Energy-Efficiency Maximization of Hybrid Massive MIMO Precoding with Random-Resolution DACs via RF Selection

Energy-Efficiency Maximization of Hybrid Massive MIMO Precoding with Random-Resolution DACs via RF Selection

Evangelos Vlachos, and John Thompson, *Fellow, IEEE.*

***Abstract*—Energy-efficiency (EE) is identified as a key 5G metric and will have a major impact on the hybrid beamforming system design. The most promising system designs include a reduced number of radio-frequency (RF) chains with digital-to-analog converters (DACs) of lower sampling resolution. However, naive reduction of beamformer components to reduce power consumption typically leads to significant loss of spectral-efficiency (SE). In this paper, we focus on the transmit beamforming (precoding) and we introduce an architecture with low-end components that maximizes the EE while minimizing the effects on SE. This is achieved by the novel design of the analog part of the precoder, where the number of the RF chains is not reduced a priori, but deactivated based on an optimization algorithm. Thus, the problem becomes a subset selection one, where only the RF chains with the optimal SE-EE performance are being activated. The selection algorithm not only determines the optimal number of RF chains to activate but also selects optimally between DACs of randomly-allocated resolution. Through simulations, we verify that the proposed architecture exhibits improved performance when compared with baseline precoding techniques which use a predefined number of RF chains with low-resolution DACs.**

***Index Terms*—Energy-efficiency maximization, low-resolution digital-to-analog converter (DAC), millimeter wave (mmWave), massive MIMO, hybrid beamforming.**

I. Introduction

The conventional microwave frequency spectrum between the range of 300 MHz and 10 GHz is increasingly occupied which highlights the need to use the available spectrum for future wireless communication systems. Adopting millimeter-wave (mmWave) spectrum can resolve this issue [1], [2] and provide improved rate performance [3]. Some more advantages of using a mmWave frequency band are increased capacity, improved coverage, lower latency, high mobility and reliability, and lower infrastructure costs [4]–[6]. However, the propagation characteristics associated with a mmWave spectrum induce some challenges such as high path loss. This can be compensated by using large-scale antenna arrays leading to a massive multiple-input multiple-output (MIMO) system. The use of large-scale antennas and wide bandwidths at mmWave MIMO systems makes it hard to implement one radio frequency (RF) chain and associated digital-to-analog/analog-to-digital converter (DAC/ADC) components per antenna [7]. The analog-only beamforming approach [8], [9] hardly supports multi-user communication and cannot implement multi-stream communication with a single RF chain only. To provide higher data rates, the 5G New Radio technology [10] will support Hybrid BeamForming (HBF) [11] with up to four spatial streams for communication. This technology capitalizes on both analog and digital signal processing to enable a large number of antenna elements to be connected to a much smaller number of RF chains.

Lately, the high fidelity specifications considering the sampling resolution of the DACs/ADCs have been alleviated, and designs with lower resolution DACs/ADCs are being proposed [12]. In particular, since DACs/ADCs components have large power consumption, lowering their resolution significantly reduces the overall system power consumption, at the expense of the introduced distortion to the transmitted/received signal. The effect of low-resolution ADCs on channel capacity has been studied for MIMO channels in [13]–[15]. An approximated linear model to capture the effect of the low-resolution DACs/ADCs is the additive quantization model (AQNM) [16]. Although it is only a rough approximation of the non-linear effect, due to its simplicity it has been adopted by many authors in the massive MIMO literature [16]–[18].

The idea of a mixed-ADC architecture has been first investigated in [17] for massive MIMO systems. The authors developed an optimal detector to provide a minimum mean-squared error (MMSE) estimate on data symbols with the use of the combination of low and high-resolution ADCs. Reference [16] employs AQNM for the case of a point-to-point mmWave MIMO system, while [18] and [19] for the case of mmWave fading channels. In [20], the energy-efficiency (EE) and spectral efficiency (SE) for low-resolution ADC HBF architecture are studied, showing the advantages of a low-resolution HBF with few RF chains over an infinite resolution ADC structures. In [21], the SE of uplink massive MIMO with low-resolution ADCs are studied, showing that few ADC bits are enough to achieve almost the same SE of unquantized MIMO.

In [22], joint optimization of the ADC resolution and the number of antennas is studied, where a particle swarm optimization algorithm has been proposed to obtain a near-optimal non-uniform bit allocation for the ADCs. The results showed that having optimized, higher than 1-bit bit resolution for the ADCs, is beneficial in terms of EE for the case of MIMO systems. In [23], a closed-form solution has been proposed for a minimum mean square quantization error problem subject to a constraint on the total ADC power, showing that allocating more bits to the RF chain with stronger channel gain is

Evangelos Vlachos is with the Industrial Systems Institute, Athena Research Centre, Patras, Greece. (e-mail: evlachos@isi.gr). John Thompson is with the Institute for Digital Communications, The University of Edinburgh, U.K. (e-mail: j.s.thompson@ed.ac.uk).

beneficial. Theoretical analysis of these results is provided in [24]. For the transmitter side, reference [25] uses low-resolution DACs for a single user MIMO system while [26] employs low-resolution DACs at the base station (BS) for a narrowband multi-user MIMO system. In [27], a single user MIMO system with quantized hybrid transmit beamforming including the RF quantized noise term is considered while evaluating EE and SE performance.

A. Motivation and Contributions

To meet the increased capacity requirements of next-generation mobile networks, novel designs have to be employed. For this reason, even fully-digital massive MIMO beamforming designs for the mmWave frequencies [28] have appeared, despite the high hardware complexity and power demands. However, for power-limited scenarios, e.g., unmanned aerial vehicle (UAV) aided communications, there is a need for novel designs that will provide both high energy and spectral efficiency. Align with this direction are dynamic HBF architectures that will be able to change their properties depending on the operating environment requirements. Steps to this direction are designs with dynamic structure [29]–[31] and dynamic bit allocation [32].

In this paper, we build on the direction of dynamic HBF structures, to jointly maximize the EE and SE of a mmWave massive MIMO system with hybrid precoding. To do so, we introduce a novel precoding architecture, where the RF chains can be activated/deactivated based on an optimization algorithm on a frame-by-frame basis. A preliminary work related to the proposed technique has been presented in [31], where a subset selection approach is introduced to minimize the required power over a constraint of the minimum information rate. However, only equal bit allocation is considered and the fractional EE optimization problem was not solved optimally. In this work, we consider that the implemented RF chains are equipped with DACs of randomly-allocated resolution. Each RF chain is equipped with a low-resolution DAC. The key contributions of this paper can be summarized as follows:

1) We investigate joint EE and SE maximization for the scenario of power-limited BS with a large antenna array.
2) We introduce a novel dynamic analog precoding architecture, which employs a large number of RF chains with DACs of randomly-allocated resolution.
3) We develop a selection algorithm that activates only a subset of the RF chains in an optimal manner.
4) We express the optimization problem via fractional programming and we introduce *random bit allocation* for the DACs to transform it into a subset selection one.

The SE-EE performance of the proposed technique is investigated through extensive simulation results, achieving increased energy efficiency compared to other baseline techniques with predefined DAC resolution and RF chains, as well as an exhaustive search-based approach which finds the optimum number of RF chains and bit resolution which is common for all DACs.

TABLE I
THE NOTATIONS OF THIS PAPER.

$a, \mathbf{a}$ and $\mathbf{A}$	Scalar, vector, and matrix
$j \triangleq \sqrt{-1}$	The imaginary unit
$\mathbf{A}^T$ and $\mathbf{A}^H$	Matrix transpose and Hermitian transpose
$\mathbf{A}^{-1}$ and $\mathbf{A}^\dagger$	Matrix inverse and pseudo-inverse
$[\mathbf{A}]_{i,j}$	Matrix element at the i-th row and j-th column
$[\mathbf{a}]_i$	The i-th vector element
$\mathbf{A}$ and $\hat{\mathbf{A}}$	Actual and estimated matrix
$\mathbf{I}_N$	$N \times N$ identity matrix
$\mathbf{0}_{N\times K}$	$N \times K$ matrix with zeros
$\mathbf{I}_{N\times K}$	Column concatenated matrix $[\mathbf{I}_N \ \mathbf{0}_{N\times K}]$
$\mathbf{\Omega}$	Matrix containing 0's and 1's
$\|\cdot\|_F$	Matrix Frobenius norm
$\times$	Scalar multiplication
$\circ$	Element-wise (Hadamard) matrix product
$\otimes$	Kronecker product
$\mathrm{tr}(\mathbf{A})$	The trace of matrix $\mathbf{A}$
$\mathrm{vec}(\mathbf{A})$	Vectorization of $\mathbf{A}$
$\mathrm{unvec}(\mathbf{A})$	Inverse operation of $\mathrm{vec}(\cdot)$
$\mathrm{diag}(\mathbf{a})$	Diagonal matrix with $\mathbf{a}$ on the main diagonal

B. Notations and Organization

A summary of the notation used throughout this manuscript can be found in Table I.

The paper is organized as follows: Section II presents the channel and system models where the channel model is based on the mmWave channel properties and the system model defines the low-resolution quantization at the transmitter. Section III presents the problem formulation for the proposed technique and the solution to obtain an energy-efficient transmitter. Section IV verifies the proposed technique through the spectral efficiency and energy efficiency plots. Section V sums up the outcomes of the proposed work in the paper.

II. HYBRID MMWAVE MIMO SYSTEM

A. Millimeter wave channel model

Considering the point-to-point mmWave channel of the k-th user, with N_{T} transmitting and $N_{\mathrm{R_k}}$ receiving antennas, the mmWave channel matrix can be written as follows [7]:

$$\mathbf{H}_k = \sqrt{\frac{N_{\mathrm{T}} N_{\mathrm{R_k}}}{N_{\mathrm{cl}} N_{\mathrm{ray}}}} \sum_{\mathrm{i}=1}^{N_{\mathrm{cl}}} \sum_{\mathrm{l}=1}^{N_{\mathrm{ray}}} \alpha_{\mathrm{k,i,l}} \mathbf{a}_{\mathrm{r}}(\phi^r_{\mathrm{k,i,l}}) \mathbf{a}^H_{\mathrm{t}}(\phi^t_{\mathrm{k,i,l}}), \quad (1)$$

with $k = 1, \ldots, K$, N_{cl} is the number of the clusters and N_{ray} is the number of the propagation paths in each cluster for the narrowband channel; $\alpha_{\mathrm{k,i,l}} \in \mathcal{CN}(0, \sigma^2_{\alpha,\mathrm{i}})$ is the gain term with $\sigma^2_{\alpha,i}$ being the average power of the i^{th} cluster. Furthermore, $\mathbf{a}_{\mathrm{t}}(\phi^t_{\mathrm{k,i,l}})$ and $\mathbf{a}_{\mathrm{r}}(\phi^r_{\mathrm{k,i,l}})$ represent the normalized transmit and receive array response vectors, where $\phi^t_{\mathrm{k,i,l}}$ and $\phi^r_{\mathrm{k,i,l}}$ denote the azimuth angles of departure and arrival, respectively. We use uniform linear array antennas for simplification and model the antenna elements at the transmitter as ideal sectored elements. We assume the transmit antenna gains to be unity over the sectors defined by $\phi^t_{\mathrm{k,i,l}} \in [\phi^t_{\min}, \phi^t_{\max}]$. For signal wavelength λ and inter-element spacing d, and considering azimuth angles

of departure, the transmit array response vector can be written as follows [33]:

$$\mathbf{a}_{\mathrm{t}}(\phi^{t}_{\mathrm{k,i,l}}) = \frac{1}{N_{\mathrm{T}}}\left[1, e^{j\frac{2\pi}{\lambda}d\sin(\phi^{t}_{\mathrm{k,i,l}})}, ..., e^{j(N_{\mathrm{T}}-1)\frac{2\pi}{\lambda}d\sin(\phi^{t}_{\mathrm{k,i,l}})}\right]^{T} \tag{2}$$

and similarly for the receive array response vector $\mathbf{a}_{\mathrm{r}}(\phi^{r}_{\mathrm{k,i,l}})$.

