

# Energy Efficiency Maximization in Millimeter Wave Hybrid MIMO Systems for 5G and Beyond

Aryan Kaushik, John Thompson and Evangelos Vlachos

Institute for Digital Communications, The University of Edinburgh, United Kingdom

Email: {a.kaushik, j.s.thompson, e.vlachos}@ed.ac.uk

(Invited Paper)

**Abstract**—At millimeter wave (mmWave) frequencies, the higher cost and power consumption of hardware components in multiple-input multiple output (MIMO) systems do not allow beamforming entirely at the baseband with a separate radio frequency (RF) chain for each antenna. In such scenarios, to enable spatial multiplexing, hybrid beamforming, which uses phase shifters to connect a fewer number of RF chains to a large number of antennas is a cost effective and energy-saving alternative. This paper describes our research on fully adaptive transceivers that adapt their behaviour on a frame-by-frame basis, so that a mmWave hybrid MIMO system always operates in the most energy efficient manner. Exhaustive search based brute force approach is computationally intensive, so we study fractional programming as a low-cost alternative to solve the problem which maximizes energy efficiency. The performance results indicate that the resulting mmWave hybrid MIMO transceiver achieves significantly improved energy efficiency results compared to the baseline cases involving analogue-only or digital-only signal processing solutions, and shows performance trade-offs with the brute force approach.

**Index Terms**—energy efficiency, hybrid beamforming, MIMO, millimeter wave, 5G and beyond.

## I. INTRODUCTION

Fifth generation (5G) technology is set to address the consumer demands and performance enhancements for mobile communication in 2020 and beyond [1]. There will be 28.5 billion networked devices and connections by 2022 [2] and 8.9 billion mobile subscriptions by the end of 2024 [3]. For such large scale use of mobile devices through 5G and beyond 5G services, the communication systems would require increased capacity, high data rates, improved coverage and also reduced energy consumption. We currently use the microwave frequency spectrum for communication which is congested with a large number of consumer devices raising the demand for an unused and available spectrum. This increased demand on bandwidth and capacity can be resolved by the use of millimeter wave (mmWave) frequency spectrum which ranges from 30-300 GHz [4]. This is beneficial as the larger spectral channels at mmWave would lead to higher data rates. Moreover, the large scale antenna arrays such as the multiple-input multiple-output (MIMO) systems can reduce the high path loss at mmWave frequencies [5], [6]. However, it would be difficult to use one radio frequency (RF) chain per antenna leading to a least energy efficient and highly complex system. Thus, using digital beamforming which needs a dedicated RF

chain per antenna is not very practical from energy efficiency (EE) and hardware complexity perspectives. To save power and reduce complexity, analogue beamforming can be used where a network of analogue phase shifters connects the antennas to a single RF chain [7], but multi-stream and multi-user communication can not be supported.

A mmWave MIMO system with hybrid beamforming (HBF) architecture can save power and reduce hardware complexity using fewer number of RF chains than the large number of antennas, and support multi-stream communication with high spectral efficiency (SE) [8]–[12]. Such systems can also be optimized to achieve high EE gains [13] but this has not been widely studied for EE maximization with low complexity. Low resolution sampling can be implemented to save power such as in [14] we discuss EE maximization with low resolution digital-to-analogue converters (DACs) at the transmitter (TX), in [15] with low resolution analogue-to-digital converters (ADCs) at the receiver (RX) and in [16] with low resolution sampling at both the DACs and the ADCs. However, the existing literature mostly considers fixed number of RF chains for high SE performance [8]–[12] and RF chains consume a lot of power which increases the cost of MIMO systems [17]. Reference [13] provides an exhaustive search based brute force (BF) approach where a full precoder design is evaluated for all possible combinations of RF chains, in order to select the number of RF chains that maximizes EE but this is a computationally inefficient solution. Moreover, lower complexity solutions can be implemented to design the HBF matrices than in [8], [13].

**Contribution:** This paper describes different approaches to performing *dynamic adaptation* of a mmWave hybrid MIMO system on a frame-by-frame basis. Our idea exploits the beam training phase in the communication system to learn the propagation conditions. Based on this, we can choose to adapt the behaviour of the transceiver in order to optimize a performance metric of interest, such as EE. Maximizing EE is challenging mathematically because it is a ratio of two important parameters, namely data rate (or SE) and power. In our recent research, we use the Dinkelbach method (DM) [18] to replace this ratio function by an iterative sequence of problems based on the difference of the numerator and denominator. In this work, we discuss different ways to optimize the transceivers, particularly in relation to the number of activated

Notations	Description
$a$	Scalar
$\mathbf{a}$	Vector
$\ \mathbf{a}\ _0$	$l_0$ -norm of $\mathbf{a}$
$\mathbf{A}$	Matrix
$ \mathbf{A} $	Determinant of $\mathbf{A}$
$\mathbf{A}^T$	Transpose of $\mathbf{A}$
$\mathbf{A}^H$	Complex conjugate transpose of $\mathbf{A}$
$\mathbf{A}^{(i)}$	$i$ -th column of $\mathbf{A}$
$\ \mathbf{A}\ _F$	Frobenius norm of $\mathbf{A}$
$\mathcal{CN}(\mathbf{a}; \mathbf{A})$	Complex Gaussian vector; mean $\mathbf{a}$ , covariance $\mathbf{A}$
$\mathbb{C}^{A \times B}$	To represent matrix of size $A \times B$ with complex entries
$\mathbb{E}\{\cdot\}$	Expectation operator
$\mathbf{I}_N$	Identity matrix with size $N \times N$
$\mathbb{R}^+$	Set of positive real numbers
$\mathbb{R}\{\cdot\}$	Real part
$\text{tr}(\mathbf{A})$	Trace of $\mathbf{A}$
$\mathbf{X} \in \mathbb{C}^{A \times B}$	Complex-valued matrix $\mathbf{X}$ of size $A \times B$
$\mathbf{X} \in \mathbb{R}^{A \times B}$	Real-valued matrix $\mathbf{X}$ of size $A \times B$

TABLE I: List of notations and their description.

RF chains and the sample rate of the system. As a practical example, we present a more detailed discussion of how the Dinkelbach's approach can be used to optimize the EE and simultaneously achieve a low complexity alternative to the exhaustive search based BF approach in [13]. An attractive feature of our approach is that we only need to compute the HBF matrices once, after the number of RF chains is determined by the DM based solution.

