Energy Efficiency Maximization of Millimeter Wave Hybrid MIMO Systems with Low Resolution DACs

Aryan Kaushik, Evangelos Vlachos and John Thompson
Institute for Digital Communications, The University of Edinburgh, United Kingdom.
Email: {a.kaushik, e.vlachos, j.s.thompson}@ed.ac.uk

***Abstract*—This paper proposes an energy efficient millimeter wave (mmWave) hybrid multiple-input multiple-output (MIMO) beamformer with low resolution digital to analog converters (DACs) at the transmitter. We consider the case where all DACs have the same sampling resolution for each radio frequency (RF) chain and select the best subset of the active RF chains and the DAC resolution. A novel technique based on the Dinkelbach method and subset selection optimization is proposed to maximize the energy efficiency (EE) given a predefined power budget for transmission. We also implement an exhaustive search approach to serve as an upper bound on the EE performance and show the performance trade-offs. The simulation results verify that the proposed technique exhibits EE performance similar to the optimal exhaustive search technique while requiring lower computational complexity.**

***Index Terms*—energy efficiency maximization, low resolution DACs, mmWave MIMO, hybrid beamforming.**

I. Introduction

Millimeter Wave (mmWave) technology can meet the needs of the fifth generation (5G) wireless communication systems and provide improved rate and capacity [1], [2]. The higher path loss associated with moving up in frequency from widely used cellular microwave bands can be compensated using large-scale antennas. The use of both antenna arrays and wide bandwidth frequencies at mmWave multiple-input multiple-output (MIMO) systems make it hard to implement one radio frequency (RF) chain and associated digital-to-analog/analog-to-digital converter (DAC/ADC) components per antenna [3]. The analog/digital hybrid beamforming architectures reduce the hardware complexity through fewer RF chains and support multi-stream communication with good capacity performance [4]–[6]. Moreover, implementing low resolution quantization in hybrid MIMO systems further improves the energy efficiency (EE) of such systems [3].

The existing literature mostly discusses low resolution DACs/ADCs with a large or full number of RF chains or full or high resolution sampling with a small number of RF chains. As the power consumption of DACs/ADCs increases exponentially with the number of bits, to further reduce the power consumption one can consider a combined analog and digital hybrid structure with small number of RF chains and low resolution DACs/ADCs. A hybrid beamforming system with low resolution sampling has been analyzed for channel estimation in [7]. To observe the effect of low resolution ADCs, an additive quantization model (AQNM) is considered in [8] for the case of a point-to-point mmWave MIMO system and [9] for the case of mmWave fading channels. Reference [10] assumes fully digital precoding at the transmitter, and baseband and RF combining with low resolution sampling at the receiver. Reference [11] works on the idea of a mixed-ADC architecture where a better energy-rate trade off is achieved with the use of a combination of low and high resolution ADCs than using only full resolution or low resolution systems. Most of the literature studies the use of low resolution sampling only at the receiver side, assuming fully digital or hybrid transmitters with high resolution DACs. Given the use of wide bandwidths in typical mmWave systems at the transmitter, employing low resolution DACs at transmitters can help to reduce the power consumption. So EE approaches that are mainly focused on ADCs at receiver can also be applied to the DACs at transmitter considering the transmitter specific system model parameters. Reference [12] uses low resolution DACs which can be implemented to reduce the power consumption for a hybrid MIMO architecture. Reference [13] employs low resolution DACs at the base station for a narrowband multi-user MIMO system. References [14], [15] consider the EE optimization problem for hybrid transceivers but with full resolution sampling at the DACs/ADCs.

Contributions: We consider a analog/digital hybrid transmit beamformer with low resolution DACs. The analog and digital parts are connected with a predefined number of RF chains which can be in active or inactive state. Assuming that the power consumption of the transmitter is determined mainly by the DACs of the RF chains, deactivating specific RF chains in an intelligent manner would increase the EE of the beamformer. Therefore, in this paper, we derive an optimal approach in terms of EE maximization, which selects the best subset between the available RF chains. We implement an iterative method to overcome the non-convexity of the fractional programming optimization problem. The proposed approach capitalizes from sparse-based subset selection techniques to provide an efficient solution to the problem. We also implement an exhaustive search approach (for example, in [14]) which expresses the upper bound for EE maximization and clearly shows the performance trade-offs.

Notation: $\mathbf{A}$, $\mathbf{a}$, and a denote a matrix, a vector, and a scalar, respectively. The complex conjugate transpose, and transpose of $\mathbf{A}$ are denoted as $\mathbf{A}^H$ and $\mathbf{A}^T$; tr($\mathbf{A}$) and $|\mathbf{A}|$ represent the trace and determinant of $\mathbf{A}$, respectively; $\mathbf{I}_N$ represents $N \times N$ identity matrix; $\mathbf{X} \in \mathbb{C}^{A\times B}$ and $\mathbf{X} \in \mathbb{R}^{A\times B}$ denote $A\times B$ size $\mathbf{X}$ matrix with complex and real entries, respectively;

978-1-5386-8088-9/19/$31.00 ©2019 IEEE

$\mathcal{CN}(\mathbf{a}, \mathbf{A})$ denotes a complex Gaussian vector having mean $\mathbf{a}$ and covariance matrix $\mathbf{A}$; $[\mathbf{A}]_k$ denotes the k-th column of matrix $\mathbf{A}$ and $[\mathbf{A}]_{kl}$ is the matrix entry at the k-th row and l-th column.