Note that, although the narrowband channel is a simplified model for mmWave communications, in this work it is being used to provide a proof of concept for the proposed technique. An investigation of the wideband case is being left as future work.

B. System model

We consider a HBF transmitter (TX) with N_{T} TX antenna elements, sending N_{s} data streams.. This transmitter serves K user terminals at very high data rates, where each terminal has N_{R} receiving antenna elements. In this work, we focus on the design of the TX for EE maximization but also for high SE requirements. To achieve high data rates along with energy efficiency, we assume that the TX is equipped with a large number of RF chains $L_{\mathrm{T}} \leq N_{\mathrm{T}}$ [34]. Each RF chain i, with $i = 1, \ldots, L_{\mathrm{T}}$, has a DAC which transforms the infinite precision digital input to an analog output. Practically quantization in DACs is necessary since the precision of the input values has to be reduced to be transformed into the analog domain. Additionally, to drastically reduce the power consumption, we consider low precision for the DACs at the TX RF chains, i.e., the number of the bits for the quantization of the i-th DAC is $b_i \in [1, 8]$. It is known that by reducing the quantization resolution b at the DAC, the consumed power is exponentially reduced, i.e., $P_{\mathrm{DAC}} \sim e^{b}$. However, this power reduction via low-resolution DACs introduces quantization noise to the transmitted signal, thus, a trade-off design criterion between power efficiency and signal fidelity is obtained.

The TX operation is decomposed into three main parts, as shown in Fig. 1:

1) the baseband precoder $\mathbf{F}_{\mathrm{BB}} \in \mathbb{C}^{L_{\mathrm{T}} \times N_s}$, where N_s is the number of independent transmission streams,
2) the L_{T} RF chains with a tunable resolution DAC at each chain with $\mathcal{Q}(\cdot)$ the quantization function, i.e.,
$$\mathcal{Q}(\cdot) : \mathbb{C}^{L_{\mathrm{T}}\times 1} \rightarrow \{q_1, \ldots, q_M\}^{L_{\mathrm{T}}\times 1}, \tag{3}$$
where $q_j \in \mathbb{R}$, for $j = 1, \ldots, M$, denote the M quantization levels, and
3) the analog transmit precoder $\mathbf{F}_{\mathrm{RF}} \in \mathcal{F}^{N_{\mathrm{T}}\times L_{\mathrm{T}}}$, where $\mathcal{F}^{N_{\mathrm{T}}\times L_{\mathrm{T}}}$ denotes the set of matrices with entries of constant modulus.

At the receiver side, we define the matrix $\mathbf{W}_k \in \mathbb{C}^{N_{\mathrm{R_k}}\times N_s}$ as the receiver combiner of the k-th user. Given that the vector $\mathbf{s} \in \mathbb{C}^{N_s\times 1}$ represents the vector with the independent transmission symbols, then the TX broadcasts the vector

$$\mathbf{t} \triangleq \mathbf{F}_{\mathrm{RF}}\mathcal{Q}(\mathbf{F}_{\mathrm{BB}}\mathbf{s}), \tag{4}$$

with $\mathbf{t} \in \mathbb{C}^{N_{\mathrm{T}}\times 1}$ and $\mathcal{E}\{\mathbf{s}\mathbf{s}^{*}\} = \frac{P_{\mathrm{TX}}}{N_{\mathrm{s}}}\mathbf{I}_{N_s}$, where P_{TX} represents the transmit power. The discrete-time received signal at each user terminal k is expressed as:

$$\mathbf{y}_k = \mathbf{W}_k^{H}\mathbf{H}_k\mathbf{F}_{\mathrm{RF}}\mathcal{Q}(\mathbf{F}_{\mathrm{BB}}\mathbf{s}) + \mathbf{W}_k^{H}\mathbf{n}, \tag{5}$$

where $\mathbf{y}_k \in \mathbb{C}^{N_s\times 1}$ and $\mathbf{n} \in \mathbb{C}^{N_{\mathrm{R_k}}\times 1}$ is a noise vector with entries which are i.i.d. and follow the complex Gaussian distribution, i.e., $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2\mathbf{I}_{N_{\mathrm{R_k}}})$.

The concatenated received signal $\mathbf{y} \in \mathbb{C}^{N_s\times 1}$ for all user terminals, $k = 1, \ldots, K$, can be expressed as:

$$\mathbf{y} = \mathbf{W}^{H}\mathbf{H}\mathbf{t} + \mathbf{W}^{H}\mathbf{n}, \tag{6}$$

where

$$\mathbf{W} \triangleq [\mathbf{W}_1 \cdots \mathbf{W}_K]^{T} \in \mathbb{C}^{N_{\mathrm{R}}\times N_s},$$
$$\mathbf{H} \triangleq [\mathbf{H}_1^{T} \cdots \mathbf{H}_K^{T}]^{T} \in \mathbb{C}^{N_{\mathrm{R}}\times N_{\mathrm{T}}},$$

with $\sum_{k=1}^{K} N_{\mathrm{R_k}} = N_{\mathrm{R}}$.

C. Approximation model for quantization

Let us consider the well-known Bussgang's model [35] for the linear approximation of the introduced distortion of the quantization noise [36]. Given that $Q(\cdot)$ denotes a uniform scalar quantizer then for the scalar input s we have that

$$Q(s) \approx \delta s + \epsilon, \tag{7}$$

where $\delta = \sqrt{1 - \frac{\pi\sqrt{3}}{2}2^{-2b}}$ is the multiplicative quantization parameter for bit resolution equal to b, which characterizes the fidelity of the quantized output [25], [37]. Parameter ϵ is the additive quantization noise with $\epsilon \sim \mathcal{CN}(0, \sigma_\epsilon^2)$. The quantization noise variance σ_ϵ^2 is equal to $\mathcal{E}[|\epsilon|^2]$. Note that $\delta \in (0, 1]$ where the upper bound is achieved at the limit of $b \rightarrow \infty$. Therefore, the higher the bit resolution, the lower distortion (or higher fidelity) is introduced into the quantized values.

Remark: In this work we focus on modeling the nonlinearity effect of DACs. However, the approximation (7) could also include the nonlinearity effect of power amplifiers (PAs) [36]. This extension is left for future work.

Extending this scalar model to the HBF MIMO case we have:

$$Q(\mathbf{F}_{\mathrm{BB}}\mathbf{x}) \approx \mathbf{\Delta}\mathbf{F}_{\mathrm{BB}}\mathbf{x} + \boldsymbol{\epsilon}, \tag{8}$$

where $Q(\mathbf{F}_{\mathrm{BB}}\mathbf{x}) \in \mathbb{C}^{L_{\mathrm{T}}\times 1}$ and $\mathbf{F}_{\mathrm{BB}}$ is the baseband transmit beamforming matrix, $\mathbf{\Delta} \in \mathbb{R}^{L_{\mathrm{T}}\times L_{\mathrm{T}}}$ is a diagonal matrix with values depending on the DAC resolution b_i, for $i = 1, \ldots, L_{\mathrm{T}}$. Specifically, each diagonal entry of $\mathbf{\Delta}$ is given by:

$$[\mathbf{\Delta}]_{ii} = \sqrt{1 - \frac{\pi\sqrt{3}}{2}2^{-2b_i}}. \tag{9}$$

Following the analysis in [27] and [24], the second term of (8) expresses the additive quantization noise for all RF chains, with $\boldsymbol{\epsilon} \in \mathcal{CN}(\mathbf{0}, \mathbf{C}_\epsilon)$ with

$$[\mathbf{C}_\epsilon]_{ii} = \sqrt{1 - \frac{\pi\sqrt{3}}{2}2^{-2b_i}}\sqrt{\frac{\pi\sqrt{3}}{2}2^{-2b_i}}. \tag{10}$$

This leads us to the following linear approximation for the transmitted signal, as seen at the output of the HBF TX:

$$\tilde{\mathbf{t}} \triangleq \mathbf{F}_{\mathrm{RF}}\left(\mathbf{\Delta}\mathbf{F}_{\mathrm{BB}}\mathbf{x} + \boldsymbol{\epsilon}\right) = \mathbf{F}_{\mathrm{RF}}\mathbf{\Delta}\mathbf{F}_{\mathrm{BB}}\mathbf{x} + \mathbf{F}_{\mathrm{RF}}\boldsymbol{\epsilon}, \tag{11}$$

where $\tilde{\mathbf{t}} \in \mathbb{C}^{N_{\mathrm{T}}\times 1}$ is the output signal at the TX based on the linear approximation model for the quantization function.

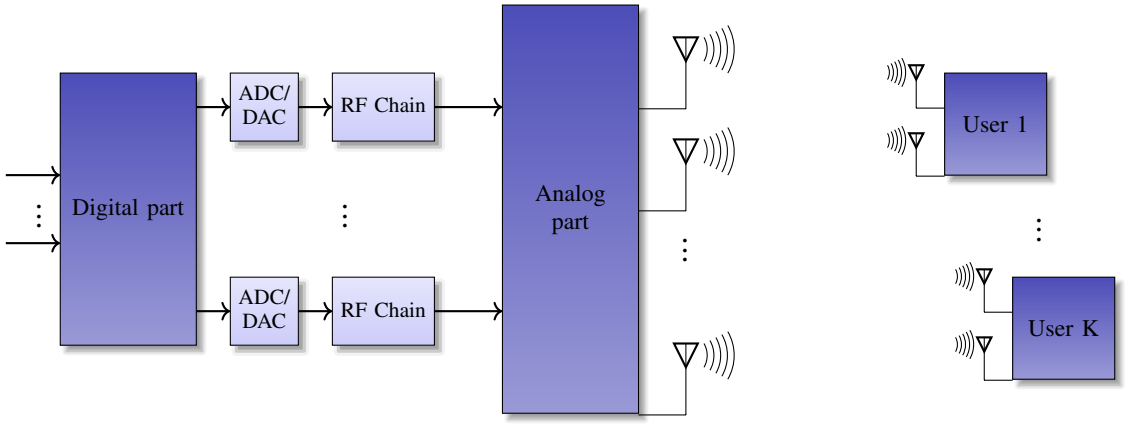

Fig. 1. Block diagram of a hybrid analog/digital (A/D) transmit beamforming MIMO system with N_{T} TX antenna elements and L_{T} number of RF chains. Each user terminal is equipped with N_{R} antennas.