*Notations and Organization:* Table I provides a list of notations used in this paper along with their description. The remainder of the paper is structured as follows: Section II describes the channel model and HBF architecture that is used in the paper. Section III describes the EE maximization problem and we describe different approaches that we have studied to address this problem. In Section IV, we discuss in more detail how the DM can be applied to select the optimal number of RF chains. Section V presents simulation results to show the performance improvements of the DM and finally Section VI presents conclusions to the paper.

## II. MMWAVE MIMO SYSTEM WITH HBF

### A. MmWave Channel

We use a narrowband clustered channel model due to different channel settings at mmWave such as the number of multipaths, amplitudes, etc. [6]. We consider  $N_{cl}$  clusters with  $N_{ray}$  paths related to each cluster and for a single user system we have  $N_T$  TX antennas transmitting  $N_s$  data streams to  $N_R$  RX antennas. This mmWave channel can be expressed as

$$\mathbf{H} = \sum_{i=1}^{N_{cl}} \sum_{l=1}^{N_{ray}} \alpha_{il} \mathbf{a}_R(\phi_{il}^r) \mathbf{a}_T(\phi_{il}^t)^H, \quad (1)$$

where  $\alpha_{il} \in \mathcal{CN}(0, \sigma_{\alpha,i}^2)$  is the gain term with  $\sigma_{\alpha,i}^2$  being the average power of the  $i^{th}$  cluster. The vectors  $\mathbf{a}_T(\phi_{il}^t)$  and  $\mathbf{a}_R(\phi_{il}^r)$  denote the normalized array response vectors at the TX and the RX, respectively [6], with  $\phi_{il}^t$  being the azimuth angles of departure and  $\phi_{il}^r$  being the azimuth angles

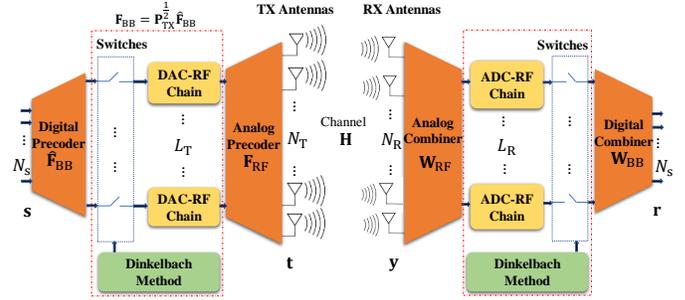


Fig. 1: A mmWave MIMO system with HBF architecture and the proposed DM framework.

of arrival. We assume the transmit and receive arrays are uniform linear arrays (ULAs) of antennas, which are modelled as ideal sectored elements [19].

### B. MIMO System with HBF Architecture

Fig. 1 shows the system model considered in this paper where  $L_T$  is the number of available RF chains at the TX and  $L_R$  at the RX. Based on MIMO communication with HBF, we follow the conditions  $N_s \leq L_T \leq N_T$  and  $N_s \leq L_R \leq N_R$ . The symbol vector  $\mathbf{s} \in \mathbb{C}^{N_s \times 1}$  at the TX is such that  $\mathbb{E}\{\mathbf{s}\mathbf{s}^H\} = \frac{1}{N_s} \mathbf{I}_{N_s}$ . The digital precoder matrix right before the DAC-RF chain blocks is  $\mathbf{F}_{BB} \in \mathbb{C}^{L_T \times N_s} = \mathbf{P}_{TX}^{\frac{1}{2}} \hat{\mathbf{F}}_{BB}$  where  $\hat{\mathbf{F}}_{BB}$  is the digital precoder matrix before the switches and  $\mathbf{P}_{TX} \in \mathbb{R}^{L_T \times L_T}$  is a diagonal matrix with entries of power allocation values. We have  $\text{tr}(\mathbf{P}_{TX}) = P_{\max}$ , where  $P_{\max}$  is the maximum allocated power. The entries of the analogue precoder matrix  $\mathbf{F}_{RF} \in \mathbb{C}^{N_T \times L_T}$  are of constant modulus and this matrix models the phase shifting network which is only able to adjust the phase of the incoming signals, not the amplitude [8]. Note that the power constraint at the TX is satisfied by  $\|\mathbf{F}_{RF} \mathbf{F}_{BB}\|_F^2 = P_{\max}$ . The matrices  $\mathbf{W}_{BB} \in \mathbb{C}^{L_R \times N_s}$  and  $\mathbf{W}_{RF} \in \mathbb{C}^{N_R \times L_R}$  denote the digital combiner and the analogue combiner at the RX, respectively. The analogue combiner matrix is also constant modulus.

We assume the channel state information (CSI) to be known at both the TX and the RX. Then the signal received at the RX antennas  $\mathbf{y} \in \mathbb{C}^{N_R \times 1}$  can be written as

$$\mathbf{y} = \mathbf{H} \mathbf{F}_{RF} \mathbf{F}_{BB} \mathbf{s} + \mathbf{n}, \quad (2)$$

where  $\mathbf{n} \in \mathbb{C}^{N_R \times 1} = \mathcal{CN}(0, \sigma_n^2)$  represents independent and identically distributed complex additive noise. After the analogue combiner and digital combiner units, the RX output signal can be expressed as

$$\mathbf{r} = \mathbf{W}_{BB}^H \mathbf{W}_{RF}^H \mathbf{y} = \mathbf{W}_{BB}^H \mathbf{W}_{RF}^H \mathbf{H} \mathbf{F}_{RF} \mathbf{F}_{BB} \mathbf{s} + \mathbf{W}_{BB}^H \mathbf{W}_{RF}^H \mathbf{n}. \quad (3)$$

The mechanism to select only required number of RF chains  $L_T^{opt}$  out of the available  $L_T$  RF chains is implemented during the baseband processing. The proposed DM based solution drives this selection mechanism, which uses dynamic power allocation to decide on how many RF chains should be active

during each channel realization. In the next section, we derive a fractional programming problem from the problem which maximizes EE and implement the Dinkelbach's approach to obtain the number of RF chains optimally at the TX/RX.

### III. OVERVIEW OF EE MAXIMIZATION

In terms of the SE  $R$  (bits/s/Hz) and the power consumption  $P$  (W), the EE can be written as

$$\text{EE}(\mathbf{P}_{\text{TX}}) \triangleq \frac{R(\mathbf{P}_{\text{TX}})}{P(\mathbf{P}_{\text{TX}})} \quad (\text{bits/Hz/J}). \quad (4)$$

In (4),  $\mathbf{P}_{\text{TX}} \in \mathcal{D}^{L_T \times L_T}$  represents a square matrix whose diagonal entries contain the transmission power of each data stream at the output of the digitally-computer precoder matrix, while all non-diagonal entries are zero. The notation  $\mathcal{D}^{L_T \times L_T} \subset \mathbb{R}^{L_T \times L_T}$  represents the set of possible choices for  $L_T \times L_T$  matrices, given the existence of a maximum transmit power constraint.