II. Hybrid mmWave MIMO

A. MmWave channel and system model

MmWave channels can be modeled by a narrowband clustered channel model due to different channel settings such as number of multipaths, amplitudes, etc., with N_{cl} clusters and N_{ray} propagation paths in each cluster [3], [4]. Considering a single-user mmWave system with N_{T} antennas at the transmitter, transmitting N_{s} data streams to N_{R} antennas at receiver, the mmWave channel matrix can be written as follows:

$$\mathbf{H} = \sum_{i=1}^{N_{\text{cl}}} \sum_{l=1}^{N_{\text{ray}}} \alpha_{il} \mathbf{a}_{\text{R}}(\phi_{il}^r) \mathbf{a}_{\text{T}}(\phi_{il}^t)^H, \tag{1}$$

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. Furthermore, $\mathbf{a}_{\text{T}}(\phi_{il}^t)$ and $\mathbf{a}_{\text{R}}(\phi_{il}^r)$ represent the normalized transmit and receive array response vectors [3], where ϕ_{il}^t and ϕ_{il}^r denote the azimuth angles of departure and arrival, respectively. We use uniform linear array (ULA) antennas for simplicity and model the antenna elements at the transmitter as ideal sectored elements [16]. However, the proposed technique is not limited to this setup and can be easily extended to the case of wideband channels and uniform planar arrays.

B. Quantization Model

We consider the linear model approximation (AQNM) to represent the introduced distortion of the quantization noise [18]. Given that $Q(\cdot)$ denotes a uniform scalar quantizer then for the scalar input s we have that,

$$Q(s) \approx \delta x + \epsilon, \tag{2}$$

where

$$\delta = \sqrt{1 - \frac{\pi\sqrt{3}}{2} 2^{-2b}} \tag{3}$$

is the multiplicative distortion parameter for bit sampling resolution equal to b and ϵ is the additive quantization noise with $\epsilon \sim \mathcal{CN}(0, \sigma_\epsilon^2)$, where

$$\sigma_\epsilon = \sqrt{1 - \frac{\pi\sqrt{3}}{2} 2^{-2b}} \sqrt{\frac{\pi\sqrt{3}}{2} 2^{-2b}} = \delta(1 - \delta^2). \tag{4}$$

C. System Model

In the analog and digital hybrid beamforming architecture, the number of transmitter RF chains L_{T} is usually smaller than the number of the transmitting antennas N_{T}, $L_{\text{T}} \leq N_{\text{T}}$, and similarly for the receiver, the number of RF chains $L_{\text{R}} \leq N_{\text{R}}$ (the number of receiving antennas). After the RF or analog precoding, each phase shifter is connected to all the antenna elements. Fig. 1 shows the system setup.

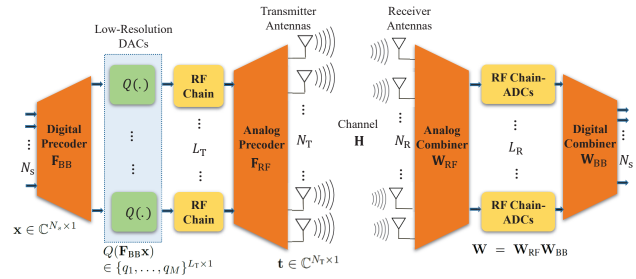

Fig. 1. A mmWave hybrid MIMO system with low resolution DACs.

Let $\mathbf{x} \in \mathbb{C}^{N_s \times 1}$ is the normalized data vector, then based on the AQNM the vector containing the complex output of all the DACs can be expressed as:

$$Q(\mathbf{F}_{\text{BB}}\mathbf{x}) \approx \delta \mathbf{F}_{\text{BB}}\mathbf{x} + \boldsymbol{\epsilon}, \tag{5}$$

where $Q(\mathbf{F}_{\text{BB}}\mathbf{x}) \in \mathbb{C}^{L_{\text{T}} \times 1}$ and $\mathbf{F}_{\text{BB}} \in \mathbb{C}^{L_{\text{T}} \times N_s}$ is the baseband part of transmit beamformer. The second term of (5) expresses the additive quantization noise for all RF chains with $\boldsymbol{\epsilon} \in \mathcal{CN}(\mathbf{0}, \sigma_\epsilon^2 \mathbf{I}_{L_{\text{T}}})$. This leads us to the following expression for the transmitted signal, as seen at the output of the analog and digital hybrid transmitter:

$$\mathbf{t} = \mathbf{F}_{\text{RF}} \left(\delta \mathbf{F}_{\text{BB}}\mathbf{x} + \boldsymbol{\epsilon} \right) = \delta \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \mathbf{x} + \mathbf{F}_{\text{RF}} \boldsymbol{\epsilon}, \tag{6}$$

where $\mathbf{F}_{\text{RF}}$ is the analog precoding matrix at the transmitter.

