Throughput vs. Latency: QoS-centric Resource Allocation for Multi-User Millimeter Wave Systems

Miltiades C. Filippou[1], Danilo De Donno[2], Camila Priale[1], Joan Palacios[2, 3], Domenico Giustiniano[2], and Joerg Widmer[2]

[1]Intel Germany GmbH, Neubiberg, Germany, [2]IMDEA Networks Institute, Madrid, Spain, [3]Universidad Carlos III de Madrid, Spain
E-mail: {firstname.lastname}@{intel.com, imdea.org}

***Abstract*—Millimeter wave (mm-wave) communication is a topic of intensive recent study, as it allows to significantly boost data rates of future 5G networks. In this paper, we focus on a mm-wave system consisting of a single Access Point (AP) and two User Equipments (UEs), where one UE requires high throughput, while the other is characterized by a low latency demand. Given that setup, we aim at optimally allocating the available AP hardware resources for the beam training phase and data communication, in order to efficiently serve both UEs via hybrid analog-digital beamforming. We evaluate an optimization framework with the objective to maximize the expected rate of one UE, for a given latency constraint set by the other UE. The optimal data rates are illustrated for different latency constraints and for different strategies of exploiting the full RF chain set at the AP side. We observe that our proposed access schemes outperform the basic TDMA approach by up to 22 %.**

***Index Terms*—Millimeter wave, QoS, resource allocation, beam-training, hybrid beamforming.**

I. Introduction

In recent years, the evolution of new wireless applications, such as media-on-demand, smart offices, tactile internet and others, has inevitably led towards an exponentially growing need for high throughput and low latency, among other requirements [1], [2]. Hence, with the aim of achieving multi-gigabit-per-second data rates, millimeter wave (mm-wave) communications have been recently proposed as a promising solution [3]–[5].

However, the advantages of mm-wave systems come at the cost of certain drawbacks stemming from the particular propagation characteristics of the mm-wave channel. High attenuation and shadowing losses appear at these frequencies, which only make the mm-wave air interface conditionally attractive for short range deployments. The condition is that, since a large number of antenna elements can be used in a small form factor, extremely directional transmissions can take place, hence making up for the mentioned losses. Nevertheless, the existence of highly directional beamformed transmissions/receptions can lead to the occurrence of beam misalignment events, which degrade the Quality-of-Service (QoS) at the User Equipment (UE) side, when the latter is expressed by means of throughput and by means of latency.

Such a trade-off between maximizing the received power at a UE via a directional transmission, on one hand, and minimizing the beam-training time till a connection between the Access Point (AP) and the UE is established, on the other hand, has been a popular topic of research, and a multitude of beam-training mechanisms, as well as hybrid analog-digital beamforming (BF) schemes, have been proposed. More concretely, works such as [6]–[10] focus on the design of beam-training mechanisms. However, in these works, only a single UE system model is considered, with the aim of either accelerating the beam training procedure [6]–[9], or by proposing an opportunistic, QoS-based scheme [10]. Also, in other works, such as in [11]–[14], multiple UE scenarios are taken into consideration, nevertheless, the performance metric is the achievable UE sum rate, in the absence of divergent QoS-related constraints. Finally, it is worth noting that most of the literature on mm-wave beam training via hybrid BF proposes high complexity solutions built upon unrealistic hardware assumptions. For instance, one of these assumptions is the availability of RF phase shifters with large (or even infinite) numbers of quantization bits [7], [10], [12], [14].

Motivated by the above described situation, in this paper, we consider a mm-wave system consisting of a single AP, featuring hybrid analog-digital BF, and aim at optimally allocating its available hardware resources in order to efficiently serve two UEs, which are characterized by conflicting QoS requirements. In further detail, our contributions are the following:

- We formulate, for the first time, an optimization problem, according to which, the hardware resources at the AP, namely, the available Radio Frequency (RF) chains, are allocated, such that the expected data rate of one UE is maximized, given a latency constraint set by the other UE. More specifically, two different allocation strategies are considered in the paper: *joint access* and *separate access*, where the term of "access" refers to both beam-training and data transmission phases. According to the first strategy, the full RF chain set at the AP is used to concurrently beam-train and then transmit data to both UEs, whereas, according to the second one, two separate RF chain subsets are formed at the beginning of each Medium Access Control (MAC) frame and each subset is destined towards accessing a specific UE;
- Having defined the problem space, we apply an exhaustive search to illustrate the achievable average rates

of the throughput-demanding UE, for different latency requirements posed by the delay-critical UE, considering a variety of system setups. Simulation results reveal that: (a) the joint access strategy is the optimally performing one in most of the cases, (b) strict latency constraints posed by one UE result in poor throughput performance experienced by the other UE, even when fully-digital BF is applied, and (c) high rate values at one UE can be achieved and sustained, even subject to a strict latency constraint set by the other UE, when the angular separation of the two devices is above a certain threshold.