In order to represent the selection mechanism for RF chains at the digital precoder, we consider  $[\mathbf{P}_{\text{TX}}]_{kk} \in [0, P_{\max}] \forall k = 1, \dots, L_T$ . The diagonal entries of  $\mathbf{P}_{\text{TX}}$  with a zero value means an open switch in the selection mechanism shown in Fig. 1. This means that the non-zero diagonal entries of the matrix  $\mathbf{P}_{\text{TX}}$  determine the number of the active RF chains currently selected at the TX side, i.e.,  $L_T^{\text{opt}} = \|\mathbf{P}_{\text{TX}}\|_0$ . We may achieve high SE by increasing the number of RF chains, however, it increases power consumption as well. Thus, maximizing EE in (4) given suitable constraints on the solution provides us with a practical method for selecting the TX/RX configuration with the best performance trade-off.

The optimization problem in (4) has inspired us to study several different approaches to optimize the performance of a mmWave hybrid MIMO transceiver. As shown in Fig. 2, we deal with two phases in a single communication frame where we assume that at the start of each data frame, a beam training phase provides information to both the TX and RX about the current channel matrix  $\mathbf{H}$  and there are  $L_T$  active RF chains. Based on this knowledge it is possible to adapt the behaviour of the TX and RX before the main data communication phase, where in this paper, the DM based solution is applied to activate only required number of RF chains, i.e.,  $L_T^{\text{opt}}$ , which is obtained from the solution of EE maximization problem. In the process, the HBF matrices can be designed through an Euclidean distance minimization problem [8] as discussed in the next section and we also propose a low complexity alternative to design the HBF matrices. Next, we discuss the approaches which we implemented to adapt the behaviour of the TX and RX in order to achieve maximum EE.

1) **RF Chain Selection:** In Fig. 1, the analogue precoder and the analogue combiner may connect every RF chain to every TX/RX antenna, which is termed as a *fully-connected* structure. Alternatively, in a structure which is termed as *partially-connected*, each RF chain may only be connected to a subset of all the antennas. In the latter case, we have explored an optimization technique to select the best set of RF chains for data transmission in [20]. A key feature of this



Fig. 2: Single communication frame with two phases process: beam training and data communications.

approach is that we use a low signal-to-noise ratio (SNR) approximation of the data rate to simplify the optimization approach. A sparse solution for the RF chains is desired and this is obtained by minimizing the number of non-zero entries in the matrix  $\mathbf{P}_{\text{TX}}$ . This is achieved practically by using a technique called convex relaxation which allows the optimization to be performed efficiently. However, there is lack of research in literature dealing with the selection of RF chains. In a hardware setup, whether its fully-connected or partially-connected, when HBF is implemented on a field-programmable gate array (FPGA) chip, switching on only the needed RF chains would save a lot of power leading to an energy efficient communication system. Following that approach, in [18] we consider a fully-connected structure (as shown in Fig. 1) and the Dinkelbach's approach selects only that number of RF chains which maximizes EE and the complexity is kept minimum. More details of this approach are presented in Section IV below.

2) **Sampling Rate Selection:** A number of papers recently have shown that using limited resolution digital-to-analogue or analogue-to-digital converters in the TX or RX can improve communications efficiency [21]. The reason for this is that the power consumed by a sampling device scales in an exponential manner with the number of quantization bits that are used. The limitation of using limited resolution sampling is that it can limit the overall data rate at high SNR values. However, limited resolution sampling can be particularly attractive for low or medium SNR values where the SE is lower. Reference [14] extends the RF chain selection approach of [18] to the case where the TX uses the fully-connected structure and each RF chain uses fixed resolution DACs at the TX. In that paper, a linear model is used to describe the impact of quantization, through a scaling factor and the addition of a noise term which represents the quantization noise. Similarly, the partially-connected case is with limited resolution sampling studied in [20]. We have recently extended this work to consider the joint optimization of both the HBF matrices design and the bit level resolution of each RF chain [15], [16]. This involves a complex model where the effect of the quantization noise on the data throughput is explicitly modelled and the bit level resolution can be adjusted to optimize the resulting EE. We introduce a novel matrix decomposition that is applied to the HBF matrices at both the TX and RX, i.e., the joint decomposition of a matrix representing analogue beamforming matrix, a second matrix modelling the impact of bit resolution on receiver noise and a third matrix that models digital baseband beamforming. Moreover, we address the joint TX-RX problem

unlike in the existing literature and the optimization approach we follow requires the use of the alternating direction method of multipliers to find the best solution for both the HBF matrices and the required bit resolutions at the TX and RX in order to maximize EE.

Next, we describe the Dinkelbach's approach for selecting the number of RF chains optimally and show how this leads to a low-cost solution to EE maximization.

#### IV. RF CHAIN SELECTION FOR MAXIMUM EE

##### A. RF Chain Selection Formulation

For MIMO with HBF and point-to-point communication, the SE  $R$  given the active number of RF chains is

$$R(\mathbf{P}_{\text{TX}}, \mathbf{P}_{\text{RX}}) = \log \left| \mathbf{I}_{N_s} + \frac{1}{\sigma_n^2} \mathbf{W}_{\text{BB}}^H \mathbf{P}_{\text{RX}}^{\frac{1}{2}} \mathbf{W}_{\text{RF}}^H \mathbf{H} \mathbf{F}_{\text{RF}} \times \right. \\ \left. \mathbf{P}_{\text{TX}}^{\frac{1}{2}} \hat{\mathbf{F}}_{\text{BB}} \hat{\mathbf{F}}_{\text{BB}}^H \mathbf{P}_{\text{TX}}^{\frac{1}{2}} \mathbf{F}_{\text{RF}}^H \mathbf{H}^H \mathbf{W}_{\text{RF}} \mathbf{P}_{\text{RX}}^{\frac{1}{2}} \mathbf{W}_{\text{BB}} \right|, \quad (5)$$

where the real valued  $L_T \times L_T$  matrix  $\mathbf{P}_{\text{TX}}$  is the diagonal matrix allocating power at the TX side. At the RX, instead we use the  $L_R \times L_R$  real-valued diagonal matrix  $\mathbf{P}_{\text{RX}}$  with entries from  $\{0, 1\}$ , since this matrix represents the activated RF chains, thus,  $L_R^{\text{opt}} = \|\mathbf{P}_{\text{RX}}\|_0$ .