After the effect of the wireless mmWave channel and after the RF processing at the RX, the received signal based on the linear approximation model is given by:

$$\tilde{\mathbf{y}} = \mathbf{W}^H \mathbf{H} \tilde{\mathbf{t}} + \mathbf{W}^H \mathbf{n} \tag{12}$$

$$= \mathbf{W}^H \mathbf{H} \mathbf{F}_{\mathrm{RF}} \mathbf{\Delta} \mathbf{F}_{\mathrm{BB}} \mathbf{x} + \underbrace{\mathbf{W}_k^H \mathbf{H} \mathbf{F}_{\mathrm{RF}} \boldsymbol{\epsilon} + \mathbf{W}_k^H \mathbf{n}}_{\triangleq \boldsymbol{\eta}}, \tag{13}$$

where $\tilde{\mathbf{y}} \in \mathbb{C}^{N_s \times 1}$ and $\boldsymbol{\eta}_k \in \mathbb{C}^{N_s \times 1}$ is the combined effect of the Gaussian and quantization noise for all users, with $\boldsymbol{\eta} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_\eta)$. The Hermitian symmetric matrix $\mathbf{R}_\eta \in \mathbb{C}^{N_s \times N_s}$ is the combined noise covariance matrix with

$$\mathbf{R}_\eta = \mathbf{W}^H \mathbf{H} \mathbf{F}_{\mathrm{RF}} \mathbf{C}_\epsilon \mathbf{C}_\epsilon^H \mathbf{F}_{\mathrm{RF}}^H \mathbf{H}^H \mathbf{W} + \sigma_n^2 \mathbf{W}^H \mathbf{W}. \tag{14}$$

Remark: In this work, we assume perfect CSI at the TX and frequency synchronization (as in, e.g., [23], [24], [27]). Additionally, we consider that the PAs operate in their linear regime, thus, they are not further distorting the TX signal. However, our formulation makes no assumptions on the structure of noise covariance matrix $\mathbf{R}_\eta$. Thus, the effects of phase noise and PAs distortion may be incorporated into the noise term, $\boldsymbol{\eta}$.

When the linear approximation model for the quantization is applied at the transmitting signal, the achievable spectral efficiency (ASE) for k-th user is expressed as function of the fidelity parameter, $\mathbf{\Delta}$, i.e.,

$$R(\mathbf{\Delta}) \triangleq \log_2 |\mathbf{I}_{N_s} + K/N_s \mathbf{R}_\eta^{-1} \mathbf{W}^H \mathbf{H} \mathbf{F}_{\mathrm{RF}} \mathbf{\Delta} \mathbf{F}_{\mathrm{BB}} \mathbf{F}_{\mathrm{BB}}^H \mathbf{\Delta} \mathbf{F}_{\mathrm{RF}}^H \mathbf{H}^H \mathbf{W}| \ \text{(bits/s/Hz)}, \tag{15}$$

where $\mathbf{R}_\eta$ is defined in (14).

D. DACs power consumption model

An exact computation of the power consumption of a HBF TX is a complex task since it depends significantly on the adopted structure and hardware technology [38]. In this work, the focus is on the optimization over the active RF chains L_{T} and the resolution of the corresponding DACs. Thus, the provided framework aims to be generic and applicable to any HBF structure that uses a given number of RF chains.

The power consumption of the i-th RF chain with a DAC of b_i-bits resolution can be expressed as [27]:

$$P_i \triangleq P_{\mathrm{C}} \gamma 2^{b_i}, \tag{16}$$

where γ is a parameter that depends on the sampling frequency and the DAC hardware specifications, while P_{C} is the required power of the other components that exist in the RF chain, e.g., low-power-amplifier, filters. Note that in our analysis we assume that $\gamma = 1$. Given that all RF chains have similar components, the overall power consumption can be expressed as:

$$P_{\mathrm{RF}}(\{b_i\}_i^{L_{\mathrm{T}}}) = \sum_{i=1}^{L_{\mathrm{T}}} P_i = P_{\mathrm{C}} \sum_{i=1}^{L_{\mathrm{T}}} 2^{b_i}, \tag{17}$$

which is a function of the bit resolution b_i for each DAC.

Alternatively, (16) can be expressed based on the i-th diagonal entry of $\mathbf{\Delta}$ of (9), i.e.,

$$P_i = P_{\mathrm{C}} \left(\frac{\pi\sqrt{3}}{2(1 - [\mathbf{\Delta}]_{ii}^2)} \right)^{1/2}, \tag{18}$$

and (17) is expressed as:

$$P_{\mathrm{RF}}(\mathbf{\Delta}) = P_{\mathrm{C}} \sum_{i=1}^{L_{\mathrm{T}}} \left(\frac{\pi\sqrt{3}}{2(1 - [\mathbf{\Delta}]_{ii}^2)} \right)^{1/2}. \tag{19}$$

Using (9), the bit resolution given the fidelity of the i-th DAC is given as:

$$\hat{b}_i = \left\lfloor -\frac{1}{2} \log_2 \left(\frac{2}{\pi\sqrt{3}} (1 - [\mathbf{\Delta}]_{ii}^2) \right) \right\rfloor + 1. \tag{20}$$

E. Beamformer codebook design

Let us consider the design of the digital and the analog beamformers that maximizes the ASE R. Assuming normalized equal-power allocation, the digital transmit beamformer (DBF) is given by:

$$\mathbf{F}_{\mathrm{DBF}} = \frac{1}{\sqrt{N_s}} \mathbf{V}_{N_s} \tag{21}$$

and the digital combiner is given by:

$$\mathbf{W}_{\mathrm{DBF}} = \mathbf{U}_{N_s} \in \mathbb{C}^{N_{\mathrm{R}} \times 1}, \tag{22}$$

where $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ is the singular value decomposition of the channel matrix and $\mathbf{V}_{N_s}$, $\mathbf{U}_{N_s}$ represent the N_s columns of the respective singular value matrices.

It follows that, the hybrid transmit beamforming and combining matrices can be obtained by solving the following optimization problem [7]:

$$\begin{aligned} &\min_{\mathbf{F}_{\mathrm{BB}},\mathbf{F}_{\mathrm{RF}}} \|\mathbf{F}_{\mathrm{DBF}} - \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_2^2 \\ &\text{subject to } \mathbf{F}_{\mathrm{RF}} \in \mathcal{F} \text{ and } (\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}})^H(\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}) = \mathbf{I}, \end{aligned} \tag{23}$$

where $\mathcal{F}$ is the set of feasible matrices for the analog transmit beamformer.

III. EE Maximization via RF Selection

In this section, we investigate the problem of EE maximization by selecting the minimal subset of RF chains to utilize. First, we present the problem formulation in the case where the goal is to jointly optimize the subset of the RF chains along with their DAC resolution. Then, we describe a random bit allocation scheme for the available DACs. The problem becomes a fractional subset selection one, which is solved by a Dinkelbach approximation approach. Finally, we propose a projected Newton-type method to solve the near-optimal derived formulation.

A. Problem formulation

The energy efficiency (EE) is defined as the ratio of the ASE R and the total power at the TX P [39],

$$\mathrm{EE} \triangleq \frac{R}{P} \quad \text{(bits/Joule)}. \tag{24}$$

The ASE R is computed based on (15) using the obtained hybrid beamforming matrices from (23). The power $P \triangleq P_{\mathrm{TX}} + P_{\mathrm{RF}}$ depends on the transmission power P_{TX} as well as the DACs power consumption P_{RF}. The transmission power is expressed as $P_{\mathrm{TX}} = \mathrm{tr}(\mathbf{F}) = \mathrm{tr}(\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}})$, while the DACs power consumption via (17). The analytical expression of EE is given by:

$$\mathrm{EE} = \frac{\log|\mathbf{I} + \frac{K}{N_s}\mathbf{R}_\eta^{-1}\mathbf{W}^H\mathbf{H}\mathbf{F}_{\mathrm{RF}}\mathbf{\Delta}\mathbf{F}_{\mathrm{BB}}\mathbf{F}_{\mathrm{BB}}^H\mathbf{\Delta}\mathbf{F}_{\mathrm{RF}}^H\mathbf{H}\mathbf{W}|}{\mathrm{tr}(\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}) + P_{\mathrm{C}}\sum_{i=1}^{L_{\mathrm{T}}}\left(\frac{\pi\sqrt{3}}{2(1-[\mathbf{\Delta}]_{ii}^2)}\right)^{1/2}}. \tag{25}$$

Note that, in this work, we are not investigating the optimal codebook design for the beamformers, i.e., $\mathbf{F}_{\mathrm{RF}}, \mathbf{F}_{\mathrm{BB}}\mathbf{W}$. Instead, we provide a method that utilizes the available codebook to maximize the energy efficiency at the transmitter by determining the best DAC resolution for each RF chain and by selecting the best subset of RF chains to activate.

Our goal is to jointly optimize the bit allocation as well as the active RF chains. This can be represented by extending the feasible values of the diagonal entries of $\mathbf{\Delta}$ with zero, i.e., $\mathbf{\Delta} \in \mathcal{D} \cup \{0\}$. Therefore, a zero diagonal entry of $\mathbf{\Delta}$ corresponds to an inactive RF chain, while the non-zero values of $\mathbf{\Delta} \in \mathbb{R}^{L_{\mathrm{T}} \times L_{\mathrm{T}}}$ represent the DAC fidelity for each RF chain.

In the case where ASE R and power P are expressed with respect to $\mathbf{\Delta}$, we formulate the following EE problem:

$$\begin{aligned} &\max_{\mathbf{\Delta}\in\mathcal{D}\cup\{0\}} \frac{R(\mathbf{\Delta})}{P(\mathbf{\Delta})} \\ &\text{subject to } P(\mathbf{\Delta}) \le P_{\max}, \end{aligned} \tag{26}$$

where $P_{\max}$ is the maximum power budget, and $\mathcal{D}$ is the set of feasible values. This joint problem can also be seen as a combinatorial one, where the solution is obtained via exhaustive search over all possible bit allocations, augmented by one more state which represents the RF chain is active or not. Therefore, for an increased number of RF chains, e.g., $L_{\mathrm{T}} > 8$, the combinatorial approach becomes impractical due to the exponential increase in computational complexity, i.e., $(b_{\max} - b_{\min} + 2)^{L_{\mathrm{T}}}$.