We use the following notation along the paper. $\mathcal{CN}(\mu, \sigma^2)$ denotes the complex normal distribution with mean μ and variance σ^2, $\mathbf{A}$ is a matrix, and $\mathbf{a}$ is a vector with Euclidean norm $\|\mathbf{a}\|_2$. Finally, $\|\mathbf{A}\|_F$, $\mathbf{A}^H$, and $\mathbf{A}^T$ represent the Frobenius norm, the Hermitian matrix, and the transpose matrix of $\mathbf{A}$, respectively, while $\mathbb{E}[\cdot]$ denotes the expectation operator.

II. System Model

The considered mm-wave system is illustrated in Fig. 1. It is composed of a single AP featuring hybrid BF transceiver architecture with M antennas connected to $N_{\text{RF}} < M$ RF chains through a network of analog phase shifters. As in [8], we consider that, in order to reduce hardware complexity and power consumption, the phase shifters have a resolution of only two quantization bits without amplitude adjustment. The availability of multiple RF chains allows the AP to communicate via N_{S} data streams, with $N_{\text{S}} < N_{\text{RF}} < M$, and serve multiple users at the same time. In the vicinity of the AP, two mobile UEs equipped with a single omnidirectional antenna, each, are randomly located in the two-dimensional space: UE 1, which is characterized by increased data-rate requirements and UE 2, which is a latency-intolerant device[1].

Depending on whether blockage events occur or not, mm-wave channels are characterized by the potential existence of a Line-of-Sight (LoS) path, along with the potential presence of a number of Non-LoS (NLoS) resolvable paths due to reflection effects [5]. Focusing on downlink communication, the Multiple-Input-Single-Output (MISO) link between the AP and UE i, $i = \{1, 2\}$, can be effectively described by a geometric (multipath) channel model. According to this model, a number of L_i scatterers is used to represent the existence of the various propagation paths between the AP and UE i. Hence, by assuming that the AP is equipped with a Uniform Linear Array (ULA), the widely adopted geometry-based stochastic channel model [7]–[10], [12] is obtained

$$\mathbf{h}_i = \frac{1}{\sqrt{L_i}} \sum_{\ell=1}^{L_i} \alpha_{i,\ell} \mathbf{u}_{\text{AP}}^H(\phi_{i,\ell}), \quad i = \{1, 2\}, \tag{1}$$

where $\alpha_{i,\ell} \sim \mathcal{CN}\left(0, \sigma^2_{\alpha_{i,\ell}}\right)$ denotes the complex gain of the ℓ-th path, $\sigma^2_{\alpha_{i,\ell}}$ stands for its average attenuation, which includes the distance-based deterministic path loss, and

[1]Multi-antenna UEs could be also considered, however, channel acquisition needed for initial access would require the application of a bi-directional beam-training scheme, which is not the main focus of this paper.

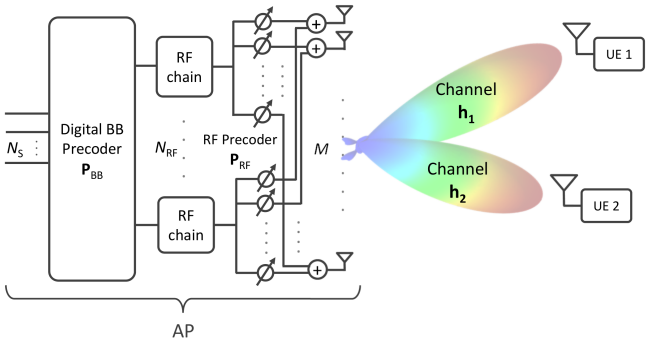


Fig. 1. Overview of the mm-wave AP transceiver architecture for hybrid BF and topology of the examined two-UE mm-wave system.

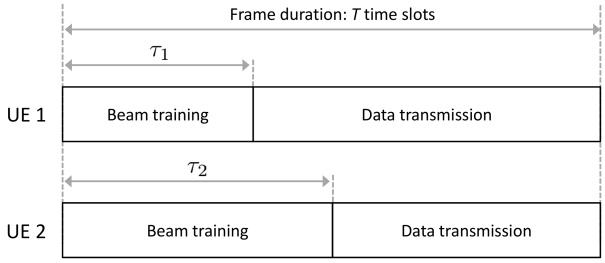


Fig. 2. Considered MAC frame structure encompassing both beam training and data transmission phases.