After the effect of the mmWave channel and the RF processing at the receiver, the received signal is expressed as:

$$\mathbf{y} = \mathbf{W}^H \mathbf{H} \mathbf{t} + \mathbf{W}^H \mathbf{n} \tag{7}$$

$$= \underbrace{\delta \mathbf{W}^H \mathbf{H} \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}}_{\mathbf{H}_{\text{eff}}(L_{\text{T}}, \delta)} \mathbf{x} + \underbrace{\mathbf{W}^H \mathbf{H} \mathbf{F}_{\text{RF}} \boldsymbol{\epsilon} + \mathbf{W}^H \mathbf{n}}_{\boldsymbol{\eta}}, \tag{8}$$

where $\mathbf{H}_{\text{eff}}(L_{\text{T}}, \delta)$ is the effective channel which is a function of the number of the RF chains L_{T} and the distortion δ, $\mathbf{W} \in \mathbb{C}^{N_{\text{R}} \times N_{\text{s}}}$ is the receiver combining matrix, $\boldsymbol{\eta}$ is the combined effect of the Gaussian and quantization noise with $\boldsymbol{\eta} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_\eta)$, while $\mathbf{R}_\eta$ is the combined noise covariance matrix with,

$$\mathbf{R}_\eta(L_{\text{T}}, \delta) = \sigma_\epsilon^2 \mathbf{W}^H \mathbf{H} \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{RF}}^H \mathbf{H}^H \mathbf{W} + \sigma_{\text{n}}^2 \mathbf{W}^H \mathbf{W}, \tag{9}$$

which is also a function of the number of the RF chains L_{T} and the distortion δ. Note that unlike what is common in the existing literature, in this work we also take into account the cross-terms of the noise covariance matrix $\mathbf{R}_\eta$. We believe this is a more realistic scenario since it can also incorporate system impairments such as phase noise into the problem formulation.

III. Energy Efficiency Maximization

The EE of a point-to-point MIMO system is defined as the ratio of the information rate and the total consumed power [22]. Since these quantities depend on the distortion of the DACs δ and the number of the RF chains L_{T}, EE is expressed as

$$\text{EE}(L_{\text{T}}, \delta) \triangleq \frac{R(L_{\text{T}}, \delta)}{P(L_{\text{T}}, \delta)} \text{ (bits/Joule)}. \tag{10}$$

Exploiting the linearity property of the quantization model in (5), the information rate $R(L_{\mathrm{T}},\delta)$ is expressed as:

$$R(L_{\mathrm{T}},\delta)=\log_2|\mathbf{I}_{N_s}+\frac{1}{N_s}\mathbf{R}_\eta^{-1}\mathbf{H}_{\mathrm{eff}}\mathbf{H}_{\mathrm{eff}}|\ \text{(bits/s/Hz)},\quad (11)$$

where the values of L_{T} and δ will affect the noise covariance matrix $\mathbf{R}_\eta(L_{\mathrm{T}},\delta)$ and the effective channel $\mathbf{H}_{\mathrm{eff}}(L_{\mathrm{T}},\delta)$.

Concerning the power consumption model, we consider that the total power consumption $P(L_{\mathrm{T}},\delta)$ is proportional to:

$$P(L_{\mathrm{T}},\delta)\propto L_{\mathrm{T}}\left[P_{\mathrm{DAC}}\left(\frac{\pi\sqrt{3}}{2(1-\delta^2)}\right)^{1/2}+N_{\mathrm{T}}P_{\mathrm{PS}}\right]\ (W) \quad (12)$$

where P_{DAC} and P_{S} depend upon the DAC and phase-shifter power consumption values, respectively.

Given the expressions (11) and (12), we can now define the EE maximization problem as a fractional programming problem:

$$\arg\max_{L_{\mathrm{T}},\delta}\ \mathrm{EE}(L_{\mathrm{T}},\delta)\ \text{subject to}\ P(L_{\mathrm{T}},\delta)\leq P_{\max},\qquad (13)$$

where $P_{\max}$ is the maximum available power budget. Our goal, by solving (13), is to obtain the number of RF chains and bit resolution in an optimal manner. To obtain a solution to (13) we have developed an iterative procedure that approximates the initial fractional problem with a convex-concave optimization, using Dinkelbach approximation [20] and subset selection. Dinkelbach approach makes an iterative approximation of the fractional problem with a sequence of non-fractional but constrained optimization ones. Although simpler, each one of these problems is still non-convex. However, by decomposing the contribution of each RF chain to the EE performance of the system, we can employ subset selection methods which minimize the number of the RF chains by solving an ℓ_1 approximation to the non-convex problem.

Before proceeding with the description of the proposed technique, we derive a technique based on exhaustive search for EE maximization, which will serve as an upper bound for comparison with the proposed method.