B. RF selection with random bit allocation

In the following, we introduce the proposed selection mechanism that describes the active/inactive RF chains. For this scope, let us define the *binary* matrix $\mathbf{S} \in \{0,1\}^{L_{\mathrm{T}} \times L_{\mathrm{T}}}$. Specifically, $\mathbf{S}$ is a diagonal matrix composed by zeros and ones, with $[\mathbf{S}]_{kk} \in \{0,1\}$ and $[\mathbf{S}]_{kl} = 0$ for $k \neq l$. A physical interpretation of $\mathbf{S}$ is possible by considering that the diagonal entries of this matrix represent hypothetical switches which activate or deactivate the DACs/RF chains. However note that the implementation of this matrix is in the digital domain, as part of the baseband precoder. Hence, there is no need for additional hardware when considering the implementation of the proposed architecture. Using $\mathbf{S}$, the joint optimization of bit allocation and RF chain selection can be described by:

$$\begin{aligned} &\max_{\mathbf{S}\in\{0,1\}^{L_{\mathrm{T}}\times L_{\mathrm{T}}},\mathbf{\Delta}\in\mathcal{D}} \frac{R(\mathbf{\Delta}\mathbf{S})}{P(\mathbf{\Delta}\mathbf{S})} \\ &\text{subject to } P(\mathbf{\Delta}\mathbf{S}) \le P_{\max}, \end{aligned} \tag{27}$$

where the extended set $\mathcal{D} \cup \{0\}$ of (26) is represented using the product $\mathbf{\Delta}\mathbf{S}$ of the unknown matrices $\mathbf{\Delta}$ and $\mathbf{S}$. The problem (27) belongs to the class mixed-integer fractional programming, thus, there is no global solution. To overcome this difficulty, in this work, we consider the case where $\mathbf{\Delta}$ is predefined, thus, the only unknown is $\mathbf{S}$. Therefore, EE is maximized by solving a *subset selection* problem, where a non-zero diagonal value of $\mathbf{S}$ will select the respective RF chain.

Let us consider a random allocation for the DACs bit resolution. In this manner, the bit resolution of each DAC is randomly selected from the set $\{b_{\min}, \ldots, b_{\max}\} \subseteq [1, 8]$. A simple approach is to adopt the uniform distribution, so each entry is generated as:

$$[\mathbf{\Delta}_{ii}] \sim \mathcal{U}(b_{\min}, b_{\max}). \tag{28}$$

However, we stress that the proposed technique can be extended to any distribution of the available bits. Then, the problem becomes equivalent to finding only the selection vector, $\mathrm{diag}(\mathbf{S}) \in \{0,1\}^{L_{\mathrm{T}}\times 1}$, where each unity value represents one active RF chain with a predefined resolution, while the a zero value represents an inactive RF chain.

Incorporating the selection matrix into our formulation, the received signal $\mathbf{y} \in \mathbb{C}^{N_s \times 1}$ at the k-th user is expressed as:

$$\hat{\mathbf{y}} = \underbrace{\mathbf{W}^H \mathbf{H} \mathbf{F}_{\mathrm{RF}} \mathbf{\Delta} \mathbf{S} \mathbf{F}_{\mathrm{BB}}}_{\triangleq \mathbf{H}_{\text{eff}}} \mathbf{x} + \boldsymbol{\eta}, \tag{29}$$

where the effective channel,

$$\mathbf{H}_{\text{eff}} \triangleq \mathbf{W}^H \mathbf{H} \mathbf{F}_{\mathrm{RF}} \mathbf{\Delta} \mathbf{S} \mathbf{F}_{\mathrm{BB}}, \tag{30}$$

can be expressed as a sum of L_{T} terms (see Appendix VI-A), namely:

$$\mathbf{H}_{\text{eff}} = \sum_{i=1}^{L_{\mathrm{T}}} [\mathbf{S}]_{ii} \mathbf{a}_i \mathbf{b}_i^T. \tag{31}$$

where

$$\mathbf{a}_i \triangleq [\mathbf{W}^H \mathbf{H} \mathbf{F}_{\mathrm{RF}} \mathbf{\Delta}]_i \in \mathbb{C}^{N_s \times 1}, \tag{32}$$

and

$$\mathbf{b}_i \triangleq [\mathbf{F}_{\mathrm{BB}}^T]_i \in \mathbb{C}^{N_s \times 1}. \tag{33}$$

Therefore, it can be seen that (31) is the sum of L_{T} rank-one matrices $\mathbf{a}_i \mathbf{b}_i^T$, where $\mathbf{a}_i$ and $\mathbf{b}_i$ are vectors of dimension $N_s \times 1$. The rank of the effective channel matrix $\mathbf{H}_{\text{eff}} \in \mathbb{C}^{N_s \times N_s}$ can be at most N_s. However, (31) sums $L_{\mathrm{T}} \geq N_s$ terms, indicating that if $L_{\mathrm{T}} > N_s$ some of the rank-one matrices will contribute to the same subspace. The contribution of each rank-one matrix depends on the binary diagonal value of $\mathbf{S}$.

Based on (31), the received signal for all users can be expressed as the following measurement vector:

$$\hat{\mathbf{y}} = \sum_{i=1}^{L_{\mathrm{T}}} [\mathbf{S}]_{ii} \mathbf{a}_i (\mathbf{b}_i^T \mathbf{x}) + \tilde{\boldsymbol{\eta}}, \tag{34}$$

where $\tilde{\boldsymbol{\eta}} \triangleq \mathbf{S}[\boldsymbol{\eta}_1 \cdots \boldsymbol{\eta}_K]$, and $\tilde{\boldsymbol{\eta}} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_{\tilde{\eta}})$, with the noise covariance matrix $\mathbf{R}_{\tilde{\eta}} \in \mathbb{C}^{N_s \times N_s}$ expressed with respect to the selection matrix, i.e.,

$$\mathbf{R}_{\tilde{\eta}} = \mathbf{W}^H \mathbf{H} \mathbf{F}_{\mathrm{RF}} \mathbf{C}_\epsilon \mathbf{S} \mathbf{F}_{\mathrm{BB}} \mathbf{F}_{\mathrm{BB}}^H \mathbf{S} \mathbf{C}_\epsilon^H \mathbf{F}_{\mathrm{RF}}^H \mathbf{H}^H \mathbf{W} + \sigma_n^2 \mathbf{W}^H \mathbf{W}. \tag{35}$$

Similarly to (15), the ASE for the case of the proposed architecture with the selection matrix $\mathbf{S} \in \{0,1\}^{L_{\mathrm{T}} \times L_{\mathrm{T}}}$, can be expressed as:

$$R(\mathbf{S}) = \log_2 \left| \mathbf{I}_{N_s} + \frac{1}{N_s} \mathbf{H}_{\text{eff}}^H \mathbf{R}_{\tilde{\eta}}^{-1} \mathbf{H}_{\text{eff}} \right|, \quad \text{(bits/s/Hz)} \tag{36}$$

where $\mathbf{H}_{\text{eff}} \in \mathbb{C}^{N_s \times N_s}$ is defined in (30) and $\mathbf{R}_{\tilde{\eta}} \in \mathbb{C}^{N_s \times N_s}$ in (35) and they both are functions of $\mathbf{S}$. This property will enable the decomposition of the ASE into terms that describe the contribution of each RF chain.

To proceed, first note that the ASE in (36) is equivalent to:

$$R(\mathbf{S}) = \log_2 \left| \mathbf{I}_{N_s} + \frac{1}{N_s} \mathbf{H}_{\text{eff}}^H \mathbf{R}_{\tilde{\eta}}^{-\frac{1}{2}} \mathbf{R}_{\tilde{\eta}}^{-\frac{1}{2}} \mathbf{H}_{\text{eff}} \right| \tag{37}$$

$$= \log_2 \left| \mathbf{I}_{N_s} + \frac{1}{N_s} \mathbf{R}_{\tilde{\eta}}^{-\frac{1}{2}} \mathbf{H}_{\text{eff}} \left(\mathbf{R}_{\tilde{\eta}}^{-\frac{1}{2}} \mathbf{H}_{\text{eff}} \right)^H \right|. \tag{38}$$

Utilizing the decomposition of the effective channel, as it is given in (31), we have that:

$$R(\mathbf{S}) = \log_2 \left| \mathbf{I}_{N_s} + \frac{1}{N_s} \left(\mathbf{R}_{\tilde{\eta}}^{-\frac{1}{2}} \sum_{i=1}^{L_{\mathrm{T}}} [\mathbf{S}]_{ii} \mathbf{a}_i \mathbf{b}_i^T \right) \left(\mathbf{R}_{\tilde{\eta}}^{-\frac{1}{2}} \sum_{j=1}^{L_{\mathrm{T}}} [\mathbf{S}]_{jj} \mathbf{a}_j \mathbf{b}_j^T \right)^H \right|. \tag{39}$$

By defining $\mathbf{c}_i \triangleq \mathbf{R}_{\tilde{\eta}}^{-\frac{1}{2}} \mathbf{a}_i$, (39) can be written as:

$$\begin{aligned} R(\mathbf{S}) &= \log_2 \left| \mathbf{I}_{N_s} + \frac{1}{N_s} \sum_{i=1}^{L_{\mathrm{T}}} \sum_{j=1}^{L_{\mathrm{T}}} ([\mathbf{S}]_{ii} [\mathbf{S}]_{jj}) (\mathbf{c}_i^T \mathbf{c}_j^*) (\mathbf{b}_i \mathbf{b}_j^H) \right| \\ &= \log_2 \mathbf{I}_{N_s} + \frac{1}{N_s} \Bigg| \underbrace{\sum_{i=1}^{L_{\mathrm{T}}} [\mathbf{S}]_{ii} (\mathbf{c}_i^H \mathbf{c}_i) (\mathbf{b}_i \mathbf{b}_i^H)}_{\triangleq \mathbf{Q}} \\ &\quad + \underbrace{\sum_{i \neq j} ([\mathbf{S}]_{ii} [\mathbf{S}]_{jj}) (\mathbf{c}_i^T \mathbf{c}_j^*) (\mathbf{b}_i \mathbf{b}_j^H)}_{\triangleq \bar{\mathbf{Q}}} \Bigg|. \end{aligned} \tag{40}$$

Essentially, the second term of (40) (i.e., the matrix $\mathbf{Q}$) expresses *the contribution of each RF chain* to the overall ASE, while the third term (i.e., the matrix $\bar{\mathbf{Q}}$) expresses the portion of the ASE that occurs due to cross-correlation between different RF chains. The second term provides a linear expression over $\mathbf{S}$, while the third term is quadratic. Nevertheless, $\bar{\mathbf{Q}} = \mathbf{0}$ since $\mathbf{b}_i^H \mathbf{b}_j = 0$ for $i \neq j$, as it is proven by the following proposition.

