XXXVII SIMPÓSIO BRASILEIRO DE TELECOMUNICAÇÕES E PROCESSAMENTO DE SINAIS - SBrT2019, 29/09/2019–02/10/2019, PETRÓPOLIS, RJ

# Hybrid Beamforming Design for Massive MIMO Systems with Limited CSI Feedback

Francisco Hugo Costa Neto, Daniel Costa Araújo, and Tarcisio Ferreira Maciel

***Abstract*— In this study, we investigate a limited channel state information (CSI) feedback scheme to allow the implementation of hybrid beamforming based on joint spatial division and multiplexing (JSDM). We propose the usage of beam sweeping to provide channel measurements that are used to feed a CSI report scheme that allows an adequate estimation of the channel characteristics with short pilot sequences and very few precoder matrix indicators (PMIs), thus significantly reducing the required signaling. The hybrid beamforming scheme exploits successfully the provided CSI and achieves satisfactory total data rates in comparison to the full CSI report approach, as shown in our simulation results.**

***Keywords*— hybrid beamforming, massive MIMO**

## I. Introduction

Massive multiple input multiple output (MIMO) systems, by means of the large number of antenna elements, can provide high multiplexing and array gains. However, such advantages can be obtained if the precoding process is performed properly which imposes some technological challenges. The traditional fully digital precoding requires a large amount of radio frequency (RF) chains (one for each antenna element), which in this new context is impracticable due to hardware costs, energy consumption and implementation complexity [1].

Hybrid precoding is considered a promising solution to overcome these issues. It replaces the fully digital precoding at baseband by a combination of the analog beamforming at RF domain and digital precoding at baseband. The analog beamforming is implemented by applying an analog phase shifter network to control the signal phase on each antenna element. It employs power efficient phase shifters to reduce the utilization of RF chain an consequently the hardware cost and the implementation complexity [2]. Several hybrid precoding algorithms have been proposed in the literature, such as [3], [4]. However, most of these schemes assume the availability of the downlink (DL) CSI at the base station (BS). In fact, the determination of CSI between transmit and receive antennas is another challenge of massive MIMO systems. There is a signaling overhead for CSI acquisition at the BS, which is impracticable due to the limitation of spectral resources.

F. H. Costa Neto, e-mail: hugo@gtel.ufc.br; Daniel C. Araújo, e-mail: araujo@gtel.ufc.br; Tarcisio F. Maciel, e-mail: maciel@gtel.ufc.br, Wireless Telecommunications Research Group, Department of Teleinformatics Engineering, Center of Technology, Federal University of Ceará. This work was supported by Ericsson Research, Sweden, and Ericsson Innovation Center, Brazil, under UFC.44 Technical Cooperation Contract Ericsson/UFC. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. Francisco Hugo Costa Neto was supported by CAPES under the grant 88882.183549/2018-01. T. F. Maciel would like to acknowledge CNPq for its financial support under the grants 426385/2016-0 and 308621/2018-2.

The authors of [5] proposed a hybrid beamforming with small training and feedback overhead based on beam sweeping, i.e., a codebook based beam training. This method simplified the CSI acquisition process and reduced the training overhead, but still requires an exhaustive search of all candidate beams.

In this study, we develop a low-complexity hybrid beamforming algorithm for DL of a massive MIMO mmWave system and assumes the availability of only a limited feedback of CSI between the user equipments (UEs) and BS. The main contributions of the work can be summarized as (i) evaluation of a channel measurement technique based on beam sweeping aiming at reducing the training and feedback overhead, and (ii) assessment of a hybrid precoding algorithm for a mmWave system with limited CSI feedback signaling. Our work considers a BS that performs hybrid analog/digital precoding where the number of RF chains is limited by the number of elements of the space-division multiple access (SDMA) groups.

The remainder of this work is organized as follows. We present in Section II the assumptions about our system model. Section III presents the proposed channel measurement technique based on beam sweeping. Section IV discusses the proposed hybrid precoding scheme with limited CSI feedback. Performance results are discussed in Section V, and the main conclusions are drawn in Sections VI.

