + \frac{\bar{\mathbf{N}}/2}{B} \frac{m_{gpp}^o}{2B \bar{\mathbf{G}}(1 + m_{gpp}^o)^2}.
\]

In addition, \(\mathbf{m}_g = [m_{gvp}, m_{gvh}]^T\), and \(\mathbf{m}_{gpp} = [m_{gppvp}, m_{gppvh}]^T\) given by

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

where

\[
\mathbf{J} = \frac{\bar{\mathbf{N}}}{2B} \left[ \begin{array}{cc} \frac{\text{tr}(\bar{\mathbf{R}}_{gvp}, \mathbf{T}_g \mathbf{R}_{gvp}, \mathbf{T}_g)}{B(1 + m_{gpp}^o)^2} & \frac{\text{tr}(\bar{\mathbf{R}}_{gvp}, \mathbf{T}_g \mathbf{R}_{gvp}, \mathbf{T}_g)}{B(1 + m_{gpp}^o)^2} \\
\frac{\text{tr}(\bar{\mathbf{R}}_{gvp}, \mathbf{T}_g \mathbf{R}_{gvp}, \mathbf{T}_g)}{B(1 + m_{gpp}^o)^2} & \frac{\text{tr}(\bar{\mathbf{R}}_{gvp}, \mathbf{T}_g \mathbf{R}_{gvp}, \mathbf{T}_g)}{B(1 + m_{gpp}^o)^2} \end{array} \right], \quad \mathbf{v}_g = \frac{1}{B} \left[ \begin{array}{c} \text{tr}(\bar{\mathbf{R}}_{gvp}^2 \mathbf{T}_g) \\
\text{tr}(\bar{\mathbf{R}}_{gvp} \mathbf{T}_g) \end{array} \right], \quad \mathbf{v}_{gpp} = \frac{1}{B} \left[ \begin{array}{c} \text{tr}(\bar{\mathbf{R}}_{gpp} \mathbf{T}_g \mathbf{R}_{gpp} \mathbf{T}_g) \\
\text{tr}(\bar{\mathbf{R}}_{gpp} \mathbf{T}_g \mathbf{R}_{gpp} \mathbf{T}_g) \end{array} \right].
\]

with \(\bar{\mathbf{R}}_{gpp}\) defined in (17). In addition, \(\mathbf{m}_{gpl} = [m_{gplvp}, m_{gplvh}]^T\) given by

\[
\mathbf{m}_{gpl} = (\mathbf{I}_2 - \mathbf{J})^{-1} \mathbf{v}_{gpl},
\]

where

\[
\mathbf{J} = \frac{\bar{\mathbf{N}}}{2B} \left[ \begin{array}{cc} \frac{\text{tr}(\bar{\mathbf{R}}_{gvp}, \mathbf{T}_g \mathbf{R}_{gvp}, \mathbf{T}_g)}{B(1 + m_{gpp}^o)^2} & \frac{\text{tr}(\bar{\mathbf{R}}_{gvp}, \mathbf{T}_g \mathbf{R}_{gvp}, \mathbf{T}_g)}{B(1 + m_{gpp}^o)^2} \\
\frac{\text{tr}(\bar{\mathbf{R}}_{gvp}, \mathbf{T}_g \mathbf{R}_{gvp}, \mathbf{T}_g)}{B(1 + m_{gpp}^o)^2} & \frac{\text{tr}(\bar{\mathbf{R}}_{gvp}, \mathbf{T}_g \mathbf{R}_{gvp}, \mathbf{T}_g)}{B(1 + m_{gpp}^o)^2} \end{array} \right], \quad \mathbf{v}_{gpl} = \frac{1}{B} \left[ \begin{array}{c} \text{tr}(\bar{\mathbf{R}}_{gpl}^2 \mathbf{T}_g) \\
\text{tr}(\bar{\mathbf{R}}_{gpl} \mathbf{T}_g) \end{array} \right], \quad \mathbf{v}_{gpp} = \frac{1}{B} \left[ \begin{array}{c} \text{tr}(\bar{\mathbf{R}}_{gpp} \mathbf{T}_g \mathbf{R}_{gpp} \mathbf{T}_g) \\
\text{tr}(\bar{\mathbf{R}}_{gpp} \mathbf{T}_g \mathbf{R}_{gpp} \mathbf{T}_g) \end{array} \right].
\]
Proof: See Appendix A.