Proposition 1. *Consider an ideal hybrid TX with precoder decomposed as* $\mathbf{F} = \mathbf{F}_{RF}\mathbf{F}_{BB}$*, with phase-shifters of infinite-resolution and* L_T*. Then, following* (36)-(40)*, the achievable information rate (ASE) can be expressed as:*

$$R(\mathbf{S}) = \log_2 \left| \mathbf{I}_{N_s} + \frac{1}{N_s} \sum_{i=1}^{L_T} [\mathbf{S}]_{ii} (\mathbf{a}_i^H \mathbf{a}_i) (\mathbf{b}_i \mathbf{b}_i^H) \right|, \tag{41}$$

where $\mathbf{a}_i$ *and* $\mathbf{b}_i$ *are defined in* (32) *and* (33)*, respectively, while* $[\mathbf{S}]_{ii} \in \{0,1\}$ *determines the state of the* i*-th RF chain.*

Proof. see Appendix VI-B. □

Eq. (41) provides an expression for the ASE that is a function of the selection matrix $\mathbf{S}$. The inner product $\mathbf{a}_i^H \mathbf{a}_i$, weights the contribution of the i-th RF chain to the ASE, given the combined effect of the inverse of the noise covariance matrix $\mathbf{R}_{\tilde{\eta}}^{-1}$, the combiner matrix $\mathbf{W}$, the channel $\mathbf{H}$, the analog beamforming matrix $\mathbf{F}_{\mathrm{RF}}$ and the random bit allocation $\mathbf{\Delta}$. Assuming that the baseband precoder matrix $\mathbf{F}_{\mathrm{BB}}$ is normalized, the outer-product $\mathbf{b}_i \mathbf{b}_i^H$ does not contribute to the ASE of each RF chain.

Algorithm 1 Dinkelbach-based Algorithm

1: **for** $m = 1, 2, \ldots, I_{\max}$ **do**
2: Obtain $\mathbf{S}^{(m)}$ by solving (46)
3: Calculate $R^{(m)}$ and $P^{(m)}$
4: $\kappa^{(m)} = R^{(m)}/P^{(m)}$
5: **end for**

Furthermore, the power for all the RF chains is now expressed with respect to the binary diagonal matrix $\mathbf{S}$ as follows:

$$P_{\mathrm{RF}}(\mathbf{S}) = P_{\mathrm{C}} \sum_{i=1}^{L_{\mathrm{T}}} [\mathbf{S}]_{ii} \left(\frac{\pi\sqrt{3}}{2(1-[\mathbf{\Delta}]_{ii}^2)} \right)^{1/2} \tag{42}$$

$$= \boldsymbol{\beta}^T \mathrm{vec}(\mathbf{S}), \tag{43}$$

where $\mathrm{vec}(\mathbf{S}) \in \mathbb{R}^{L_{\mathrm{T}} \times 1}$ and $\boldsymbol{\beta} \in \mathbb{R}^{L_{\mathrm{T}} \times 1}$ with

$$[\boldsymbol{\beta}]_i \triangleq P_{\mathrm{C}} \left(\frac{\sqrt{3}}{2(1-[\mathbf{\Delta}]_{ii}^2)} \right)^{1/2}, \tag{44}$$

for $i = 1, 2, \ldots, L_{\mathrm{T}}$. Eq. (42) states that only the selected RF chains contribute to the total power depending on the associated weight.

C. EE maximization via Dinkelbach method

Capitalizing the random resolution DACs framework, the EE problem can be expressed as:

$$\max_{\mathbf{S} \in \mathcal{S}} \frac{R(\mathbf{S})}{P(\mathbf{S})}$$
$$\text{subject to } P(\mathbf{S}) \leq P_{\max}, \tag{45}$$

which states an integer concave-convex fractional problem. Dinkelbach (DB) algorithm [40] has been widely used for solving fractional problems. In brief, DB method is an iterative and parametric algorithm, where a sequence of easier problems converge to the global solution of (26).

To solve (45), first we alleviate the constraint for the integer values of $\mathbf{S}$, so as $\tilde{\mathbf{S}} \in [0, 1]$. Once the matrix $\tilde{\mathbf{S}}$ has been recovered, its diagonal values will be thresholded to indicate the active RF chains. Let $\kappa^{(m)} \in \mathbb{R}$, for $m = 1, 2, \ldots, I_{\max}$, then each iteration step of DB can be expressed as:

$$\tilde{\mathbf{S}}^{(m)}(\kappa^{(m)}) \triangleq \arg\max_{\tilde{\mathbf{S}} \in \tilde{\mathcal{S}}} \left\{ R(\tilde{\mathbf{S}}) - \kappa^{(m)} P(\mathbf{S}) \right\}$$
$$= \arg\min_{\tilde{\mathbf{S}} \in \tilde{\mathcal{S}}} \left\{ P(\tilde{\mathbf{S}}) - \frac{1}{\kappa^{(m)}} R(\tilde{\mathbf{S}}) \right\} \tag{46}$$

where $\tilde{\mathcal{S}}$ is the set of diagonal matrices with the feasible bit allocations which satisfy $P(\tilde{\mathbf{S}}) \leq P_{\max}$. The DB algorithm is summarized in Algorithm 1 and it can be shown that with proper selection of $\kappa^{(m)}$ it converges to the global solution of (26) [39, Proposition 3.2].

The (m)-th step of the Dinkelbach algorithm (46) can be written as:

$$\min_{\tilde{\mathbf{S}}^{(m)} \in \tilde{\mathcal{S}}} \sum_{i=1}^{L_{\mathrm{T}}} [\tilde{\mathbf{S}}^{(m)}]_{ii} \left(\frac{\pi\sqrt{3}}{2(1-[\mathbf{\Delta}]_{ii}^2)} \right)^{1/2}$$
$$- \frac{1}{\kappa^{(m)}} \log_2 \left| \mathbf{I}_{N_s} + \frac{1}{N_s} \sum_{i=1}^{L_{\mathrm{T}}} [\tilde{\mathbf{S}}^{(m)}]_{ii} \|\mathbf{a}_i\|_2^2 (\mathbf{b}_i \mathbf{b}_i^H) \right| \tag{47}$$

The optimization problem of (47) is a form of sparse subset selection [41], where given the set of L_{T} measurement vectors, namely $(\mathbf{b}_i^T \mathbf{x})\mathbf{a}_i$, we have to determine the subset that will provide us with the best results based on the optimality criterion. This criterion can be also expressed in terms of the eigenvalues of the matrix $\mathbf{Q} \in \mathbb{C}^{N_s \times N_s}$ defined as:

$$\mathbf{Q} \triangleq \sum_{i=1}^{L_{\mathrm{T}}} [\tilde{\mathbf{S}}]_{ii} \|\mathbf{c}_i\|_2^2 \mathbf{B}_i. \tag{48}$$

with

$$\mathbf{B}_i \triangleq \mathbf{b}_i \mathbf{b}_i^H \in \mathbb{C}^{N_s \times N_s}. \tag{49}$$

An equivalent way to express (47) is by using the trace of the matrix, i.e.,

$$\min_{\tilde{\mathbf{S}}^{(m)} \in \tilde{\mathcal{S}}} \sum_{i=1}^{L_{\mathrm{T}}} [\tilde{\mathbf{S}}^{(m)}]_{ii} \left(\frac{\pi\sqrt{3}}{2(1-[\mathbf{\Delta}]_{ii}^2)} \right)^{1/2}$$
$$- \frac{1}{\kappa^{(m)}} \mathrm{tr} \left(\mathbf{I}_{N_s} + \frac{1}{N_s} \mathbf{Q} \right). \tag{50}$$

The optimization problem in (50) for the m-th iteration of the DB method is a standard semi-definite one which can be solved in polynomial time using interior-point methods, e.g., the projected Newton-type method (PNM) [42, p. 619]. There are several available software packages of implemented interior-point methods (e.g., CVX [18], Mosek). However, in the next subsection, we will implement our interior-point based technique, which is customized for our case.

D. Projected Newton-Type Method

Let us describe the PNM expressed for our problem where (46) can be seen as the log-determinant barrier function, i.e.,

$$\arg\min_{\mathbf{s} \in \mathcal{S}} \kappa^{(m)} \boldsymbol{\beta}^T \mathbf{s} - \log_2 |\frac{1}{N_s} \mathbf{Q} - \gamma \mathbf{I}|, \tag{51}$$

where $\mathbf{s} \triangleq \mathrm{vec}(\tilde{\mathbf{S}})$ and $\gamma \in \mathbb{R}^+$ is a predefined weighting parameter. The i-th step of the projected Newton's update equation is given by:

$$\mathbf{s}^{(i+1)} = \mathcal{P}_{\mathcal{S}} \left(\mathbf{s}^{(i)} - \alpha^{(i)} \mathbf{\Xi}^{-1} \mathbf{g}^{(i)} \right), \tag{52}$$

for $i = 1, \ldots, L_{\mathrm{T}}$, where $\mathcal{P}_{\mathcal{S}}$ is a thresholding operation, which is performed element-wise on the input vector, with

$$\mathcal{P}(x) = \begin{cases} 1 & \text{if } x > \rho \\ 0 & \text{otherwise} \end{cases}, \tag{53}$$

$\rho > 0$ is a small positive value, e.g., $\rho = 10^{-5}$, and $\boldsymbol{\beta}$ is defined in (44). The scalar $\alpha^{(i)}$ represents the step-length