## II. System Model

We consider the DL of a massive multi-user (MU) multiple input multiple output (MU-MIMO) system employing orthogonal frequency division multiple access (OFDMA). The system is composed of one BS and $J$ UEs. The BS is equipped with a uniform planar array (UPA) that is composed of $M_\text{v}$ vertical and $M_\text{h}$ horizontal antenna elements, where the total number of antennas $M = M_\text{v} M_\text{h}$, and has a total of $R$ RF chains, so that $R \leq M$, available. We consider that the UEs are equipped with a single omnidirectional antenna. In each time slot, $K$ UEs out of the entire set of $J$ UEs are selected to compose a SDMA group $\mathcal{G}$ and transmit sharing a the same frequency-time resource block (RB) in space. In each resource, the transmitter uses its $M$ antennas to send a single data stream to each of the selected UEs. We consider a frequency division duplex (FDD) system, where the BS selects UEs to transmit and determine proper precoding matrices to mitigate the MU interference.

Prior to a transmission to the $k$th selected UE, for a given RB and transmission time interval (TTI), the symbol $x_k$ is filtered by the precoding vector $\mathbf{f}_k \in \mathbb{C}^{M\times 1}$. Input symbols $x_k$

are assumed to have unitary average power allocation among all UEs, i.e., their power is considered normalized so that $\mathbb{E}\{x_k x_k^H\} = 1, \forall k$.

The DL channel matrix between the BS and the receive antenna of the $k$th selected UE is $\mathbf{H}_k \in \mathbb{C}^{M_v \times M_h}$, which can be written as $\mathbf{h}_k = \mathbf{vec}(\mathbf{H}_{k,n}) \in \mathbb{C}^{M_v M_h \times 1}$, which represents the 3D channel response. The channel coefficient of a given RB is associated with the middle subcarrier and the first transmitted orthogonal frequency division multiplexing (OFDM) symbol in a TTI. We consider that the channel coefficients remain constant during resource allocation in a TTI.

We consider a two-stage precoder composed of an analog precoding matrix $\mathbf{F}_{\text{RF},k} \in \mathbb{C}^{M \times R}$ and a digital precoding vector $\mathbf{f}_{\text{BB},k} \in \mathbb{C}^{R \times 1}$. The analog beamforming matrix is assumed to have elements of equal magnitude, i.e., we consider its implementation via analog phase shifters. Thus, the precoding vector can be written as $\mathbf{f}_k = \mathbf{F}_{\text{RF},k}\mathbf{f}_{\text{BB},k}$. The filtered symbol is then transmitted through the channel associated with the RB.

Therefore, the prior-filtering receive symbol $y_k$ at the $k$th selected UE is

$$y_k = \mathbf{h}_k^T \mathbf{f}_k \sqrt{p_k}\, x_k + \sum_{i \neq k}^{K} \mathbf{h}_k^T \mathbf{f}_i \sqrt{p_i}\, x_i + z_k, \tag{1}$$

where $p_k$ is the transmit power allocated to the stream associated to the $k$th UE; the second term on the right-hand side of (1) represents the MU interference; $z_k$ is the additive white gaussian noise (AWGN), statistically distributed as $\mathcal{CN}(0, \sigma_{\mathbf{z}}^2)$, with standard deviation $\sigma_z$.

The total power constraint is enforced by normalizing the digital and analog filters, such that $\| \mathbf{F}_{\text{RF},k}\mathbf{f}_{\text{BB},k}\sqrt{p_k} \|_F^2 = \frac{P\mu_{\text{active}}}{\mu_{\text{total}}}$, where $P$ is the total transmit power of the BS, $\mu_{\text{active}}$ is the number of active RB and $\mu_{\text{total}}$ is the total number of RBs of the system.

## III. Channel Measurement

In order to determine the precoders appropriately, the BS must dispose of CSI available. In our work, we consider a limited CSI feedback scheme based on channel quality indicator (CQI) and PMI reports [6]. The motivation behind this approach is the signaling overhead that arises with large number of antennas in FDD systems. We deal with overhead by using a set beams, each one associated to a PMI, that transmits channel state information - reference signal (CSI-RS) for measurement purposes at the UE. More specifically, each UE estimates the CSI based on the received CSI-RS and determines the most appropriate set of precoding vectors from a given codebook. We assume that the BS has a codebook of $M$ possible beamforming vectors, whose patterns have non-overlapping main lobes that cover a sector of a single cell. In this study, the codebook is modeled as a 2D discrete Fourier transform (DFT) matrix $\mathbf{W} \in \mathbb{C}^{M \times M}$.