A. Upper Bound on EE via Exhaustive Search

To obtain an upper bound, we consider the case where $L_{\mathrm{T}}=N_{\mathrm{T}}$. This simplifies the computation of the beamformers at the receiver and the receiver, by using the singular value decomposition of the channel (SVD). However, since we change the number of the RF chains/antennas, the channel and its SVD, has to be updated at each time. Specifically, an exhaustive search approach is needed to obtain the optimum EE over all possible values of $(L_{\mathrm{T}},\delta)\in\{1,\ldots,b_{\max}\}\times\{1,\ldots,L_{\mathrm{T}}\}$. For each set value (L_{T},δ), the singular value decomposition (SVD) of the effective channel has to be obtained, i.e.,

$$\mathbf{H}_{\mathrm{eff}}(L_{\mathrm{T}},\delta)=\delta\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H,\qquad (14)$$

where $\mathbf{U}\in\mathbb{C}^{N_{\mathrm{R}}\times N_{\mathrm{R}}}$ and $\mathbf{V}\in\mathbb{C}^{N_{\mathrm{T}}\times N_{\mathrm{T}}}$ are unitary matrices, and $\mathbf{\Sigma}\in\mathbb{R}^{N_{\mathrm{R}}\times N_{\mathrm{T}}}$ is a rectangular matrix of singular values in decreasing order whose diagonal elements are non-negative real numbers and whose non-diagonal elements are zero. We assume that the rank of the channel is r.

Algorithm 1: Brute-force approach

Input: $b_{\max}$, $\mathbf{H}$
Begin:
1. **for** $b=1,\ldots,b_{\max}$
2. Compute $\delta(b)$ based on (3)
3. **for** $l_{\mathrm{t}}=1,\ldots,N_{\mathrm{T}}$
4. Compute the SVD of $\mathbf{H}_{\mathrm{eff}}(l_t,\delta(b_i))$ based on (14)
5. Compute $\mathrm{EE}(l_{\mathrm{t}},\delta(b))$ based on (11) and (12)
6. **end**
7. **end**
8. Find the L_{T}^{opt} and b^{opt} such as $\mathrm{EE}(L_{\mathrm{T}}^{\mathrm{opt}},\delta(b^{\mathrm{opt}}))>\mathrm{EE}(l_t,\delta(b))\quad\forall(b,l_t)$

Output: $L_{\mathrm{T}}^{\mathrm{opt}}$ and b^{opt}

Hence, the rate expression in (11) becomes:

$$\begin{aligned}R(L_{\mathrm{T}},\delta)&=\log_2|\mathbf{I}_{N_s}+\frac{\delta^2}{N_s}\mathbf{R}_\eta^{-1}\mathbf{W}^H\mathbf{H}\mathbf{F}\mathbf{F}^H\mathbf{H}^H\mathbf{W}|\\&=\log_2|\mathbf{I}_{N_s}+\frac{\delta^2}{N_s}\mathbf{R}_\eta^{-1}\mathbf{\Sigma}\mathbf{\Sigma}^H|\\&=\sum_{i=1}^{r}\log_2(1+\frac{\delta^2}{N_s}[\mathbf{R}_\eta^{-1}]_{ii}[\mathbf{\Sigma}\mathbf{\Sigma}^H]_{ii}),\end{aligned}\qquad (15)$$

where $\mathbf{R}_\eta$ becomes a diagonal matrix with entries $[\mathbf{R}_\eta]_{ii}=\sigma_\epsilon^2[\mathbf{\Sigma}\mathbf{\Sigma}^H]_{ii}+\sigma_n^2$. Based on (15), the rate expression is decomposed into the singular values domain, thus, the number of the rank r represents the *virtual* number of RF chains. So, the goal here is to reduce the number of virtual RF chains r, alongside with the distortion δ which depends on the bit resolution b.

Algorithm 1 shows the exhaustive search approach (similar to [14]), called the Brute-force technique, thus, it provides the solution to achieve the optimal number of RF chains and the optimal number of associated DAC bits at each channel realization. It makes a search of all the possible number of RF chains/antennas, i.e., $l_{\mathrm{t}}=\{1,\ldots,N_{\mathrm{T}}\}$ and over the available bit resolution, i.e., $b=1,\ldots,b_{\max}$, where $b_{\max}$ is the highest achievable resolution. It then finds the best EE out of all the efficiencies and chooses the corresponding optimal number of active RF chains L_{T}^{opt} and optimal resolution sampling b^{opt} for the transmitter. This method provides the best possible energy efficiency performance assuming that the SVD of $\mathbf{H}$ is perfectly known at the transmitter.

B. Proposed Method

Let us now consider an optimal design where we seek the sampling resolution for each DACs and the optimal number of active RF chains L_{T} that will maximize the EE of the transmitter. We consider a variable number of RF chains, i.e., by using switches to activate/deactivate each one independently [19], then the problem becomes:

$$\arg\max_{\mathbf{S},\delta}\frac{R(\mathbf{S},\delta)}{P(\mathbf{S},\delta)}\ \text{subject to}\ P(\mathbf{S},\delta)\leq P_{\max},\qquad (16)$$

where $\mathbf{S} \in \{0,1\}^{L_\mathrm{T} \times L_\mathrm{T}}$ is a diagonal binary matrix representing switches which activate or deactivate the RF chains. Hence, the resulting optimization problem of (16) has two unknown quantities to be recovered, the matrices $\mathbf{S}$ and δ. We transform the problem into a subset selection based problem considering sparse optimization and compressive sampling.