Algorithm 2 Proposed algorithm based on projected Newton's method

Input: $\mathbf{a}_i, \mathbf{B}_i, \boldsymbol{\beta}_i, I_{\max}$
Output: $\mathbf{S}^{I_{\max}}$
Initialization: $\mathbf{S}^{(1)} = \mathbf{0}_{L_{\mathrm{T}} \times L_{\mathrm{T}}}$
for $m = 1, 2, \ldots, I_{\max}$ **do**
  $\mathbf{s}^{(1)} = \mathrm{vec}(\mathbf{S}^{(m)})$.
  for $i = 1, 2, \ldots, L_{\mathrm{T}}$ **do**
    Update the Hessian matrix using (55).
    Obtain the gradient vector using (56).
    Obtain $\mathbf{s}^{(i+1)}$ from (52).
  end for
  Reshape the obtained vector into matrix $\mathbf{S}^{(m)} = \mathrm{unvec}(\mathbf{s}^{(L_{\mathrm{T}})})$.
  Calculate $R^{(m)}(\mathbf{S}^{(m)})$ and $P^{(m)}(\mathbf{S}^{(m)})$.
  $\kappa^{(m)} = R^{(m)}/P^{(m)}$.
end for

chosen by line-search. The (k,l)-th entry of the Hessian matrix $\boldsymbol{\Xi}(\mathbf{s}^{(i)}) \in \mathbb{C}^{L_{\mathrm{T}} \times L_{\mathrm{T}}}$ is given by:

$$[\boldsymbol{\Xi}]_{k,l}(\mathbf{x}^{(i)}) = \frac{\partial^2}{\partial[\mathbf{s}]_k \partial[\mathbf{s}]_l} \left(\boldsymbol{\beta}^T \mathbf{s} - \log_2 |\mathbf{Q} - \gamma \mathbf{I}| \right) \tag{54}$$

$$= \|\mathbf{a}_k\|^2 \|\mathbf{a}_l\|^2 \mathrm{tr}\left((\mathbf{Q} - \gamma \mathbf{I})^{-1} \mathbf{B}_k (\mathbf{Q} - \gamma \mathbf{I})^{-1} \mathbf{B}_l \right) \tag{55}$$

and the k-th entry of the gradient vector is given by

$$[\mathbf{g}^{(i)}]_k = \kappa^{(m)} [\boldsymbol{\beta}]_k + \|\mathbf{a}_k\|^2 \mathrm{tr}((\mathbf{Q} - \gamma \mathbf{I})^{-1} \mathbf{B}_k). \tag{56}$$

Thus, Algorithm 2 provides the PNM solution of (50), after a fixed maximum number of iterations $I_{\max}$. Considering computational complexity, the formation of the matrix $\mathbf{Q}$ has computational cost of the order $\mathcal{O}(L_{\mathrm{T}} N_s^3)$, while the computation of the Cholesky factorization to obtain $(\mathbf{Q} - \gamma \mathbf{I})^{-1} \mathbf{B}_k \, \forall k$ costs $\mathcal{O}(L_{\mathrm{T}} N_s^3)$. Additionally, the computation of the Hessian matrix costs $\mathcal{O}(L_{\mathrm{T}}^2 N_s^2)$ and the Cholesky factorization for the Newton step $\mathcal{O}(L_{\mathrm{T}}^3)$. Therefore, the overall computational complexity per iteration of the PNM is $\mathcal{O}(L_{\mathrm{T}}^3)$.

IV. Simulation Results

In this section, we evaluate the performance of the proposed technique via computer simulation results using MATLAB$^{\mathrm{TM}}$. All the results are averaged over 500 Monte-Carlo (MC) realizations.

A. Setup

Before proceed, let us define the parameters and the system characteristics. We assume that the transmitter employs a hybrid transmit beamforming with N_{T} antennas. The number of RF chains is $N_s \leq L_{\mathrm{T}} \leq N_{\mathrm{T}}$ and each one is equipped with a b_i-bit DAC, for $i = 1, \ldots, L_{\mathrm{T}}$, with $b_i \in [1, 8]$. Each transmission broadcasts a zero-mean random Gaussian vector with $\mathbf{x} \in \mathbb{C}^{N_s \times 1}$ and $\mathcal{E}\{\mathbf{x}\mathbf{x}^H\} = \frac{P_{\mathrm{TX}}}{N_s} \mathbf{I}_{N_s}$. We assume uniform linear arrays (ULAs) at both TX and RX sides and operating over a 28 GHz outdoor mmWave channel [43]. The azimuth angles of arrival and departure are generated from the Laplace distribution with a standard deviation of 50°. The number of clusters set to $N_{\mathrm{cl}} = 4$ and the number of rays per cluster to $N_{\mathrm{rays}} = 5$. We assume digital combining is performed at the user equipment with $N_{\mathrm{R}} = N_s$ and the combiner matrix $\mathbf{W}$ is given by (22). The power of the RF chain components P_{C} is set to 0.01 and $P_{\max} = 1$.

Algorithm 3 Exhaustive search algorithm

Input: $\mathbf{F}_{\mathrm{RF}}, \mathbf{F}_{\mathrm{BB}}, \mathcal{C}$
Output: $\mathbf{S}_{\mathrm{opt}}$
for $i = 1, \ldots, |\mathcal{C}|$ **do**
  Compute the EE for $\mathbf{S}_i \in \mathcal{C}$ via (36) and (42)
end for
Find $\mathbf{S}_{\mathrm{opt}} = \arg\max_{\mathbf{S}_i} \mathrm{EE}$

In this work, we investigate the case of selecting the best active subset of RF chains in terms of EE maximization. Ideally, the optimal subset can be obtained via an exhaustive search over all possible combinations. Let the set $\mathcal{C}$ represents all possible combinations for the state (active/inactive) of the virtual switches for the L_{T} RF chains. Then, the exhaustive search algorithm has to compute the EE for all combinations $|\mathcal{C}|$ and select the one with the highest EE. The exhaustive search is summarized in Algorithm 3. Since the number of the combinations $|\mathcal{C}|$ increases exponentially with the number of active RF chains.

B. Energy and spectral efficiency performance

For the evaluation of the proposed technique in terms of energy efficiency (EE) and the spectral efficiency (SE) performance, we have considered the following cases for the TX:

1) **DBF**: digital beamforming architecture ($N_{\mathrm{T}} = L_{\mathrm{T}}$) with high resolution DACs (8-bits), which represents the optimum from the achievable SE perspective,
2) **b-bit-HBF**: hybrid transmit beamforming with L_{T} RF chains with common b-bits DACs resolution, where $b \in \{1, 8\}$,
3) **random-bit-HBF**: hybrid transmit beamforming with L_{T} RF chains with mixed resolution DACs. The bit resolutions are generated using the uniform distribution $\mathcal{U}(1, 8)$,
4) **common-exhaustive-HBF**: hybrid transmit beamforming with minimum number of RF chains L_{T} and common bit resolution for all DACs.
5) **proposed HBF**: Proposed technique with $I_{\max} = 5$ and the bit resolutions are generated using the uniform distribution $\mathcal{U}(1, 8)$. The number of RF chains was set equal to the number of transmitting antennas $L_{\mathrm{T}} = N_{\mathrm{T}}$.

In Figs. 2 and 3 we plot the EE and SE with respect to the transmit power P_{TX} over a wide range. Note that 24 dBm is the maximal output of the mobile user equipment (class 3), while a typical value for macro cell base station is 43 dBm at the antenna connector. Future wireless networks will have to achieve high SE performance with limited power, e.g., UAV-aided communications. Thus, a reasonable TX power range would be between 24-40 dBm. From Fig. 2 it is evident that for this range of P_{TX} values, the proposed technique outperforms

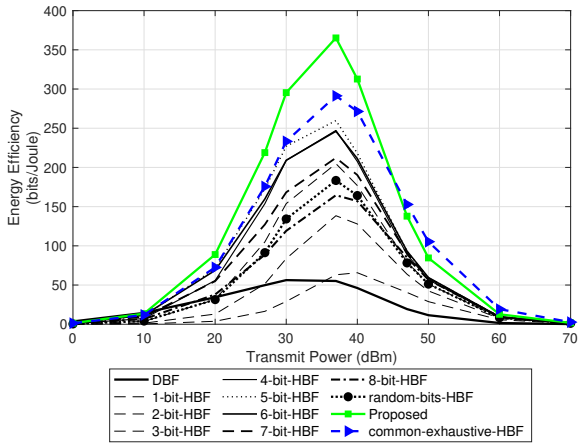

Fig. 2. Energy efficiency as a function of the transmit power P_{TX} for $N_{\mathrm{T}} = 128$, $N_{\mathrm{R}} = 8$, and $\sigma_n = 0.562$.

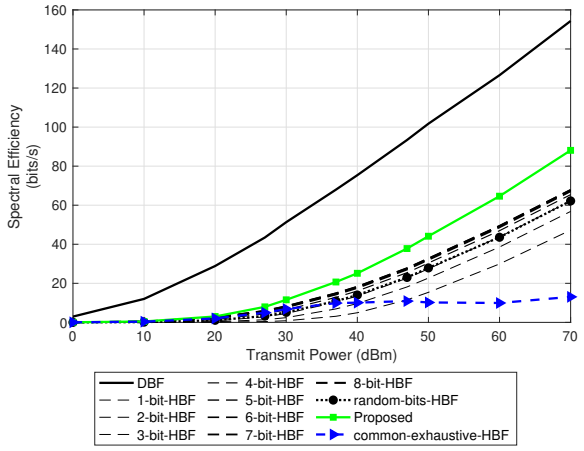

Fig. 3. Spectral-efficiency as a function of the transmit power for $N_{\mathrm{T}} = 128$, $N_{\mathrm{R}} = 8$, and $\sigma_n = 0.562$.

the other baselines reaching over 300 bits/Joule. The common-exhaustive-HBF also achieves high EE performance at the expense of exponential computational complexity. Considering EE performance for the fixed b-bit-HBF cases, the best performance is achieved by the 5-bit-HBF case. However, the best DAC resolution is directly related to P_{C}. Specifically, for $P_{\mathrm{C}} = 0.1$, the 4-bit-HBF achieves the best performance.

In Fig. 3, the SE performance of the considered techniques is shown. DBF achieves the higher SE, at the expense of high power consumption, and computational and hardware complexity. The SE performance of the b-bit-HBF cases is fully analogous to the bit resolution b. Thus, among the HBF techniques, the highest performance is achieved by the 8-bit case and the lowest by the 1-bit case.

The proposed technique outperforms the other baselines, thus, achieving both high SE and EE. This is possible since it can search over the whole N_{T} codebook space. To achieve this performance, the analog part (e.g., phase-shifters) of the BS beamformer have been increased, based on the proposed design. Thus, a trade-off between the hardware complexity

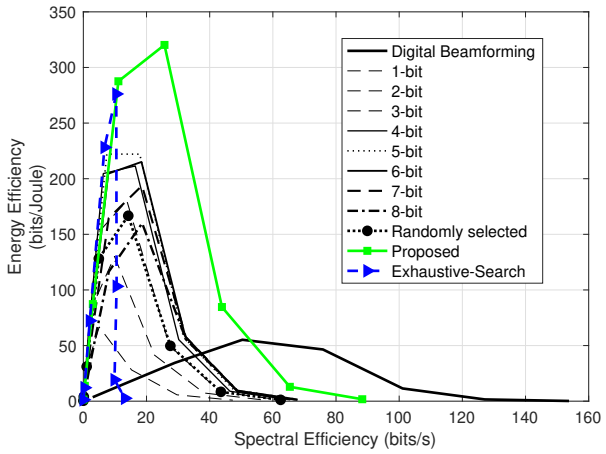

Fig. 4. Energy efficiency versus spectral efficiency for $N_{\mathrm{T}} = 128$, $N_{\mathrm{R}} = 8$ and $\sigma_n = 0.562$.

and the SE performance is possible. Note that in this work, we consider that the power consumption of each phase-shifter is negligible compared to the power consumption of each RF chain [44]. Indeed, the proposed structure for the analog part can be realized by using very energy-efficient elements, e.g., passive phase-shifters, Butler matrix [45].