Following [8], we assume that  $\hat{\mathbf{F}}_{\text{BB}} \hat{\mathbf{F}}_{\text{BB}}^H \approx \mathbf{I}_{L_T}$  and  $\mathbf{W}_{\text{BB}} \mathbf{W}_{\text{BB}}^H \approx \mathbf{I}_{L_R}$ , then the SE can be written as

$$R(\mathbf{P}_{\text{TX}}, \mathbf{P}_{\text{RX}}) = \log \left| \mathbf{I}_{L_R} + \frac{1}{\sigma_n^2} \mathbf{P}_{\text{RX}}^{\frac{1}{2}} \mathbf{W}_{\text{RF}}^H \mathbf{H} \mathbf{F}_{\text{RF}} \times \right. \\ \left. \mathbf{P}_{\text{TX}} \mathbf{F}_{\text{RF}}^H \mathbf{H}^H \mathbf{W}_{\text{RF}} \mathbf{P}_{\text{RX}}^{\frac{1}{2}} \right|. \quad (6)$$

The problem in (6) can be simplified by considering the TX side and the RX side separately. To compute the matrix  $\mathbf{P}_{\text{TX}}$  it is assumed that the RX has activated all its RF chains, so that  $\mathbf{P}_{\text{RX}} = \mathbf{I}_{L_R}$ . In that case, the SE can be expressed as

$$R(\mathbf{P}_{\text{TX}}) = \log \left| \mathbf{I}_{L_R} + \frac{1}{\sigma_n^2} \mathbf{W}_{\text{RF}}^H \mathbf{H} \mathbf{F}_{\text{RF}} \mathbf{P}_{\text{TX}} \mathbf{F}_{\text{RF}}^H \mathbf{H}^H \mathbf{W}_{\text{RF}} \right|. \quad (7)$$

Once the matrix  $\mathbf{P}_{\text{TX}}$  is obtained, the matrix  $\mathbf{P}_{\text{RX}}$  can be computed via the following SE expression:

$$R(\mathbf{P}_{\text{RX}}) = \log \left| \mathbf{I}_{L_R} + \frac{1}{\sigma_n^2} \mathbf{P}_{\text{RX}}^{\frac{1}{2}} \mathbf{W}_{\text{RF}}^H \mathbf{H} \mathbf{F}_{\text{RF}} \times \right. \\ \left. \mathbf{P}_{\text{TX}} \mathbf{F}_{\text{RF}}^H \mathbf{H}^H \mathbf{W}_{\text{RF}} \mathbf{P}_{\text{RX}}^{\frac{1}{2}} \right|. \quad (8)$$

Next, we focus on how to maximize the EE for the TX in order to select the optimal number of RF chains  $L_T^{\text{opt}}$ . The alternative of trying to solve (8) to maximize EE at the RX results leads to a complex integer programming optimization problem. In this paper, we will assume that the number of TX and RX spatial streams are the same, so that  $L_R^{\text{opt}} = L_T^{\text{opt}}$ .

Following [5], the total consumed power  $P$  for a HBF MIMO communication system can be expressed as

$$P = \beta \text{tr}(\mathbf{P}_{\text{TX}}) + 2P_{\text{CP}} + N_T P_T + N_R P_R + L_T^{\text{opt}} \times \\ (P_{\text{RF}} + N_T P_{\text{PS}}) + L_R^{\text{opt}} (P_{\text{RF}} + N_R P_{\text{PS}}) \quad (\text{W}), \quad (9)$$

where the power terms  $P_{\text{CP}}$ ,  $P_{\text{RF}}$ ,  $P_{\text{PS}}$ ,  $P_T$  and  $P_R$  represent the power required by the circuit components, the power required by each RF chain, the power required by each phase shifter, the consumed power for each antenna at the TX and that required for each RX antenna, respectively. The parameter  $\beta$  is the reciprocal of amplifier efficiency.

Let us delete the subscript "TX" from  $\mathbf{P}_{\text{TX}}$  in order to write simplified expressions. Hence, the EE maximization problem in (4) can be expressed with respect to  $\mathbf{P} \in \mathbb{R}^{L_T \times L_T}$  as

$$\max_{\mathbf{P} \in \mathcal{D}^{L_T \times L_T}} \frac{R(\mathbf{P})}{P(\mathbf{P})} \quad \text{s. t. } P(\mathbf{P}) \leq P'_{\text{max}} \quad \& \quad R(\mathbf{P}) \geq R_{\text{min}}. \quad (10)$$

Note that the power constraint in (10) provides an upper limit on the power required for the HBF MIMO communication system, i.e.,  $P'_{\text{max}} = \beta P_{\text{max}} + 2P_{\text{CP}} + N_T P_T + N_R P_R + L_T \times (P_{\text{RF}} + N_T P_{\text{PS}}) + L_R (P_{\text{RF}} + N_R P_{\text{PS}})$ . Next, we proceed with the proposed Dinkelbach's approach to obtain both the number of RF chains and the data streams optimally.

##### B. Dinkelbach's Approach to EE Maximization

In order to obtain a solution to (10) which is a fractional programming problem, we can implement the DM based solution. Dinkelbach's algorithm was first introduced in [22] and it appears to be an efficient algorithm to solve fractional problems. This is verified by the simulation results presented in Section V where we can observe that the Dinkelbach's approach achieves good performance. We can replace the EE ratio in (10) with an iterative sequence of difference-based optimizations as follows:

$$\max_{\mathbf{P}^{(m)} \in \mathcal{D}^{L_T \times L_T}} \left\{ R(\mathbf{P}^{(m)}) - \nu^{(m)} P(\mathbf{P}^{(m)}) \right\} \\ \text{s. t. } P(\mathbf{P}) \leq P'_{\text{max}} \quad \text{and} \quad R(\mathbf{P}) \geq R_{\text{min}}. \quad (11)$$

The DM involves a sequence of iterations where the constant  $\nu^{(m)}$  is updated at each iteration based on the SE and power values estimated during the previous iteration which is equal to the ratio  $R(\mathbf{P}^{(m-1)})/P(\mathbf{P}^{(m-1)}) \in \mathbb{R}^+$ , for  $m = 1, 2, \dots, I_{\text{max}}$ , where  $I_{\text{max}}$  denotes the maximum number of iterations. In order to reduce complexity compared to the BF method, we wish to use a SE expression that does not depend explicitly on the RF and baseband processing matrices. This avoids the need to compute the HBF matrices each time the number of selected RF chains is updated.