Corollary 1: When the spatial covariance matrix is the same for both polarization, i.e., infinitesimally small dual-polarized antennas are co-located (see footnote 2), the asymptotic SINR $\gamma_{gk}^{BD,o}$ in (35) can be written as in a simpler form as

$$\gamma_{gk}^{BD,o} = \frac{P}{N} (\xi_{g}^{o})^{2} (1 - \tau^{2}) (m_{g}^{o})^{2} \left( \xi_{g}^{o} \frac{2}{g} (1 - \tau^{2} (1 - (1 + m_{g}^{o})^{2}))) + (1 + \sum_{l \neq g} (\xi_{l}^{o}) \frac{2}{g} (1 + m_{g}^{o})^{2}) \right),$$

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

$$m_{g}^{o} = \frac{1}{B} tr(\bar{R}_{g}^{o} T_{g}), \quad T_{g} = \left( \frac{N}{B} \frac{R_{g}}{1 + m_{g}^{o}} + \alpha I_{B} \right)^{-1},$$

$$\Psi_{g}^{o} = \frac{1}{B} \frac{P}{G} \frac{m_{g}^{o}}{(1 + m_{g}^{o})^{2}}, \quad m_{g}^{o} = \frac{1}{B} \frac{1}{1 - \frac{N}{B} tr(R_{g}^{o} T_{g} R_{g}^{o} R_{g}^{o} T_{g})},$$

$$\Upsilon_{gl}^{o} = \frac{N - 1}{P} \frac{m_{gl}^{o}}{N (1 + m_{g}^{o})^{2}}, \quad \Upsilon_{gl}^{o} = \frac{P}{N} \frac{m_{gl}^{o}}{B (1 + m_{g}^{o})^{2}},$$

$$m_{gl}^{o} = \frac{1}{B} \frac{1}{1 - \frac{N}{B} tr(R_{g}^{o} T_{g} R_{g}^{o} T_{g})}, \quad m_{gl}^{o} = \frac{1}{B} \frac{1}{1 - \frac{N}{B} tr(R_{g}^{o} T_{g} R_{g}^{o} T_{g})},$$

where $R_{g} = \frac{1}{2} R_{g}$ and $R_{g}^{o} = \frac{1}{2} R_{g}$. Here, $R_{g}$ and $R_{g}^{o}$ are defined in (7) and (17).

Proof: From (7) and (17), we can see that $m_{gh}^{o} = m_{gv}^{o} = \frac{1}{B} tr(\frac{1}{2} R_{g} T_{g})$ in (37). Accordingly, by letting $m_{g}^{o} = \frac{1}{B} tr(\frac{1}{2} R_{g} T_{g})$, $T_{g}$ in (37) can be rewritten as that in (44). Furthermore, because, in (40),

$$tr(R_{g} T_{g}^{2}) = tr(R_{g} T_{g}^{2}) = \frac{1}{2} tr(R_{g} T_{g}^{2}),$$

$$tr(R_{g} T_{g}^{2}) + tr(R_{g} T_{g}^{2}) = \frac{1}{4} tr(R_{g} T_{g}^{2}).$$

$m_{gl}^{o} = m_{gh}^{o} = m_{gv}^{o}$ as in (45). Similarly, we can prove that the parameters $m_{glp}^{o}$ are given as (47). By substituting $m_{g}^{o}$, $m_{g}^{o}$, $m_{gl}^{o}$ into $\Psi_{g}$, $\Upsilon_{gvp}^{o}$, and $\Upsilon_{glp}^{o}$ of (37) and (38), we can prove that $\Psi_{g}^{o}$, $\Upsilon_{gvp}^{o}$, 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 $\gamma_{gk}^{BD,o} = \gamma_{g}^{BD,o}$ for $k = 1, ..., N$, the asymptotic sum rate can be approximated as

$$R_{BD}^{o} \approx \sum_{g=1}^{G} \sum_{p \in \{v,h\}} \sum_{k=1}^{N/2} \log_{2}(1 + \gamma_{gk}^{BD,o}) = \sum_{g=1}^{G} \sum_{k=1}^{N/2} \log_{2}(1 + \gamma_{g}^{BD,o}).$$

Remark 2: We note that, when $\tau = 0$, $\gamma_{gk}^{BD,o}$ in Corollary 1 is analogous to the asymptotic SINR of the dual precoding with BD derived in [8] under the perfect CSIT and unipolarized system. That
is, when the infinitesimally small dual-polarized antennas are co-located, the asymptotic SINRs of the MSs in the same group are the same which is independent with their antenna deployment, i.e., vertical or horizontal polarization. In addition, the effective covariance matrix is given by $R_g' = \frac{1}{2}R_g$. 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 $I_2 \otimes R_g^s$ and the effective transmit power of BS is reduced from $P$ to $\frac{1+\chi}{2}P$.

