Energy Efficient ADC Bit Allocation and Hybrid Combining for Millimeter Wave MIMO Systems

Aryan Kaushik[1], Christos Tsinos[2], Evangelos Vlachos[1], John Thompson[1]
[1]Institute for Digital Communications, The University of Edinburgh, United Kingdom.
[2]Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg, Luxembourg.
Emails: {a.kaushik, e.vlachos, j.s.thompson}@ed.ac.uk, christos.tsinos@uni.lu

Abstract—**Low resolution analog-to-digital converters (ADCs) can be employed to improve the energy efficiency (EE) of a wireless receiver since the power consumption of each ADC is exponentially related to its sampling resolution and the hardware complexity. In this paper, we aim to jointly optimize the sampling resolution, i.e., the number of ADC bits, and analog/digital hybrid combiner matrices which provides highly energy efficient solutions for millimeter wave multiple-input multiple-output systems. A novel decomposition of the hybrid combiner to three parts is introduced: the analog combiner matrix, the bit resolution matrix and the baseband combiner matrix. The unknown matrices are computed as the solution to a matrix factorization problem where the optimal, fully digital combiner is approximated by the product of these matrices. An efficient solution based on the alternating direction method of multipliers is proposed to solve this problem. The simulation results show that the proposed solution achieves high EE performance when compared with existing benchmark techniques that use fixed ADC resolutions.**

Index Terms—**energy efficient design, optimal bit resolution and hybrid combining, mmWave MIMO.**

I. INTRODUCTION

The analog/digital (A/D) hybrid beamforming architectures for millimeter wave (mmWave) multiple-input multiple-output (MIMO) systems reduce the hardware complexity and the power consumption through fewer radio frequency (RF) chains and support multi-stream communication with good capacity performance [1]–[3]. Designing such systems for high energy efficiency (EE) gains would leverage their significance [4], [5]. An alternative solution to reduce the power consumption and hardware complexity is by reducing the resolution sampling [6]. Some approaches have been applied in hybrid mmWave MIMO systems for EE maximization and low complexity with full resolution [7] and low resolution [8].

The existing literature mostly discusses full or high resolution analog-to-digital converters (ADCs) with a small number of RF chains or low resolution ADCs with a large number of RF chains: either way only the fixed resolution ADCs are taken into account. References [4], [5] consider EE optimization problems for A/D hybrid transceivers but with fixed and high resolution digital-to-analog converters (DACs)/ADCs. Reference [8] proposes a novel EE maximization transmission technique with subset selection optimization to find the best subset of the active RF chains and DAC resolution, which can be extended to low resolution ADCs at the receiver (RX). Reference [9] suggests implementing fixed and low resolution

ADCs with few RF chains. Reference [10] studies the idea of a mixed-ADC architecture where a better energy-rate trade off is achieved by using mixed resolution ADCs but still with a fixed resolution for each ADC and it does not consider A/D hybrid beamforming. A hybrid beamforming system with fixed and low resolution ADCs has been analyzed for channel estimation in [11]. Varying resolution ADCs can be implemented at the RX [12] which may provide a better solution than fixed and low resolution ADCs. Extra care is needed when deciding the range of number of ADC bits as the total ADC power consumption can be dominated by only a few high resolution ADCs. Thus, a good trade-off between power consumption and performance is to consider the range of 1-8 bits for the varying number of ADC bits.

Contributions: This paper designs an optimal EE solution for a mmWave A/D hybrid receiver MIMO system by introducing the novel decomposition of the A/D hybrid combiner to three parts representing the analog combiner matrix, the bit resolution matrix and baseband combiner matrix. Our aim is to minimize the distance between this decomposition, which is expressed as the product of three matrices, and the fully digital combiner matrix. The joint problem is decomposed into a series of sub-problems which are solved using an alternating optimization framework, i.e., alternating direction method of multipliers (ADMM) is developed to obtain the unknown matrices. The proposed design has high flexibility, given that the analog combiner is codebook-free, thus there is no restriction on the angular vectors and different bit resolutions can be assigned to each ADC. Our proposed solution optimizes the resolution on a packet-by-packet basis for each one of the ADCs unlike existing approaches that are based on fixed resolution sampling. We also implement an exhaustive search approach [4] for comparison which provides the upper bound for EE maximization.