$\phi_{i,\ell} \in [0, 2\pi]$ denotes the corresponding Angle-of-Departure (AoD). Moreover, $\mathbf{u}_{\text{AP}}(\phi) \in \mathbb{C}^{M \times 1}$ represents the steering vector of the AP antenna array, which, in the case of a ULA, is given by what follows

$$\mathbf{u}_{\text{AP}}(\phi) = \left[1, \; e^{j\frac{2\pi}{\lambda}\Delta\sin(\phi)}, \cdots, e^{j(M-1)\frac{2\pi}{\lambda}\Delta\sin(\phi)}\right]^T, \tag{2}$$

where, Δ is the distance between two adjacent antenna elements and λ is the wavelength of the transmitted signal. Note that, in order to simplify the notation, we consider the AP implementing horizontal (2-D) beamforming only, which implies that all scattering occurs in the azimuthal domain. Extensions to 3-D beamforming are straightforward.

Focusing on the MAC frame, we consider the structure in Fig. 2. It is assumed that each frame, associated to UE i, $i = \{1, 2\}$, consists of T time slots, where, a number of them (i.e., τ_i time slots), are exploited for beam-training and the remaining $T - \tau_i$ ones for data transmission, once beam-training is terminated. It is also assumed that the MAC frames of the two UEs are synchronized in time and that the involved mm-wave channels remain fixed during a MAC frame.

In the next section, we describe two different strategies of handling the total number of RF chains at the AP, in a way that, by means of hybrid BF, the tasks of beam-training and data transmission are carried out for both UEs, taking into account their dissimilar QoS demands. As one would expect, the performance of these strategies depends on the beam-training protocol for initial access, hence, we also provide a description of it.

III. QoS-centric resource allocation

We envision two potential strategies of exploiting the available RF chains at the AP, in view of accessing the two UEs. The two strategies are described in what follows.

Joint access strategy. According to this strategy, the AP synthesizes multi-beam antenna patterns for training and transmitting data to both UEs simultaneously, without any decoupling of the RF chain set. More concretely, referring to the architecture in Fig. 1 and assuming that $N_S = 2$ streams (i.e., one per UE), the AP applies the joint digital baseband precoder $\mathbf{P}_{\text{BB,jnt}} \in \mathbb{C}^{N_{\text{RF}} \times 2}$ followed by the joint RF precoder $\mathbf{P}_{\text{RF,jnt}} \in \mathbb{C}^{M \times N_{\text{RF}}}$ to the vector of transmitted discrete-time symbols $\mathbf{s}(t) \in \mathbb{C}^{2\times 1}$. Focusing on the latter symbol vector, its two elements, i.e., $s_1(t)$ and $s_2(t)$, represent the symbols intended for UE 1 and UE 2 respectively. The transmit power constraint is ensured by imposing the normalization $\|\mathbf{P}_{\text{RF,jnt}}\mathbf{P}_{\text{BB,jnt}}\|_F^2 = 1$. The discrete-time symbol received by UE i, $i = \{1, 2\}$, is, hence, given by:

$$\mathrm{y}_i(t) = \sqrt{\frac{P_t}{2}}\mathbf{h}_i\mathbf{P}_{\text{jnt}}\mathbf{s}(t) + n(t), \quad i = \{1, 2\}\,, \tag{3}$$

where P_t is the transmit power (equally split over the two streams), $n(t) \sim \mathcal{CN}(0, \sigma^2)$ is the noise, and $\mathbf{P}_{\text{jnt}} \in \mathbb{C}^{M\times 2}$ is the joint hybrid precoder given by

$$\mathbf{P}_{\text{jnt}} = \mathbf{P}_{\text{RF,jnt}}\mathbf{P}_{\text{BB,jnt}}\,. \tag{4}$$

In order to serve the two UEs simultaneously, the joint hybrid precoder $\mathbf{P}_{\text{jnt}}$ should be designed to provide multi-beam antenna patterns. The problem can be formulated as that of finding $\mathbf{P}_{\text{RF,jnt}}$ and $\mathbf{P}_{\text{BB,jnt}}$ that minimize $||\mathbf{P}_{\text{jnt}}^{\text{opt}} - \mathbf{P}_{\text{jnt}}||_2$, where $\mathbf{P}_{\text{jnt}}^{\text{opt}}$ is the optimal, multi-beam precoder synthesized via fully-digital beamforming. To solve this problem, we here adopt the greedy geometric strategy proposed in our prior work [9, Algorithm 1], which, in contrast to state-of-the-art Orthogonal Matching Pursuit (OMP) schemes [7], [8], [10], [12], is dictionary-free, requires much lower computational complexity, and provides higher accuracy. Our strategy iterates over the number of RF chains with the objective to minimize the residual between the provisional hybrid precoder selected in the current iteration and the ideal, fully-digital precoder. To do that, we adopt a residual update mechanism that tries to share the effort of approaching the ideal precoder almost equally among all the available RF chains. This is quite different from what is done in OMP algorithms, which find the best solution for the current iteration, without optimizing for later iterations.