In Fig. 4, the EE versus the SE performance is illustrated. The proposed technique covers much larger area of the EE-SE region, thus, enables high flexibility for the design of a hybrid HBF with low-resolution DACs by providing higher EE and SE.

In Figs. 6 and 5, we plot the SE and EE with respect to the number of transmission streams N_s, respectively. For the proposed technique we consider that the maximum number of active RF chains is constrained by N_s, while for the rest of techniques we assume that $L_{\mathrm{T}} = N_s$. The SE increases as the number of the streams increases, exploiting the available degrees-of-freedom provided by the channel, e.g., $N_{\mathrm{cl}}N_{\mathrm{ray}} = 20$. EE also increases, but as the number of N_s, and thus L_{RF}, increases, the RF chains power consumption sets an upper bound.

It is important to stress out that the performance of the proposed algorithm depends on the ratio of L_{T} and N_s which represents the sparsity level of the problem. Specifically, the problem expressed by (47) describes a sparse subset selection one. Via experimentation, we have observed that to maintain the SE performance of the proposed technique as N_s increases, the sparsity level has to be $\frac{N_s}{L_{\mathrm{T}}} \leq 0.1$.

C. Active RF chains and bit-allocation

To get a better insight into how the proposed technique works, in Fig. 7 we show the average active RF chains with respect to the transmit power, for the proposed technique and common-exhaustive-HBF. In the same figure, we also plot the minimum and maximum values for active RF chains for each technique over the MC realizations. The proposed technique activates 8 RF chains for almost all the transmit power

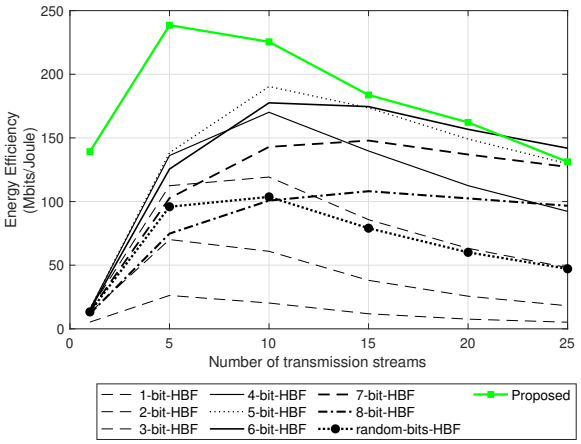

Fig. 5. Energy-efficiency versus the number of transmission streams N_s.

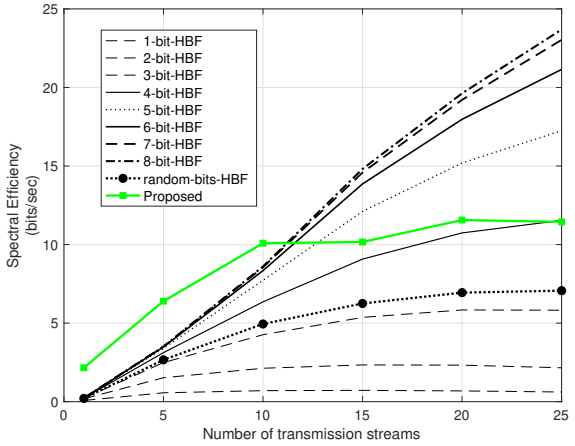

Fig. 6. Spectral-efficiency versus the number of transmission streams N_s.

values, which belong to the range $[-20, 30]$ dBm. However, the variance of the number of active RF chains is rather wide, indicating the dynamic characteristics of the proposed technique. On the other side, the common-exhaustive-HBF reduces the average number of active RF chains as transmit power increases, to maximize the EE.

In Fig. 8 we show the sum of the bit resolution of the DACs for the active RF chains. Although different RF chains are activated at each channel realization, the sum of the DACs resolution seems not to vary significantly, over the transmit power. On the contrary, in the case of common-exhaustive-HBF, the common DACs resolution decreases as transmit power increases. In this manner common-exhaustive-HBF achieves high EE compared to the other techniques with fixed-resolution. Therefore, by selecting the best subset of active RF chains with the proper DAC resolution, via the proposed technique, provides high EE and SE performance.

D. Convergence of Dinkelbach iterations

The convergence of the Dinkelbach iterative algorithm can be investigated via the behaviour of the $\kappa^{(m)}$ variable over

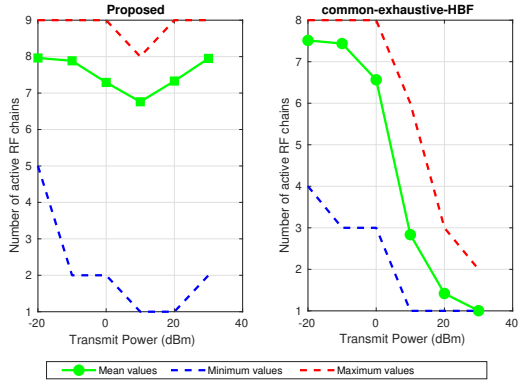

Fig. 7. Average number of active RF chains over transmit power.

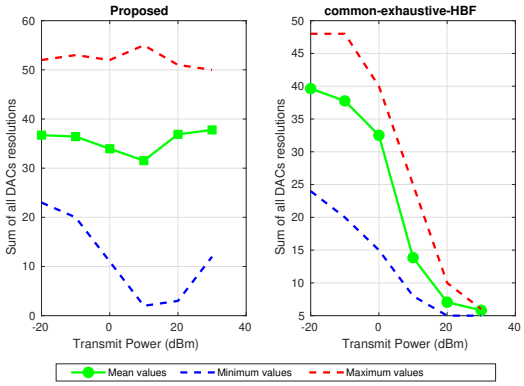

Fig. 8. The sum of DACs resolutions for all active RF chains over transmit power.

iteration m. Recall that κ represents an approximation for the achievable spectral efficiency. Thus, it is expected to be increasing with each iteration, till an upper bound value which indicates that the technique has converged. In Fig. 9, we plot the $\kappa^{(m)}$ value with respect to the iteration number m of the proposed algorithm. We consider three cases for the antenna array size N_{T}. The number of the RF chains is set to the half for each case, i.e., $L_{\mathrm{T}} = N_{\mathrm{T}}/2$. We note that in all cases DB requires only a small number of iterations to converge to the optimal value of the parameter $\kappa^{(m)}$. Also, the convergence behaviour depends on the DACs resolution. We consider three cases, namely: (a) $b_{\min} = b_{\max} = 1$, (b) $b_{\min} = 1, b_{\max} = 8$, and (c) $b_{\min} = b_{\max} = 8$. When the DACs resolution is mixed, case (b), the proposed algorithm requires twice the number of iterations to reach the upper floor. That is 6 iterations instead of 3 which are required for the other two cases, i.e., (a) and (b).

V. Conclusion

This paper proposes a selection framework to jointly allocate the best active RF chains that maximize the energy-efficiency performance, along with the spectral-efficiency, of the transmitter. To cope with the non-convexity of this problem, we employ random bit allocation for the DACs and we formulate a binary fractional programming problem that is

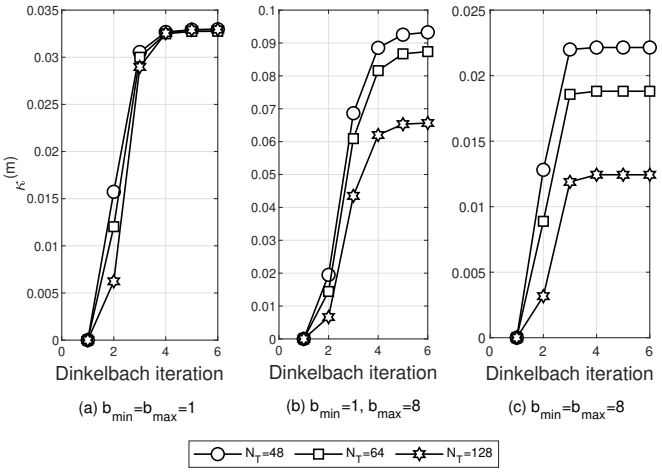

Fig. 9. Dinkelback convergence curves.

solved via the Dinkelbach approximation and subset selection. The proposed algorithm mostly outperforms all the baselines such as fully digital beamforming, fixed DACs bit resolution. The convergence of the proposed fractional programming algorithm is fast as it typically requires only 3-4 iterations to converge.

VI. Appendix

A. Derivation of (31)

It is known that for matrices $\mathbf{A} \in \mathbb{C}^{N\times M}$, $\mathbf{B} \in \mathbb{C}^{M\times T}$, and $\mathbf{S} \in \mathbb{C}^{M\times M}$ a diagonal matrix, the following equality holds:

$$\mathbf{ASB} = \sum_{i=1}^{M} [\mathbf{S}]_{ii}\mathbf{a}_i\mathbf{b}_i, \tag{57}$$

where $\mathbf{a}_i$ is the i-th column of $\mathbf{A}$ and $\mathbf{b}_i$ is the i-row of $\mathbf{B}$. Similarly, by defining $\mathbf{A} \triangleq \mathbf{W}^H\mathbf{H}\mathbf{F}_{\mathrm{RF}}\mathbf{\Delta}$ and $\mathbf{B} \triangleq \mathbf{F}_{\mathrm{BB}}^T$, the eq. (31) follows, i.e.,

$$\mathbf{H}_{\mathrm{eff}} = \sum_{i=1}^{L_{\mathrm{T}}} [\mathbf{S}]_{ii}[\mathbf{W}^H\mathbf{H}\mathbf{F}_{\mathrm{RF}}\mathbf{\Delta}]_i[\mathbf{F}_{\mathrm{BB}}^T]_i^T. \tag{58}$$

B. Proof of Proposition 1

Essentially, Proposition 1 is based on the observation that, under idealized conditions for the analog precoder, namely, phase-shifters with infinite resolution and $L_{\mathrm{T}} = N_{\mathrm{T}}$, the analog and the digital matrices are orthonormal, i.e., $\mathbf{F}_{\mathrm{RF}}^H\mathbf{F}_{\mathrm{RF}} = \mathbf{I}_{L_{\mathrm{T}}}$ and $\mathbf{F}_{\mathrm{BB}}^H\mathbf{F}_{\mathrm{BB}} = \mathbf{I}_{N_s}$. To prove the proposition we have to show that $\mathbf{b}_i^H\mathbf{b}_j = 0$ for $i \neq j$, with $\mathbf{b}_i \triangleq [\mathbf{F}_{\mathrm{BB}}^T]_i$. First recall that the digital beamformer $\mathbf{F}_{\mathrm{opt}}$ is composed by the N_s columns of the orthonormal matrix $\mathbf{V}$ with the right singular vectors of the channel impulse response, i.e., $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$. Also, based on the solution of (23), ideally we have that $\mathbf{F} = \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}$. Hence,

$$\mathbf{F}_{\mathrm{opt}}^H\mathbf{F}_{\mathrm{opt}} = \mathbf{I}_{N_s} \tag{59}$$

$$\Rightarrow (\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}})^H(\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}) = \mathbf{I}_{N_s} \tag{60}$$

$$\Rightarrow \mathbf{F}_{\mathrm{BB}}^H\mathbf{F}_{\mathrm{RF}}^H\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}} = \mathbf{I}_{N_s} \tag{61}$$

$$\Rightarrow \mathbf{F}_{\mathrm{BB}}^H\mathbf{I}_{L_{\mathrm{T}}}\mathbf{F}_{\mathrm{BB}} = \mathbf{I}_{N_s}, \tag{62}$$

since $\mathbf{F}_{\mathrm{RF}}^H\mathbf{F}_{\mathrm{RF}} = \mathbf{I}_{L_{\mathrm{T}}}$ with $L_{\mathrm{T}} = N_{\mathrm{T}}$.