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

$$\gamma_{gpk}^{BDS} = \gamma_{gpk}^{BDS,o} + M \rightarrow \infty \rightarrow 0,$$

(50)

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

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

(51)

where

$$IN_{gpk}^o = (\xi_{g}^o)^2 \Upsilon_{gpp}^o (1-\tau^2(1-(1+m_{gp}^o)^2)) + (1 + \sum_{q \neq p} (\xi_{g}^o)^2 \Upsilon_{gqq}^o + \sum_{l \neq g} \sum_{q} (\xi_{l}^o)^2 \Upsilon_{glq}^o)(1+m_{gp}^o)^2,$$

(52)

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

$$m_{gp}^o = \frac{2}{B} tr(\tilde{R}_{gp}T_{gp}), \quad T_{gp} = \left( \frac{N}{B} \tilde{R}_{gp} + \alpha I_B / 2 \right)^{-1}, \quad \Psi_{g}^o = \frac{1}{B G (1+m_{gp}^o)^2},$$

(53)

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

(54)

with

$$m_{gp}^\prime = \frac{2}{B} tr(\tilde{R}_{gp}T_{gp}^2), \quad m_{gpp}^\prime = \frac{2}{B} tr(\tilde{R}_{gp}T_{gp}\tilde{R}_{gp}T_{gp})$$

(55)

$$m_{glq}^\prime = \frac{2}{B} tr(\tilde{R}_{lq}T_{lq}B_{lq}\tilde{R}_{gp}B_{lq}T_{gp}),$$

(56)

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

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

February 18, 2014 DRAFT
From Theorem 3, the asymptotic SINR is independent of MS index \( k \). Furthermore, because \( X \) in (5) is symmetric, \( \gamma_{BDS,o}^{gpk} = \gamma_{BDS,o}^{ghk} \). Accordingly, by letting \( \gamma_{BDS,o}^{gpk} = \gamma_{BDS,o}^{ghk} \equiv \gamma_{BDS,o}^{g} \), the asymptotic sum rate can be approximated as

\[
R_{BDS}^{o} \approx \sum_{g=1}^{G} \sum_{p \in \{v,h\}} \frac{\tilde{N}}{2} \log_{2}(1 + \gamma_{BDS,o}^{gpk}) = \sum_{g=1}^{G} \tilde{N} \log_{2}(1 + \gamma_{BDS,o}^{g}).
\]  

(57)

C. Asymptotic performance analysis as a function of the XPD parameter \( \chi_{0} \)

In this section, we investigate the effect of the XPD parameter \( \chi_{0} \) 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 \( \gamma_{BDS,o}^{g} \) of \( \chi \), then \( \gamma_{BDS,o}^{g} \approx \gamma_{BDS,o}^{g}(0) \).

**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 \( \gamma_{BDS,o}^{g} \) of \( \chi \) given by

\[
\gamma_{BDS,o}^{g}(\chi) \approx \frac{\gamma_{BDS,o}^{g}(0)}{1 + c_{0}\chi}.
\]

(58)

where

\[
c_{0} = \frac{E_{g,p}}{\left(\xi_{gp}(0)\right)^{2}Y_{gpp}(0)} + \frac{1}{\left(1+m_{gp}^{o}(0)\right)^{2}} \left(\tau^{2}\right) \left(1 + m_{gp}^{o}(0)\right)^{2} - 1 + 1 + 1
\]

(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 \( P_{gp} \) 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 \( \gamma_{BDS,o}^{g}(0) = \gamma_{BDS,o}^{g}(0) \).
V. A NEW DUAL STRUCTURED PRECODING/FEEDBACK

A. A new dual structured precoding/feedback

Even though the performance of the dual precoding with BDS is degraded according to the XPD parameter, 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. When random vector quantization (RVQ) with $N_B$ bits is utilized [21], the quantization error for the short-term CSIT in the dual precoding with BD (i.e., the columns of $G_g$ in (4)) is upper bounded as

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

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

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

Because the bound is tight for a large $N_B$ [21], 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

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

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

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

where

$$A_0 = \frac{B_0^2}{X} (\xi_{gp}^o(0))^2 (m_{gp}^o(0))^2, \quad B_0 = (\xi_{gp}^o(0))^2 \gamma_{gpp}^o(0), \quad D_0 = (1 + m_{gp}^o(0))^2 - 1, \quad E_0 = \sum_{l \neq g} (\xi_{lp}^o(0))^2 \gamma_{gpp}^o(0).$$

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 $\gamma_{gpk,BD}^o(\chi) \geq \gamma_{gpk,BD}^o(\chi)$, by letting

$$\tau_{BD}^2 \triangleq \tau^2 = 2^{-\frac{N_B}{2r-1}}(\approx \tau_{BDS}), \quad (65)$$

we have

$$\frac{1 - \tau^4}{(B_0(1 + D_0 \tau^4) + (1 + E_0)(D_0 + 1))(1 + c_0 \chi)} \geq \frac{(1 - \tau^2)}{B_0(1 + D_0 \tau^2) + (1 + E_0)(D_0 + 1)}.$$
After a simple calculation, we have
\[
\chi \leq \frac{(B_0 D_0 + B_0 + (1 + E_0)(D_0 + 1)) \tau^2}{c_0 (B_0 D_0 \tau^4 + B_0 + (1 + E_0)(D_0 + 1))}.
\] (66)

