# Downlink Precoding for Massive MIMO Systems Exploiting Virtual Channel Model Sparsity

Thomas Ketseoglou, *Senior Member, IEEE*, and Ender Ayanoglu, *Fellow, IEEE*

***Abstract*— In this paper, the problem of designing a forward link linear precoder for massive multiple-input multiple-output (MIMO) systems in conjunction with quadrature amplitude modulation (QAM) is addressed. A challenge in such system design is to consider finite alphabet inputs, especially with larger constellation sizes, such as $M \geq 16$. First, we employ a novel and efficient methodology that allows for an orthogonal, sparse representation of multiple users and groups in a fashion similar to joint spatial division and multiplexing (JSDM), thus offering an extension of JSDM to finite alphabet data symbols. We term the new approach JSDM for finite alphabets (JSDM-FA). JSDM-FA employs virtual channel model beams in order to explore the orthogonality between different groups. Then, we present a natural generalization of JSDM-FA to the frequency-selective case with orthogonal frequency-division multiplexing (OFDM) and also enhance OFDM with combined frequency and spatial division and multiplexing. This configuration offers high flexibility in Massive MIMO systems, as it is capable of offering separate decoding of each user data within a group or increase the spectral efficiency of spatially overlapping groups without sacrificing the overall cell spectral efficiency. The proposed methodology is next applied jointly with the complexity-reducing Per-Group Processing within groups technique, on a per user group basis, in conjunction with QAM modulation and in simulations, for constellation size up to $M = 64$. We show by numerical results that the precoders developed offer significantly better performance than the configuration with no precoder or the plain beamformer and with $M \geq 16$.**

***Index Terms*— Massive MIMO, precoding, finite alphabet inputs, joint spatial division and multiplexing, per-group precoding, uniform linear arrays, uniform planar arrays.**

## I. Introduction

MASSIVE MIMO employs a very large number of antennas and enables very high spectral efficiency [1]–[3]. For Massive MIMO to be capable of offering its full benefits, accurate and instantaneous channel state information is required at the base station (BS). Within Massive MIMO research, the problem of designing an optimal linear precoder toward maximizing the mutual information between the input and output on the downlink in conjunction with a finite input alphabet modulation and multiple antennas per user has not been considered in the literature, due to its complexity. There are techniques proposed for downlink linear precoding in a multi-user MIMO scenario, e.g., Joint Spatial Division and Multiplexing (JSDM) [4]–[6], but their implementation has been challenging so far. In addition, there has been a lack of publications on how to realistically integrate OFDM in Massive MIMO with success and without sacrificing the spectral efficiency of the system. On the other hand, the problem of finite-alphabet input MIMO linear precoding has been extensively studied in the literature. Globally optimal linear precoding techniques were presented [7], [8] for scenarios employing channel state information available at the transmitter (CSIT)[1] with finite-alphabet inputs, capable of achieving mutual information rates much higher than the previously presented Mercury Waterfilling (MWF) [9] techniques by introducing input symbol correlation through a unitary input transformation matrix in conjunction with channel weight adjustment (power allocation). In addition, more recently, [10] has presented an iterative algorithm for precoder optimization for sum rate maximization of Multiple Access Channels (MAC) with Kronecker MIMO channels. Furthermore, more recent work has shown that when only Statistical Channel State Information (SCSI)[2] is available at the transmitter, in asymptotic conditions when the number of transmitting and receiving antennas grows large, but with a constant transmitting to receiving antenna number ratio, one can design the optimal precoder by looking at an equivalent constant channel and its corresponding adjustments as per the pertinent theory [13], and applying a modified expression for the corresponding ergodic mutual information evaluation over all channel realizations. This development allows for a precoder optimization under SCSI in a much easier way [13]. Finally, [14] and [15] present for the first time results for mutual information maximizing linear precoding with large size MIMO configurations and QAM constellations. Such systems are particularly difficult to analyze and design when the inputs are from a finite alphabet, especially with QAM constellation sizes, $M \geq 16$.

In this paper, we present near-optimal linear precoding techniques for Massive MIMO, suitable for QAM with con-

Manuscript received June 21, 2017; revised September 21, 2017 and November 30, 2017; accepted December 27, 2017. Date of publication January 8, 2018; date of current version May 15, 2018. This work was partially supported by NSF grant 1547155. The associate editor coordinating the review of this paper and approving it for publication was C. Yuen. *(Corresponding author: Thomas Ketseoglou.)*

T. Ketseoglou is with the Electrical and Computer Engineering Department, California State Polytechnic University, Pomona, CA 91768 USA (e-mail: tketseoglou@cpp.edu).

E. Ayanoglu is with the Center for Pervasive Communications and Computing, Department of Electrical Engineering and Computer Science, University of California at Irvine, Irvine, CA 92697 USA (e-mail: ayanoglu@uci.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCOMM.2018.2790402

[1]Under CSIT the transmitter has perfect knowledge of the MIMO channel realization at each transmission.

[2]SCSI pertains to the case in which the transmitter has knowledge of only the MIMO channel correlation matrices [11], [12] and the thermal noise variance.

stellation size $M \geq 16$ and CSIT. Two types of antenna arrays are considered for the Base Station (BS), Uniform Linear Arrays (ULA) and Uniform Planar Arrays (UPA). In the UPA case, we consider arrays deployed either over the $x$, $y$ direction or the $z$, $x$ one. We show that by projecting the per user antenna uplink channels on the DFT based angular domain, called virtual channel model [16][3] (VCM) herein, a sparse representation is possible for the channels [17]–[19]. Then, by dividing spatially "distant" users into separate spatial sectors, we show that the spatial virtual channel representations between these users become approximately orthogonal. We then show that the concept of JSDM [4] can be easily applied in the sparse virtual channel model domain, and that linear precoding on the downlink using Per-Group Precoding within groups[4] (PGP-WG) in conjunction with the Gauss-Hermite approximation in MIMO [15], [20] attains high gains. Then, by generalizing the presented approach to the frequency-selective (FS) channel case and applying OFDM, we show that much more flexibility and gains are available by the techniques presented. We further show that when OFDM is integrated in JSDM-FA, in conjunction with Combined Frequency and Spatial Division and Multiplexing (CFSDM), the system can easily decode different users' data within a group at the cost of slightly smaller achieved utilization, or increase the utilization of groups with significant spatial overlapping without sacrificing the overall cell utilization. In all examples presented, we show high gains are achievable by the proposed downlink precoding approach. More specifically, the present paper makes the following contributions in Massive MIMO:

1) It provides an analytical framework that allows to reliably identify spatially separated user groups that are approximately orthogonal and thus to require independent per-group precoding beams from the base station for finite alphabet data inputs and for a realistic and general channel model that applies to both ULA and UPA.
2) It allows the users to employ multiple receiving antennas with ease.
3) It generalizes the approach to include OFDM under frequency-selective channel conditions in a flexible way.
4) It introduces CFSDM, a flexible way to separately decode intra-group data with a small loss in spectral efficiency or increase group throughput.
5) It shows very significant gains in conjunction with PGP-WG [15] and QAM modulation with significant reduction in the transmitter and receiver complexity.

The paper is organized as follows: Section II presents the system model and problem statement. Then, in Section III, we present a novel virtual channel approach which allows for efficient downlink precoding in a JSDM fashion for ULA and UPA narrowband and frequency selective channels. In addition, in Section III the CFSDM conecpt is introduced. In Section IV, we present numerical results for optimal downlink precoding on the system proposed that implements the Gauss-Hermite approximation in the block coordinate gradient ascent method in conjunction with the complexity reducing PGP methodology [15]. Finally, our conclusions are presented in Section V.

*Notation:* We use small bold letters for vectors, capital bold letters for matrices. $\mathbf{A}^T$, $\mathbf{A}^H$, and $\mathbf{A}^*$, denote the transpose, Hermitian conjugate, and complex conjugate of matrix $\mathbf{A}$, respectively. $\mathbf{S}^T$ denotes a selection matrix, i.e., of size $k \times n$ with $k < n$ that consists of rows equal to different unit row vectors $\mathbf{e}_i$ where the row vector element $i$ is equal to 1 in the $i$th position and is equal to 0 in all other positions. $\mathbf{F}_N$ denotes the DFT matrix of order $N$. The Kronecker delta function is denoted as $\delta[n]$. $\mathbf{A} \otimes \mathbf{B}$ represents the Kronecker product of matrices $\mathbf{A}$ and $\mathbf{B}$. Many different notations for different channels are employed within the paper. We use $\mathbf{h}_{u,g,k,n}$ for the uplink channel of user $k$'s antenna $n$ in group $g$. $\mathbf{H}_g$ is the uplink channel of group $g$, while $\tilde{\mathbf{H}}_g$ is its projection to the VCM basis. The overall virtual channel representation of $\mathbf{H}_g$ is denoted by $\mathbf{H}_{g,v}$. $\mathbf{H}_{d,g}$ represents the overall downlink channel for group $g$, i.e., due to time division duplex reciprocity $\mathbf{H}_{d,g} = \mathbf{H}_g^H$. For a ULA or a UPA, $\mathbf{H}_{u,g,k,n}$ represents the uplink channel of user $k$'s antenna $n$ in group $g$, then $\tilde{\mathbf{H}}_{u,g,k,n}$ is its projection to the two DFT matrices (VCM bases), $\tilde{\mathbf{h}}_{u,g,k,n}$ is its corresponding vectorized form. Finally, $\mathbf{H}_{u,g,k,v}^{(q)}$ is in the frequency selective fading case the uplink virtual channel of user $k$ in group $g$ at sub-carrier $q$.

## II. System Model and Problem Statement

Consider the downlink precoding equation on a narrowband (flat-fading) Massive MIMO system with a single cell and JSDM [4]

$$\mathbf{y}_d = \mathbf{H}_u^H \mathbf{P} \mathbf{x}_d + \mathbf{n}_d, \tag{1}$$

where $\mathbf{y}_d$ is the downlink received vector of size $\sum_{g=1}^{G} N_{d,g} \times 1$, $\mathbf{x}_d$ is the $N_u \times 1$ vector of transmitted symbols drawn independently from a QAM constellation, where the downlink channel matrix $\mathbf{H}_d = \mathbf{H}_u^H$, where $\mathbf{H}_u = [\mathbf{H}_1, \cdots, \mathbf{H}_G]$ is the $N_u \times K_{eff}$ uplink channel matrix from all $K$ users, employing $N_u$ receiving antennas at the base, with $K_{eff} = \sum_g N_{d,g}$, where $N_{d,g}$ is the total number of antennas of all users in group $g$. Users have been divided into $G$ groups with $K_g$ users in group $g$ ($1 \leq g \leq G$), with user $k$ of group $g$ denoted as $k^{(g)}$ and employing $N_{d,k^{(g)}}$ transmitting antennas, with ($\sum_{g=1}^{G} K_g = K$), $\mathbf{H}_g = [\mathbf{H}_{g^{(1)}} \cdots \mathbf{H}_{g^{(K_g)}}]$ being group $g$'s uplink channel matrix of size $N_u \times N_{d,g}$, with $N_{d,g}$ comprising the total number of antennas in the group, i.e., $N_{d,g} = \sum_{k^{(g)}} N_{d,k^{(g)}}$, where $\mathbf{n}_d$ represents the independent, identically distributed (i.i.d.) complex circularly symmetric Gaussian noise of variance per component $\sigma_u^2 = \frac{1}{\mathrm{SNR}_{s,d}}$, where $\mathrm{SNR}_{s,d}$ is the channel symbol signal-to-noise ratio (SNR) on the downlink. The uplink symbol vector of size $\mathbf{x}_u$ of size $\sum_g^G N_{d,g} \times 1$ has i.i.d. components drawn from a QAM constellation of order $M$. We assume that Time

[3]We need to stress that this is not our propagation-related channel model, but it is a virtual angular orthonormal basis that we use to express the actual channel in an efficient way toward deriving a JSDM type of decomposition. Thus, the VCM model is used as a basis to which we project the actual channel and thus this representation is reversible.

[4]In this paper, we use PGP-BG to refer to the PGP approach between different groups, as proposed in [4], and PGP-WG to refer to PGP within the same group, as proposed in [20], respectively.

Division Duplexing (TDD) is employed in the system, to be able to exploit the reciprocity between the uplink and downlink channels. The optimal CSIT precoder $\mathbf{P}$ needs to satisfy

$$\begin{aligned} &\underset{\mathbf{P}}{\text{maximize}}\ I(\mathbf{x}_d; \mathbf{y}_d) \\ &\text{subject to } \mathrm{tr}(\mathbf{P}\mathbf{P}^H) = N_u, \end{aligned} \tag{2}$$

where the constraint is due to keeping the total power emitted from the $N_u$ antennas constant after precoding.