In order to proceed with the DM based solution, let us first update the SE and power expressions. For that, we consider channel's singular value decomposition (SVD) as  $\mathbf{H} = \mathbf{U}_H \mathbf{\Sigma}_H \mathbf{V}_H^H$ , where  $\mathbf{U}_H \in \mathbb{C}^{N_R \times N_R}$  and  $\mathbf{V}_H \in \mathbb{C}^{N_T \times N_T}$  are unitary matrices, and  $\mathbf{\Sigma}_H \in \mathbb{R}^{N_R \times N_T}$  represents a matrix which is rectangular in nature where the diagonal entries contain the singular values of the channel matrix and all the other entries are zero. Considering the SVD of the channel, (7) is written as

$$R(\mathbf{P}) = \log \left| \mathbf{I}_{N_R} + \frac{1}{\sigma_n^2} \mathbf{W}_{\text{RF}}^H \mathbf{U}_H \mathbf{\Sigma}_H \mathbf{V}_H^H \mathbf{F}_{\text{RF}} \times \right. \\ \left. \mathbf{P} \mathbf{F}_{\text{RF}}^H \mathbf{V}_H \mathbf{\Sigma}_H^H \mathbf{U}_H^H \mathbf{W}_{\text{RF}} \right|. \quad (12)$$

Using the approach given in [8], it can be shown that  $\mathbf{V}_H^H \mathbf{F}_{\text{RF}} \approx [\mathbf{I}_{L_T} \mathbf{0}_{(N_T-L_T) \times L_T}^T]^T$  and  $\mathbf{U}_H^H \mathbf{W}_{\text{RF}} \approx [\mathbf{I}_{L_R} \mathbf{0}_{(N_R-L_R) \times L_R}^T]^T$ , hence,

$$R(\mathbf{P}) = \log \left| \mathbf{I}_{N_R} + \frac{1}{\sigma_n^2} \bar{\Sigma}^2 \mathbf{P} \right|, \quad (13)$$

where the  $L_R \times L_T$  matrix  $\bar{\Sigma}$  has diagonal entries  $[\bar{\Sigma}]_{kk} = [\Sigma_H]_{kk}$  for  $k = 1, \dots, L_T$ , assuming  $L_T = L_R$ . Again, the remaining entries of this matrix are zero. In (13) all of the matrices are diagonal, so it is possible to decompose the SE calculation into  $L_T$  parallel and orthogonal channels as

$$R(\mathbf{P}) \approx \sum_{k=1}^{L_T} \log \left( 1 + \frac{1}{\sigma_n^2} [\bar{\Sigma}^2]_{kk} [\mathbf{P}]_{kk} \right) \quad (\text{bits/s/Hz}). \quad (14)$$

The number of available RF chains at the TX  $L_T$  and at the RX  $L_R$  are determined by the hardware setup of the transceiver. For the TX side, the power values in the matrix  $\mathbf{P}$  can be written as

$$P_{\text{TX}}(\mathbf{P}) = P_{\text{static}} + \sum_{k=1}^{L_T} (\beta [\mathbf{P}]_{kk} + P_{\text{RF}} + N_T P_{\text{PS}}) \quad (15)$$

$$\implies P_{\text{TX}}(\mathbf{P}) = P_{\text{static}} + \sum_{k=1}^{L_T} \beta' [\mathbf{P}]_{kk} \quad (\text{W}), \quad (16)$$

where the value of  $P_{\text{static}} \triangleq P_{\text{CP}} + N_T P_T$  does not depend on the entries of the matrix  $\mathbf{P}$  and  $\beta' \triangleq \beta + \frac{P_{\text{RF}} + N_T P_{\text{PS}}}{P_{\text{max}}}$ . Simplifying (15) into the form given in (16) is possible as  $\sum_{k=1}^{L_T} [\mathbf{P}]_{kk} = \text{tr}(\mathbf{P}) = P_{\text{max}}$ .

Following (14)-(16), the  $m$ -th DM step can be written as

$$\begin{aligned} \{\mathbf{P}^{(m)}, \nu^{(m)}\} &= \arg \max_{\mathbf{P}^{(m)} \in \mathcal{D}^{L_T \times L_T}} \mathcal{G}(\mathbf{P}^{(m)}, \nu^{(m)}), \\ \text{s. t. } P(\mathbf{P}) &\leq P'_{\text{max}} \text{ and } R(\mathbf{P}) \geq R_{\text{min}}, \end{aligned} \quad (17)$$

where  $\mathcal{G}(\mathbf{P}^{(m)}, \nu^{(m)}) \triangleq \sum_{k=1}^{L_T} \log \left( 1 + \frac{1}{\sigma_n^2} [\bar{\Sigma}^2]_{kk} [\mathbf{P}^{(m)}]_{kk} \right) - \nu^{(m)} \sum_{k=1}^{L_T} \beta' [\mathbf{P}^{(m)}]_{kk}$ . Note that (17) is generally not convex given the constraint associated with  $\mathbf{P}^{(m)}$ , i.e.,  $\mathbf{P}^{(m)} \in \mathcal{D}^{L_T \times L_T}$ . Indeed, in the case where the set  $\mathcal{D}$  also contains the zero value, the problem (17) is a mixed-integer programming one. To proceed, we alleviate this constraint on  $\mathbf{P}^{(m)}$  first, so that (17) can be solved using a standard interior-point method, e.g., using CVX [23]. A theoretical analysis of DM convergence is presented in [26].

In order to explain the steps of Algorithm 1, it begins with the maximum number of RF chains  $L_T$ . Step 4 shows that we solve (17) to update  $\mathbf{P}^{(m)}$  using CVX after alleviating the constraint as mentioned above. Then we apply the constraint again as highlighted in Step 5 of Algorithm 1. This is achieved by setting the values  $\mathbf{P}^{(m)}$  to zero when they fall below the tolerance value  $\epsilon_{\text{th}}$  (see Table II for  $\epsilon_{\text{th}}$  value). Step 6 shows that counting the non-zero values of  $\mathbf{P}_{\text{th}}^{(m)}$  determines the number of activated RF chains. The DM method keeps updating these values within the loop and finally computes  $\|\mathbf{P}_{\text{th}}^{(m)}\|_0$  when the loop ends. Step 7 determines the SE  $R(\mathbf{P}^{(m)})$  and the power  $P_{\text{TX}}(\mathbf{P}^{(m)})$ , and in Step 8  $\mathcal{G}(\mathbf{P}^{(m)}, \nu^{(m)})$  is