The vector $\mathbf{s} \in \mathbb{C}^{1 \times L}$ defines the CSI-RS pilot sequences whose length is $L$. Its transmission is repeated over $M$ TTIs, and every time slot is associated to a given beam direction. This process is called beam sweeping. The received signal at the $k$th UE of the $b$th beam during the beam sweeping can be written as

$$\mathbf{r}_{k,b} = \underbrace{\mathbf{h}_k^T \mathbf{w}_b^*}_{\alpha_{k,b}} \mathbf{s} + \mathbf{n}_k, \tag{2}$$

where $\alpha_{k,b}$ is the projection of the channel vector $\mathbf{h}_k$ onto the conjugate codebook column $\mathbf{w}_b^*$ and $\mathbf{n}_k \in \mathbb{C}^{1 \times L}$ is an AWGN vector distributed as $\mathcal{CN}(0, \sigma_{\mathbf{n}}^2 \mathbf{I})$. The UE is interested in measuring $\alpha_{k,b}, \forall\, b = \{1, 2, \ldots, M\}$, because it gives the UE the knowledge to decide which beams are the most appropriate.

Given the knowledge of CSI-RS, the estimated channel vector projection $\alpha_{k,b}$ of the $k$th UE onto the $b$th beam $\mathbf{w}_b$ can be written as

$$\hat{\alpha}_{k,b} = \mathbf{r}_{k,b}\mathbf{s}^H. \tag{3}$$

The performance of the beam sweeping channel measurement is evaluated in terms of the normalized mean square error (NMSE), which is defined as

$$\text{NMSE}(\alpha) = \frac{1}{C} \sum_{c=1}^{C} \frac{\| \boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}} \|_2^2}{\| \hat{\boldsymbol{\alpha}} \|_2^2}, \tag{4}$$

where $C$ is the number of Monte Carlo runs.

The $k$th UE measures the projection of its channel $\mathbf{h}_k$ on the beam patterns configured by the DFT-based codebook $\mathbf{W}$ and, among those beam patterns, recommends a suitable set for subsequent data transmission. In the proposed scheme, each UE determines the set of beams with highest weights $\hat{\alpha}_{k,b}$, since this is the set of beams that provide more information about the estimated instantaneous CSI of the UE. Afterwards, the set of selected beams $\mathcal{B}$ is fed back to the BS as a vector of PMIs. Furthermore, the $k$th UE also sends the corresponding set $\mathcal{A}$ of weights $\alpha_{k,b}$ associated to the beams' indices in $\mathcal{B}$ also as a vector.

## IV. Hybrid Beamforming with Limited CSI Feedback

We consider a hybrid beamforming scheme based on joint spatial division and multiplexing (JSDM) [7], on which the BS splits all UEs of the system into subsets accordingly to their channel compatibility. Therefore, each subset $C_i$, hereinafter referred to as cluster, has elements with similar DL channel characteristics. In the following, the BS schedules UEs from different clusters in order to minimize MU interference. The number of clusters is determined by the number of RF chains available in the BS antenna.

The first step of the hybrid beamforming based on JSDM consists in the partioning of UEs into clusters. The BS will receive a CSI feedback from each UE and, based on that information, establish the different clusters. The authors in [7], [8] consider a CSI feedback scheme that reports to the BS the channel covariance matrix $\mathbf{R}_k$ of each UE, i.e., a high-dimensional channel second order statistic. Given the $j$th UE, its channel vector on the $t$th TTI is $\mathbf{h}_{t,j} \in \mathbb{C}^{M \times 1}$. Thus, the transmit covariance matrix $\mathbf{R}_j \in \mathbb{C}^{M \times M}$ can be written as

$$\mathbf{R}_j = \frac{1}{T} \sum_{t=1}^{T} \mathbf{h}_{t,j}\mathbf{h}_{t,j}^H, \tag{5}$$

where $T$ is the TTI window size, i.e., indicates the number of channel vector $\mathbf{h}_{t,j}$ samples considered in the averaging process.