The problem in (2) results in exponential complexity at both transmitter and receiver, and it becomes especially difficult for QAM with constellation size $M \geq 16$ or large MIMO configurations. There are two major difficulties in (2): a) There are $N_u$ input symbols in (2) where $N_u$ is very large, thus making the design of the precoder and its optimization practically impossible, and b) The decoding operation at the receiver needs to be performed by jointly employing all elements of $\mathbf{y}_d$ simultaneously, another impossible demand due to the users being distributed over the entire cell. In order to circumvent these difficulties, the ground-breaking JSDM concept was proposed in [4]. JSDM divides users into approximately orthogonal groups assuming a Gaussian channel [4], based on approximately equal channel covariance matrices in each group. Furthermore, JSDM employs Gaussian data inputs. Because of the orthogonality between different groups, $I(\mathbf{x}_d; \mathbf{y}_d) = \sum_{g=1}^{G} I(\mathbf{x}_g; \mathbf{y}_g)$, where $\mathbf{x}_g$, $\mathbf{y}_g$ represent the data symbols, and received data of group $g$, respectively (see (11), (12) below and [4]). Thus, the problem in (2) becomes equivalent to the one that maximizes the sum of the group information rates, i.e., the total sum-rate of the system. Under this premise, the downlink precoding problem is divided into two parts: a) A pre-beamforming matrix that comprises the square root of each group's channel covariance matrix, and b) A CSIT Multi-User MIMO (MU-MIMO) optimal precoder. In addition, [4] shows that when the number of base antennas grows asymptotically to infinity, the (optimal) pre-beamforming matrix approaches a DFT matrix. JSDM helps reduce the complexity inherent in the downlink precoding design tremendously due to the two-stage design optimal approach, plus the introduction of PGP-BG technique. Thus, it represents a major breakthrough advancement toward downlink precoding optimization. However, a major impediment to JSDM in practice has been the lack of a simple way that identifies the different groups of users with ease. Furthermore, [4] has employed Gaussian input symbols, an assumption that can lead to discrepancies as far as the precoder performance is concerned, especially in high SNR [7], [21]. Finally, JSDM deals with the overall downlink precoding on a per group basis, i.e., no methodology for individual user equipment (UE) to separately decode their own information (without intra-group joint decoding) is presented so far. In this paper, a methodology that employs the virtual channel model decomposition, based on the DFT channel angular domain is employed in order to facilitate the group selection problem in JSDM and then the methodology of PGP-WG technique [15] is employed in order to allow for the design of an optimal overall precoder on a per group basis. In addition, we present different ways to distribute the group received information to the multiple UEs. Finally, we also extend the JSDM-FA concept to the case of frequency-selective channels which leads to a very natural introduction of OFDM to JSDM-FA. Our system model allows for multiple receiving antennas per user and multiple data symbols per user with ease, as well as multiple data streams per user, with separate channel coding per stream, in a fashion similar to the methodology in, e.g., [7].

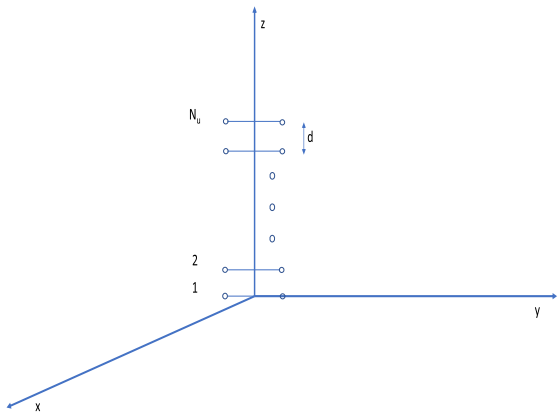



Fig. 1. A ULA deployed across the $z$ axis, together with the projection of a tentative transmission point on the $x$, $y$ plane.

## III. Narrowband and Wideband System Analysis

### A. The Narrowband System Description Under the Virtual Channel Model Representation

*1) ULA at the Base With Flat Fading:* We begin with a ULA deployed at the BS along the $z$ direction as depicted in Fig. 1 and for flat fading, i.e., $B < B_{COH}$, where $B$, $B_{COH}$ are the RF signal bandwidth and the coherence bandwidth of the channel, respectively. Each user group on the uplink transmits from the same "cluster" of elevation angles $\theta_g \in [\bar{\theta}_g - \Delta\theta, \bar{\theta}_g + \Delta\theta]$, with $\bar{\theta}_g$ being the mean of $\theta_g$, distributed uniformly in the support interval, thus each user's $k^{(g)}$ of group $g$, ( $1 \leq k^{(g)} \leq K_g$ and $1 \leq g \leq G$) transmitting antenna $n$ channel, $\mathbf{h}_{u,g,k,n} = \frac{1}{\sqrt{L}} \sum_{l=1}^{L} \beta_{lgkn} \mathbf{a}(\theta_{lgkn})$, where $\mathbf{a}(\theta_{lgkn}) = [1, \exp(-j2\pi D \cos(\theta_{lgkn})), \cdots, \exp(-j2\pi D (N_u - 1)\cos(\theta_{lgkn}))]^T$ is the array response vector, where each $\theta_{lgkn}$ is independently selected and uniformly distributed in its group's angular support $[\bar{\theta}_g - \Delta\theta, \bar{\theta}_g + \Delta\theta]$, with $D = d/\lambda$ representing the normalized distance of successive array elements, $\lambda$ being the wavelength, $\theta_{lgkn}$ is the elevation (arrival) angle of the $l$ path of group $g$ $k$ user's $n$ receiving antenna, and the path gains $\beta_{lgkn}$ are independent complex Gaussian random variables with zero mean and variance 1. This channel model is similar to the one in [22]. Note that the channel model adopted here is not Gaussian if $L$ is relatively small [22]. The VCM representation, presented in [16], is formed by projecting the original channel $\mathbf{H}_u$ to the $N_u$ dimensional space formed by the $N_u \times N_u$ DFT matrix $\mathbf{F}_{N_u}$, with row $k$, column $l$ ($1 \leq k,\ l \leq N_u$) element equal to $\exp(-j\frac{2\pi}{N_u}(k-1)(l-1))$. For Massive MIMO systems, i.e., when $N_u \gg 1$, the following Lemma 1 and 2 as well as Theorem 1 are true.

*Lemma 1: By employing VCM for a ULA at the BS and under flat fading, the number of non-zero components of the VCM representation is small, i.e., the number of non-zero or significant elements in the channels of each group g VCM representation, $|\mathcal{S}_g|$, satisfy $|\mathcal{S}_g| \ll N_u$. Thus, in the VCM domain, a sparse overall group channel representation results. Note that Fifth Generation cellular wireless (5G) millimeter wavelength channels are known to present sparsity, e.g., [6], [23], [24], however by using the VCM DFT basis vectors, we can capitalize on the VCM sparsity-induced orthogonality to design an efficient downlink precoder for a quite general family of channels as presented below.*

*Proof:* By projecting each group channel $\mathbf{H}_g$ on the DFT virtual channel space [16], we get

$$\tilde{\mathbf{H}}_g = \mathbf{F}_{N_u}^H \mathbf{H}_g, \tag{3}$$

where $\mathbf{F}_{N_u}$ is the DFT matrix of order $N_u$. Since each group attains the same angular behavior, over all users and antennas in the group, only a few, consecutive elements of $\tilde{\mathbf{H}}_g$ will be significant [25]. This comes as a result of the fact that significant angular components need to be in the main lobe of the response vector, i.e., the condition

$$|\cos(\theta_{lgkn}) - \frac{p}{DN_u}| \leq \frac{1}{DN_u}, \tag{4}$$

with $D = \frac{d}{\lambda}$, needs to be satisfied for angular component in the VCM $p$ ($1 \leq p \leq N_u$) to be significant, i.e., with power $> 1$. From (4), we can easily see that the corresponding condition over the significant components becomes

$$DN_u \cos(\theta_{lgkn}) - 1 \leq p \leq DN_u \cos(\theta_{lgkn}) + 1, \tag{5}$$

i.e., there are 3 significant non-zero components in the VCM representation for each channel's path. Since each path contains a different angle, due to the ULA model presented above, this number will be increased, but will be upper-bounded by $DN_u|\cos(\bar{\theta}_g + \Delta\theta) - \cos(\bar{\theta}_g - \Delta\theta)| + 3 = 3 + 2DN_u|\sin(\bar{\theta}_g)\sin(\Delta\theta)| \approx 3 + 2DN_u|\sin(\bar{\theta}_g)|(\Delta\theta)$, where $\Delta\theta$ is in radians. For a typical scenario, $N_u = 100$, $D = 1/2$, $\bar{\theta}_g = 30°$, and $\Delta\theta = 4° = 0.0698$ radian, then the maximum number of non-zero (significant) paths is upper-bounded by 7. □

*Lemma 2: Within the premise of the previous lemma, if $\cos(\bar{\theta}_g - \Delta\theta) < \cos(\bar{\theta}_{g'} + \Delta\theta) - \frac{2}{DN_u}$, where g and $g'$ represent two different groups ($g \neq g'$) and with $\bar{\theta}_g > \bar{\theta}_{g'}$ and $0 \leq \bar{\theta}_g, \bar{\theta}_{g'} \leq 90°$, then the support sets for each group are mutually exclusive, thus their corresponding virtual channel model beams (VCMB) become orthogonal. A similar relationship holds in the remaining quadrants.*

*Proof:* When $\bar{\theta}_g > \bar{\theta}_{g'}$ and $0 \leq \bar{\theta}_g, \bar{\theta}_{g'} \leq 90°$, since the $\cos(\cdot)$ function is decreasing in this quadrant, we can easily see that the two support sets for the two groups, $\mathcal{S}_g$, $\mathcal{S}_{g'}$, will be disjoint. This comes from the fact that the assumed condition is equivalent to

$$\cos(\bar{\theta}_g - \Delta\theta) + \frac{1}{DN_u} < \cos(\bar{\theta}_{g'} + \Delta\theta) - \frac{1}{DN_u}, \tag{6}$$

which means that the two support sets are not overlapping, by virtue of (4). We can develop similar conditions for all remaining quadrants. Thus, by assuming adequate spatial separation between groups, we can ensure that the support sets of each group in the virtual channel representation do not overlap. Then, due to the non-overlapping of the support sets, there exists orthogonality between the components of each group in the virtual channel model, as it is next shown. □

*Theorem 1: By employing VCM for a ULA at the BS and under flat fading, provided user groups are sufficiently geographically apart, as per the previous lemma, the channel model of the entire downlink channel can be expressed in a fashion that is fully suitable for JSDM type of processing where different groups become orthogonal and the downlink precoder is designed on a per group basis employing the virtual channel model representation alone. In the resulting JSDM type of decomposition, the corresponding group channel matrices are the virtual channel matrices of the group VCM projections and the group pre-beamforming matrices are the group's non-zero (significant) VCM beamforming directions.*

*Proof:* By employing a size $|\mathcal{S}_g| \times N_u$ selection matrix[5]

$$\mathbf{H}_{g,\upsilon} = \mathbf{S}_g^T \tilde{\mathbf{H}}_g = \mathbf{S}_g^T \mathbf{F}_{N_u}^H \mathbf{H}_g, \tag{7}$$

where the group $g$ virtual channel matrix is a reduced size, $r_g \times N_{d,g}$, matrix, with $r_g = |\mathcal{S}_g|$ being the number of significant angular components in group $g$, due to the sparsity available in the angular domain. We can then write for the uplink group $g$ channel matrix $\mathbf{H}_g$,

$$\mathbf{H}_g = \mathbf{F}_{N_u}\mathbf{S}_g\mathbf{S}_g^T\mathbf{F}_{N_u}^H\mathbf{H}_g = \mathbf{F}_{N_u,\mathcal{S}_g}\mathbf{H}_{g,\upsilon}, \tag{8}$$

where $\mathbf{F}_{N_u,\mathcal{S}_g}$ represents the selected columns of $\mathbf{F}_{N_u}$ due to its sparse representation in the angular domain. It is important to stress that the above equation is true, although $\mathbf{S}_g\mathbf{S}_g^T$ is not an identity matrix. The reason for the validity of (8) is due to the fact that because of sparsity, the columns of $\mathbf{H}_g$ only have components for the columns of $\mathbf{F}_{N_u}$ defined by $\mathcal{S}_g$, thus these columns only are needed in the representation of $\mathbf{H}_g$ over $\mathbf{F}_{N_u}$. We can then write that due to non-overlapping supports in groups $g$, $g'$, $\mathcal{S}_g \cap_{g \neq g'} \mathcal{S}_{g'} = \emptyset$, that

$$\mathbf{H}_g^H \mathbf{F}_{N_u,\mathcal{S}_{g'}} = \mathbf{0}, \tag{9}$$

for $g \neq g'$. By TDD channel reciprocity, the group $g$ downlink channel matrix is given as

$$\mathbf{H}_{d,g} = \mathbf{H}_g^H = \mathbf{H}_{g,\upsilon}^H \mathbf{F}_{N_u,\mathcal{S}_g}^H. \tag{10}$$

Since each group attains its non-zero virtual channel representation at non-overlapping positions, we can then use pre-beamforming matrices provided by the matrix $\mathbf{B} = [\mathbf{F}_{N_u,\mathcal{S}_1} \cdots \mathbf{F}_{N_u,\mathcal{S}_G}]$. As we show below these pre-beamforming matrices are optimal for the type of JSDM presented here. Finally, due to non-overlapping of the support sets, i.e., $\mathcal{S}_n \cap_{m \neq n} \mathcal{S}_m = \emptyset$, we see that the system becomes approximately orthogonal inter-group wise,