Separate access strategy. According to this strategy, the full set of RF chains is decomposed into two disjoint subsets[2] from the beginning of each MAC frame. The RF chain subset related to UE i, $i = \{1, 2\}$, consists of $N_{\text{RF},i} \in \{1, \cdots, N_{\text{RF}} - 1\}$ RF chains, such that $N_{\text{RF},1} + N_{\text{RF},2} = N_{\text{RF}}$, and a hybrid, analog-digital BF scheme is then applied both for training, as well as for transmitting data to that UE.

[2]It is assumed that the resulting subsets are such that no RF chain remains unused after the decoupling takes place.

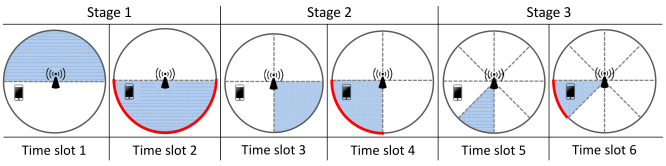


Fig. 3. Example of angular domain partitioning performed by the AP during the first three stages of adaptive beam training and mapping with MAC frame training time slots. The sectors with red border are the ones providing maximum SNR at the end of each stage.

By applying this strategy, the hybrid precoder (or, equivalently, beam pattern) for accessing UE i, namely $\mathbf{p}_i \in \mathbb{C}^{M\times 1}$, can be synthesized separately using [9, Algorithm 1]:

$$\mathbf{p}_i = \mathbf{P}_{\text{RF},i}\mathbf{p}_{\text{BB},i}, \quad i = \{1, 2\}\,, \tag{5}$$

where $\mathbf{P}_{\text{RF},i} \in \mathbb{C}^{M\times N_{\text{RF},i}}$ and $\mathbf{p}_{\text{BB},i} \in \mathbb{C}^{N_{\text{RF},i}\times 1}$ represent the RF precoder and the baseband precoder for accessing UE i, respectively. The combined hybrid precoder $\mathbf{P}_{\text{sep}} \in \mathbb{C}^{M\times 2}$, to be used in place of $\mathbf{P}_{\text{jnt}}$ in Eq. 3 is given by

$$\mathbf{P}_{\text{sep}} = \mathbf{P}_{\text{RF,sep}}\mathbf{P}_{\text{BB,sep}}\,, \tag{6}$$

where $\mathbf{P}_{\text{RF,sep}} \in \mathbb{C}^{M\times N_{\text{RF}}}$ and $\mathbf{P}_{\text{BB,sep}} \in \mathbb{C}^{N_{\text{RF}}\times 2}$ are defined as:

$$\mathbf{P}_{\text{RF,sep}} \triangleq [\mathbf{P}_{\text{RF},1}, \mathbf{P}_{\text{RF},2}],\ \mathbf{P}_{\text{BB,sep}} \triangleq \begin{bmatrix} \mathbf{p}_{\text{BB},1} & 0 \\ 0 & \mathbf{p}_{\text{BB},2} \end{bmatrix}. \tag{7}$$

A. Beam training protocol

Given that either the joint or the separate access strategy is applied at the AP, data communication can be established with each UE upon the accomplishment of a beam training procedure. Focusing on UE i, $i \in \{1, 2\}$, we suppose that $N_{\text{RF},i} \in \{1, ..., N_{\text{RF}}-1\}$ RF chains undertake the task of beam-training it, when the separate access strategy is considered, while all the N_{RF} RF chains are used to beam-train both UEs when the joint access strategy is applied. We consider the adaptive protocol used in [8] and originally proposed in [7]. During each stage of that protocol, the AP transmits training sequences over different sectors (see Fig. 3), while the UE performs Signal-to-Noise Ratio (SNR) measurements and notifies the AP of the sector with the highest received SNR[3]. The protocol proceeds from stage to stage according to the bisection concept, that is, at each stage, the selected sectors are further divided into two sub-sectors. The process terminates when a stopping criterion is met, e.g., when the desired signal strength is reached, when the required angular resolution is achieved, or as soon as a given latency constraint is satisfied. At the end of beam training, the AP steers its antennas according to the selected directions and initiates data transmission.

[3]In this work, for brevity, we assume that the interference received by each UE is treated as noise and that the AP is informed of the best-beam selection with zero delay (implicit feedback).