From (59) and the fact that \(E_0 \ll 1\) due to the BD, \(c_0 \approx \frac{B_0(D_0+1)}{B_0 D_0 \tau^4 + B_0 + D_0 + 1}\) and we have
\[
\chi \leq \left(1 + \frac{D_0}{B_0(D_0 + 1)}\right) \tau^2 = \left(1 + \frac{(1 + m_{gp}(0))^2 - 1}{(\xi_{gp}(0))^2 \Upsilon_{ggpp}(0)(1 + m_{gp}(0))^2}\right) \tau^2,
\] (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 antennas. In addition, from (62), when \(\chi \to 0\), the dual precoding with BDS always exhibits better performance than that with BD. Note that for high SNR (i.e., \(m_{gp}(0) \gg 1\)), the expectation term in (62) is approximated as \(E_{g,p} \left(\frac{1}{(\xi_{gp}(0))^2 \Upsilon_{ggpp}(0)}\right)\), 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.

Fig. 2. Block diagram of a new dual precoding/feedback scheme.
Motivated by 3D beamforming [8], [18], 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 $h_{gl}$ be the $2M_AM_E \times 1$ vectorized channel of the $k$th MS in the $g$th group of the $l$th region, it can be written as

$$h_{gl} = ((I_2 \otimes U_{glA}) \otimes U_{lE})(A_{glA}^{\frac{1}{2}} \otimes A_{lE}^{\frac{1}{2}})g_{gl},$$

where $A_{glA}$ and $A_{lE}$ are the $r_{glA} \times r_{glA}$ and $r_{lE} \times r_{lE}$ diagonal matrices with non-zero eigenvalues of the spatial correlation matrices $R_{glA}$ and $R_{lE}$ over the azimuth and elevation directions, respectively, and $U_{glA}$ and $U_{lE}$ are the matrices of the associated eigenvectors. Here, $g_{gl} = [g_{gl1v}, \chi_0 g_{gl1h}]^T$ (resp. $g_{gl} = [\chi_0 g_{gl1v}, g_{gl1h}]^T$) for vertically (resp. horizontally) polarized MSs and $g_{gl,lp}$ 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

$$x = \sum_{l=1}^{L} (\sum_{g=1}^{G} V_{gl}(d_{gl}) \otimes q_l),$$

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

$$y_{gl} = \sum_{g'=1}^{G} G_{gl}^H (A_{glA}^{\frac{1}{2}} \otimes A_{lE}^{\frac{1}{2}}) ((I_2 \otimes U_{glA})^H V_{g'l}g_{l}^t) \otimes (U_{lE}^H q_l) + n_{gl}$$

$$= \sum_{g'=1}^{G} \sqrt{\lambda_l} G_{gl}^H A_{glA}^{\frac{1}{2}} (I_2 \otimes U_{glA})^H V_{g'l}d_{gl} + n_{gl},$$

where $\lambda_l = q_l^H R_{lE}^s q_l$. 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 $\lambda_l$ for the $l$th elevation region.
Fig. 3. Block diagram of a 3D dual polarized multi-user downlink system.

VI. SIMULATION RESULTS

Throughout the simulations, we consider the one-ring model for the spatially correlated channel [26], [27]. The correlation between the channel coefficients of antennas $1 \leq m, n \leq M$ is given by

$$[R^s_{g}]_{m,n} = \frac{1}{2\Delta_g} \int_{-\Delta_g}^{\Delta_g} e^{-j\pi \lambda_0^{-1} \Omega(\alpha + \theta_g)(r_m - r_n)} d\alpha,$$

(71)

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, $r_m = [x_m, y_m]^T$ is the position vector of the $m$th antenna, 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 antennas in multi-user massive MIMO system: To see the suitability of the dual polarized antennas in multi-user massive MIMO system, we have compared the performance of the linear array antennas with dual polarized antennas and single polarized antennas, 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 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 antennas 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, $\frac{N}{2}$ vertically polarized
and $\frac{\bar{N}}{2}$ horizontally polarized MSs coexist. 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 $R_g^s$, $g = 1, ..., G$. In addition, $\Delta_1 = ... = \Delta_4 = \Delta = \frac{8}{180}\pi$ and $\theta_g = -\frac{\pi}{4} + \frac{\pi}{6}(g - 1)$ for $g = 1, ..., 4$. Fig. 4 shows the sum-rate curves for the arrays with dual polarized antennas, $d_s = \frac{\lambda_0}{2}$ and with single polarized antennas, $d_s = \{\frac{\lambda_0}{2}, \frac{\lambda_0}{4}\}$, 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 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 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 $R_g^s$, $g = 1, ..., G$. In addition, $R_g^s$ is generated by (71) with $d_s = \frac{\lambda_0}{2}$, $\Delta = \frac{\pi}{12}$, and $\theta_g = -\frac{\pi}{4} + \frac{\pi}{6}(g - 1)$ for $g = 1, ..., 4$. 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.