[5] A selection matrix $\mathbf{S}^T$ of size $k \times n$ with $k < n$ consists of rows equal to different unit row vectors $\mathbf{e}_i$ where the row vector element $i$ is equal to 1 in the $i$th position and is equal to 0 in all other positions. Such a matrix has the property that $\mathbf{S}^T\mathbf{S} = \mathbf{I}$.

i.e., $\sum_{m\neq g}\mathbf{H}_{d,g}\mathbf{H}_{d,m}^{H}\approx 0$. Then,

$$\mathbf{y}_d = \begin{bmatrix} \mathbf{H}_{1,v}^{H}\mathbf{F}_{N_u,\mathcal{S}_1}^{H} \\ \mathbf{H}_{2,v}^{H}\mathbf{F}_{N_u,\mathcal{S}_2}^{H} \\ \vdots \\ \mathbf{H}_{G,v}^{H}\mathbf{F}_{N_u,\mathcal{S}_G}^{H} \end{bmatrix} \begin{bmatrix} \mathbf{F}_{N_u,\mathcal{S}_1} & \mathbf{F}_{N_u,\mathcal{S}_2} & \cdots & \mathbf{F}_{N_u,\mathcal{S}_G} \end{bmatrix} \times \begin{bmatrix} \mathbf{P}_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_2 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{P}_3 & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{P}_{G-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{P}_G \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_G \end{bmatrix} + \mathbf{n}, \tag{11}$$

where for $1 \leq g \leq G$, $\mathbf{H}_{g,v}^{H}$ is a size $N_{d,g} \times |\mathcal{S}_g|$ matrix, $\mathbf{F}_{N_u,\mathcal{S}_G}$ is a size $|\mathcal{S}_g| \times N_u$ matrix, $\mathbf{P}_g$ is a size $|\mathcal{S}_g| \times |\mathcal{S}_g|$ matrix, and $\mathbf{x}_g$ is the group $g$ downlink symbol vector of size $|\mathcal{S}_g| \times 1$. At this point, we would like to stress a main difference between JSDM and JSDM-FA: In (12) we observe that the groups are formed based on the virtual channels $\mathbf{H}_{g,v}^{H}$ $(1 \leq g \leq G)$, while in the original JSDM, because Gaussian channels are assumed, the groups are formed based on the assumed common channel correlation matrix $\mathbf{R}_{h,g}$ common eigenvector matrices $\mathbf{U}_g$ for each user in each group [4]. This is not the case in JSDM-FA, because group formation is performed by employing the projections of users' channels to the VCM DFT matrix basis using the channel model employed here. Due to these differences, the JSDM-FA prebeamformer is significantly different than the JSDM one. In other words, while the JSDM prebeamformer is determined by the eigenvectors of each group's channel correlation matrix, the JSDM-FA prebeamformer is determined by the columns of the DFT matrix that correspond to $\mathcal{S}_g$. Therefore, if one attempts to employ the JSDM beamformer in the JSDM-FA model presented here, by, e.g., determining the covariance matrix of each group's channels, there will be loss of orthogonality between groups and thus severe loss in the system performance. The reason for this is due to the fact that the channels employed in the paper are not Gaussian, thus the eigenvectors of the channel matrices do not in general provide for the actual basis of the group's channel. In other words, this difference represents a very important distinction between the two methods.

Now due to orthogonality, we can write equivalently

$$\mathbf{y}_d = \begin{bmatrix} \mathbf{H}_{1,v}^{H} \\ \mathbf{H}_{2,v}^{H} \\ \vdots \\ \mathbf{H}_{G,v}^{H} \end{bmatrix} \begin{bmatrix} \mathbf{P}_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_2 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{P}_3 & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{P}_{G-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{P}_G \end{bmatrix} \times \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_G \end{bmatrix} + \mathbf{n} = \begin{bmatrix} \mathbf{H}_{1,v}^{H}\mathbf{P_1}\mathbf{x}_1 \\ \mathbf{H}_{2,v}^{H}\mathbf{P_2}\mathbf{x}_2 \\ \vdots \\ \mathbf{H}_{G,v}^{H}\mathbf{P_G}\mathbf{x}_G \end{bmatrix} + \mathbf{n}, \tag{12}$$

where we can set $\mathbf{y} = [\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_G]^T$, with $\mathbf{y}_g$ $(1 \leq g \leq G)$ representing the received downlink data of group $g$. Since each group's precoding becomes independent of other groups, the overall downlink precoding becomes much easier and less complex for both the transmitter and the receiver. In addition, the introduction of the pre-beamforming matrices in the form of VCM beamforming directions also reduces the number of RF chains [4]. Finally, it is worth stressing that the optimal prebeamformers developed herein, being the columns of a DFT matrix, can be implemented as RF (analog) phase-shifters resulting in further simplification due to this RF domain implementation [19], with the rest of the precoder relating to the VCM decomposition being implemented in baseband (BB), resulting in hybrid precoding [19]. As in JSDM, in JSDM-FA, the maximization of (2) is equivalent to maximizing the sum of the group rates, i.e., the total sum-rate. The individual precoding of each group becomes now the optimization of a $|\mathcal{S}_g| \times |\mathcal{S}_g|$ precoding matrix $\mathbf{P}_g$, as per the next theorem. □

*Theorem 2: For each group g in the VCM representation, the equivalent optimum precoder, $\mathbf{P}_{g,v}$ needs to satisfy*

$$\begin{aligned} &\underset{\mathbf{P}_{g,v}}{\text{maximize}}\ I(\mathbf{x}_{d,g};\mathbf{y}_{d,g}) \\ &\text{subject to } \mathrm{tr}(\mathbf{P}_{g,v}\mathbf{P}_{g,v}^{H}) = N_{d,g}, \end{aligned} \tag{13}$$

*where the group g reception model becomes*

$$\mathbf{y}_{d,g} = \mathbf{H}_{g,v}^{H}\mathbf{P}_g\mathbf{x}_g + \mathbf{n}_g, \tag{14}$$

*$\mathbf{H}_{g,v}^{H}$ is the VCM group's downlink matrix of size $N_{d,g} \times |\mathcal{S}_g|$, $\mathbf{y}_{d,g}$ is the group's size $N_{d,g}$ reception vector, and $\mathbf{n}_g$ is the corresponding noise. This per group precoding problem is equivalent to a precoding problem within the original group channel model, i.e., the VCM transformation does not result in mutual information gain loss in the precoding process. In other words, employing $\mathbf{F}_{N_u,\mathcal{S}_g}$ as beamforming matrix per each group g $(1 \leq g \leq G)$ and optimizing the precoder in the VCM domain, is optimal from a maximization of input-output mutual information standpoint.*

*Proof:* From (11) the original precoding problem for group $g$ $(1 \leq g \leq G)$ becomes the solution to the following optimization problem

$$\begin{aligned} &\underset{\mathbf{P}_g}{\text{maximize}}\ I(\mathbf{x}_{d,g};\mathbf{y}_{d,g}) \\ &\text{subject to } \mathrm{tr}(\mathbf{P}_g\mathbf{P}_g^{H}) = N_{d,g}, \end{aligned} \tag{15}$$

where the group $g$ reception model is

$$\mathbf{y}_{d,g} = \mathbf{H}_{d,g}\mathbf{P}_g\mathbf{x}_g + \mathbf{n}_g. \tag{16}$$

From (10), $\mathbf{H}_{d,g} = \mathbf{H}_{g,v}^{H}\mathbf{F}_{N_u,\mathcal{S}_g}^{H}$. Let the Singular Value Decomposition (SVD) of $\mathbf{P}_g$, $\mathbf{H}_{g,v}$, and $\mathbf{H}_{d,g}$ be $\mathbf{P}_g = \mathbf{U}_g\mathbf{\Sigma}_g\mathbf{V}_g^{H}$, $\mathbf{H}_{g,v} = \mathbf{U}_{g,v}\mathbf{\Sigma}_{g,v}\mathbf{V}_{g,v}^{H}$, and $\mathbf{H}_{d,g} = \mathbf{U}_{d,g}\mathbf{\Sigma}_{d,g}\mathbf{V}_{d,g}^{H}$, respectively. From [7] it is known that the optimal precoder for finite inputs for this model has $\mathbf{U}_g$ equal to the right singular vector matrix of $\mathbf{H}_{d,g}$, i.e., $\mathbf{U}_g = \mathbf{U}_{d,g}$.

Because $\mathbf{H}_{d,g} = \mathbf{H}_{g,v}^{H}\mathbf{F}_{N_u,\mathcal{S}_g}^{H}$, and $\mathbf{F}_{N_u,\mathcal{S}_g}^{H}\mathbf{F}_{N_u,\mathcal{S}_g} = \mathbf{I}$, with $N_u > N_{d,g}$, $N_u > |\mathcal{S}_g|$, $\mathbf{U}_g = \mathbf{U}_{g,v}^{H}\mathbf{F}_{N_u,\mathcal{S}_g}^{H}$, with $\mathbf{U}_{g,v}$ the left

singular vector matrix of $\mathbf{H}_{g,\upsilon}$. In addition, one can write

$$\mathbf{H}_{d,g} = \mathbf{H}_{g,\upsilon}^H \mathbf{F}_{N_u,\mathcal{S}_g}^H = \mathbf{V}_{g,\upsilon} \mathbf{\Sigma}_{g,\upsilon} \mathbf{U}_{g,\upsilon}^H \mathbf{F}_{N_u,\mathcal{S}_g}^H. \tag{17}$$

Thus, the solution of (15) becomes equivalent to the solution of the following problem

$$\begin{aligned} &\underset{\mathbf{\Sigma}_g, \mathbf{V}_g}{\text{maximize}}\ I(\mathbf{x}_{d,g}; \tilde{\mathbf{y}}_{d,g}) \\ &\text{subject to } \operatorname{tr}(\mathbf{\Sigma}_g \mathbf{\Sigma}_g^H) = N_{d,g}, \end{aligned} \tag{18}$$

where $\tilde{\mathbf{y}}_{d,g} = \mathbf{V}_{g,\upsilon}^H \mathbf{\Sigma}_{g,\upsilon} \mathbf{\Sigma}_g \mathbf{V}_g \mathbf{x}_g + \mathbf{n}_g$. However, this is exactly the same SVD-equivalent [7] optimization problem with the one employing the virtual channel $\mathbf{H}_{g,\upsilon}$ SVD-equivalent, as one can see by the relationship (17) and the fact that the optimal precoder has a left singular matrix equal to the channel right singular matrix. □

Note that after establishing the optimality of the employed prebeamformer in the presented model, one needs to determine the optimal precoder for each group based on optimizing (13) over $\mathbf{P}_g$ for finite alphabet inputs. Assume that optimal precoder is $\mathbf{P}_g^*$. The overall optimal precoder for group $g$ ($1 \leq g \leq G$), including the prebeamformer is denoted by $\tilde{\mathbf{G}}_g^* = \mathbf{F}_{N_u,\mathcal{S}_g} \mathbf{P}_g^*$. Without a complexity-simplifying method the determination of a globally optimal $\mathbf{P}_g^*$ is impossible for $N_{d,g}$, $|\mathcal{S}_g| \geq 3$ for QAM with $M \geq 16$. However, efficient solutions to this problem which are suboptimal for $N_{d,g}$, $|\mathcal{S}_g| \geq 3$, but have been proven to be near-optimal, have been presented in [14] and [20]. They are both using the PGP-WG concept which divides the inputs in independent data groups, the PGP-WG groups, and precodes independently each group in the SVD domain. The PGP-WG groups are selected based on their corresponding singular values [20]. It is important to stress that, within each PGP-WG group an exact optimal precoder is developed based on the framework of [7] and that for each group the corresponding constraint on the power is satisfied by, e.g., using an additional power constraint factor in the power optimizing loop of the steepest ascent method employed [7], so that overall the constraint in (13) remains valid. The globally optimal implementation within PGP-WG groups is performed by employing two backtracking line searches [26], one for updating $\mathbf{W} = \mathbf{U}_{g,\upsilon} \mathbf{\Sigma}_{g,\upsilon}^2 \mathbf{U}_{g,\upsilon}$ (the matrix $\mathbf{U}_{g,\upsilon}$ update loop) and another one for $\mathbf{\Sigma}_G^2$, in conjunction with enforcing the power constraint (the per antenna power update loop) during each iteration.