The amount of information associated to the CSI feedback is determined by the number of antenna elements at the BS $M$ and by the number of UEs in the system, being proportional to $M^2J$ for reporting the whole $\mathbf{R}_j$ matrices. If we consider a scenario with a massive arrangement of antennas at the BS and a high number of UEs, this feedback scheme implies at a huge signaling overhead.

To provide a JSDM implementation with reduced CSI signaling overhead, we consider herein to report to the BS the two vectors associated with $\mathcal{A}$ and $\mathcal{B}$ described in the previous section. The combination of the information contained in these vectors with the predefined codebook $\mathbf{W}$ provides the means to compute an estimation of the UEs' channels. Thus, the first vector derives from the set of PMIs $\mathcal{B}_j$ of each UE $j$. Each PMI represents a selected column of the codebook $\mathbf{W} = \begin{bmatrix}\mathbf{w}_1 & \cdots & \mathbf{w}_M\end{bmatrix}$. Therefore, the vector of PMIs can be written as $\mathbf{b}_j = \begin{bmatrix}\beta_{j,1} & \cdots & \beta_{j,N}\end{bmatrix} \in \mathbb{C}^{N\times 1}$, where $\beta_{j,w} \in \{1, \ldots, M\}$. The second vector derives from the set of weights $\mathcal{A}_j$ of each UE $j$, so that the vector $\mathbf{a}_j = \begin{bmatrix}\hat{\alpha}_{j,1} & \cdots & \hat{\alpha}_{j,N}\end{bmatrix} \in \mathbb{C}^{N\times 1}$ can be defined. This vector indicates the weight associated to the projection of each selected column of the codebook on the instantaneous channel $\mathbf{h}_k$.

In this case, the signaling amount is determined by the number of reported PMIs $N$, rather than the number of antennas elements, and by the number of UEs. Therefore, the proposed CSI feedback scheme provides a signaling load proportional to $2NJ$. Since $N \ll M$, the proposed approach provides a CSI signaling load lower than the conventional second order channel statistic feedback. However, it can be less robust as the device does not estimate the full channel or second order channel statistics.

Different UE partitioning methods have been proposed as extension to JSDM, such as k-means [8] and fixed quantization [4]. In our study, we consider the partitioning of UEs based on a strategy initially proposed in [8]. Our approach considers the use of the k-means++ algorithm. Proposed in [9], it replaces the poor initialization step of k-means with a more sophisticated one.

The k-means++ algorithm aims to cluster UEs into subsets so that each UE belongs to the cluster with the nearest central characteristic. Thereby, UEs with similar transmit channel characteristics are put on the same cluster.

Conventionally, the average channel covariance eigenspace is the channel characteristic considered in the UE partition strategies. The eigendecomposition of the transmit covariance matrix $\mathbf{R}_j$ can be written as

$$\mathbf{R}_j = \mathbf{E}_j \Delta_j \mathbf{E}_j^{-1}, \tag{6}$$

where $\mathbf{E}_j \in \mathbb{C}^{M\times M}$ defines the matrix composed of eigenvectors and $\Delta_j \in \mathbb{C}^{M\times M}$ is the diagonal matrix of eigenvalues.

Therefore, we define the characteristic associated to each UE in the clustering algorithm as $\tilde{\mathbf{h}}_j \in \mathbb{C}^{M\times 1}$ , the dominant eigenvector associated to highest eigenvalue of $\mathbf{R}_j$.

In our work, we consider that the information used by the BS to partition UEs into clusters, defined as $\hat{\mathbf{h}}_k \in \mathbb{C}^{M\times 1}$, is the combination of the projection of the instantaneous channel vector on the codebook, which can be written as

$$\hat{\mathbf{h}}_k = \sum_{n=1}^{N} \hat{\alpha}_n \mathbf{w}_{\beta_n}. \tag{7}$$

The k-means++ algorithm employs an iterative approach that aims to find a clustering that minimizes the distance among elements that belongs to each cluster to the central characteristic of that cluster, also known as centroid. Each UE has a characteristic $\mathbf{d}_j \in \mathbb{C}^{M\times 1}$, which can be associated to $\tilde{\mathbf{h}}_k$ or $\hat{\mathbf{h}}_k$, namely the dominant eigenvector of the channel covariance matrix or the sum of columns of the codebook that best represent the instantaneous channel vector, respectively.