B. Problem formulation

The optimization problem to optimally allocate the available AP resources, by means of maximizing the expected rate of UE 1, for a given latency constraint which is provided by UE 2, can be mathematically expressed as

$$\begin{aligned} \max_{N_{\text{RF},1},N_{\text{RF},2},\mathcal{S},\tau_1} \quad & \mathbb{E}\left[R_1\right] \\ \text{s.t.} \quad & \tau_2 = \tilde{\tau}_2 < T\,, \\ & N_{\text{RF},1}, N_{\text{RF},2} \in \mathcal{N}_{\text{sep}}\,, \\ & \mathcal{S} \in \{\mathcal{S}_{\text{jnt}}, \mathcal{S}_{\text{sep}}\}\,, \\ & 0 < \tau_1 < T,\ \tau_1 = 2k_1,\ \tilde{\tau}_2 = 2k_2\,, \\ & k_1, k_2 \in \mathbb{N}^+\,, \end{aligned} \tag{8}$$

where, $\mathcal{S}_{\text{jnt}}$ and $\mathcal{S}_{\text{sep}}$ denote the joint and separate access strategies, respectively, $\mathcal{N}_{\text{sep}} = \{N_{\text{RF},1}, N_{\text{RF},2} \in \mathbb{N}^+ : N_{\text{RF},1} + N_{\text{RF},2} = N_{\text{RF}} | \mathcal{S} = \mathcal{S}_{\text{sep}}\}$, $\mathbb{E}[R_1]$ denotes the expected rate of UE 1 and $\tilde{\tau}_2$ stands for the training time required by UE 2 (latency constraint). It should be noted that the expectation operator refers to the rate distribution of UE 1, when certain values of the optimization variables and parameter τ_2 are considered. Given that our ultimate goal is to show the performance of the different resource allocation schemes which constitute the space of Problem (8), we aim at conducting an exhaustive search over the domains of the optimization variables for different values of $\tilde{\tau}_2$.

IV. Numerical results

In this section, with the aim of evaluating the different design options, we perform intensive Monte Carlo simulations, focusing on the single AP, two-UE mm-wave system in Fig. 1. Two different scenarios are investigated where, depending on the scenario, the system performance is monitored as a function of 1) the relative (azimuthal) positions of the two UEs and 2) the latency requirement of UE 2. We divide the simulation into two phases as indicated by the MAC frame structure in Fig. 2. In the first phase, the AP synthesizes multi-beam antenna patterns and performs the adaptive beam training procedure in §III-A, applying either the joint or separate access strategy, to estimate the most suitable antenna sectors for accessing the two UEs. In the second phase, the selected antenna sectors are used for data communication, always following one of the two strategies. In order to take into account the impact of the beam training time on the system performance, we consider the normalized rate at each UE, which is obtained by scaling the maximum achievable rate to the effective duration of the data transmission phase. In further detail, for each MAC frame/ channel realization, we compute the normalized rate R_i of UE i, $i = \{1, 2\}$, as in Eq. 9, where, $\text{SNR}_{i,t=\tau}$ corresponds to the received SNR at UE i, $i = \{1, 2\}$, during the τ-th time-slot, while, j denotes the complementary index of i, i.e., $j = i \mod 2 + 1$.

All simulations assume the presence of a dominant LoS path and additional scattered paths for each AP-UE channel. The AoD of the LoS path is determined based on the relative AP-UE positions, while the AoDs of the scattered paths are uniformly distributed in $[0, 2\pi]$. Also, for each AP-UE channel, the power is distributed between LoS and scattered paths according to K-factor[4] values from field measurements [15]. The main simulation parameters used in this section are summarized in Table I.

TABLE I
Simulation parameters

Carrier frequency	60 GHz
System bandwidth	500 MHz
AP transmit power, P_t	30 dBm
MAC frame duration, T [time slots]	100
AP antenna elements, M	16
RF chains at the AP, N_{RF}	6
Channel paths for the AP-UE link, L_1, L_2	3
K factor	3
Number of simulated channel realizations	5000