\footnote{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 large enough MSs in a cell but we do not consider the user selection (multi-user) diversity, which is out of scope of this paper.}
Fig. 5. Sum rates of the dual precoding with BD and BDS under the perfect CSIT.

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., $\tau^2_{BD} = \tau^2$ and $\tau^2_{BDS} = \tau$ 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 derivation of analytic results is based on Theorem 1 with $z = -\alpha$ and the approximation error bound is proportional to $\frac{1}{\alpha} = \frac{BP}{N}$, 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 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.

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

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 $\chi = \{0.1, 0.2\}$. 
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. In Fig. 10, the performances of the three schemes are provided as a function of $\chi$ when $SNR = 25dB$. Here, we assumed that $\tau^2$ is uniformly distributed on $[0, 1]$. Again, for small $\chi$ the dual precoding with BDS exhibits better performance than that with BD, but the performance of the dual precoding with BDS is degraded as $\chi$ increases. The proposed dual precoding in Section VII outperforms the two other schemes regardless of $\chi$. 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. In Fig. 11 we have evaluated the 3D dual structured precoding in Section VII-B when $10 \times 50$ uniform planar array is deployed at BS with a height of $60m$ and there are 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. 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 $P_{loss,l} = \frac{1}{1 + \left(\frac{d_{gl}}{60}\right)^2}$. 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). We can find that in 3D spatial domain, the proposed dual structured scheme also outperforms two other dual precoding schemes.

**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 performance 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 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, when the feedback bits are not enough to describe the short-term CSIT accurately, 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

Fig. 10. Sum rates of dual precodings with BD and BDS and a proposed dual precoding over $\chi$.

Fig. 11. Sum rates of the 3D dual precodings over $\chi$. 
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 only difference is that we have the intergroup interferences due to the preprocessing and the co-polarized MSs in the same subgroup have the same spatial correlation as in (7). We first consider the normalization factor \( \xi_g^2 \) in (20). By letting \( \Psi = \frac{P}{N} tr(\hat{H}_g^H \hat{K}_g \hat{K}_g^T \hat{H}_g) \), \( \xi_g^2 \) can be rewritten as

\[
\xi_g^2 = \frac{P}{G} \frac{1}{\Psi}.
\]

(72)

By substituting \( \hat{K}_g \) in (19), due to the matrix inversion lemma [28], \( \Psi \) can be written as

\[
\Psi = \frac{P}{N} \sum_{k=1}^{\hat{N}} \hat{h}_{gk}^H B g \left( \hat{H}_g \hat{H}_g^H + B \alpha I_B \right)^{-2} B_g^H \hat{h}_{gk} = \frac{P}{N B} \sum_{k=1}^{\hat{N}} \left( 1 + \hat{h}_{gk}^H B g \left( A_g^{-k} + \alpha I_B \right) \right)^{-1} \frac{B_g^H \hat{h}_{gk}}{\left( B_g^H \hat{h}_{gk} \right)^2},
\]

(73)

where \( A_g^{-k} = \frac{1}{B} \hat{H}_g^{-k} (\hat{H}_g^{-k})^H \) and \( \hat{H}_g^{-k} = [\hat{h}_{g1}, \ldots, \hat{h}_{gk-1}, \hat{h}_{gk+1}, \ldots, \hat{h}_{gN}] \). From Lemma 4 in [7], we have

\[
\Psi - \frac{P}{N B} \sum_{p \in \{v, h\}} \sum_{k=1}^{\hat{N}/2} \frac{1}{B} tr(B_g^H \hat{R}_{gp} B_g (A_g^{-k} + \alpha I_B)^{-2}) - \frac{M \rightarrow \infty}{\alpha} 0.
\]

(74)

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

\[
\Psi - \frac{\hat{N} P}{2 N B} \sum_{p \in \{v, h\}} \frac{1}{B} tr(B_g^H \hat{R}_{gp} B_g (A_g + \alpha I_B)^{-2}) - \frac{M \rightarrow \infty}{\alpha} 0,
\]

(75)

where \( A_g = \frac{1}{B} \hat{H}_g \hat{H}_g^H \). Then, we define \( m_{gp}(z) \triangleq \frac{1}{B} tr(B_g^H \hat{R}_{gp} B_g (A_g - z I_B)^{-1}) \), which is analogous to (32) in Theorem 1 by setting \( S = 0 \) and \( Q = B_g^H \hat{R}_{gp} B_g (= \hat{R}_{gp} \) from (17)). The trace in the denominator of (75) is then equal to \( m_{gp}(-\alpha) \) and, from Theorem 1,

\[
\frac{1}{B} tr(B_g^H \hat{R}_{gp} B_g (A_g + \alpha I_B)^{-1}) - \frac{1}{B} tr(\hat{R}_{gp} T_g) \xrightarrow{M \rightarrow \infty} 0,
\]