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 $\mathbf{A}^T$; $|a|$ represents the determinant of a; $\mathbf{I}_N$ 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; $\mathcal{CN}(\mathbf{a}, \mathbf{A})$ denotes a complex Gaussian vector having mean $\mathbf{a}$ and covariance matrix $\mathbf{A}$; $[\mathbf{A}]_{kl}$ is the matrix entry at the k-th row and l-th column. The indicator function $\mathbb{1}_{\mathcal{S}}\{\mathbf{A}\}$ of a set $\mathcal{S}$ that acts over a matrix $\mathbf{A}$ is defined as $0 \ \forall \ \mathbf{A} \in \mathcal{S}$ and $\infty \ \forall \ \mathbf{A} \notin \mathcal{S}$.

II. A/D Hybrid MmWave MIMO System

A. MmWave Channel 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 [1]. Considering a single user mmWave system with N_T antennas at the transmitter (TX), transmitting N_s data streams to N_R antennas at the RX, the mmWave channel matrix can be written as follows:

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

where $\alpha_{il} \in \mathcal{CN}(0, \sigma_{\alpha,i}^2)$ is the gain term with $\sigma_{\alpha,i}^2$ being the average power of the ith cluster. Furthermore, $\mathbf{a}_T(\phi_{il}^t)$ and $\mathbf{a}_R(\phi_{il}^r)$ represent the normalized transmit and receive array response vectors [1], 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 RX as ideal sectored elements [13]. However, the proposed technique is not limited to this setup and can be easily extended to the case of wideband channels and uniform planar/circular arrays.

B. A/D Hybrid MIMO System Model

Based on the A/D hybrid beamforming scheme in the large-scale mmWave MIMO communication systems, the number of RX RF chains L_R follows the limitation $N_s \leq L_R \leq N_R$ [1], [2]. The matrices $\mathbf{W}_{RF} \in \mathbb{C}^{N_R \times L_R}$ and $\mathbf{W}_{BB} \in \mathbb{C}^{L_R \times N_s}$ denote the analog combiner and baseband (or digital) combiner matrices, respectively. The analog combiner matrix $\mathbf{W}_{RF}$ is based on phase shifters, i.e., the elements that have unit modulus and continuous phase. Thus, $\mathbf{W}_{RF} \in \mathcal{W}^{N_R \times L_R}$ where the set $\mathcal{W}$ represents the set of possible phase shifts in $\mathbf{W}_{RF}$ and for a variable a, is defined as, $\mathcal{W} = \{a \in \mathbb{C} \mid |a| = 1\}$. At the TX, with L_T RF chains, the analog precoder matrix is denoted as $\mathbf{F}_{RF} \in \mathbb{C}^{N_T \times L_T}$ and the baseband precoder matrix is denoted as $\mathbf{F}_{BB} \in \mathbb{C}^{L_T \times N_s}$. The received signal $\mathbf{y} \in \mathbb{C}^{N_R \times 1}$ can be expressed as:

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

where $\mathbf{x} \in \mathbb{C}^{N_s \times 1}$ is the transmit symbol vector and $\mathbf{n} \in \mathbb{C}^{N_R \times 1}$ is a noise vector with independent and identically distributed entries and follow the complex Gaussian distribution with zero mean and σ_n^2 variance, i.e., $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2 \mathbf{I}_{N_R})$.