*2) UPA at the Base:* The concept generalizes easily to Uniform Planar Arrays (UPA), both for arrays formed in the $z$, $x$ plane as well as in the $x$, $y$ plane, a UPA deployed along the $x$, $y$ plane is shown in Fig. 2. The theory behind planar arrays results in a Kronecker product of two virtual channels, one channel per array dimension, as shown below. UPAs result in a three-dimensional spatial representation, thus offering higher user capacity per cell. Among the two UPA possibilities, we present the analysis for an $x$, $z$ direction deployed UPA, as the analysis for an $x$, $y$ direction deployed UPA is very similar. For a UPA formed on $x$, $z$ directions, each group $g$'s uplink channel, $\mathbf{H}_g$, corresponds to the combination of $N_{u,x}$ uniform linear arrays deployed along the $x$ direction with $N_{u,z}$ uniform linear elements deployed in the $z$ direction. Without loss in generality, we assume that the normalized distances are the same for each direction and equal to $D$. A UPA introduces a new level of control and additional degrees of freedom for MIMO [27]. UPAs result in a two-dimensional (matrix) antenna response matrix per user, group, and antenna expressed as

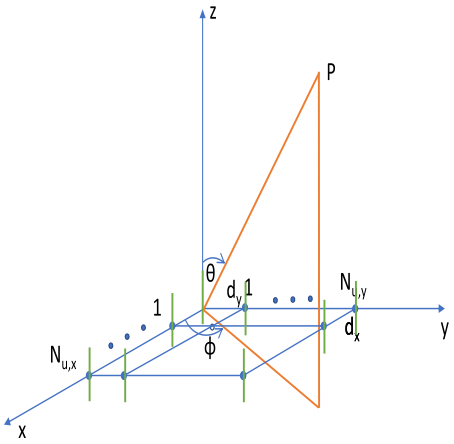



Fig. 2. A UPA deployed across the $x$, $y$ axes, together with the projection of a tentative transmission point on the $x$, $y$ plane.

$$\mathbf{H}_{u,g,k,n} = \sum_{l=1}^{L} \beta_{lgkn} \mathbf{a}_x(\theta_{lgkn}, \phi_{lgkn}) \mathbf{a}_z^T(\theta_{lgkn}), \tag{19}$$

where the path gain $\beta_{lgkn}$ is as in the ULA case, $\theta_{lgkn}$ is the elevation angle for the $z$-element, same as in the ULA case, and $\phi_{lgkn}$ is the azimuth angle of user $k$'s $n$ antenna for group $g$, assumed to be a uniform r.v. in the interval $[\bar{\phi}_g - \Delta\phi, \bar{\phi}_g + \Delta\phi]$, where $\bar{\phi}_g$ is the mean of $\phi_g$. In (19), the two spatial vectors $\mathbf{a}_x(\theta_{lgkn}, \phi_{lgkn})$, $\mathbf{a}_z(\theta_{lgkn})$ are given as

$$\begin{aligned} \mathbf{a}_z(\theta_{lgkn}) = [&1, \exp(-j2\pi D \cos(\theta_{lgkn})), \\ &\cdots, \exp(-j2\pi D(N_{u,z} - 1)\cos(\theta_{lgkn}))]^T, \end{aligned} \tag{20}$$

and

$$\begin{aligned} \mathbf{a}_x(\theta_{lgkn}, \phi_{lgkn}) = [&1, \exp(-j2\pi D \sin(\theta_{lgkn}) \cos(\phi_{lgkn})), \\ &\cdots, \exp(-j2\pi D(N_{u,x} - 1)\sin(\theta_{lgkn}) \\ &\times \cos(\phi_{lgkn}))]^T, \end{aligned} \tag{21}$$

respectively.

By projecting the channel matrix $\mathbf{H}_{u,g,k,n}$ to both angular directions, i.e., on VCM for $z$, $x$ directions, we get

$$\tilde{\mathbf{H}}_{u,g,k,n} = \sum_{l=1}^{L} \beta_{lgkn} \mathbf{F}_{N_{u,x}}^H \mathbf{a}_x(\theta_{lgkn}, \phi_{lgkn}) \mathbf{a}_z^T(\theta_{lgkn}) \mathbf{F}_{N_{u,z}}^*. \tag{22}$$

By taking the vector form of both sides in (22) and using identities from [28], e.g., $\mathbf{a} \otimes \mathbf{b} = \operatorname{vec}(\mathbf{b}\mathbf{a}^T)$ and $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BD})$, we can write for the vector of $\tilde{\mathbf{H}}_{u,g,k,n}$, $\tilde{\mathbf{h}}_{u,g,k,n} \doteq \operatorname{vec}(\tilde{\mathbf{H}}_{u,g,k,n})$, the following equation

$$\tilde{\mathbf{h}}_{u,g,k,n} = \sum_{l=1}^{L} \beta_{lgkn} \tilde{\mathbf{a}}_{z,x}(\theta_{lgkn}, \phi_{lgkn}), \tag{23}$$

where

$$\tilde{\mathbf{a}}z, x(\theta_{lgkn}, \phi_{lgkn}) = \left(\mathbf{F}_{N_u,z} \otimes \mathbf{F}_{N_u,x}\right)^H \left(\mathbf{a}_z(\theta_{lgkn}) \otimes \mathbf{a}_x(\theta_{lgkn}, \phi_{lgkn})\right). \quad (24)$$

The behavior in (24) is similar with the ULA case, i.e., sparsity is achieved and different groups occupy different support sets in the angular domain. The expansion basis matrix now for the VCM becomes the Kronecker product of the two DFT matrices $\mathbf{F}_{N_{u,x}}$ and $\mathbf{F}_{N_{u,z}}$. The downlink reception model stays within the same premise, but the new Kronecker product basis is employed. Due to the Kronecker product, the group sparsity presents some periodicity with period equal to $N_{u,z}$. In other words, the reception model now becomes

$$\mathbf{y}_d = \begin{bmatrix} \mathbf{H}_{1,v}^H \left([\mathbf{F}_{N_{u,z}} \otimes \mathbf{F}_{N_{u,x}}])_{\mathcal{S}_1}\right)^H \\ \mathbf{H}_{2,v}^H \left([\mathbf{F}_{N_{u,z}} \otimes \mathbf{F}_{N_{u,x}}]_{\mathcal{S}_2}\right)^H \\ \vdots \\ \mathbf{H}_{G,v}^H \left([\mathbf{F}_{N_u,z} \otimes \mathbf{F}_{N_u,x}]_{\mathcal{S}_G}\right)^H \end{bmatrix} \times \begin{bmatrix} \left(\mathbf{F}_{N_{u,z}} \otimes \mathbf{F}_{N_{u,x}}\right)_{\mathcal{S}_1} & \left(\mathbf{F}_{N_{u,z}} \otimes \mathbf{F}_{N_{u,x}}\right)_{\mathcal{S}_2} \\ \cdots & \left(\mathbf{F}_{N_{u,z}} \otimes \mathbf{F}_{N_{u,x}}\right)_{\mathcal{S}_G} \end{bmatrix} \times \begin{bmatrix} \mathbf{P}_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_2 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{P}_3 & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{P}_{G-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{P}_G \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_G \end{bmatrix} + \mathbf{n}, \quad (25)$$

where the notation $(\mathbf{A})_{\mathcal{S}_g}$ means the matrix resulting from selecting the columns of $\mathbf{A}$ that belong to $\mathcal{S}_g$. The case of a UPA over $x$, $y$ dimensions can be treated in a similar way by invoking

$$\tilde{\mathbf{a}}_{y,x}(\theta_{lgkn}, \phi_{lgkn}) = \left(\mathbf{F}_{N_{u,y}} \otimes \mathbf{F}_{N_{u,x}}\right)^H \left(\mathbf{a}_y(\theta_{lgkn}, \phi_{lgkn}) \otimes \mathbf{a}_x(\theta_{lgkn}, \phi_{lgkn})\right), \quad (26)$$

with

$$\mathbf{a}_y(\theta_{lgkn}) = [1, \exp(-j2\pi D \sin(\theta_{lgkn}) \sin(\phi_{lgkn})), \cdots, \exp(-j2\pi D(N_{u,y}-1)) \sin(\theta_{lgkn}) \sin(\phi_{lgkn}))]^T, \quad (27)$$

and

$$\mathbf{a}_x(\theta_{lgkn}, \phi_{lgkn}) = [1, \exp(-j2\pi D \sin(\theta_{lgkn}) \cos(\phi_{lgkn})), \cdots, \exp(-j2\pi D(N_{u,x}-1) \sin(\theta_{lgkn}) \times \cos(\phi_{lgkn}))]^T. \quad (28)$$

The sparsity in the UPA case is due to the behavior of both angles, i.e., the elevation and the azimuth ones. The corresponding conditions to Lemma 1 are posted in the next lemma.

*Lemma 3: In the UPA over $z$, $x$ dimensions, when $N_u \doteq N_{u,z}N_{u,x} \gg 1$, then the significant components of the channel for group $g$, i.e., the support set $\mathcal{S}_g$, are found through the following two conditions*

$$|\cos(\theta_{lgkn}) - \frac{p}{DN_{u,z}}| < \frac{1}{DN_{u,z}}, \quad (29)$$

*and*

$$|\sin(\theta_{lgkn})\cos(\phi_{lgkn}) - \frac{p}{DN_{u,x}}| < \frac{1}{DN_{u,x}}. \quad (30)$$

*Proof:* The proof stems from generalizing the condition in (4) to the geometries of the UPA array. For the $z$ direction the equation remains unchanged, while for the $x$ direction the factor $\cos(\theta_{lgkn})$ needs to be substituted by $\sin(\theta_{lgkn})\cos(\phi_{lgkn}) - \frac{p}{DN_{u,x}}$. For significant factors to exist, both conditions need to be satisfied simultaneously, because the composite array factor is the product of the two individual ones. This completes the proof of the lemma. □

In comparison to the ULA channel case sparsity behavior though, it is important to stress that UPA channels present a repetitive, semi-periodic sparsity structure, due to the Kronecker product that exists in the vectorized form of the channel vectors. This behavior is further contrasted to the ULA one in Section V where numerical results are used to depict differences between ULA and UPA behavior with regards to sparsity in the VCM representation.

### *B. The Frequency-Selective System Description Under the Virtual Channel Model Representation*

Here we present a generalization of JSDM-FA for frequency-selective fading with OFDM. The presentation looks at a UPA deployed over the $z$, $x$ directions. However, similar descriptions can be found for ULA and for different directions of deploying the array.

Using the Tap Delay Line (TDL) model of an FS channel [29], we can write for the uplink channel response of the UPA[6] in time domain at discrete time $m$

$$\mathbf{h}_{u,g,k,n}^{(t)}[m] = \frac{1}{\sqrt{L}} \sum_{l=1}^{L} \beta_{lgkn} \mathbf{a}_y(\theta_{lgkn}, \phi_{lgkn}) \otimes \mathbf{a}_x(\theta_{lgkn}, \phi_{lgkn}) \delta[\frac{m}{B} - \frac{l-1}{B}], \quad (31)$$

where $\delta(\cdot)$ represents the Dirac delta function, $B$ is the system bandwidth, with $B \gg B_{COH}$ and where $B_{COH}$ is the coherence bandwidth of the channel. We can then write for the frequency response of the channel

$$\tilde{\mathbf{H}}_{u,g,k,n}^{(f)} = \frac{1}{\sqrt{L}} \sum_{l=1}^{L} \beta_{lgkn} \left(\mathbf{F}_{N_u,y} \otimes \mathbf{F}_{N_u,x}\right)^H \mathbf{a}_y(\theta_{lgkn}, \phi_{lgkn}) \otimes \mathbf{a}_x(\theta_{lgkn}, \phi_{lgkn}) \mathbf{f}_{L,Q,l} = \left(\mathbf{F}_{N_u,y} \otimes \mathbf{F}_{N_u,x}\right)^H \mathbf{M}_h \mathbf{F}_{L,Q}, \quad (32)$$

where $\mathbf{F}_{L,Q}$ is the last $Q - L$ row-truncated DFT matrix of order $Q$, i.e., a matrix of size $L \times Q$, $\mathbf{f}_{L,Q,l}$ is its $l$th column, $\mathbf{M}_h$ is an $N_{u,x}N_{u,y} \times L$ matrix equal to $[\mathbf{a}_y(\theta_{lgkn}, \phi_{lgkn}) \otimes \mathbf{a}_x(\theta_{1gkn}, \phi_{1gkn}) \cdots \mathbf{a}_y(\theta_{Lgkn}, \phi_{Lgkn}) \otimes$

[6]Similar results are derived for any UPA or ULA configuration within the context of this paper.