A. Impact of the angular separation of the two UEs

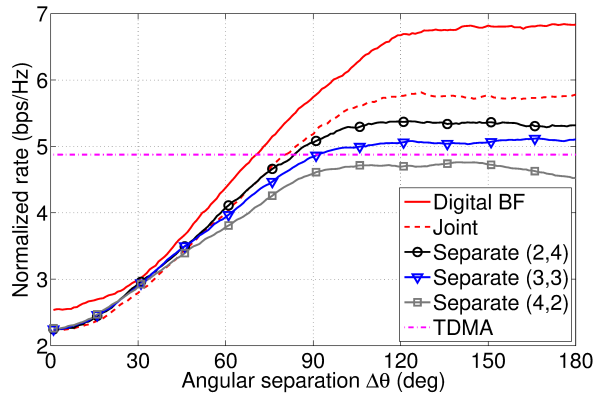
In order to evaluate the impact of concurrent transmissions on the achievable rates when varying the angular separation between the two UEs, we place UE 1 and UE 2 on a circumference of radius 10 m and center coinciding with the AP position. In Fig. 4, the optimal average data rates at UE 1 (Fig. 4(a)), in terms of solving Problem (8), and the corresponding average data rates at UE 2 (Fig. 4(b)), are depicted as a function of the angular separation of the two UEs (which we denote as $\Delta\theta$) when $\tilde{\tau}_2 = 4$ time slots. In each of the figures, the curves representing the average rate performance achieved by applying fully-digital BF, as well as hybrid BF following the joint and separate access strategies, are illustrated, along with the performance of a baseline Time Division Multiple Access (TDMA) approach. According to the TDMA scheme, a UE will be exclusively accessed (with respect to both beam-training and data transmission phases) during a MAC frame and during the following frame it will be idle in order for the other UE to be served[5]. Regarding the separate access strategy, different combinations of RF chain subsets ($N_{\text{RF},1}$, $N_{\text{RF},2}$) are investigated, i.e., (2, 4), (3, 3) and (4, 2), where, the first index refers to the number of RF chains used for accessing UE 1, $N_{\text{RF},1}$, and the second one to the number of RF chains used for accessing UE 2, $N_{\text{RF},2}$.

Focusing on both Fig. 4(a) and 4(b), the achieved normalized rates related to the TDMA baseline scheme show a constant behavior regardless of the angular separation between the two UEs. The TDMA approach enables interference-free access both in terms of beam-training and data transmission, due to the orthogonal allocation of time resources to the two UEs. As a result, this feature explains the observed behavior. In comparison to the proposed access schemes, it can be seen that the TDMA approach outperforms them only for values of

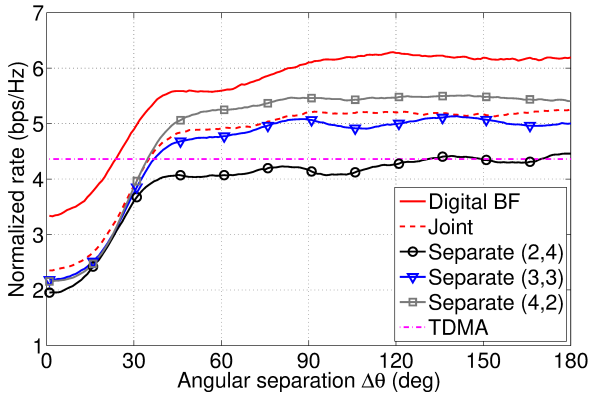
[4]The K factor is defined as the ratio of signal power of the LoS component, P_{LoS}, over the local-mean scattered power, P_{scatt}, i.e., $K = 10\log_{10}\left(\frac{P_{\text{LoS}}}{2P_{\text{scatt}}}\right)$.

[5]Of course, in this case, the latency constraint is always violated, however, it is an existing approach which serves as a performance benchmark.

$$R_i = \begin{cases} \frac{1}{T} \sum\limits_{\tau=\tau_i}^{\tau_j} \log_2\left(1 + \text{SNR}_{i,t=\tau}\right) + \left(1 - \frac{\tau_j}{T}\right) \log_2\left(1 + \text{SNR}_{i,t=\tau_j}\right), & \text{if } \tau_i < \tau_j \\ \left(1 - \frac{\tau_i}{T}\right) \log_2\left(1 + \text{SNR}_{i,t=\tau_i}\right), & \text{if } \tau_i \geq \tau_j \end{cases} \tag{9}$$



(a) Average normalized rate at UE 1.



(b) Average normalized rate at UE 2.

Fig. 4. Average normalized rate at both UEs as a function of their relative angular separation, $\Delta\theta$: comparison between digital BF and hybrid BF (applying the joint and separate access strategies) when $\tilde{\tau}_2 = 4$ time slots.

$\Delta\theta$ smaller than around 35° for UE 2 and around 80° for UE 1. This occurs because, unlike the TDMA approach, the proposed access schemes suffer from interference mainly related to the used beam patterns for small values of $\Delta\theta$, whereas, when the angular separation increases, the data rate of each UE can be improved through appropriate allocation of time resources.