References

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

[2] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-Wave Cellular Wireless Networks: Potentials and Challenges," *Proc. IEEE*, vol. 102, no. 3, pp. 366–385, Mar. 2014.

[3] T. Bai and R. W. Heath, "Coverage and rate analysis for millimeter-wave cellular networks," in *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, Feb. 2015, pp. 1100–1114.

[4] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[5] T. S. Rappaport, Shu Sun, R. Mayzus, Hang Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[6] F. Boccardi, R. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[7] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, 2014.

[8] O. El Ayach, R. W. Heath, S. Abu-Surra, S. Rajagopal, Z. Pi, O. E. Ayach, R. W. Heath, S. Abu-Surra, S. Rajagopal, and Z. Pi, "The capacity optimality of beam steering in large millimeter wave MIMO systems," in *2012 IEEE 13th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, Jun. 2012, pp. 100–104.

[9] D. J. Love and R. W. Heath, "Equal gain transmission in multiple-input multiple-output wireless systems," *IEEE Trans. Commun.*, vol. 51, no. 7, pp. 1102–1110, Jul. 2003.

[10] 3GPP, "Study on new radio access technology; radio interface protocol aspects," *3GPP TR 38.804*, vol. Release 15, 2017.

[11] 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 Commun. Mag.*, vol. 55, no. 9, pp. 134–141, 2017.

[12] S. Jacobsson, G. Durisi, M. Coldrey, U. Gustavsson, and C. Studer, "Throughput Analysis of Massive MIMO Uplink With Low-Resolution ADCs," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 4038–4051, Jun. 2017.

[13] A. Mezghani and J. A. Nossek, "On Ultra-Wideband MIMO Systems with 1-bit Quantized Outputs: Performance Analysis and Input Optimization," *2007 IEEE International Symposium on Information Theory*, pp. 1286–1289, 2007.

[14] J. Mo and R. W. Heath, "Capacity Analysis of One-Bit Quantized MIMO Systems With Transmitter Channel State Information," *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5498–5512, Oct. 2015.

[15] S. Jacobsson, G. Durisi, M. Coldrey, U. Gustavsson, and C. Studer, "One-bit massive MIMO: Channel estimation and high-order modulations," in *2015 IEEE International Conference on Communication Workshop (ICCW)*. IEEE, Jun. 2015, pp. 1304–1309.

[16] O. Orhan, E. Erkip, and S. Rangan, "Low power analog-to-digital conversion in millimeter wave systems: Impact of resolution and bandwidth on performance," in *2015 Information Theory and Applications Workshop, ITA 2015 - Conference Proceedings*. IEEE, Feb. 2015, pp. 191–198.

[17] T. C. Zhang, C. K. Wen, S. Jin, and T. Jiang, "Mixed-ADC Massive MIMO Detectors: Performance Analysis and Design Optimization," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7738–7752, Nov. 2016.

[18] L. Fan, S. Jin, C. K. Wen, and H. Zhang, "Uplink achievable rate for massive MIMO systems with low-resolution ADC," *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2186–2189, Dec. 2015.

[19] W. B. Abbas, F. Gomez-Cuba, and M. Zorzi, "Millimeter wave receiver efficiency: A comprehensive comparison of beamforming schemes with low resolution ADCs," in *IEEE Trans. Wireless Commun.*, vol. 16, no. 12, Dec. 2017, pp. 8131–8146.

[20] J. Mo, A. Alkhateeb, S. Abu-Surra, and R. W. Heath, "Hybrid Architectures with Few-Bit ADC Receivers: Achievable Rates and Energy-Rate Tradeoffs," *IEEE Trans. Wireless Commun.*, vol. 16, no. 4, pp. 2274–2287, Apr. 2017.

[21] J. Zhang, L. Dai, S. Sun, and Z. Wang, "On the Spectral Efficiency of Massive MIMO Systems with Low-Resolution ADCs," *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 842–845, May 2016.

[22] A. M. Q. Bai and J. A. Nossek, *On the Optimization of ADC Resolution in Multi-antenna Systems*. [VDE Verlag], 2013.

[23] J. Choi, B. L. Evans, and A. Gatherer, "ADC bit allocation under a power constraint for mmWave massive MIMO communication receivers," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017, pp. 3494–3498.

[24] J. Choi, B. L. Evans, and A. Gatherer, "Resolution-Adaptive Hybrid MIMO Architectures for Millimeter Wave Communications," *IEEE Trans. Signal Process.*, vol. 65, no. 23, pp. 6201–6216, Dec. 2017.

[25] A. Mezghani, R. Ghiat, and J. A. Nossek, "Transmit processing with low resolution D/A-converters," in *2009 16th IEEE International Conference on Electronics, Circuits and Systems, ICECS 2009*. IEEE, Dec. 2009, pp. 683–686.

[26] S. Jacobsson, G. Durisi, M. Coldrey, T. Goldstein, and C. Studer, "Quantized Precoding for Massive MU-MIMO," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 4670–4684, Nov. 2017.

[27] L. N. Ribeiro, S. Schwarz, M. Rupp, and A. L. De Almeida, "Energy Efficiency of mmWave Massive MIMO Precoding with Low-Resolution DACs," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 2, pp. 298–313, May 2018.

[28] B. Yang, Z. Yu, J. Lan, R. Zhang, J. Zhou, and W. Hong, "Digital beamforming-based massive mimo transceiver for 5g millimeter-wave communications," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 7, pp. 3403–3418, July 2018.

[29] A. Kaushik, J. Thompson, E. Vlachos, C. Tsinos, and S. Chatzinotas, "Dynamic RF Chain Selection for Energy Efficient and Low Complexity Hybrid Beamforming in Millimeter Wave MIMO Systems," *IEEE Trans. Green Commun. Netw.*, pp. 1–1, 2019.

[30] H. Li, M. Li, and Q. Liu, "Hybrid Beamforming With Dynamic Subarrays and Low-Resolution PSs for mmWave MU-MISO Systems," *IEEE Transactions on Communications*, vol. 68, no. 1, pp. 602–614, 2020.

[31] E. Vlachos, A. Kaushik, and J. Thompson, "Energy Efficient Transmitter with Low Resolution DACs for Massive MIMO with Partially Connected Hybrid Architecture," in *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*. IEEE, Jun. 2018, pp. 1–5.

[32] A. Kaushik, E. Vlachos, and J. Thompson, "Energy efficiency maximization of millimeter wave hybrid mimo systems with low resolution dacs," in *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, May 2019, pp. 1–6.

[33] C. A. Balanis, *Antenna theory: analysis and design*. Wiley-Interscience, 2005.

[34] T. E. Bogale, L. B. Le, A. Haghighat, and L. Vandendorpe, "On the number of rf chains and phase shifters, and scheduling design with hybrid analog-digital beamforming," *IEEE Transactions on Wireless Communications*, vol. 15, no. 5, pp. 3311–3326, May 2016.

[35] J. J. Bussgang and J. J. Bussgang, "Crosscorrelation functions of amplitude-distorted gaussian signals," *Tech. Rep. 216, Research Lab. Electron*, 1952.

[36] M. Fozooni, M. Matthaiou, E. Bjornson, and T. Q. Duong, "Performance limits of mimo systems with nonlinear power amplifiers," in *2015 IEEE Global Communications Conference (GLOBECOM)*, 2015, pp. 1–7.

[37] A. Gersho and R. M. Gray, *Vector quantization and signal compression*. Kluwer Academic Publishers, 1992.

[38] R. Mendez-Rial, C. Rusu, N. Gonzalez-Prelcic, A. Alkhateeb, and R. W. Heath, "Hybrid MIMO Architectures for Millimeter Wave Communications: Phase Shifters or Switches?" *IEEE Access*, vol. 4, pp. 247–267, 2016.

[39] A. Zappone and E. Jorswieck, "Energy Efficiency in Wireless Networks via Fractional Programming Theory," *Foundations and Trends® in Communications and Information Theory*, vol. 11, no. 3-4, pp. 185–396, 2015.

[40] J. P. Crouzeix and J. A. Ferland, "Algorithms for generalized fractional programming," *Mathematical Programming*, vol. 52, no. 1-3, pp. 191–207, May 1991.

[41] S. P. Chepuri and G. Leus, "Sparsity-promoting sensor selection for non-linear measurement models," *IEEE Transactions on Signal Processing*, vol. 63, no. 3, pp. 684–698, 2015.

[42] S. P. Boyd and L. Vandenberghe, "Convex Optimization," in *Communication Networking*. Cambridge University Press, 2004, pp. 803–818.

[43] A. Forenza, D. J. Love, and R. W. Heath, "Simplified spatial correlation models for clustered MIMO channels with different array configurations," *IEEE Trans. Veh. Technol.*, vol. 56, no. 4 II, pp. 1924–1934, 2007.

[44] P. V. Amadori and C. Masouros, "Low RF-complexity millimeter-wave beamspace-MIMO systems by beam selection," *IEEE Transactions on Communications*, vol. 63, no. 6, pp. 2212–2223, 2015.

[45] J. Li, L. Xiao, X. Xu, X. Su, and S. Zhou, "Energy-efficient Butler-matrix-based hybrid beamforming for multiuser mmWave MIMO system," *Science China Information Sciences*, vol. 60, no. 8, 2017.