(76)
where $T_g$ is given by (37) from (33) and (34). Furthermore, the trace in the numerator of (75) is the derivative of $m_{gp}(z)$ at $z = -\alpha$, i.e., $m'(\alpha)$. Therefore,
\[
\frac{1}{B} tr(B^H T_g \bar{R}_g B + \alpha I) - \frac{1}{B} tr(R_g T_g') M \to \infty 0,
\]
(77)
where
\[
T_g' = T_g \left( \frac{\bar{N}}{2B} \sum_{p \in \{v,h\}} \frac{R_{gp} m'_{gp}}{(1 + m_{gp}(\alpha))^2} + I_B \right) + T_g.
\]
(78)
Then, by putting $T_g'$ in (78) into (79), $m'$ can be obtained as in (39) and $\Psi$ is converged as
\[
\Psi - \frac{P\bar{N}}{2NB} \sum_{p \in \{v,h\}} \frac{m'_{gp}}{(1 + m_{gp})^2} M \to \infty 0.
\]
(80)

For the signal power component $\sum_{p \in \{v,h\}} |h_{gk}^H \hat{B}_g \hat{K}_k \hat{h}_{gk}|^2$ and the intra-group interference component $\sum_{p \in \{v,h\}} |h_{gk}^H \hat{B}_g \hat{K}_k \hat{h}_{gk}|^2$ in (21), because $G_g$ and $Z_g$ in (8) are independent, by taking a similar approach described in (73) and (74) (See also Appendix II.B and C in [7]), we can have the following relations:
\[
|h_{gk}^H \hat{B}_g \hat{K}_k \hat{h}_{gk}|^2 \to (1 - \tau^2) (1 + m_{gp})^2 M \to \infty 0,
\]
(81)
\[
\sum_{j \neq k} |h_{gj}^H \hat{B}_g \hat{K}_j \hat{h}_{gj}|^2 = \frac{\gamma_{gpp}^k}{1 + m_{gp}^2 (1 - (1 + m_{gp}^2))} M \to \infty 0,
\]
(82)
where
\[
\gamma_{gpp}^k = \frac{1}{B} \sum_{p \in \{v,h\}} g_{gj}^H R_{gp}^{1/2} B_g \hat{K}_g \hat{R}_g B_g \hat{K}_g \hat{R}_g \bar{R}_{gp} g_{gj}
\]
\[
+ \frac{1}{B} \sum_{j=1}^{N/2} g_{gj}^H R_{gp}^{1/2} B_g \hat{K}_j \hat{R}_g B_g \hat{K}_g \hat{R}_g \bar{R}_{gp} g_{gj}.
\]
(83)
Note that by taking a similar approach described in (73) and (74),
\[
\gamma_{gpp}^k = -\frac{N}{2} - \frac{1}{B} tr(\bar{R}_{gp} (A_g + \alpha I_B)^{-1} \bar{R}_{gp} (A_g + \alpha I_B)^{-1})
\]
\[
- \frac{1}{2B} tr(\bar{R}_{gp} (A_g + \alpha I_B)^{-1} \bar{R}_{gp} (A_g + \alpha I_B)^{-1}) M \to \infty 0.
\]
(84)
By defining $m_{gpp}(z) = \frac{1}{B} tr(\bar{R}_{gq} (A_g + \alpha I_B - z \bar{R}_{gp})^{-1})$, the trace in the numerator of (84) is the derivative of $m_{gpp}(z)$ at $z = 0$. Furthermore, by using Theorem 1,
\[
m_{gpp}(z) - \frac{1}{B} tr(\bar{R}_{gq} T_g(z)) M \to \infty 0,
\]
(85)
where
\[
T_{gp}(z) = \left( \frac{\hat{N}}{2B} \sum_{q \in \{v, h\}} \frac{\hat{R}_{gq}}{1 + m_{gqp}(z)} + \alpha I_B - z \hat{R}_{gp} \right)^{-1}.
\] (86)

Accordingly, we have
\[
T'_{gp}(z) = T_{gp}(z) \left( \frac{\hat{N}}{2B} \sum_{q \in \{v, h\}} \frac{\hat{R}_{gq} m'_{gqp}(z)}{(1 + m_{gqp}(z))^2} + \hat{R}_{gp} \right) T_{gp}(z).
\] (87)

Note that \(T_{gp}(0) = T_g\) in (37) and from (85), \(m_{gqp}(0) \rightarrow 0\). Together with \(m'_{gqp}(z) - \frac{1}{B} tr (\hat{R}_{gq} T'_{gp}(z)) \rightarrow 0\), \(m'_{gqp}\) can be obtained as (39). Accordingly,
\[
\Upsilon^k_{gqp} - \Upsilon^o_{gqp} \rightarrow 0,
\] (88)

where \(\Upsilon^o_{gqp}\) is defined in (38).