$\mathbf{a}_x(\theta_{Lgkn}, \phi_{Lgkn})]\text{diag}[\beta_{1gkn} \cdots \beta_{Lgkn}]$, where diag[·] is the diagonal matrix of the vector in the brackets. Thus, $\tilde{\mathbf{H}}^{(f)}_{u,g,k,n}$ is of size $N_{u,x}N_{u,y} \times Q$, with only a few non-zero entries on each column, all of them on the same row numbers. The $q$th column of $\tilde{\mathbf{H}}^{(f)}_{u,g,k,n}$ is the uplink channel impulse response at sub-carrier $q$ denoted as $\mathbf{h}^{(q)}_{u,g,k,n}$. By recalling the fact that the spatial channel is sparse when projected to the virtual angles, exploiting the virtual channel domain representation, and after using the channel reciprocity between uplink and downlink due to TDD, for each sub-carrier, we can rewrite the downlink channel of user's $k$, antenna $n$, sub-carrier $q$, and group $g$ as $(\mathbf{h}^{(q)}_{u,g,k,n,v})^H$. We can then write for the downlink channel over all sub-carriers, $\mathbf{H}^{(f)}_{d,g,k,n}$,

$$\mathbf{H}^{(f)}_{d,g,k,n} = \begin{bmatrix} (\mathbf{h}^{(0)}_{u,g,k,n,v})^H \\ (\mathbf{h}^{(1)}_{u,g,k,n,v})^H \\ \vdots \\ (\mathbf{h}^{(Q-1)}_{u,g,k,n,v})^H \end{bmatrix} \Big[ \left( [\mathbf{F}_{N_u,y} \otimes \mathbf{F}_{N_u,x}]_{\mathcal{S}_g} \right)^H \Big], \quad (33)$$

then by stacking together all antennas for user $k$, we get

$$\mathbf{H}^{(f)}_{d,g,k} = \begin{bmatrix} (\mathbf{H}^{(0)}_{u,g,k,v})^H \\ (\mathbf{H}^{(1)}_{u,g,k,v})^H \\ \vdots \\ (\mathbf{H}^{(Q-1)}_{u,g,k,v})^H \end{bmatrix} \Big[ \left( \mathbf{F}_{N_u,y} \otimes \mathbf{F}_{N_u,x} \right)^H_{\mathcal{S}_g} \Big], \quad (34)$$

where $\mathbf{H}^{(q)}_{u,g,k,v} = \left[ \mathbf{H}^{(q)}_{u,g,k,1,v} \cdots \mathbf{H}^{(q)}_{u,g,k,N_{d,k(g)},v} \right]$, a size $|\mathcal{S}_g| \times N_{d,k(g)}$ matrix. Each group, $g$ ($1 \le g \le G$) can be considered independently due to JSDM, as explained above. We can then employ different sub-carriers for different users within a group or between different groups, which is explained in more detail next.

### C. Combined Frequency and Spatial Division and Multiplexing (CFSDM)

In certain scenarios, user co-ordination-related issues within each group data receiver might make JSDM difficult to deploy, in general. In addition, in some cases, one might desire to use all VCMBs available to a group, although they experience significant spatial overlapping with other groups, in order to *increase the group's delivered throughput*, without sacrificing the cell overall spectral efficiency. One promising solution to mitigate this problem, without sacrificing the overall system capacity, is proposed herein by virtue of a novel introduction of the concept of CFSDM. This idea is described below.

In CFSDM between groups (CFSDM-BG), group support sets with common VCMBs are assigned different OFDM sub-carriers. We present here some theoretical justification for the benefits of CFSDM-BG, while numerical results and comparisons are presented in the next section. Let us assume that a Massive MIMO system within the premise of our model presented previously is employing JSDM-FA with OFDM, as described above. Assume that $N_c$ carriers are deployed without CFSDM. Denote the VCMBs of each group by $\mathcal{S}_1, \ \mathcal{S}_2, \cdots, \mathcal{S}_G$, respectively. In addition, denote by $\mathcal{S}_f = \{\mathcal{S}_1, \ \mathcal{S}_2, \cdots, \mathcal{S}_G\}$, i.e., the set of all VCMB groups. As illustrated in the next section, there will be partial spatial overlapping between adjacent groups for both ULA and UPA. In general, UPAs tend to experience more spatial overlapping for similar parameter values. In order to apply JSDM-FA without CFSDM, one needs to release (neutralize) all overlapping VCMBs between groups, thus reducing each group's spectral efficiency. We call the technique of releasing common VCMBs Neutralization of Overlapping VCMBs (NOV). For simplicity, we assume here analysis in high SNR, where there are groups with large numbers of multiple transmitting and receiving antennas. Then, each group can achieve a spectral efficiency (utilization), $E_g = |\mathcal{S}_{g,no}| \log_2(M)$ over all $N_c$ sub-carriers employed, where $\mathcal{S}_{g,no}$ denotes the resulting non-overlapping VCMBs available for group $g$, with $\mathcal{S}_g \ge \mathcal{S}_{g,no}$. We have assumed that all groups employ the same modulation size, $M$, without loss of generality. Then, the overall spectral efficiency of the system for the entire cell using a total of $N_c$ sub-carriers and with near-optimal precoding becomes[7]

$$E_{NCFSDM} = \sum_{g=1}^{G} E_g = \log_2(M) \sum_{g=1}^{G} |\mathcal{S}_{g,no}|. \quad (35)$$

Now, let us apply CFSDM-BG in the same scenario. It is evident that the optimal strategy from the spectral efficiency point of view should be to use different sub-carriers to the groups which lost the most VCMBs due to the process of switching off overlapping VCMBs above, because these groups experience the highest loss of VCMBs. For example, group $g$ ($1 \le g \le G$) has experienced a $\log_2(M)(|\mathcal{S}_g| - |\mathcal{S}_{g,no}|) = \log_2(M)(\Delta_g)$ loss in group utilization, where we defined $\Delta_g = |\mathcal{S}_g| - |\mathcal{S}_{g,no}|$ to be the VCMB loss of group $g$. Let us calculate the corresponding utilization with CFSDM-BG after assigning a set of $N_c$ different sub-carriers to groups in $\mathcal{S}_{f,diff} = \{\mathcal{S}_{n_1}, \cdots, \mathcal{S}_{n_h}\}$ ($n_k$ with $1 \le k \le h$ are different integers with $1 \le n_k \le G$) are the CFSDM selected groups with high overlapping, while the rest of the groups use the same set of the previous $N_c$ sub-carriers. This way, the groups in $\mathcal{S}_{f,diff}$ can employ the entirety of their VCMBS, i.e., all $|\mathcal{S}_{n_k}|$ VCMBs. The CFSDM-BG utilization is then

$$\begin{aligned} E_{CFSDM} &= \frac{\log_2(M)}{2N_c} \Big( N_c \cdot \sum_{g \text{ in } \mathcal{S}_g - \mathcal{S}_{f,diff}} |\mathcal{S}_{g,no}| \\ &\quad + N_c \cdot \sum_{g \text{ in } \mathcal{S}_{f,diff}} |\mathcal{S}_g| \Big) \\ &\ge \log_2(M) \frac{\left( \frac{E_{NCFSDM}}{\log_2(M)} + \sum_{g \text{ in } \mathcal{S}_{f,diff}} \Delta_g \right)}{2}, \end{aligned} \quad (36)$$

where the reason for the inequality is due to the fact that by employing a different sub-carrier in $\mathcal{S}_{f,diff}$, there will in general exist more VCMBs available without overlapping in the set $\mathcal{S}_g - \mathcal{S}_{f,diff}$, however the lower bound suffices for our illustration. Thus, if $\sum_{g \in \mathcal{S}_{f,diff}} \Delta_g \ge \sum_{g=1}^{G} |\mathcal{S}_{g,no}|$, we see that $E_{CFSDM} \ge E_{NCFSDM}$. It is important to stress that, under this condition, the overall cell utilization is increased, but also the group utilization is increased, while

[7]This analysis assumes Type-II channel behavior (see next section).

simultaneously the $\mathcal{S}_{f,diff}$ groups enjoy highly increased utilization, i.e., significantly higher data rates. In fact, it is easily seen that if $\sum_{g \in \mathcal{S}_{f,diff}} \Delta_g \geq \sum_{g=1}^{G} |\mathcal{S}_{g,no}|$, then $\mathcal{S}_{f,diff}$ will get at least twice as much utilization than the one without CFSDM. This property shows clearly the potential of CFSDM-BG in offering high throughput to highly impacted groups by NOV. Furthermore, the CFSDM-BG concept can be generalized to using more than two separate sub-carrier frequency sets. In the next section, numerical examples are presented in order to further illustrate the CFSDM-BG concept.

In CFSDM within groups (CFSDM-WG), users with multiple antennas within each group are also assigned different OFDM sub-carriers. Finally, for users with a single antenna on the downlink, offering multiple sub-carriers is the only possibility toward higher data rates. The novelty of combining JSDM based on the VCM decomposition as proposed here and OFDM is due to the fact that it helps mitigate interference issues associated with intra-group co-ordination during the data reception. Due to the orthogonality among the sub-carriers in OFDM, it becomes feasible for different users within a group, by assigning different sub-carriers to each user, that each user receives its data on a separate sub-carrier, utilizing its own receiving antennas only, thus obliterating the requirement for user co-ordination at the receiver, while the receiver complexity is dramatically reduced. Specifically, let's look at a system with FS and OFDM as described in the previous subsection. Assume the system groups are as in Section I and that the OFDM component contains $Q$ orthogonal sub-carriers, for some "high enough" number, $Q$ (e.g., $Q \geq 64$). First, let's assume that there is overlapping of the VCMBs between groups $g$ and $g'$, i.e., $\mathcal{S}_g \cap \mathcal{S}_{g'} \neq \emptyset$. The system then assigns these groups to different sub-carrier groups, say $\mathcal{S}_{g,q}$, $\mathcal{S}_{g',q'}$, which will be defined explicitly after the user sub-carriers are assigned. Since there are $K_g$ users in group $g$, there is a need to assign $K_g$ sub-carriers for group $g$ and $K_{g'}$ for group $g'$, if no coordination exists between users in the groups. In order for the two groups to employ all spatial capability available to them, the two groups need to avoid interference over the common VCMBs, thus in total the two groups need $K_g + K_{g'}$ different sub-carriers assigned to them. Within each group, say for group $g$, user $k^{(g)}$ employing sub-carrier $q_{g,k}$, there will be a PGP-WG precoder employed in the sub-carrier domain pertaining to the following receiver model

$$\begin{aligned}\mathbf{y}_{d,k^{(g)}}^{(q)} &= \left[ \ (\mathbf{H}_{u,k^{(g)},v}^{(q)})^H \right] \left[ \ \left( \ [\mathbf{F}_{N_u,y} \otimes \mathbf{F}_{N_u,x}]_{\mathcal{S}_g} \right)^H \right] \\ &\quad \times \left(\mathbf{F}_{N_u,y} \otimes \mathbf{F}_{N_u,x}\right)_{\mathcal{S}_g} \mathbf{P}_{g,k^{(g)}}^{(q)} \mathbf{c}_{g,k}^{(q)} + \mathbf{n}_{g,k^{(g)}}^{(q)} \\ &= \left[ \ (\mathbf{H}_{u,g,k,v}^{(q)})^H \right] \mathbf{P}_{g,k^{(g)}}^{(q)} \mathbf{c}_{g,k}^{(q)} + \mathbf{n}_{g,k}^{(q)}. \end{aligned} \tag{37}$$

Now, precoding is performed on a per user and sub-carrier basis, without the need for user co-operation within the group. This CFSDM approach allows for more flexible data rate allocations on a per user basis as well as helps in overcoming issues associated with spatial overlapping between groups. The following lemma also helps simplify the precoder design when the number of group antennas $N_{d,g}$ is smaller than the number of available spatial dimensions $|\mathcal{S}_g|$.

*Lemma 4: When all users in a group have the same number of antennas and with $L \ll \sqrt{Q}$ and sub-carrier pairs $q$, $q'$ assigned within a group satisfying $|q - q'| \ll \sqrt{Q}$, then all sub-carrier virtual downlink channel matrices, i.e., for all $q = 1, 2, \cdots, Q$, $(\mathbf{H}_{u,g,k,v}^{(q)})^H$ have the same singular values. Thus, the optimal precoder over all sub-carriers is the same.*

*Proof:* We can easily rewrite (37) by employing Kronecker matrix products as

$$\mathbf{y}_{d,k^{(g)}}^{(q)} = \left[ \ \left( \mathbf{I}_{N_{k^{(g)}}} \otimes \mathbf{f}_{q,L}{}^H \right) (\mathbf{M}_{u,g,k,v})^H \right] \mathbf{P}_{g,k^{(g)}}^{(q)} \mathbf{c}_{g,k} + \mathbf{n}_{g,k}^{(q)}, \tag{38}$$

where $(\mathbf{M}_{u,g,k,v})$ is a $N_{k^{(g)}} L \times |\mathcal{S}_g|$ virtual channel matrix derived from (32) and $\mathbf{f}_{q,L}$ represents the $q$th column of the matrix $\mathbf{F}_{L,Q}$. Now, based on the assumptions of the lemma, for any different sub-carriers assigned to the group and for all $1 \leq l \leq L$, we have $\exp(j2\pi \frac{(q-q')l}{Q}) \approx 1$, from which we see that the matrices are approximately $\left( \mathbf{I}_{N_{k^{(g)}}} \otimes \mathbf{f}_{q,L}{}^H \right) (\mathbf{M}_{u,g,k,v})^H$ equal for all users in the group, thus they possess approximately equal singular values. □

Note that for a massive MIMO system a large $Q$ will be needed. In addition, in the millimeter wavelength channels envisaged for 5G cellular wireless systems, the assumption of $L \ll \sqrt{Q}$ will also be valid, since $L$ is small [30]. Thus, by assigning contiguous frequency sub-carriers to different users within groups we can achieve the conditions of the above lemma. Based on the premise of this lemma, the optimal downlink precoder in the group is the same, independently of the sub-carrier employed. This is due to the fact that for CSIT optimal precoding, the optimal precoder only depends on the singular values of the channel matrix [7], [20]. Thus, if many sub-carriers are deployed to offer higher data rates, the precoding complexity stays the same.