The vector $\mathbf{c}_i \in \mathbb{C}^{M\times 1}$ that describes the central characteristic of the $i$th cluster $C_i$ is called centroid, which can be written as

$$\mathbf{c}_i = \frac{1}{|C_i|} \sum_{j\in C_i} \mathbf{d}_j. \tag{8}$$

The first step is the initialization of centroids. The k-means algorithm determines the initial centroids from the characteristic $d_j$ from $K$ UEs randomly sorted out of the entire set of UEs in the system. This process is simple, but may cause convergence of the centroids to the local optima. To overcome this issue, the k-means++ considers a more refined initialization of centroids. In the beginning, an arbitrary first centroid is randomly chosen from the entire set of UEs in the system. The selection of the others initial centroids is established by a weighted probability function $\psi_j$ based on the distance of the UEs to the closest centroid $\mathbf{c}_i$, which can be written as

$$\psi_j = \frac{\| \mathbf{d}_j - \mathbf{c}_i \|_2^2}{\sum\limits_{j'\neq j} \| \mathbf{d}_{j'} - \mathbf{c}_i \|_2^2}. \tag{9}$$

Then, the UE with highest $\psi$ is selected as centroid, which means that we choose UEs which are furthest away from the already selected centroids. After the initialization step, the algorithm follows the same steps of the k-means. Each iteration of the algorithm consists in a clustering assignment followed by a centroid update. Given the clusters centroids $\mathbf{c}_i$ provided in the first iteration ($t = 0$), in the group assignment step, each UE $j$ is assigned to the cluster $C_i$ with smallest distance to the centroid as

$$C_i = \min_i \| \mathbf{d}_j - \mathbf{c}_i \|_2^2 . \tag{10}$$

We define a threshold $\epsilon > 0$ and test at every iteration $t$ if there is no significant change of the centroids in comparison to the previous iteration, i.e.,

$$\sum_{i=1}^{K} \| \mathbf{c}_i^{(t)} - \mathbf{c}_i^{(t-1)} \|_2^2 \leq \epsilon. \tag{11}$$

In the centroid update step, new centroids $\mathbf{c}_i$ are computed for each cluster from the subset of UEs in $C_i$ using (8). The assignment and centroid update steps are carried out until we reach a convergence. The output of the algorithm is a clustering of the UEs into $K$ disjoint clusters and a set of vectors $\{\mathbf{c}_1, \cdots, \mathbf{c}_K\}$ obtained as the centroids of the clusters.

The analog precoder is assumed to have elements of equal magnitude, i.e., only phase shifting is performed in the analog domain. Therefore, we only consider the phases of the elements of the centroids. Considering this assumption, the analog precoder $\mathbf{F}_{\text{RF}} \in \mathbb{C}^{M \times K}$ is defined from the centroids of the clusters and can be written as

$$\mathbf{F}_{\text{RF}} = [\mathbf{c}_1 \quad \mathbf{c}_2 \quad \ldots \quad \mathbf{c}_K]. \tag{12}$$

In the following, given the clustering of UEs, we select one UE $j$ from each cluster $C_i$ to create an SDMA group $\mathcal{G}$ containing $K$ UEs. In our study, we consider the weight vector $\mathbf{a}$ associated to each UE and selects the $j$th UE from each cluster $C_i$ that presents the highest sum of weights. Therefore, this approach schedules the UEs which presents the highest estimated instantaneous channel gain. This criterion to choose the UE $j$ from cluster $C_i$ can be written as

$$g_i = \max_{j \in C_i} \sum_{n=1}^{N} \hat{\alpha}_{j,n}. \tag{13}$$

We define the group channel matrix $\mathbf{H}_{\mathcal{G}} \in \mathbb{C}^{K \times M}$ based on the channel vectors of the scheduled UE that compose the SDMA group. It can be written as

$$\mathbf{H}_{\mathcal{G}} = \left[\mathbf{h}_1 \quad \mathbf{h}_2 \quad \ldots \quad \mathbf{h}_K\right]^{\text{T}}. \tag{14}$$