Now let us consider the inter-group interference component \(\Upsilon^k_{glp} \triangleq \sum_{j \mid h \in H_g} |h^H B_j \hat{K}_l B_l^H \hat{h}_{lq}|^2\) in the denominator of (21) for \(l \neq g\). Then, \(\Upsilon^k_{glp}\) can be rewritten as
\[
\Upsilon^k_{glp} = \frac{\hat{N}}{2B} \sum_{q \in \{v, h\}} \frac{1}{2} \sum_{j=1}^{N/2} \hat{g}_{lj}^H R_{lq}^{1/2} B_j \hat{K}_l B_l^H R_{gq} B_j \hat{K}_l B_l^H R_{lq}^{1/2} \hat{g}_{lj}.
\] (89)

Note that (89) has a similar form with (83). Therefore, similarly done in (84), we have
\[
\Upsilon^k_{glp} - \frac{\hat{N}}{2B} \sum_{q \in \{v, h\}} \frac{1}{2} tr (\hat{R}_{lq} (A_l + \alpha I_B)^{-1} B_l^H R_{gq} B_l (A_l + \alpha I_B)^{-1}) \rightarrow 0.
\] (90)

Furthermore, by following the steps similarly done in (85)-(87), we can obtain
\[
\Upsilon^k_{gl} - \Upsilon^o_{gl} \rightarrow 0,
\] (91)

where \(\Upsilon^o_{gl}\) is defined in (38). By substituting (80), (81), (82), (88), and (91) into (21), we can have (35).

**APPENDIX B**

**Proof of Proposition 1**

Let \(R_g(\chi) = (1 + \chi) \begin{bmatrix} R^s_g & 0 \\ 0 & R^s_g \end{bmatrix}\) from (7). Then, we have \(R_g(\chi) = (1 + \chi) R_g(0)\) (respectively, \(\hat{R}_g(\chi) = (1 + \chi) \hat{R}_g(0)\)) and, from Corollary 1, \(\gamma_{BD, o}(\chi)\) can be evaluated by using (43) with \(\frac{1}{2} \hat{R}_g(\chi)\).

From (44), we define
\[
T_g(\chi) = \left( \frac{\hat{N}}{2B} \frac{\hat{R}_g(\chi)}{1 + m^o_g(\chi) + \alpha I_B} \right)^{-1},
\] (92)

\[
m^o_g(\chi) = \frac{1}{2B} tr (\hat{R}_g(\chi) T_g(\chi)).
\] (93)
First, let us consider the high SNR regime (i.e., small $\alpha$). Because

$$tr(R(\beta R + \alpha I)^{-1}) = \frac{1}{\beta} tr(R(R + \alpha/\beta I)^{-1}) \approx \frac{1}{\beta} tr(R (R + \alpha I)^{-1}),$$

(94)

for any Hermitian nonnegative definite matrix $R$ and small $\alpha$, by substituting (92) into (93), we have

$$m_0^g(\chi) = \frac{1}{2B} tr((1 + \chi)Rg(0) \left( \frac{\bar{N}}{2B} \left( \frac{Rg(0)}{1 + m_0^g(\chi)} + \alpha I_B \right) \right)^{-1}),$$

$$\approx \frac{1}{2B} tr(Rg(0) \left( \frac{\bar{N}}{2B} \left( \frac{Rg(0)}{1 + m_0^g(\chi)} + \alpha I_B \right) \right)^{-1}),$$

(95)

which implies that

$$m_0^g(\chi) = m_0^g(0).$$

(96)

Furthermore, by substituting (96) into (92),

$$T_g(\chi) = \frac{1}{1 + \chi} \left( \frac{\bar{N}}{2B} \frac{Rg(0)}{1 + m_0^g(\chi)} + \frac{\alpha}{1 + \chi} I_B \right)^{-1}.$$  

(97)

By using (94), (96) and (97), we can also derive

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

(98)

which implies that

$$\Psi_0^g(\chi) \approx \frac{1}{1 + \chi} \Psi_0^g(0), \quad \Upsilon_0^g(\chi) \approx \Upsilon_0^g(0), \quad \Upsilon_0^g(\chi) \approx \Upsilon_0^g(0), \quad \xi_0^g(\chi) \approx (1 + \chi) \xi_0^g(0).$$

(99)

Then, based on (96) and (99) together with (43), $\gamma_{BD,o}^{gpk}(\chi) \approx \gamma_{BD,o}^{gpk}(0)$.

For low SNR regime (i.e., large $\alpha$), in (92), $T_g(\chi) \approx \frac{1}{\alpha} I_B$ and accordingly, $m_0^g(\chi) \approx (1 + \chi)m_0^g(0) \ll 1$. Furthermore,

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

(100)

and

$$\Psi_0^g(\chi) \approx (1 + \chi) \Psi_0^g(0), \quad \Upsilon_0^g(\chi) \approx (1 + \chi)^2 \Upsilon_0^g(0),$$

$$\Upsilon_0^g(\chi) \approx (1 + \chi)^2 \Upsilon_0^g(0), \quad \xi_0^g(\chi) \approx \frac{1}{1 + \chi} \xi_0^g(0).$$

(101)

Accordingly, in the low SNR regime, we can have that $\gamma_{BD,o}^{gpk}(\chi) \approx \gamma_{BD,o}^{gpk}(0)$.  