---

### Algorithm 1 Dinkelbach Method (DM)

---

- 1: **Initialize:**  $\mathbf{P}^{(0)}$ , choose tolerance  $\epsilon$ ,  $L_T$  and set  $\nu^{(0)}$  with  $\mathcal{G}(\mathbf{P}^{(0)}, \nu^{(0)}) \geq 0$ .
  - 2: Start Iteration Step  $m = 0$ .
  - 3: **while**  $|\mathcal{G}(\mathbf{P}^{(m)}, \nu^{(m)})| > \epsilon$  **do**
  - 4: Alleviate the constraint on  $\mathbf{P}^{(m)}$  and solve (17).
  - 5: Threshold the entries of  $\mathbf{P}^{(m)} \rightarrow$  obtain  $\mathbf{P}_{\text{th}}^{(m)}$ .
  - 6: Count non-zero values of  $\mathbf{P}_{\text{th}}^{(m)} \rightarrow$  update  $L_T^{\text{opt}}$ .
  - 7: Calculate  $R(\mathbf{P}^{(m)})$  and  $P_{\text{TX}}(\mathbf{P}^{(m)})$  using (14)-(16).
  - 8: Compute  $\mathcal{G}(\mathbf{P}^{(m)}, \nu^{(m)})$ .
  - 9: Update the value  $\nu^{(m)}$  as  $R(\mathbf{P}^{(m)})/P_{\text{TX}}(\mathbf{P}^{(m)})$ .
  - 10: Update  $m = m + 1$  for next iteration.
  - 11: **end while**
  - 12: Compute  $L_T^{\text{opt}}$  as the value  $\|\mathbf{P}_{\text{th}}^{(m)}\|_0$ .
- 

computed based on its given expression above, where  $\nu^{(m)} = R(\mathbf{P}^{(m-1)})/P(\mathbf{P}^{(m-1)}) \in \mathbb{R}^+$ . Step 9 is used to update  $\nu^{(m)}$  according to the current value  $R(\mathbf{P}^{(m)})/P_{\text{TX}}(\mathbf{P}^{(m)})$ . The loop terminates when  $|\mathcal{G}(\mathbf{P}^{(m)}, \nu^{(m)})|$  is lower than the specified value  $\epsilon$ , which is determined empirically (see Table II for  $\epsilon$  value). The number of spatial streams is then set to be equal to the optimal number of RF chains, i.e.,  $N_s = L_T^{\text{opt}}$ .

Once we obtain  $L_T^{\text{opt}}$ ,  $L_R^{\text{opt}} (= L_T^{\text{opt}})$  and  $N_s$ , we can design the HBF matrices  $\mathbf{F}_{\text{RF}}$ ,  $\mathbf{F}_{\text{BB}}$ ,  $\mathbf{W}_{\text{RF}}$  and  $\mathbf{W}_{\text{BB}}$ . We assume that as in [8], the matrices  $\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}$  can be designed to yield a good approximation of the fully digital precoder  $\mathbf{F}_{\text{DBF}}$ . Note that the precoder matrix  $\mathbf{F}_{\text{DBF}} = \mathbf{V}_{\text{H1}} \mathbf{P}_{\text{TX}}^{(1/2)}$  where the matrix  $\mathbf{V}_{\text{H1}} \in \mathbb{C}^{N_T \times N_s}$  consists of the  $N_s$  columns of the matrix  $\mathbf{V}_{\text{H}}$  which contains the right singular eigenvectors [8] with  $\|\mathbf{F}_{\text{DBF}}\|_F^2 = \text{tr}(\mathbf{P}_{\text{TX}}) = P_{\text{max}}$ . Following [8], the problem to compute the hybrid precoder decomposition  $\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}$  through Euclidean distance minimization can be transformed to a sparse approximation problem. To solve that, we use gradient pursuit (GP) algorithm [24] which is implemented as an alternative to the most commonly used orthogonal matching pursuit (OMP) algorithm for HBF design. The GP algorithm has same performance as the OMP algorithm, but it uses only one matrix vector multiplication per iteration to avoid matrix inversion, leading to faster approximation and low complexity [9]. At the RX, the hybrid combiner can be designed with a similar mathematical formulation as at the TX except there is no power constraint. Following the steps in [8], we compute the fully digital combiner matrix  $\mathbf{W}_{\text{DBF}}$  and the Euclidean distance minimization problem for the combiner design is transformed to the sparse approximation problem likewise at the TX. The sparse approximation problem at the RX can then be solved by the GP algorithm [9] in order to obtain the hybrid combiner decomposition  $\mathbf{W}_{\text{RF}} \mathbf{W}_{\text{BB}}$ .

*Computational Complexity:* The computation for the DM based solution requires only  $\mathcal{O}(L_T^{\text{opt}})$  operations per iteration. The complexity comparison with the BF approach is provided in Section V. The complexity order in computing beamforming weights for the GP algorithm is  $\mathcal{O}((L_T^{\text{opt}})^3 N_T)$  and for the OMP algorithm equals  $\mathcal{O}((L_T^{\text{opt}})^4) + \mathcal{O}((L_T^{\text{opt}})^3)$  - the GP

System Parameter	Value
Number of clusters	$N_{cl} = 2$
Number of rays	$N_{ray} = 10$
Angular spread	$7.5^\circ$
Average power for each cluster	$\sigma_{\alpha, i} = 1$
Mean angles (azimuth domain)	$60^\circ - 120^\circ$
Mean angles (elevation domain)	$80^\circ - 100^\circ$
Normalized system bandwidth	1 Hz
SNR	$1/\sigma_n^2$
Amplifier efficiency	$1/\beta = 0.4$
Minimum desired SE in (10)	$R_{min} = 1$ bits/s/Hz
Tolerance values	$\epsilon = 10^{-4}$ and $\epsilon_{th} = 10^{-6}$
Number of available RF chains	$L_T = L_R = \text{length}(\text{eig}(\mathbf{H}\mathbf{H}^H))$
Spacing between antenna elements	$d = \lambda/2$ (e.g., $\lambda = 1/28$ GHz [13])

(a) Values of the system parameters.

Power Term	Value
Power required by all circuit components	$P_{CP} = 10$ W
Power required by each RF chain	$P_{RF} = 100$ mW
Power required by each phase shifter	$P_{PS} = 10$ mW
Power per TX/RX antenna element	$P_T = P_R = 100$ mW
Maximum allocated power	$P_{max} = 1$ W

(b) Values of the power terms in (9) [25].

TABLE II: Values of the system parameters and power terms used in the simulations.