## IV. Numerical Results

In this section, we present our numerical results based on ULA and UPA Massive MIMO systems with $N_u = 100$ antennas at the base station. The systems employ QAM with size $M = 16, 64$. We present results for both systems with and without OFDM. We have used an $L = 3$ Gauss-Hermite approximation [15] which results in $3^{2N_r}$ total nodes in the Gauss-Hermite approximation due to MIMO in order to facilitate results with near-optimal precoding in conjunction with QAM modulation. The implementation of the globally optimizing methodology is performed by employing two backtracking line searches, one for $\mathbf{W}$ and another one for $\mathbf{\Sigma}_G^2$ at each iteration, in a fashion similar to [7]. For the results presented, it is worth mentioning that only a few iterations (e.g., typically $< 8$) are required to converge to the near-optimal solution results as presented in this paper. We apply the complexity reducing method of PGP-WG [20] to each JSDM-FA formed group which offers near-optimal results under exponentially lower transmitter and receiver complexity [20]. PGP-WG divides the transmitting and receiving antennas into independent sub-groups, within each JSDM-FA group, thus achieving a much simpler detector structure while the precoder search is also dramatically reduced as well. We divide this section into three parts, the first part looks at the VCM sparse

channel representations for ULA and UPA systems, the second one examines the performance of linear precoding for Massive MIMO without OFDM, while the third one studies systems with OFDM. We use $N_{t,v}, N_{r,v}$ to denote the number of data symbol inputs, and the number of of antenna outputs, respectively, in the virtual domain. By employing PGP-WG, one can trade in higher values of $N_{t,v}$, $N_{r,v}$ for higher overall throughput, albeit at a slightly increased complexity at the transmitter and receiver, as explained in detail in some of the examples below. Alternatively, one can employ a smaller number of $N_{t,v}$, $N_{r,v}$, in order to achieve higher throughput, but at significantly lower complexity. In all cases, it is stressed that the actual number of transmission and reception antennas stays the same, while all physical antennas are employed always. The details of these techniques are omitted here due to space limitation.

It is worthwhile mentioning that for precoding methods with finite inputs, two types of channels are regularly present in the literature [7], [13], [15], [20], [32], and [33]: a) Type-I channels in which the precoder offers gain in the lower SNR regime, and b) Type-II channels in which the precoder offers gain in the high SNR regime. Our results herein fully corroborate this type of behavior in all cases considered.

### A. VCM Channel Sparsity for ULA and UPA Scenarios

First, we present results for the sparse behavior of the VCM representation in the ULA case. We randomly create 5 groups of channels as per the ULA model presented. The base ULA is deployed along the $z$ direction with $N_u = 100$ elements spaced at a normalized distance $D = 0.5$. There are $L = 5$ paths in each channel (a smaller number of $L$ results in sparser representations). The elevation angles for groups $G_1$, $G_2$, $G_3$, $G_4$, $G_5$ are at 5°, 33°, 61°, 89°, and 117°, respectively. In addition, the groups possess 16, 2, 4, 4, and 6 antennas, respectively. The angular spread for all groups is taken to be $\pm$ 4° around the elevation angle of each group. The channels are projected to the VCM space, then only components greater than 1 in absolute square power are selected. In all cases considered, this selection process results in more than 94% of the total power of each channel selected. The corresponding, non-overlapping support sets are as follows (the numbers of each set correspond to the numbered components of the VCM representation vector, i.e., the significant VCMBs):

$$\begin{aligned}
\mathcal{S}_1 &= [56,\ 57,\ 58,\ 59,\ 60,\ 61,\ 62,\ 63,\ 64,\ 65],\\
\mathcal{S}_2 &= [38,\ 39,\ 40,\ 41,\ 42,\ 43,\ 44],\\
\mathcal{S}_3 &= [27,\ 28,\ 29,\ 30,\ 31,\ 32,\ 33,\ 34],\\
\mathcal{S}_4 &= [1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 99,\ 100],\\
\mathcal{S}_5 &= [70,\ 71,\ 72,\ 73,\ 74,\ 75,\ 76,\ 77,\ 78,\ 79].
\end{aligned}$$

We observe that a ULA allows for easy sparse non-overlapping support sets for multiple groups.

Next we present similar results for a UPA array along the $x$, $y$ directions. In this example, there are 8 groups, $G_1$ through $G_8$, formed. The normalized distance between successive elements in both directions is $D = 0.6$, while the number of elements on each direction is equal to 10, i.e., $N_{u,x} = N_{u,y} = 10$. There are a total of 16, 1, 4, 4, 6, 1, 6, and 8 antennas available for each group. The angle spread per dimension is $\pm 2°$, while $L = 2$. The corresponding VCMBs per group are as follows:

$$\begin{aligned}
\mathcal{S}_1 &= [1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9,\ 10,\ 11,\ 12,\ 20,\ 21,\\
&\qquad 31,\ 41,\ 51,\ 61,\ 71,\ 81,\ 91],\\
\mathcal{S}_2 &= [12,\ 13,\ 14,\ 15],\\
\mathcal{S}_3 &= [2,\ 3,\ 4,\ 5,\ 6,\ 11,\ 12,\ 13,\ 14,\ 15,\ 16,\ 17,\ 23,\\
&\qquad 24,\ 34,\ 44,\ 54,\ 64,\ 74,\ 84,\ 93,\ 94\ ],\\
\mathcal{S}_4 &= [3,\ 4,\ 5,\ 6,\ 14,\ 15,\ 24,\ 34,\ 64,\ 74,\ 75,\ 84,\\
&\qquad 85,\ 91,\ 92,\ 93,\ 94,\ 95,\ 96,\ 97,\ 98,\ 99,\ 100],\\
\mathcal{S}_5 &= [74,\ 83,\ 84,\ 85,\ 94],\\
\mathcal{S}_6 &= [73,\ 83],\\
\mathcal{S}_7 &= [1,\ 2,\ 11,\ 21,\ 61,\ 71,\ 81,\ 82,\ 91,\ 92,\ 93,\\
&\qquad 94,\ 95,\ 96,\ 97,\ 98,\ 99,\ 100],\\
\mathcal{S}_8 &= [1,\ 2,\ 3,\ 4,\ 8,\ 9,\ 10,\ 11,\ 20,\ 21,\ 31,\ 41,\\
&\qquad 51,\ 61,\ 71,\ 81,\ 91,\ 92,\ 99,\ 100].
\end{aligned}$$

It is easy to see that UPA deployments offer more VCMBs per group, however at a cost to orthogonality. In addition, UPAs offer better resolution compared to ULAs, thus they could in principle offer higher capacity. An additional benefit of a UPA is the fact that one gets more VCMBs per group thus the resulting throughput with precoding is higher. However, due to the significant overlapping between different group VCMBs, there are two options when UPAs are selected for higher capacity: a) Release common VCMBs, i.e., leave the common VCMBs between groups unused, however at the expense of utilization in certain groups, or, b) Employ CFSDM-BG. The latter approach can offer higher cell and group utilization due to its capability to mitigate overlapping VCMBs in spatial domain. Both approaches are explained in more detail below.

### B. Precoding Results Without OFDM

As a first example, we present results for a ULA with 5 groups formed, shown as $G_1$, $G_2, \cdots, G_5$, respectively. They occupy the following groups of non-overlapping, i.e., disjoint VCMBs

$$\begin{aligned}
\mathcal{S}_1 &= [57,\ 58,\ 59,\ 60,\ 61,\ 62,\ 63,\ 64],\\
\mathcal{S}_2 &= [39,\ 41,\ 42, 43,\ 44,\ 45,\ 46,\ 47],\\
\mathcal{S}_3 &= [25,\ 26,\ 27,\ 28,\ 29,\ 30,\ 31,\ 32,\ 33, 34,\ 35],\\
\mathcal{S}_4 &= [1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 98,\ 99, 100],\ \text{and}\\
\mathcal{S}_5 &= [68,\ 69,\ 70,\ 71,\ 72,\ 73,\ 74,\ 75,\ 76,\ 77],
\end{aligned}$$

respectively. The groups include 4, 2, 4, 4, 6 antennas at UE, respectively. In the non-OFDM case, users within groups need to co-ordinate their downlink. Thus, the number of users within the group becomes irrelevant and only the number of antennas becomes essential. In Fig. 3 we present results for $G_4$. We observe that high gains in throughput are available for low SNR, i.e., a Type-I channel behavior. For example, at $\text{SNR}_b = -7\ dB$ there is a 33% throughput increase by using

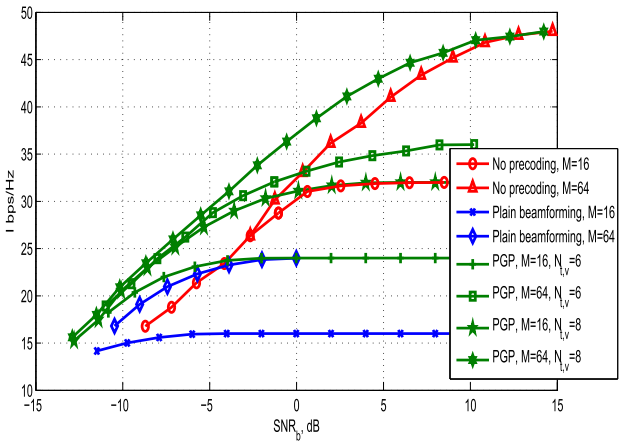



Fig. 3. $I(\mathbf{x};\mathbf{y})$ results for PGP-WG, plain beamforming, and no-precoding cases for the channel in $G_4$ in conjunction with QAM $M = 16,\ 64$ modulation.

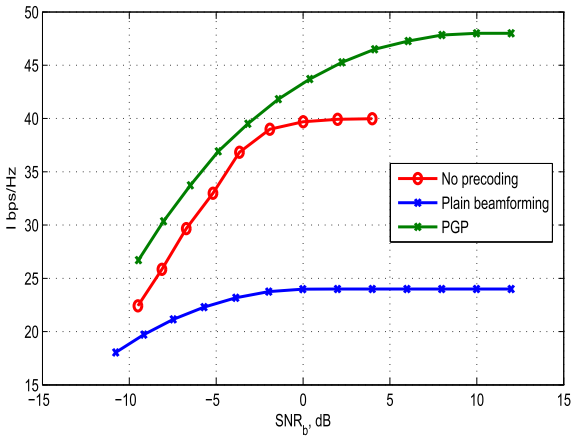



Fig. 4. $I(\mathbf{x};\mathbf{y})$ results for PGP-WG, plain beamforming, and no-precoding cases for the channel in $G_5$ in conjunction with QAM $M = 16$ modulation.

PGP-WG over the no precoding case. In addition, there is a precoding gain of $4-5\ dB$ over the low SNR regime. As far as complexity is concerned, based on the analysis of [15], the PGP-WG precoding example presented with $N_{t,v} = 6$ require a complexity (both at the transmitter and receiver) on the order of $3M^4$, while the no precoding example requires a complexity at the receiver on the order of $M^{18}$, thus PGP-WG needs a factor of $(1/3)M^{14}$ less complexity. For the $N_{t,v} = 8$ case the complexity reduction with PGP-WG over the no PGP-WG case becomes a factor of $(1/4)M^{14}$. Thus, we see that PGP-WG helps keep the UE complexity low, while it gives significant gains in throughput and SNR. In Fig. 4 we present results for $G_5$. Here, we observe high gains in throughput in high SNR regime. Here we employ $N_{t,v} = 6$. We observe that this is a Type-II channel behavior. At $\text{SNR}_b > 0$, the no precoding case throughput saturates at 40 $bps/Hz$. However, with PGP-WG we get significantly higher throughput, e.g., at $\text{SNR}_b = 10\ dB$ the throughput is 48 $bps/Hz$. Further, it takes PGP-WG a factor of $(1/6)M^{16}$ less UE complexity than the no precoding one in order to achieve this additional throughput at the UE.

For a UPA along the $z,\ x$ directions, with $N_{u,z} = N_{u,x} = 10$, $D = 0.6$, we get 8 groups with the following VCMBs:

$$\mathcal{S}_1 = [1, 2, 3, 10, 11, 12, 21, 31, 41, 71, 81, 91],$$
$$\mathcal{S}_2 = [3, 4, 11, 12, 13, 14, 15, 16, 17, 93],$$
$$\mathcal{S}_3 = [3, 4, 14, 94],$$
$$\mathcal{S}_4 = [4, 14, 24, 34, 44, 54, 64, 74, 83, 84, 85, 93, 94, 95],$$
$$\mathcal{S}_5 = [53, 63, 72, 73, 74, 83, 93],$$
$$\mathcal{S}_6 = [62, 72],$$
$$\mathcal{S}_7 = [1, 11, 21, 61, 71, 81, 91, 92, 93, 99, 100],\ \text{and}$$
$$\mathcal{S}_8 = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 21, 81, 91, 100].$$

The corresponding number of each group UE antennas is 4, 2, 4, 4, 6, 1, 6, and 8, respectively. We see that partial overlapping exists between different groups VCMBs. Without OFDM, we need to leave the common VCMBs unused to avoid primary interference between groups, i.e., employ NOV.