Therefore, given the group channel matrix $\mathbf{H}_{\mathcal{G}}$ and the analog precoder $\mathbf{F}_{\text{RF}}$, the equivalent channel matrix $\mathbf{H}_{\text{eq}} \in \mathbb{C}^{K \times K}$ is given by

$$\mathbf{H}_{\text{eq}} = \mathbf{H}_{\mathcal{G}} \mathbf{F}_{\text{RF}}. \tag{15}$$

The digital precoding matrix $\mathbf{F}_{\text{BB}} \in \mathbb{C}^{K \times K}$ for the $k$th UE is defined as the zero-forcing (ZF). The ZF precoding is conceived to decorrelate the transmit signals so that the signal at every receiver output is free of interference. The precoding matrix $\mathbf{F}_{\text{BB}}$ is defined as

$$\mathbf{F}_{\text{BB}} = \frac{\mathbf{H}_{\text{eq}}^H (\mathbf{H}_{\text{eq}} \mathbf{H}_{\text{eq}}^H)^{-1}}{\| \mathbf{H}_{\text{eq}}^H (\mathbf{H}_{\text{eq}} \mathbf{H}_{\text{eq}}^H)^{-1} \|_F}. \tag{16}$$

The total power constraint is enforced by normalizing the digital and analog filters, such that $\|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \sqrt{\mathbf{P}_{\mathcal{G}}}\|_F^2 = p_{\text{RB}}$, where $\mathbf{P}_{\mathcal{G}} \in \mathbb{R}^{K \times K}$ is the block diagonal power matrix resulting of the combination of the power matrices of each UE belonging to the SDMA group and $p_{\text{RB}}$ is the power available for a given RB. We consider that the number of clusters is equal to the number of RF chains and the total number of streams, i.e., $K$. Therefore, the dimensions of $\mathbf{F}_{\text{RF}}$ and $\mathbf{F}_{\text{BB}}$ are compatible with the dimension of $\mathbf{F}$, so that $\mathbf{F} = \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \in \mathbb{C}^{M \times K}$.

The receive vector of the group $\hat{\mathbf{y}}_{\mathcal{G}} \in \mathbb{C}^{K \times 1}$ is given by

$$\hat{\mathbf{y}}_{\mathcal{G}} = \mathbf{H}_{\mathcal{G}} \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \sqrt{\mathbf{P}_{\mathcal{G}}}\, \mathbf{x}_{\mathcal{G}} + \mathbf{z}_{\mathcal{G}}, \tag{17}$$

where $\mathbf{x}_{\mathcal{G}} \in \mathbb{C}^{K \times 1}$ is the group symbol vector and the $\mathbf{z}_{\mathcal{G}} \in \mathbb{C}^{K \times 1}$ is the group noise vector.

Defining $\mathbf{Q} = \mathbf{H}_{\mathcal{G}} \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \sqrt{\mathbf{P}_{\mathcal{G}}} \in \mathbb{C}^{K \times K}$, the average signal to interference-plus-noise ratio (SINR) perceived by the stream $i$ can be calculated as

$$\Gamma_i = \frac{|\mathbf{Q}_{ii}|^2}{\sum_{j \neq i}^{K} |\mathbf{Q}_{ij}|^2 + \sigma_z^2}, \tag{18}$$

where $\sigma_z^2$ is the noise power and $\mathbf{Q}_{ij}$ is the element of $\mathbf{Q}$ at row $i$ and column $j$.

The signal to noise ratio (SNR) is defined as

$$\Upsilon_i = \frac{|\mathbf{Q}_{ii}|^2}{\sigma_z^2}, \tag{19}$$

The data rate of stream $i$ is calculated according to Shannon capacity formula [10]. Defining the bandwidth of the RB as $B$, the data rate is defined as

$$R_i = B \log_2(1 + \Gamma_i). \tag{20}$$

## V. Results

We consider a single cell system with a carrier frequency of 28 GHz and a bandwidth of 100 MHz. According to [11, Table 2], these parameters imply a set of 125 RBs, each one composed of 12 subcarriers equally spaced of 60 kHz. Furthermore, the number of subframes per frame is 10, each subframe has 14 symbols and the TTI duration is 0.25 ms. We consider the 3D quasi deterministic radio channel generator (QuaDRiGa) channel model from [12] (version 2.0.0-664) and references therein. We assume a set of 60 UEs randomly distributed inside 4 circular regions, which have the same number of elements and are characterized by a radius of 10% of the intersite distance and away from the BS by 85% of the intersite distance. We also assume that these UE hotspots, i.e., a concentration of UEs in a relatively small area of the cell compared to the surrounding, are uniformly distributed inside a sector of 60° of the cell. The details of simulations' settings are listed in Table I.