We consider the problem to be equivalent to finding only a sparse selection vector, 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 zero value represents an inactive RF chain. It is important to note that based on the proposed architecture, the optimization problem does not consider a predefined number of active/inactive RF chains, but this quantity is an optimization variable. Incorporating this selection procedure into our formulation, the received signal $\hat{\mathbf{y}} \in \mathbb{C}^{N_\mathrm{s} \times 1}$ at the baseband receiver is expressed as:

$$\hat{\mathbf{y}} = \delta \mathbf{W}^H \mathbf{H} \mathbf{F}_\mathrm{RF} \mathbf{S} \mathbf{F}_\mathrm{BB} \mathbf{x} + \boldsymbol{\eta}, \tag{17}$$

where $\mathbf{S} \in \{0,1\}^{L_\mathrm{T} \times L_\mathrm{T}}$ is a diagonal selection matrix composed by zeros and ones, with $[\mathbf{S}]_{kk} \in \{0,1\}$ and $[\mathbf{S}]_{kl} = 0$ for $k \neq l$; $\delta \mathbf{W}^H \mathbf{H} \mathbf{F}_\mathrm{RF} \mathbf{S} \mathbf{F}_\mathrm{BB}$ is the effective channel $\hat{\mathbf{H}}_\mathrm{eff} \in \mathbb{C}^{N_\mathrm{s} \times N_\mathrm{s}}$ in this case, including hybrid transmitter precoding and receiver combining and quantization distortion. The parameter that we aim to optimize in (17) is now the entries of the diagonal selection matrix $\mathbf{S} \in \{0,1\}^{L_\mathrm{T} \times L_\mathrm{T}}$. The effective channel can be decomposed as:

$$\begin{aligned} \hat{\mathbf{H}}_\mathrm{eff} &= \delta \mathbf{W}^H \mathbf{H} \mathbf{F}_\mathrm{RF} \mathbf{S} \mathbf{F}_\mathrm{BB} & (18)\\ &= \sum_{i=1}^{L_\mathrm{T}} [\mathbf{S}]_{ii} [\delta \mathbf{W}^H \mathbf{H} \mathbf{F}_\mathrm{RF}]_i [\mathbf{F}_\mathrm{BB}^T]_i^T \\ &= \sum_{i=1}^{L_\mathrm{T}} [\mathbf{S}]_{ii} \mathbf{a}_i \mathbf{b}_i^T, & (19) \end{aligned}$$

where $\mathbf{b}_i \triangleq [\mathbf{F}_\mathrm{BB}^T]_i \in \mathbb{C}^{N_s \times 1}$, $\mathbf{a}_i \triangleq [\delta \mathbf{R}_\eta^{-\frac{1}{2}} \mathbf{W}^H \mathbf{H} \mathbf{F}_\mathrm{RF}]_i \in \mathbb{C}^{N_s \times 1}$ and where $[\mathbf{S}]_{ii} \in \{0,1\}$ determines the state of the i-th RF chain. Based on (19), the received signal can be equivalently expressed as the following measurement vector:

$$\hat{\mathbf{y}} = \sum_{i=1}^{L_\mathrm{T}} [\mathbf{S}]_{ii} \mathbf{a}_i (\mathbf{b}_i^T \mathbf{x}) + \hat{\boldsymbol{\eta}}, \tag{20}$$

where $\hat{\boldsymbol{\eta}} \triangleq \mathbf{S}\boldsymbol{\eta}$ whose noise covariance matrix can be expressed with respect to the selection matrix, i.e.,

$$\hat{\mathbf{R}}_\eta = \sigma_\epsilon^2 \mathbf{W}^H \mathbf{H} \mathbf{F}_\mathrm{RF} \mathbf{S} \mathbf{F}_\mathrm{BB} \mathbf{F}_\mathrm{BB}^H \mathbf{S} \mathbf{F}_\mathrm{RF}^H \mathbf{H}^H \mathbf{W} + \sigma_\mathrm{n}^2 \mathbf{W}^H \mathbf{W}. \tag{21}$$

The problem becomes equivalent with the estimation of $\mathbf{S}$ that maximizes the EE of the hybrid precoder. It can be shown that the rate and power equations for such scenario can be expressed as:

$$R(\mathbf{S}, \delta) = \log_2 \left| \mathbf{I}_{N_\mathrm{s}} + \frac{1}{N_\mathrm{s}} \sum_{i=1}^{L_\mathrm{T}} [\mathbf{S}]_{ii} \mathbf{a}_i^H \mathbf{a}_i \mathbf{b}_i \mathbf{b}_i^H \right|, \tag{22}$$

Algorithm 2: Proposed technique

Input: $\kappa^{(0)}$, $\mathbf{H}$
Begin:
1. **for** $b = 1, ..., b_\mathrm{max}$
2. Compute $\mathbf{H}_\mathrm{eff}(N_\mathrm{T}, \delta(b))$
3. **for** $m = 1, 2, \ldots, I_\mathrm{max}$
4. Obtain $\mathbf{S}^{(m)}$ by solving (25) given $\kappa^{(m-1)}$.
5. Calculate $R(\mathbf{S}^{(m)}, \delta^{(m)})$ and $P(\mathbf{S}^{(m)}, \delta^{(m)})$.
6. Compute $\kappa^{(m)} = R(\mathbf{S}^{(m)}, \delta^{(m)}) / P(\mathbf{S}^{(m)}, \delta^{(m)})$.
7. **end**
8. **end**
Output: Optimal ${L_\mathrm{T}}^{opt}$ and b^{opt}