We thus end up with the following revised sets:

$$\mathcal{S}_{1,no} = [31,\ 41],$$
$$\mathcal{S}_{2,no} = [13,\ 15,\ 16,\ 17],$$
$$\mathcal{S}_{3,no} = [94],$$
$$\mathcal{S}_{4,no} = [64,\ 84,\ 85,\ 95],$$
$$\mathcal{S}_{5,no} = [53,\ 63,\ 73],$$
$$\mathcal{S}_{6,no} = [62],$$
$$\mathcal{S}_{7,no} = [61,\ 92,\ 93,\ 99],\ \text{and}$$
$$\mathcal{S}_{8,no} = \emptyset.$$

We see that due to NOV, some groups are presented with an extremely low number of effective VCMBs, thus their throughput is very low. For example, $G_8$ needs to neutralize its entire set of VCMBs due to NOV. This is very undesirable as $G_8$ cannot be offered any data service on the downlink. In the next subsections we show how this undesired effect can be mitigated by virtue of CFSDM-BG. It is illustrative to consider as per Section III the number of lost VCMBs per group due to NOV. The numbers are as follows: $\Delta_1 = 10$, $\Delta_2 = 6$, $\Delta_3 = 3$, $\Delta_4 = 10$, $\Delta_5 = 4$, $\Delta_6 = 1$, $\Delta_7 = 7$, and $\Delta_8 = 15$. As per our presented analysis in Section III, one needs to start deploying different sub-carriers to the sets of maximum VCMB loss, i.e., the highest impacted groups, in order to observe maximum gains with CFSDM-BG. In the current example this is $G_8$. Failing to apply this strategy leads to significant utilization loss in general, as our results in Fig. 9, 10 show below. In addition, we present the complete analysis of the optimized CFSDM-BG for this example in the next subsection.

In Fig. 5 we present results on the $G_1$ downlink precoding where we have applied PGP-WG with two additional "fictitious" inputs, similar to [15] and see dramatic improvements on downlink throughput. We see the dramatic impact of VCMB overlapping in the case of UPA. Notice that the complexity involved in the PGP-WG is two times higher than the one on the no precoding case, due to $N_{t,v} = 4$ "fictitious" antennas being introduced, while the incurred loss in $G_1$ due to the reduction on the number of useful VCMBs is highly mitigated. This example is a Type-II channel behavior in which

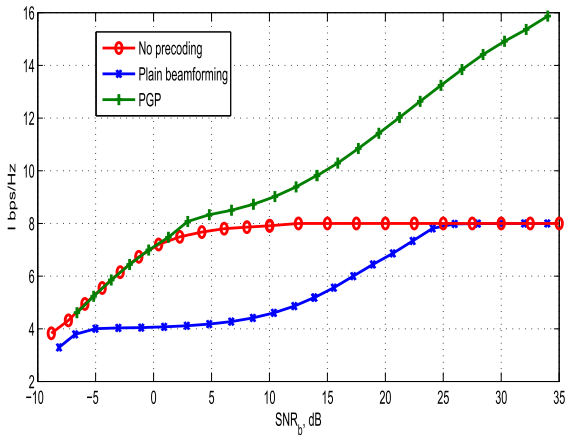



Fig. 5. $I(\mathbf{x};\mathbf{y})$ results for PGP-WG, plain beamforming, and no-precoding cases for the channel in $G_1$ in conjunction with QAM $M = 16$ modulation.

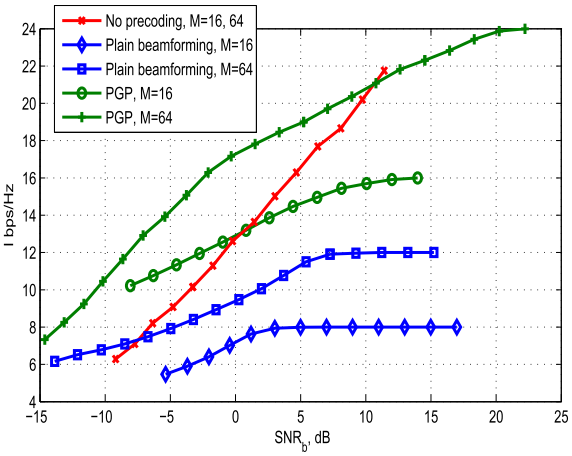



Fig. 6. $I(\mathbf{x};\mathbf{y})$ results for PGP-WG, plain beamforming, and no-precoding cases for the channel in $G_2$ in conjunction with QAM $M = 16,\ 64$ modulation.

PGP-WG achieves double the throughput in high SNR, while the corresponding UE complexity is two times higher than the no precoding one, since $N_{t,v} = 4 > N_t$. For the same system, in $G_2$ we get the results presented in Fig. 6. For the PGP-WG and plain beamforming cases we show results for both $M = 16,\ 64$. The PGP-WG and plain beamforming results use $N_{t,v} = N_t = 4$ "fictitious" antennas each, the same number as the no precoding case. We observe SNR and thoughput gains in low SNR. For example an SNR gain higher than 8 dB with PGP-WG in the SNR$_b$ over the no precoding case in low SNR, while the incurred UE receiver complexity with PGP-WG is a factor of $(1/2)M^4$ times lower than the no precoding case.

### C. Precoding Results With OFDM

We next present results with OFDM. We start with an example highlighting how $E_{CFSDM} \geq E_{NCFSDM}$ can be achieved. Let's employ the channels of the UPA example presented in Fig. 5, 6, but with CFSDM. In accordance with the optimized CFSDM-BG strategy presented in Section III, we select one set of $N_c$ sub-carriers for $G_8$, $G_6$, $G_5$, and another one for the rest of the groups. This way, from the groups presented in the UPA case above, one can see that $\mathcal{S}_8$ can employ all of its VCMBs and thus it can add a factor of

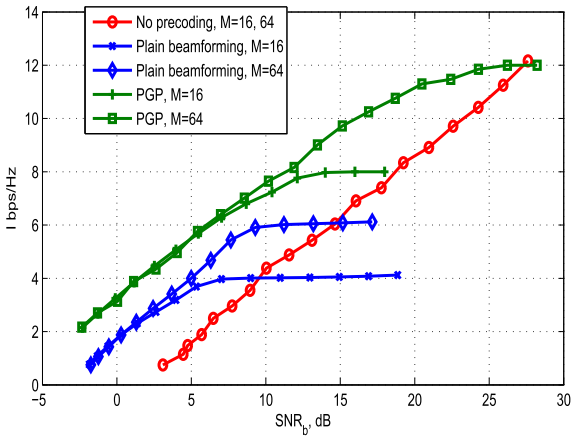



Fig. 7. $I(\mathbf{x};\mathbf{y})$ results for PGP-WG, plain beamforming, and no-precoding cases for the channel in group $G_1$, user 1 in conjunction with QAM $M = 16,\ 64$ modulation and CFSDM.

$15\log_2(M)$ to the numerator of $E_{CFSDM}$ in (36). Similarly, $G_5$, and $G_6$ can add a factor of $7\log_2(M)$, and $\log_2(M)$, respectively. Now as per equations (35), (36), we can see that $E_{CFSDM} = 21\log_2(M)$, while $E_{NCFSDM} = 19\log_2(M)$. Thus, in this case, $E_{CFSDM} > E_{NCFSDM}$. This is important because it shows that it is possible that CFSDM can achieve better cell efficiency than non-CFSDM for the entire cell. In addition, CFSDM offers much higher throughput to the selected $\mathcal{S}_{diff}$ groups as it is readily seen.

For a UPA deployed over the $x$, $y$ directions, with $N_{u,x} = N_{u,y} = 10$, an OFDM system with $Q = 64$ sub-carriers, we get 3 groups with the following VCMB's. $G_1$ has $\mathcal{S}_1 = [1,\ 2,\ 10,\ 11,\ 21,\ 81,\ 91]$, $G_2$ has $\mathcal{S}_2 = [11,\ 12,\ 13,\ 14,\ 15,\ 16,\ 17,\ 18,\ 19,\ 20]$, and $G_3$ has $\mathcal{S}_3 = [3,\ 4,\ 5,\ 12,\ 13,\ 14,\ 15,\ 24,\ 34,\ 44,\ 54,\ 64,\ 74,\ 84,\ 94]$. $G_1$ comprises 2 users with two antennas each, $G_2$ and $G_3$ comprise 2 users with 4 antennas each. There is VCMB overlapping between the groups, however by employing CFSDM we can avoid the interference coming from overlapping VCMBs. In addition, by employing different sub-carriers between the different users in each group in CFSDM, we can avoid joint decoding within the group level, i.e., the users decode their data totally independently. In this particular example we envisaged employing a total 6 OFDM sub-carriers, 2 per group for all 3 groups. In Fig. 7 and Fig. 8 we present results for user 1, user 2 of $G_1$, respectively. In both cases we see Type-I channel behavior. In this example, both users employ 2 receiving antennas. By virtue of CFSDM, the downlink can employ all VCMBs for both users, i.e., no need to partition the VCMB set. The example here applies 4 downlink pre-beamformers per user and in the PGP-WG results we use 2 groups of size $4 \times 4$ each, by extending the receiving antennas to 4, using 2 "fictitious" antennas, i.e., $N_{t,v} = 4$ in a fashion similar to [15]. Futhermore, a revised, improved version of plain beamforming is used in which only inputs with non-zero associated singular values are employed. We call this form of plain beamforming Singular Value Aware Plain Beamforming (SVAPB). We see very high throughput attained by PGP-WG over the no precoding in low SNR, and the plain beamforming case, over all shown SNR, respectively, although

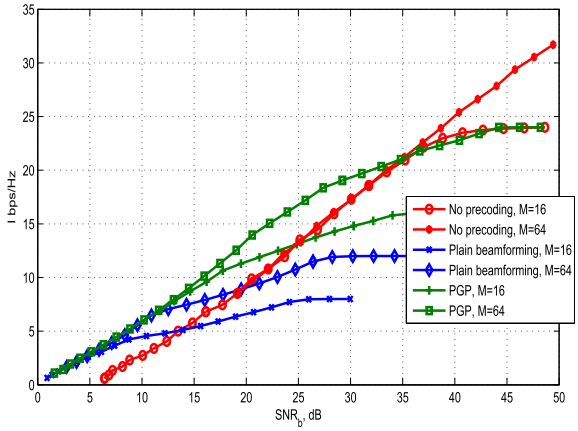



Fig. 8. $I(\mathbf{x};\mathbf{y})$ results for PGP-WG, plain beamforming, and no-precoding cases for the channel in $G_1$, user 2 in conjunction with QAM $M = 16,\ 64$ modulation and CFSDM.

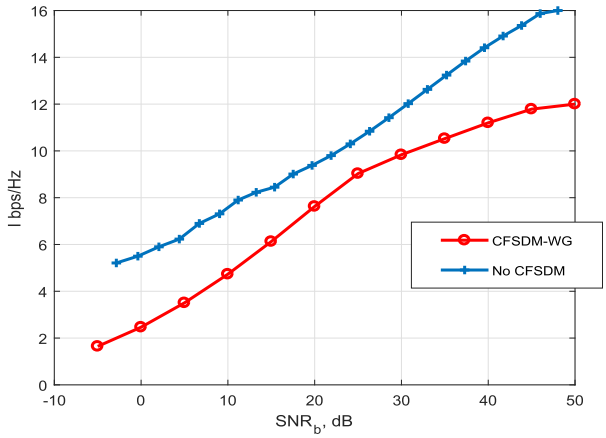



Fig. 9. $I(\mathbf{x};\mathbf{y})$ results for PGP-WG with and without CFSDM-WG for the channel in $G_1$, in conjunction with QAM $M = 16$ modulation.

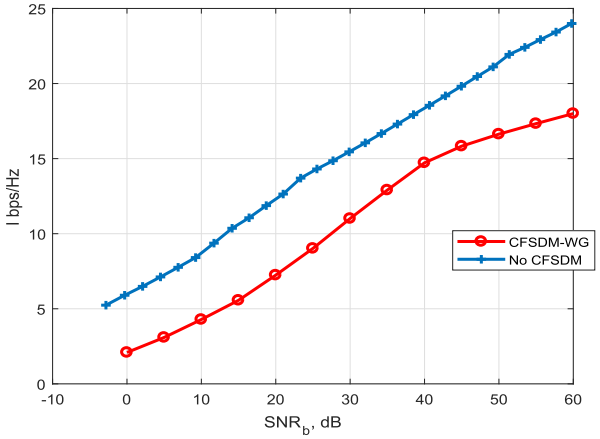



Fig. 10. $I(\mathbf{x};\mathbf{y})$ results for PGP-WG with and without CFSDM-WG for the channel in $G_1$, in conjunction with QAM $M = 64$ modulation.