We also observe that, for all proposed access strategies, the average rate of UE 1 initially increases with parameter $\Delta\theta$, up to a maximum value, unique for each strategy, corresponding to an azimuthal separation of i) 90°-100° for strategies related to hybrid BF and ii) around 120° for access via fully-digital BF, and then, when $\Delta\theta$ further increases, the average normalized rate is fixed to the maximum attained value. This happens, because, for the considered value of $\tilde{\tau}_2 = 4$ time slots, which corresponds to a training pattern beam-width of 90° for accessing UE 2 (see Fig. 3), an increase in $\Delta\theta$ implies a reduction in the interference UE 1 experiences due to the ongoing beam training process for accessing UE 2. More specifically, for small values of $\Delta\theta$, UE 1 is going to suffer from interference mostly coming from the main beam-training lobe dedicated to UE 2 and not its low interfering side lobes, hence, it will experience low data rate values. However, when $\Delta\theta$ overcomes the value of 90°, UE 1 is going to receive interference exclusively coming from the fluctuating side lobes related to the same beam training pattern, thus, its maximum feasible rate is almost fixed to a high value.

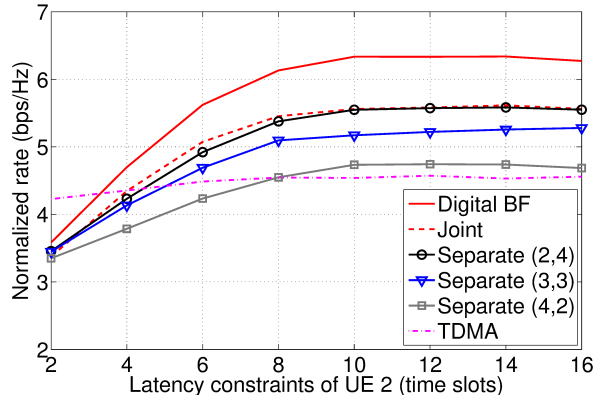
A similar argument can be used to explain why the rate gain in favor of digital BF is increased as the value of parameter $\Delta\theta$ increases. The reason lies in the fact that for large angular separations, the interference received by UE 1 when fully-digital BF is applied, is practically negligible, as interference related to side lobes is extremely weak, in comparison to the interference experienced under hybrid BF.

Additionally, it is worth noticing that when $\Delta\theta$ is equal or larger than 75°, the joint access strategy outperforms all variants of the separate access strategy, as well as that, focusing on the latter strategy, the rate improves together with the number of RF chains allocated for accessing UE 2, regardless of the reduction in the number of RF chains exploited for UE 1. Once more, the reason is that, interference leakage due to the existence of side lobes is expected to occur with higher probability when the beam training codebook of the interfering UE is tailored to a small number of RF chains. This justification point is also cross-verified by observing the achieved average rate of UE 2, depicted in Fig. 4(b).

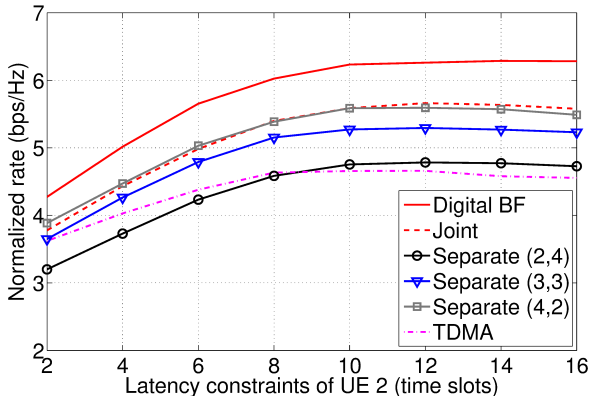
B. Impact of the latency constraint of UE 2

In a second set of simulations, the two UEs are uniformly located within a circular area of radius of 20 m around the mm-wave AP. In Fig. 5, the investigated access strategies are evaluated in terms of their achieved normalized rate, this time, as a function of the latency constraint of UE 2, $\tilde{\tau}_2$.

By observing Fig. 5(a) and, taking into consideration the arguments explained in the previous subsection, it is evident that, when the latency constraint of UE 2 is strict, then, regardless of the access strategy, UE 1 is highly likely to be exposed to interference from the wide main lobe associated with UE 2. Such a phenomenon has a detrimental effect on the rate of UE 1, even when fully-digital BF is adopted. Also, the three RF chain set decoupling combinations related to the separate access strategy perform in the same order as in the previous experiment. Nevertheless, as it is obvious in Fig. 5(b), this performance order is reverted, exactly because the rate of UE 2 can improve, given that more close-to-digital infrastructure is exploited for beam-training of UE 1.



(a) Average normalized rate at UE 1.



(b) Average normalized rate at UE 2.

Fig. 5. Average normalized rate at both UEs as a function of the latency constraint of UE 2, $\tilde{\tau}_2$: comparison among different access strategies.