and

$$P(\mathbf{S}, \delta) \propto \sum_{i=1}^{L_\mathrm{T}} [\mathbf{S}]_{ii} \left[P_\mathrm{DAC} \left(\frac{\pi \sqrt{3}}{2(1-\delta^2)} \right)^{1/2} + N_\mathrm{T} P_\mathrm{PS} \right] \tag{23}$$

$$= L_\mathrm{T} \left[P_\mathrm{DAC} \left(\frac{\pi \sqrt{3}}{2(1-\delta^2)} \right)^{1/2} + N_\mathrm{T} P_\mathrm{PS} \right]. \; (W) \tag{24}$$

The problem of maximizing EE (16) is a concave-convex fractional problem and one solution method is the Dinkelbach approximation [20]. The Dinkelbach method is an iterative and parametric algorithm, where a sequence of easier problems converge to the global solution. Let $\kappa^{(m)} = R(\mathbf{S}^{(m)}, \delta^{(m)}) / P(\mathbf{S}^{(m)}, \delta^{(m)}) \in \mathbb{R}$, for $m = 1, 2, \ldots, I_\mathrm{max}$, where I_max is the number of maximum iterations, then each iteration step of Dinkelbach can be expressed as:

$$\mathbf{S}^{(m)}(\kappa^{(m)}) \triangleq \arg\max_{\mathbf{S} \in \mathcal{S}} \left\{ R(\mathbf{S}, \delta) - \kappa^{(m)} P(\mathbf{S}, \delta) \right\}, \tag{25}$$

where $\mathcal{S}$ is the set of diagonal matrices with the feasible bit allocations which satisfy $P(\mathbf{S}, \delta) \leq P_\mathrm{max}$. Algorithm 2 summarizes the Dinkelbach algorithm via the subset selection approach where the optimal number of RF chains and associated sampling resolution is obtained.

Computational Complexity: It can be observed that the Dinkelbach method via subset selection approach requires complexity order of only $b_\mathrm{max}\mathcal{O}(L_\mathrm{T}^3)$ per iteration and the Brute-force approach requires complexity order of $b_\mathrm{max}\mathcal{O}(L_\mathrm{T}^2 N_\mathrm{T})$. Since the number of the required iterations is usually very small (as shown in Fig. 2) as $\mathbf{F}$ and $\mathbf{W}$ matrices are required to be computed in Algorithm 1 and not Algorithm 2, the overall complexity of the Dinkelbach method via subset selection approach is much less than the Brute-force approach.

IV. Simulation Results

In this section, we evaluate the performance of the proposed technique using computer simulation results. The simulations are performed with MATLAB$^\mathrm{TM}$ and all the results have been averaged over 1,000 Monte-Carlo realizations.

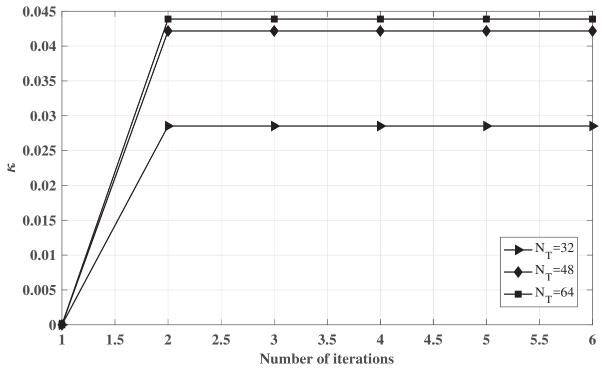

Fig. 2. Convergence of the proposed Dinkelbach method for different number of transmitter antennas at SNR = 30 dB, $N_{\mathrm{R}} = 32$, $L_{\mathrm{T}} = 32$ and $N_{\mathrm{s}} = 8$.

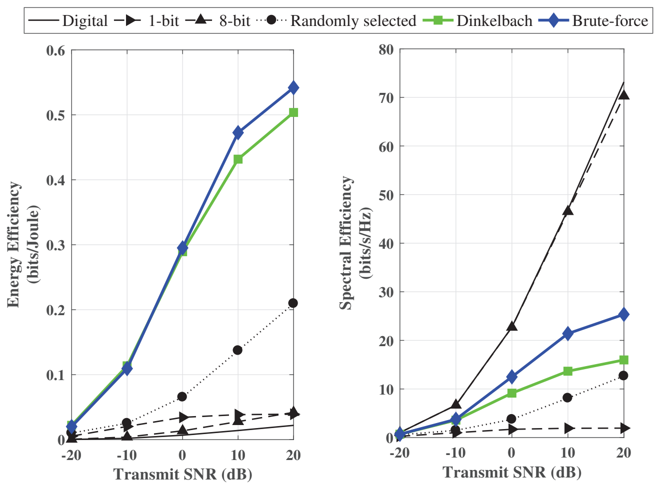

Fig. 3. Energy efficiency and spectral efficiency performance comparison w.r.t. transmit SNR (dB) at $N_{\mathrm{T}} = 64$, $N_{\mathrm{R}} = 32$, $L_{\mathrm{T}} = 32$ and $N_{\mathrm{s}} = 8$.