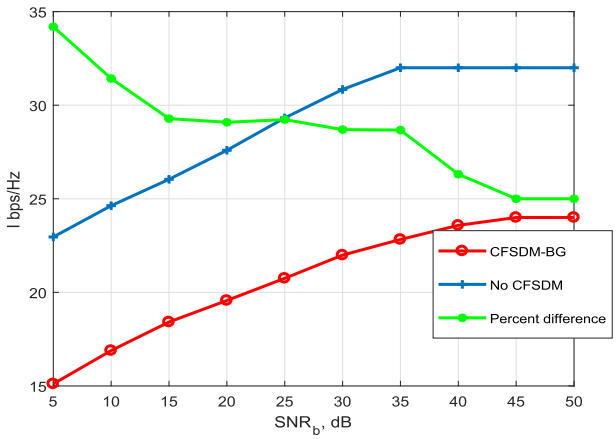



Fig. 11. $I(\mathbf{x};\mathbf{y})$ results for PGP-WG with and without CFSDM-BG for the groups $G_1$, $G_2$ from Fig. 5, 6, respectively, in conjunction with QAM $M = 16$ modulation.

the latter performs better than standard beamforming due to SVAPB. In Fig. 7 we show at $\text{SNR}_b = 5\ dB$ more than 3 times better throughput with PGP-WG than the no precoding case, while for a quite wide range of $\text{SNR}_b$ we see gains on the order of 8 $dB$ in SNR. The corresponding complexity with PGP-WG is a factor of a factor of $(1/2)M^{10}$ times lower than the no precoding one. In Fig. 8 we observe a gain in throughput of 33% at $\text{SNR}_b = 15\ dB$, while the SNR gain is on the order of 5 $dB$. The corresponding complexity with PGP-WG is same with the one in Fig. 7, i.e., $(1/2)M^{10}$ lower than the no precoding one.

An interesting comparison is presented in Fig. 9 and Fig. 10, for $M = 16,\ 64$, respectively. In these two figures a comparison with respect to the utilization between a system without CFSDM and the one with CFSDM-WG is performed. The situation presented is for users 1 and 2 in $G_1$ for the channels presented in Fig. 7, 8 above. We observe that although the system without CFSDM has generally better spectral efficiency due to employing a single carrier, the loss is limited. This is due to the fact that $4 \times 7$ channel employed in the non-CFSDM case only has 3 non-zero singular values of which only 2 have significant power. The resulting utilization loss at $\text{SNR}_b = 15\ dB$ is 25.61%, and 45.44%, for $M = 16$, and $M = 64$, respectively. We see that although CFSDM-WG is experiencing spectral efficiency loss, it still remains attractive for cases in which receiver coordination between different users in a group is not possible. In addition, for future systems in which the UEs have many antennas, CFSDM-WG may be able to perform closer to the non-CFSDM, thus increasing its attractiveness.

In Fig. 11, we show CFSDM-BG results for $G_1,\ G_2$ from the example used in Fig. 5, 6. It is important to stress that the application of CFSDM-BG in this scenario is rather inefficient as the selected group ($G_1$) for CFSDM application is not the one that lost the maximum number of VCMBs due to spatial overlapping, nor $G_1$ has a larger number of receiving antennas. By employing CFSDM-BG to $G_1$, we allow $G_1$ to employ all its VCMBs. We see that due to employing a non-optimal strategy in CFSDM deployment, employing 2 sub-carriers between these groups, there is spectral utilization loss. The spectral utilization loss by using two different sub-carriers between for $G_1,\ G_2$ ranges from 25% at high SNR to 33% at low SNR. Note that on top of the bad selection of groups for $\mathcal{S}_{f,diff}$ there is small number of receiving antennas in $G_1$, further reducing the performance of CFSDM. Thus, this scenario represents a lower bound on the performance of CFSDM-BG.

## V. Conclusions

In this paper, a novel methodology for Massive MIMO systems is presented, allowing for optimal downlink linear precoding with finite-alphabet inputs, e.g., QAM and multiple antennas per user. The methodology is based on a sparse VMC decomposition of the downlink channels, which then allows for orthogonality between different user groups, due to non-overlapping sets of VCMBs. The presented methodology extends JSDM to finite alphabet data symbols for a general family of channels. We show that JSDM-FA resorts to a DFT matrix type of prebeamformer. The methodology is applied in systems with or without OFDM and for ULA and UPA antenna configurations. By employing the PGP-WG technique to the proposed system, we show very high gains are available on the downlink. However, in the non-OFDM deployment, the users in each group need to co-ordinate their detection processes in order to achieve precoding gains. When OFDM is available, there is more flexibility in system design. For example, under CFSDM-BG users in a group can be assigned different sub-carriers, thus ameliorating the need for intra-group detection coordination, although at the cost of utilization loss. In addition, under CFSDM-BG, i.e., in cases of significant overlapping of the available VCMB sets, by carefully employing separate sub-carriers to VCMB overapping-impacted groups, the interfering groups can become completely orthogonal, and the impacted groups can attain much higher utilization, while the overall utilization gets higher. Thus, the novel combination of OFDM with the VCM JSDM system presented (CFSDM) offers many advantages, such as high throughput to users and it also obliterates the need for intragroup user decoding coordination. Thus, CFSDM prevails as quite an important element in JSDM-FA toward 5G applications. Our numerical results for both intragroup decoding coordination and CFSDM show high gains, e.g., typically higher than 60% and in some cases as high as 200% in throughput while the incurred precoding complexity is exponentially lower at both the transmitter and receiver sites.

## Acknowledgement

The authors would like to thank the anonymous reviewers for their comments which improved the quality of the paper.

## 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] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.

[3] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[4] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing—The large-scale array regime," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6441–6463, Oct. 2013.

[5] J. Nam, A. Adhikary, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing: Opportunistic beamforming, user grouping and simplified downlink scheduling," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 876–890, Oct. 2014.

[6] A. Adhikary *et al.*, "Joint spatial division and multiplexing for mm-wave channels," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1239–1255, Jun. 2014.

[7] C. Xiao, Y. R. Zheng, and Z. Ding, "Globally optimal linear precoders for finite alphabet signals over complex vector Gaussian channels," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3301–3314, Jul. 2011.

[8] M. Lamarca, "Linear precoding for mutual information maximization in MIMO systems," in *Proc. Int. Symp. Wireless Commun. Syst.*, Sep. 2009, pp. 26–30.

[9] F. Perez-Cruz, M. R. D. Rodrigues, and S. Verdu, "MIMO Gaussian channels with arbitrary inputs: Optimal precoding and power allocation," *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 1070–1084, Mar. 2010.

[10] M. A. Girnyk, M. Vehkapera, and L. K. Rasmussen, "Large-system analysis of correlated MIMO multiple access channels with arbitrary signaling in the presence of interference," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 2060–2073, Apr. 2014.

[11] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 502–513, Mar. 2000.

[12] W. Weichselberger, M. Herdin, H. Ozcelik, and E. Bonek, "A stochastic MIMO channel model with joint correlation of both link ends," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 90–100, Jan. 2006.

[13] Y. Wu, C.-K. Wen, C. Xiao, X. Gao, and R. Schober, "Linear precoding for the MIMO multiple access channel with finite alphabet inputs and statistical CSI," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 983–997, Feb. 2015.

[14] Y. Wu, D. W. K. Ng, C.-K. Wen, R. Schober, and A. Lozano, "Low-complexity MIMO precoding for finite-alphabet signals," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4571–4584, Jul. 2017.

[15] T. Ketseoglou and E. Ayanoglu, "Linear precoding gain for large MIMO configurations with QAM and reduced complexity," *IEEE Trans. Commun.*, vol. 64, no. 10, pp. 4196–4208, Oct. 2016.

[16] A. M. Sayeed, "Deconstructing multiantenna fading channels," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2563–2579, Oct. 2002.

[17] P. Schniter and A. Sayeed, "Channel estimation and precoder design for millimeter-wave communications: The sparse way," in *Proc. 48th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2014, pp. 273–277.

[18] J. Brady, N. Behdad, and A. M. Sayeed, "Beamspace MIMO for millimeter-wave communications: System architecture, modeling, analysis, and measurements," *IEEE Trans. Antennas Propag.*, vol. 61, no. 7, pp. 3814–3827, Jul. 2013.

[19] O. E. Ayach, R. W. Heath, Jr., S. Abu-Surra, S. Rajagopal, and Z. Pi, "Low complexity precoding for large millimeter wave MIMO systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2012, pp. 3724–3729.

[20] T. Ketseoglou and E. Ayanoglu, "Linear precoding for MIMO with LDPC coding and reduced complexity," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 2192–2204, Apr. 2015.

[21] T. Ketseoglou and E. Ayanoglu, "Linear precoding gain for large MIMO configurations with QAM and reduced complexity," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–7.

[22] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.

[23] J. Hoydis, C. Hoek, T. Wild, and S. ten Brink, "Channel measurements for large antenna arrays," in *Proc. Int. Symp. Wireless Commun. Syst. (ISWCS)*, Aug. 2012, pp. 811–815.

[24] X. Gao, O. Edfors, F. Rusek, and F. Tufvesson, "Linear pre-coding performance in measured very-large MIMO channels," in *Proc. IEEE Veh. Technol. Conf. (VTC Fall)*, Sep. 2011, pp. 1–5.

[25] H. Xie, F. Gao, S. Zhang, and S. Jin, "Spatial-temporal BEM and channel estimation strategy for massive MIMO time-varying systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–6.

[26] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[27] D. Ying, F. W. Vook, T. A. Thomas, D. J. Love, and A. Ghosh, "Kronecker product correlation model and limited feedback codebook design in a 3D channel model," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2014, pp. 5865–5870.

[28] A. Hjørugnes, *Complex-Valued Matrix Derivatives With Applications in Signal Processing and Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

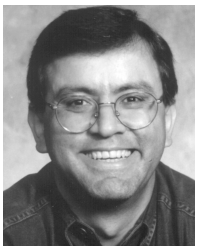
[29] J. Proakis, *Digital Communications*. New York, NY, USA: McGraw-Hill, 2001.

[30] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101–107, Jun. 2011.

[31] W. Zeng, C. Xiao, and J. Lu, "A low-complexity design of linear precoding for MIMO channels with finite-alphabet inputs," *IEEE Wireless Commun. Lett.*, vol. 1, no. 1, pp. 38–41, Feb. 2012.

[32] Y. Wu, C.-K. Wen, D. W. K. Ng, R. Schober, and A. Lozano, "Low-complexity MIMO precoding with discrete signals and statistical CSI," in *Proc. ICC*, May 2016, pp. 1–6.

**Thomas Ketseoglou** (S'85–M'91–SM'96) received the B.S. degree from the University of Patras, Patras, Greece, in 1982, the M.S. degree from the University of Maryland, College Park, MD, USA, in 1986, and the Ph.D. degree from the University of Southern California, Los Angeles, CA, USA, in 1990, all in electrical engineering. He was with the wireless communications industry, including senior level positions with Siemens, Ericsson, Rockwell, and Omnipoint. From 1996 to 1998, he participated in TIA TR45.5 (now 3GPP2) 3G standardization, making significant contributions to the cdma2000 standard. He has been an inventor and a co-inventor in several essential patents in wireless communications. Since 2003, he has been with the Electrical and Computer Engineering Department, California State Polytechnic University, Pomona, CA, USA, where he is currently a Professor. He spent his sabbatical leave in 2011 at the Digital Technology Center, University of Minnesota, Minneapolis, MN, USA, where he taught digital communications and performed research on network data and machine learning techniques. He is a part-time Lecturer at the University of California at Irvine, Irvine, CA, USA. His teaching and research interests are in wireless communications, signal processing, and machine learning, with current emphasis on MIMO, optimization, localization, and link prediction.



**Ender Ayanoglu** (S'82–M'85–SM'90–F'98) received the B.S. degree from Middle East Technical University, Ankara, Turkey, in 1980, and the M.S. and Ph.D. degrees from Stanford University, Stanford, CA, USA, in 1982 and 1986, respectively, all in electrical engineering. He was with the Communications Systems Research Laboratory, part of AT&T Bell Laboratories, Holmdel, NJ, USA, until 1996, and Bell Laboratories, Lucent Technologies, until 1999. From 1999 to 2002, he was a Systems Architect at Cisco Systems, Inc. Since 2002, he has been a Professor with the Department of Electrical Engineering and Computer Science, University of California at Irvine, Irvine, CA, USA, where he served as the Director of the Center for Pervasive Communications and Computing and held the Conexant-Broadcom Endowed Chair from 2002 to 2010. He was a recipient of the IEEE Communications Society Stephen O. Rice Prize Paper Award in 1995 and the IEEE Communications Society Best Tutorial Paper Award in 1997. From 1993 to 2014, he was an Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS and served as its Editor-in-Chief from 2004 to 2008. He is currently serving as a Senior Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS and as the founding Editor-in-Chief of the IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING. From 1990 to 2002, he served on the Executive Committee of the IEEE Communications Society Communication Theory Committee, and from 1999 to 2001, he was its Chair.