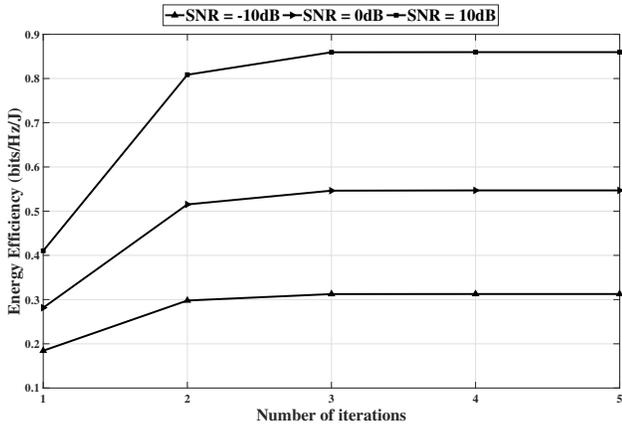


Fig. 3: EE versus number of iterations at  $N_T = 32$ ,  $N_R = 8$ ,  $N_{cl} = 2$ ,  $N_{ray} = 10$  and  $P_{max} = 16$  W.

method only makes use of matrix multiplies at each step. This reduction in complexity comes from using a gradient computation in place of a full matrix inverse calculation. Reference [9] provides a more detailed complexity comparison. Next, we present simulation results that verify the good performance of the proposed Dinkelbach approach.

## V. SIMULATION RESULTS

This section evaluates the performance of the proposed DM based solution and compares it with existing baseline cases. All results have been averaged over 1,000 Monte-Carlo realizations. In terms of the system setup, Table II (a) provides the values of all the system parameters and Table II (b) provides the values used in the simulations for the power terms in (9).

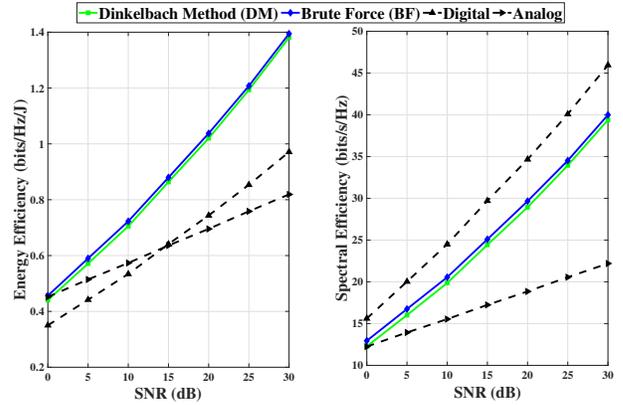


Fig. 4: EE and SE versus SNR at  $N_T = 32$ ,  $N_R = 8$ ,  $N_{cl} = 2$ ,  $N_{ray} = 10$  and  $P_{max} = 1$  W.

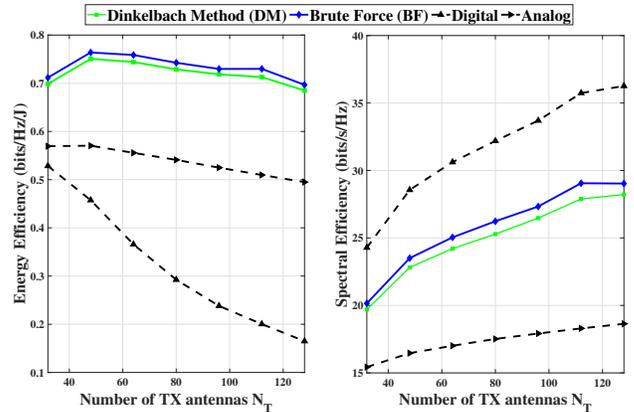


Fig. 5: EE and SE versus  $N_T$  at SNR = 10 dB,  $N_R = 8$ ,  $N_{cl} = 2$ ,  $N_{ray} = 10$  and  $P_{max} = 1$  W.

For comparison with the proposed DM based solution, following baseline cases have been considered in this paper.

1) *BF Approach*: The exhaustive search based approach in [13], i.e., the BF approach, at each realization (current channel realization), computes the EE performance by designing the beamforming matrices for each possible choice of activated RF chains, namely  $L_T = \{1, 2, \dots, N_T\}$ , and then chooses the corresponding number of RF chains corresponding to the highest EE value. In contrast, the proposed DM based solution does not need to iterate for all possible number of RF chains and then find a number of RF chains which is optimal, which reduces the complexity significantly while providing high energy efficient solution. The complexity order of the BF approach is related the number of RF chains multiplied by the total number of antennas, i.e.,  $\mathcal{O}(L_T^{opt} N_T)$  which is larger than that of the DM based solution that only requires  $\mathcal{O}(L_T^{opt})$  operations per iteration. In simulation, the BF and DM approaches uses the same HBF matrix computation.

2) *Digital Beamforming*: As mentioned above, the full digital beamforming baseline allocates one active RF chain for each antenna in all simulations, i.e.,  $L_T = N_T$  and  $L_R = N_R$ .

3) *Analogue Beamforming*: In this case, analogue beamforming only implements one active RF chain, i.e.,  $L_T = L_R = 1$ , and the HBF decomposition matrices are designed equal to phases of the first singular vectors.

Fig. 3 graphs the EE performance versus the number of iterations for SNR values of  $-10$ ,  $0$  and  $10$  dB to observe convergence of the proposed DM based solution at  $N_T = 32$ ,  $N_R = 8$ ,  $N_{cl} = 2$ ,  $N_{ray} = 10$  and  $P_{max} = 16$  W. The DM based solution converges rapidly, requiring typically about two iterations to achieve an optimal solution at each channel realization. Also, the achieved EE results increase with the SNR value, for example, after 2 iterations, the EE value at  $10$  dB SNR is  $\approx 0.55$  bits/Hz/J higher than that for  $-10$  dB SNR and  $\approx 0.3$  bits/Hz/J higher than the result for  $0$  dB SNR.

Fig. 4 shows the EE and SE performance of the DM method along with the BF approach, and both the analogue and digital baseline cases versus SNR with  $N_T = 32$ ,  $N_R = 8$ ,  $N_{cl} = 2$ ,  $N_{ray} = 10$  and  $P_{max} = 1$  W. We can observe that the DM based solution has similar EE and SE performance to the BF approach, achieving a much higher EE than the digital baseline case, and higher EE and SE results compared to the analogue baseline. At an SNR value of  $20$  dB, the DM based solution yield  $\approx 0.2$  bits/Hz/J higher EE than the digital baseline case, and  $\approx 10$  bits/s/Hz higher SE and about  $0.3$  bits/Hz/J higher EE than the analogue baseline case.