We set the following baseline parameters for simulation: $N_{\mathrm{T}} = 64$, $N_{\mathrm{R}} = 32$, $L_{\mathrm{T}} = 32$ (the number of available RF chains), $N_{\mathrm{s}} = 8$, $N_{\mathrm{cl}} = 2$, $N_{\mathrm{ray}} = 10$, and $\sigma^2_{\alpha,i} = 1$. The azimuth angles of departure and arrival are computed with uniformly distributed mean angles; each cluster follows a Laplacian distribution with mean angles equal to zero. The antenna elements in the ULA are spaced by distance $d = \lambda/2$.

Concerning the quantization model, since DACs have the same sampling resolution for each RF chain the quantization distortion parameter is the same for all DACs and the highest bit resolution $b_{\max} = 8$. The typical values of power terms for the power model in (12) of Section III are $P_{\mathrm{PS}} = 10$ mW, $P_{\mathrm{DAC}} = 0.1$ W and $P_{\max} = 1$ W. We solve the sparse approximation problem for the RF and baseband precoding matrices $\mathbf{F}_{\mathrm{RF}}$ and $\mathbf{F}_{\mathrm{BB}}$ using orthogonal matching pursuit (OMP) [4], [6], and the combiner matrix $\mathbf{W}$ is the product of $1/\sqrt{N_{\mathrm{s}}}$ and first N_{s} columns of $\mathbf{U}$ matrix.

For comparison with the proposed Dinkelbach method via subset selection solution, we have considered the digital beam-

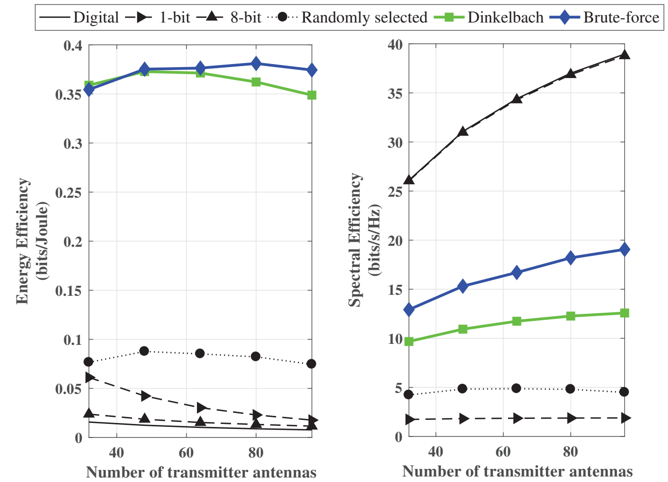

Fig. 4. Energy efficiency and spectral efficiency performance comparison w.r.t. the number of transmitter antennas at SNR = 5 dB, $N_{\mathrm{R}} = 32$, $L_{\mathrm{T}} = 32$ and $N_{\mathrm{s}} = 8$.

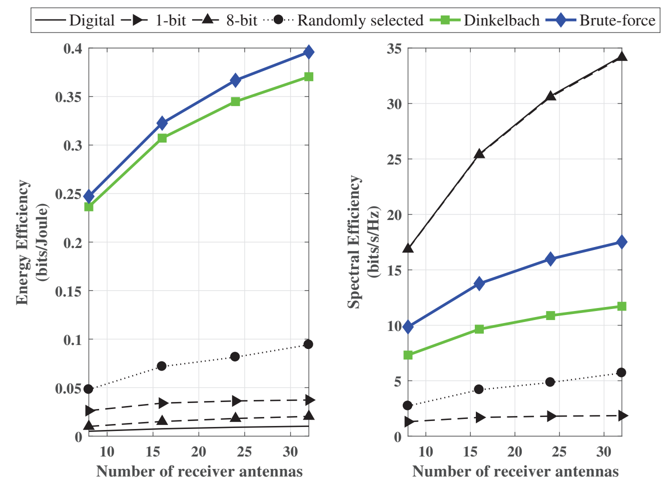

Fig. 5. Energy efficiency and spectral efficiency performance comparison w.r.t. the number of receiver antennas at SNR = 5 dB, $N_{\mathrm{T}} = 64$, $L_{\mathrm{T}} = 32$ and $N_{\mathrm{s}} = 8$.

forming architecture ($L_{\mathrm{T}} = N_{\mathrm{T}}$) with 8-bit DACs, which represents the optimum from the achievable spectral efficiency (SE) perspective, combined analog and digital hybrid precoding with L_{T} RF chains for 1-bit and 8-bit DACs, which represent the lowest and the highest SE cases. We also compare with the hybrid beamforming for L_{T} RF chains with a random resolution selected for each DAC from the range $[1, 8]$-bit, and hybrid beamforming with the optimal number of active RF chains L_{T}^{opt} and corresponding optimal sampling resolution b^{opt} obtained from the Brute-force approach.

Fig. 2 shows the convergence of the Dinkelbach method based solution as proposed in Algorithm 2 to obtain the optimal number of active RF chains and corresponding optimal sampling resolution. It can be observed that the performance curves based on the current EE κ (step 6 of Algorithm 2) for different numbers of transmitter antennas increase with respect to (w.r.t.) the number of iterations. The proposed solution converges rapidly and needs only 2-3 iterations to converge, and achieves an optimal solution at each realization.