Figure 1 shows the performance of the beam sweeping channel measurement for different pilot sequence lengths $L = \{8, 16, 32, 64\}$. The increase of the pilot sequence length provides to the UEs more knowledge about the channel characteristics. Consequently, there is a better estimation of channel vector projection onto the codebook entries. However, this improvement is limited. Above the length $L = 32$, the signaling overhead increase is not worth the observed NMSE reduction.

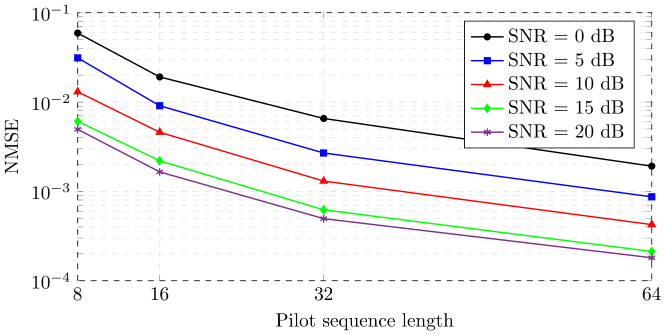


Fig. 1. Performance of the beam sweeping channel measurement in terms of the NMSE for different pilot sequence lengths.

TABLE I
SIMULATION PARAMETERS

| **Parameter** | **Value** |
|---|---|
| Channel model | QuaDRiGa [12] |
| Scenario | UMi-LOS [13] |
| Center Frequency | 28 GHz |
| Bandwidth | 100 MHz |
| Number of RBs | 125 |
| Number of subcarriers/RB | 12 |
| Subcarrier spacing | 60 KHz |
| TTI duration | 0.25 ms |
| Number of symbols/subframe | 14 |
| Simulation time | 65 ms |
| Number of simulation rounds | 250 |
| Intersite distance | 200 m |
| BS, UE heigth | 10, 1.5 m |
| BS, UE element spacing | $\lambda/2$ m |
| BS antenna model | 3GPP-mmWave [13] |
| BS vertical antennas | 8 |
| BS horizontal antennas | 8 |
| UE antennas | 1 |
| UE antenna model | omni |
| Total transmit power | 35 dBm |
| Noise figure | 9 dB |
| Noise spectral density | -174 dBm/Hz |
| Number of paths | 7 |
| Shadowing standard deviation | 3.1 dB |
| UE track | linear |
| UEs speed | 3 km/h |

In Figure 2, we show the cumulative distribution function (CDF) of the total data rate for our system considering the proposed solution of reduced feedback (Prop) compared to the conventional signaling scheme (JSDM). First, we partition UEs into $K = 4$ clusters using the k-means++ algorithm. The scheduling algorithm selects the UE from each cluster which presents the highest sum of estimated weights ($\sum_{j \in C_i} \hat{\alpha}_j$) to compose the SDMA group. The scheduling algorithm is also influenced by the limited CSI. There is an increase in total data rate when the number of PMIs changes from $N = 1$ to $N = 2$. The system total data rate is enhanced almost 30% as a consequence of a greater amount of CSI that is provided to the BS and the clustering algorithm is capable of partitioning the UEs according to their channel compatibilities more effectively and scheduling of the UEs with better estimated channel conditions. However, we observe only a marginal capacity growth when the number of PMIs is increased from $N = 2$ to

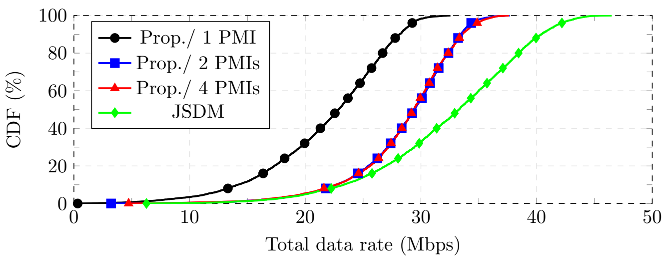


Fig. 2. Evaluation of proposed signaling scheme with different number of reported PMIs.