Fig. 5 shows the EE and SE performance versus the number of TX antennas,  $N_T$ , plotted for an SNR of  $10$  dB,  $N_R = 8$ ,  $N_{cl} = 2$ ,  $N_{ray} = 10$  and  $P_{max} = 1$  W. It is clear that as the number of antennas increases, the EE results start to decrease for both the proposed DM based solution and the existing baseline cases. For example, at  $N_T = 80$ , the EE and SE performance of the DM based solution is similar to that of the BF method. Also, the DM based solution has  $\approx 0.42$  bits/Hz/J higher EE than the digital baseline case, and  $\approx 7.5$  bits/s/Hz higher SE and about  $0.2$  bits/Hz/J higher EE than the analogue baseline case.

## VI. CONCLUSION

This paper has discussed the concept of adaptive HBF MIMO systems that adapt their behaviour on a frame-by-frame basis to optimize EE. In particular, a DM based solution has been studied to enable fractional programming to maximize the EE of the candidate transmitter and receiver architectures in a low-cost manner. The DM method described in this paper can achieve EE and SE performance similar to the exhaustive search based BF approach, while reducing the complexity significantly. Once the number of RF chains is selected, the proposed technique needs to compute the HBF matrices only once. Further, the DM solution can also provide significantly improved EE performance when compared with the existing baseline cases, e.g., at  $10$  dB SNR, it performs  $\approx 20\%$  better than the digital beamforming baseline and  $\approx 15\%$  better than the analogue beamforming case. Finally it is shown that the GP algorithm, which is used to compute the HBF matrices, is a faster and less complex algorithm in comparison to the state-of-the-art OMP algorithm.

## ACKNOWLEDGEMENT

The Engineering and Physical Sciences Research Council Grant EP/P000703/1 supported this research work.

## REFERENCES

- [1] NetWorld2020 White Paper on Research Beyond 5G, pp. 1-43, Oct. 2015.
- [2] Cisco visual networking index forecast 2017-22.
- [3] Ericsson Mobility Report, pp. 1-32, Nov. 2018.
- [4] S. Rangan et al., "Millimeter-wave cellular wireless networks: potentials and challenges," *Proc. IEEE*, vol. 102, no. 3, pp. 366385, Mar. 2014.
- [5] S. Han et al., "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 186-194, Jan. 2015.
- [6] R. W. Heath et al., "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE Journ. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436-453, Apr. 2016.
- [7] O. E. Ayach et al., "The capacity optimality of beam steering in large millimeter wave MIMO systems," *IEEE Int. Workshop Signal Process. Advances Wireless Commun. (SPAWC)*, pp. 100-104, June 2012.
- [8] O. E. Ayach et al., "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499-1513, Mar. 2014.
- [9] A. Kaushik et al., "Sparse hybrid precoding and combining in millimeter wave MIMO systems," in *Proc. IET Radio Prop. Tech. 5G*, Durham, UK, pp. 1-7, Oct. 2016.
- [10] C. G. Tsinos et al., "Hybrid analog-digital transceiver designs for multi-user MIMO mmWave cognitive radio systems," *IEEE Trans. Cognitive Commun. Netw.*, accepted, Aug. 2019.
- [11] A. Li and C. Masouros, "Hybrid analog-digital millimeter-wave mmimo transmission with virtual path selection," *IEEE Commun. Letters*, vol. 21, no. 2, pp. 438-441, Feb. 2017.
- [12] S. Payami et al., "Hybrid beamforming with a reduced number of phase shifters for massive mimo systems," *IEEE Trans. Veh. Tech.*, vol. 67, no. 6, pp. 4843-4851, June 2018.
- [13] R. Zi et al., "Energy efficiency optimization of 5G radio frequency chain systems," *IEEE Journ. Sel. Areas Commun.*, vol. 34, no. 4, pp. 758-771, Apr. 2016.
- [14] A. Kaushik et al., "Energy efficiency maximization of millimeter wave hybrid MIMO systems with low resolution DACs," *IEEE Int. Conf. Commun. (ICC)*, Shanghai, China, pp. 1-6, May 2019.
- [15] A. Kaushik et al., "Energy efficient bit allocation and hybrid combining for millimeter wave MIMO systems," *IEEE Global Commun. Conf. (GLOBECOM)*, Hawaii, USA, pp. 1-6, Dec. 2019.
- [16] A. Kaushik et al., "Joint bit allocation and hybrid beamforming optimization for energy efficient millimeter wave MIMO systems," *arXiv:1910.01479*, Oct. 2019.
- [17] E. Bjornson et al., "Optimal design of energy-efficient multi-user MIMO systems: is massive MIMO the answer?," *IEEE Trans. Wireless Commun.*, vol. 14, no.6, pp. 3059-3075, Jun. 2015.
- [18] A. Kaushik et al., "Dynamic RF chain selection for energy efficient and low complexity hybrid beamforming in millimeter wave MIMO systems," *IEEE Trans. Green Commun. Netw.*, accepted, July 2019.
- [19] S. Singh et al., "Interference analysis for highly directional 60-GHz mesh networks: the case for rethinking medium access control," *IEEE/ACM Trans. Netw.*, vol. 19, no. 5, pp. 1513-1527, Oct. 2011.
- [20] E. Vlachos et al., "Energy efficient transmitter with low resolution DACs for massive MIMO with partially connected hybrid architecture," *IEEE Veh. Tech. Conf. (VTC)-Spring*, Porto, Portugal, pp. 1-5, June 2018.
- [21] O. Orhan et al., "Low power analog-to-digital conversion in millimeter wave systems: Impact of resolution and bandwidth on performance," *Info. Theory Appl. Workshop (ITA)*, San Diego, USA, pp. 191-198, 2015.
- [22] W. Dinkelbach, "On nonlinear fractional programming," *Management Science*, vol. 13, no. 7, pp. 492-498, Mar. 1967.
- [23] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs", in *Recent Adv. Learning and Control*, Springer-Verlag Ltd., pp. 95-110, 2008.
- [24] T. Blumensath, and M. E. Davies, "Gradient pursuits", *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2370-2382, June 2008.
- [25] T. S. Rappaport et al., "Millimeter wave wireless communications," *Prentice-Hall*, Sept. 2014.
- [26] A. Zappone and E. Jorswieck, "Energy Efficiency in Wireless Networks via Fractional Programming Theory," in *Energy Efficiency in Wireless Networks via Fractional Programming Theory*, Now Foun. Trends, 2015.