It can be clearly observed from Fig. 3 that the proposed solution achieves a similar EE performance w.r.t. signal-to-noise ratio (SNR) as the Brute-force approach and outperforms hybrid 1-bit and hybrid 8-bit quantized DACs, plus the hybrid randomly selected resolution and digital beamforming with full-bit (8-bit) quantization. For example, at 10 dB SNR, EE for the proposed solution is approximating the Brute-force solution performance, about 0.3 bits/Joule better than the randomly selected resolution with hybrid beamforming, about 0.35 bits/Joule better than the hybrid 1-bit and about 0.38 bits/Joule better than the hybrid 8-bit and digital beamforming baselines. The proposed solution also achieves SE performance higher than the randomly selected and 1-bit quantization baselines. Digital beamforming and 8-bit hybrid baselines have the highest rate performance by using higher rate 8-bit quantization. For example, at 0 dB SNR, the proposed solution outperforms randomly selected quantization by about 7 bits/s/Hz, 1-bit hybrid by about 9 bits/s/Hz. Concerning the lower SE performance of the proposed technique and the Brute-Force approach, this is due to the fact that Brute-force has no constraint in the overall power consumption.

Fig. 4 shows similar performance behavior when plotting EE and SE w.r.t. the number of transmitter antennas at 5 dB SNR. For example, for $N_{\mathrm{T}} = 80$, the proposed solution has performance close to the Brute-force approach, performs about 0.3 bits/Joule and about 7.5 bits/s/Hz better than the hybrid randomly selected resolution baseline, about 0.35 bits/Joule and 10 bits/s/Hz better than the 1-bit hybrid baseline. Fig. 5 plots the performance comparison of the proposed solution with the baselines w.r.t. number of receiver antennas at 5 dB SNR. Similar to above plots, it achieves high SE and has almost the same EE performance as the Brute-force approach.

V. Conclusion

We consider a mmWave hybrid MIMO system with analog and digital parts connected with fewer number of RF chains than the transmitting antennas, while transmitter DACs operate with low-resolution sampling. We consider the case where all DACs have the same sampling resolution for each RF chain and aim to optimize the number of active RF chains and associated resolution of DACs. The proposed method achieves similar EE performance with the upper bound of the derived exhaustive search approach, while it exhibits lower computational complexity and fast convergence. Future work will include the optimization of energy efficiency with different bit resolutions for every RF chain.

Acknowledgment

The authors gratefully acknowledge partial funding of this work by EPSRC grant EP/P000703/1.

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] J. G. Andrews et al., "What will 5G be", *IEEE Journ. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065-1082, Jun. 2014.
[3] R. W. Heath et al.,"An overview of signal processing techniques for millimeter wave MIMO systems", *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436-453, Apr. 2016.
[4] 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.
[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] 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.
[7] A. Kaushik et al.,"Efficient channel estimation in millimeter wave hybrid MIMO systems with low resolution ADCs", in *IEEE Europ. Sig. Process.*, Rome, Italy, pp. 1839-1843, Sept. 2018.
[8] O. Orhan et al., "Low power analog-to-digital conversion in millimeter wave systems: Impact of resolution and bandwidth on performance", *2015 Info. Theory Appl. Workshop, San Diego, CA*, pp. 191-198, 2015.
[9] L. Fan et al., "Uplink achievable rate for massive MIMO systems with low-resolution ADC", *IEEE Commun. Letters*, vol. 19, no. 12, pp. 2186-2189, Oct. 2015.
[10] R. Mendez-Rial et al., "Hybrid MIMO architectures for millimeter wave communications: Phase shifters or switches?", *IEEE Access*, vol. 4, pp. 247-267, Jan. 2016.
[11] J. Zhang et al., "Performance analysis of mixed-ADC massive MIMO systems over Rician fading channels," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1327-1338, Jun. 2017.
[12] A. Mezghani et al., "Transmit processing with low resolution D/A-converters", in *Proc. Int. Conf. Electronics, Circuits, and Systems*, Tunisia, pp. 683-686, Dec. 2009.
[13] S. Jacobsson et al., "Quantized precoding for massive MU-MIMO", in *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 4670-4684, Nov. 2017.
[14] 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.
[15] C. G. Tsinos et al., "On the energy-efficiency of hybrid analog-digital transceivers for single- and multi-carrier large antenna array systems", in *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1980-1995, Sept. 2017.
[16] 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.
[17] C. Balanis, *Antenna Theory*, Wiley, 1997.
[18] J. Mo et al., "Channel estimation in broadband millimeter wave MIMO systems with few-bit ADCs," in *IEEE Trans. Sig. Process.*, vol. 66, no. 5, pp. 1141-1154, Mar. 2018.
[19] 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, Jun. 2018.
[20] W. Dinkelbach, "On nonlinear fractional programming", *Management Science*, vol. 13, no. 7, pp. 492-498, Mar. 1967.
[21] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs", in *Recent Adv. Learning and Control*, Springer-Verlag Ltd., pp. 95-110, 2008.
[22] 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.