$N = 4$. Hence, the CSI provided by $N = 4$ PMIs is redundant in comparison with the knowledge supplied by the report of $N = 2$ PMIs in our scenario.

We highlight, that the limited feedback of CSI reduces significantly the signaling overhead and achieves levels of the total sum rate near to the conventional signaling approach. The proposed CSI feedback scheme achieves 90% of 50th percentile total data rate observed in the conventional CSI feedback case. The JSDM feedback reports the entire channel covariance matrix $\mathbf{R}_j \in \mathbb{C}^{64\times 64}$, a signaling amount almost a thousand times higher than the report CSI using the vectors of PMIs and estimated weights ($\mathcal{B}_k, \mathcal{A}_k \in \mathbb{C}^{2\times 1}$).

## VI. CONCLUSIONS

In this study, we investigated a limited CSI feedback scheme to allow the implementation of a hybrid beamforming based on JSDM. We propose a beam sweeping channel measurement to provide CSI report scheme that allows an adequate estimation of the channel characteristics with short pilot sequences and very few PMIs. The reduction of the signaling overhead is significant and the user selection strategies could exploit successfully the CSI provided and achieved 90% of the total data rate observed at the full CSI report approach.

## REFERENCES


[1] J. Hao and H. Liu, "Fast Spatial Correlation Acquisition for Hybrid Precoding Using Sequential Compressive Sensing," in *IEEE Annual Consumer Communications Networking Conference*, Jan. 2018, pp. –.

[2] A. F. Molisch, V. V. Ratnam, S. Han, Z. Li, S. L. H. Nguyen, L. Li, and K. Haneda, "Hybrid Beamforming for Massive MIMO: A Survey," *IEEE Communications Magazine*, vol. 55, no. 9, pp. 134–141, Sep. 2017.

[3] L. Liang, W. Xu, and X. Dong, "Low-Complexity Hybrid Precoding in Massive Multiuser MIMO Systems," *IEEE Wireless Communications Letters*, vol. 3, no. 6, pp. 653–656, Dec. 2014.

[4] Y. Xu, G. Yue, N. Prasad, S. Rangarajan, and S. Mao, "User Grouping and Scheduling for Large Scale MIMO Systems With Two-Stage Precoding," in *IEEE International Conference on Communications*, Jun. 2014, pp. 5197–5202.

[5] A. Alkhateeb, G. Leus, and R. W. Heath, "Limited Feedback Hybrid Precoding for Multi-User Millimeter Wave Systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 11, pp. 6481–6494, Nov. 2015.

[6] 3GPP, "Multiplexing and channel coding," Third Generation Partnership Project, Tech. Rep. 36.212 v13.3.0 Release 13, 3 2017.

[7] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint Spatial Division and Multiplexing – The Large-Scale Array Regime," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6441–6463, Oct. 2013.

[8] J. Nam, A. Adhikary, J.-Y. Ahn, and G. Caire, "Joint Spatial Division and Multiplexing: Opportunistic Beamforming, User Grouping and Simplified Downlink Scheduling," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 876–890, Oct. 2014.

[9] D. Arthur and S. Vassilvitskii, "k-means++: The Advantages of Careful Seeding," in *ACM-SIAM Symposium on Discrete Algorithms*, 2007, pp. 1027–1035.

[10] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed., J. W. . Sons, Ed., 2006.

[11] A. A. Zaidi, R. Baldemair, H. Tullberg, H. Bjorkegren, L. Sundstrom, J. Medbo, C. Kilinc, and I. D. Silva, "Waveform and Numerology to Support 5G Services and Requirements," *IEEE Communications Magazine*, vol. 54, no. 11, pp. 90–98, Nov. 2016.

[12] S. Jaeckel, L. Raschkowski, K. Borner, and L. Thiele, "QuaDRiGa: A 3-D Multi-Cell Channel Model With Time Evolution for Enabling Virtual Field Trials," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 6, pp. 3242–3256, Jun. 2014.

[13] 3GPP, "Study on channel model for frequencies from 0.5 to 100 GHz," 3rd Generation Partnership Project (3GPP), TR 38.901, 2017.