February 18, 2014 DRAFT
APPENDIX C

PROOF OF PROPOSITION 2

From (7) and (24), we have \( \hat{R}_{g\nu}(\chi) = \hat{R}_{g\nu}(\chi) = (B_g^*)^H \hat{R}_g^s B_g \). That is, \( \hat{R}_{g\nu}(\chi) \) and \( \hat{R}_{g\nu}(\chi) \) are independent of \( \chi \) and expressed as \( \hat{R}_{g\nu}(\chi) = \hat{R}_{g\nu}(0) \). Accordingly, from (53) and (55),

\[
m_{o, gp}(\chi) = m_{o, gp}(0), \quad T_{gp}(\chi) = T_{gp}(0),
\]

\[
m'_{o, gp}(\chi) = m'_{o, gp}(0), \quad m'_{g_{gp}}(\chi) = m'_{g_{gp}}(0).
\]

Because \( B_{tq}^H \hat{R}_{g\nu}(\chi) B_{tq} = \begin{cases} B_{tq}^H \hat{R}_{g\nu}(0) B_{tq} & \text{for } q = p, \\ B_{tq}^H \hat{R}_{g\nu}(0) B_{tq} & \text{for } q \neq p \end{cases} \), from (53),

\[
m'_{g_{lpq}}(\chi) = \begin{cases} m'_{g_{lpq}}(0) & \text{for } q = p, \\ \chi m'_{g_{lpq}}(0) & \text{for } q \neq p \end{cases}.
\]

Therefore, from (51),

\[
\gamma_{gpk}^{BDS,o}(\chi) = \frac{P(N_g)^2(1 - \tau^2)(m_{o, gp}^2)}{IN_{o, gpk}(\chi)},
\]

where

\[
IN_{o, gpk}(\chi) = \frac{(\xi_{gpk}(0))^2 T_{g_{gp}}(0) (1 - \tau^2 (1 - (1 + m_{o, gp}(0))^2))}{\chi (\xi_{gpk}(0))^2 T_{g_{gp}}(0) + \sum_{l \neq q} (1 + \chi) (\xi_{gpk}(0))^2 T_{g_{lpq}}(0) (1 + m_{o, gp}(0))^2},
\]

\[
= IN_{gpk}(0) + \chi((\xi_{gpk}(0))^2 T_{g_{gp}}(0) + \sum_{l \neq q} (\xi_{gpk}(0))^2 T_{g_{lpq}}(0) (1 + m_{o, gp}(0))^2).
\]

Because of the preprocessing with BD, \( T_{g_{gp}}(0) \gg T_{g_{lpq}}(0) \) for \( g \neq l \) and

\[
IN_{gpk}(\chi) \approx IN_{gpk}(0)(1 + c_{0,gpk}),
\]

where \( c_{0,gpk} = \frac{(\xi_{gpk}(0))^2 T_{g_{gp}}(0)(1 + m_{o, gp}(0))^2}{(\xi_{gpk}(0))^2 T_{g_{gp}}(0)(\tau^2 (1 + m_{o, gp}(0))^2 - 1) + 1 + (1 + m_{o, gp}(0))^2}. \) 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. Ramprashad, “Achieving “massive MIMO” spectral efficiency with a not-so-large number of antennas,” *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3226–3239, Sept. 2012.
[4] J. Hoydis, S. T. Brink, and M. Debbah, “Massive MIMO: How many antennas do we need?” in *Proc. Annual Allerton Conference on Communication, Control, and Computing, 2011*, Sept. 2011, pp. 545–550.
[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. Stock, “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] 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.

[10] Way Forward on Aperiodic PUSCH CQI modes in Release 10, 3GPP TSG RAN WG1 No. 62 Std. R1-105 058, Aug. 2010.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] K. R. Rao and P. C. Yip, *The transform and data compression handbook*. Florida, USA: CRC Press LLC, 2001.

[18] 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.

[19] M. Coldrey, “Modeling and capacity of polarized MIMO channels,” in *Proc. IEEE Vehicular Technology Conference*, 2008, May 2008, pp. 440–444.

[20] 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.

[21] N. Jindal, “MIMO broadcast channels with finite-rate feedback,” *IEEE Trans. Inform. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.

[22] 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.

[23] D. Tse and P. Viswanath, *Fundamental of Wireless Communication*, 1st ed. Cambridge: Cambridge Univ. Press, 2005.

[24] 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.

[25] 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.

[26] D. S. Shiu, *Wireless communication using dual antenna array*. Massachusetts, USA: Kluwer Academic Publishers, 2000.

[27] 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.

[28] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd ed. Baltimore: Johns Hopkins Univ. Press, 1996.