It can be also observed that, for all proposed access schemes, the average normalized rate of UE 1 increases with higher slope when $\tilde{\tau}_2 \in [2, 10]$ time slots and shows a steady behavior, when looser latency constraints are considered for UE 2. This phenomenon occurs due to the randomness of the UE positions along with the narrow beam training patterns considered when $\tilde{\tau}_2 \geq 10$ time slots. When such a constraint is considered, the corresponding width of the main beam-training lobe for accessing UE 2 is equal to or less than 12.5°, hence, the probability of UE 1 receiving interference from such a narrow beam is quite low. Focusing on the corresponding rate of UE 2 and observing the monotonicity of the curves, it can be concluded that each access strategy addresses the trade-off between beam training and data transmission in a different way. In other words, each strategy is characterized by a specific value of $\tilde{\tau}_2$ that maximizes the average rate of UE 2.

V. Conclusion

In this work, it was shown that, when two UEs seek access from a mm-wave AP, different latency requirements by one UE lead towards different average rates experienced by the other UE. It was also observed that, the achieved rate levels highly depend on the system scenario, as well as on the resource allocation scheme applied at the AP. Numerical evaluation results indicated that a bisection-based beam training protocol impacts the performance of the system, since different access strategies lead to different levels of interference received by the UEs. The proposed hybrid BF based access schemes outperform the typical TDMA approach appearing in literature and standards by up to 22 %. Interesting extensions can be considered, focusing on more realistic channel models and dense UE deployments, when it comes to searching either for different, multi-UE beam training protocols, or, for the optimal hybrid precoders themselves.

VI. Acknowledgment

The research leading to these results received funding from the European Commission H2020 programme under grant n° 671650 (5G PPP mmMAGIC project), from the Madrid Regional Government through the TIGRE5-CM program (S2013/ICE-2919), and from the Spanish Ministry of Economy and Competitiveness through Ramon y Cajal (RYC-2012-10788) and Hyperadapt (TEC2014-55713-R) grants.

References

[1] M. Tercero *et al.*, "5G systems: The mmMAGIC project perspective on use cases and challenges between 6-100 GHz," in *2016 IEEE Wireless Communications and Networking Conference*, Apr. 2016.

[2] mmMAGIC. Millimetre-Wave Based Mobile Radio Access Network for Fifth Generation Integrated Communications, EU 5G-PPP H2020-ICT-2014-2 Project. [Online]. Available: http://5g-mmmagic.eu/

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

[4] F. Boccardi *et al.*, "Five disruptive technology directions for 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[5] T. S. Rappaport *et al.*, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[6] Y. M. Tsang, A. S. Y. Poon, and S. Addepalli, "Coding the beams: Improving beamforming training in mmWave communication system," in *IEEE Global Telecomm. Conf. (GLOBECOM 2011)*, Dec. 2011.

[7] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 831–846, Oct. 2014.

[8] D. D. Donno, J. Palacios, D. Giustiniano, and J. Widmer, "Hybrid analog-digital beam training for mmWave systems with low-resolution RF phase shifters," in *2016 IEEE International Conference on Communications Workshops (ICC)*, May 2016, pp. 700–705.

[9] J. Palacios, D. De Donno, D. Giustiniano, and J. Widmer, "Speeding up mmWave beam training through low-complexity hybrid transceivers," in *27th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sept. 2016.

[10] M. E. Eltayeb, A. Alkhateeb, R. W. Heath, and T. Y. Al-Naffouri, "Opportunistic beam training with hybrid analog/digital codebooks for mmWave systems," in *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Orlando, FL, USA, Dec. 2015.

[11] H. Shokri-Ghadikolaei, L. Gkatzikis, and C. Fischione, "Beam-searching and transmission scheduling in millimeter wave communications," in *2015 IEEE Int. Conference on Communications (ICC)*, Jun. 2015.

[12] A. Alkhateeb, R. W. Heath, and G. Leus, "Achievable rates of multi-user millimeter wave systems with hybrid precoding," in *2015 IEEE Int. Conference on Communication Workshop (ICCW)*, Jun. 2015.

[13] N. Lee, C. Jeong, J. Kim, and J. Park, "A new codebook structure for enhanced multi-user MIMO transmission in mmWave hybrid-beamforming system," in *2015 IEEE Globecom Workshops (GC Wkshps)*, Dec. 2015.

[14] D. H. N. Nguyen, L. B. Le, and T. Le-Ngoc, "Hybrid MMSE precoding for mmWave multiuser MIMO systems," in *2016 IEEE International Conference on Communications (ICC)*, May 2016.

[15] J. Hansen, "A novel stochastic millimeter-wave indoor radio channel model," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 6, pp. 1240–1246, Sep. 2006.