As widely used in the existing literature, we consider the linear additive quantization noise model (AQNM) to represent the distortion of quantization [14]. Given that $Q(\cdot)$ denotes a uniform scalar quantizer then for the scalar complex input $x \in \mathbb{C}$ that is applied to both the real and imaginary parts, we have that,

$$Q(x) \approx \delta x + \epsilon, \quad (3)$$

where $\delta = \sqrt{1 - \frac{\pi\sqrt{3}}{2} 2^{-2b}} \in [m, M]$ is the multiplicative distortion parameter for a bit resolution equal to b [15]
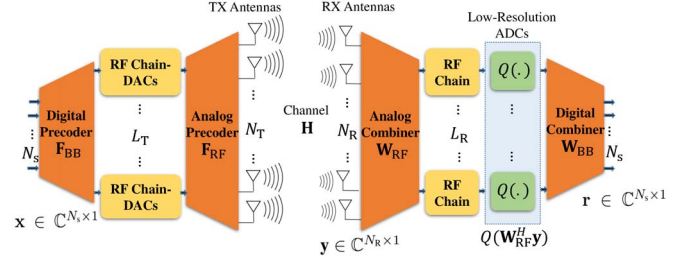

Fig. 1. A mmWave A/D hybrid MIMO system with low resolution ADCs.

where m and M denote the minimum and maximum value of the range. Note that the introduced error in the linear approximation in (3) decreases for larger resolutions. However, our proposed solution focuses on EE maximization and this linear approximation does not impact the performance significantly as observed from the simulation results in Section IV. The parameter ϵ is the additive quantization noise with $\epsilon \sim \mathcal{CN}(0, \sigma_\epsilon^2)$, where $\sigma_\epsilon = \sqrt{1 - \frac{\pi\sqrt{3}}{2} 2^{-2b}} \sqrt{\frac{\pi\sqrt{3}}{2} 2^{-2b}}$. Based on AQNM, the vector containing the complex output of all the ADCs can be expressed as follows:

$$Q(\mathbf{W}_{RF}^H \mathbf{y}) \approx \mathbf{\Delta}^H \mathbf{W}_{RF}^H \mathbf{y} + \boldsymbol{\epsilon}, \quad (4)$$

where $Q(\mathbf{W}_{RF}^H \mathbf{y}) \in \mathbb{C}^{L_R \times 1}$ and $\mathbf{\Delta} = \mathbf{\Delta}^H \in \mathbb{C}^{L_R \times L_R}$ is a diagonal matrix with values depending on the ADC resolution b_i of each ADC. Specifically, each diagonal entry of $\mathbf{\Delta}$ is given by:

$$[\mathbf{\Delta}]_{ii} = \sqrt{1 - \frac{\pi\sqrt{3}}{2} 2^{-2b_i}} \in [m, M] \, \forall \, i = 1, \ldots, L_R, \quad (5)$$

where, for simplicity, we assume that the range $[m, M]$ is the same for each one of the ADCs. The second term of (4) expresses the additive quantization noise for all RF chains, with $\boldsymbol{\epsilon} \in \mathcal{CN}(\mathbf{0}, \mathbf{C}_\epsilon)$ [8] where $\mathbf{C}_\epsilon$ is a diagonal covariance matrix with entries as follows:

$$[\mathbf{C}_\epsilon]_{ii} = \left(1 - \frac{\pi\sqrt{3}}{2} 2^{-2b_i}\right)\left(\frac{\pi\sqrt{3}}{2} 2^{-2b_i}\right) \forall \, i = 1, \ldots, L_R. \quad (6)$$

After the effect of the quantization and application of the baseband combining matrix, the output $\mathbf{r} \in \mathbb{C}^{N_s \times 1}$ at the RX can be expressed as:

$$\mathbf{r} = \mathbf{W}_{BB}^H \mathbf{\Delta}^H \mathbf{W}_{RF}^H \mathbf{y} + \mathbf{W}_{BB}^H \boldsymbol{\epsilon}. \quad (7)$$

Based on the received signal expression in (2), we can express (7) as follows:

$$\mathbf{r} = \mathbf{W}_{BB}^H \mathbf{\Delta}^H \mathbf{W}_{RF}^H \mathbf{H}\mathbf{F}_{RF}\mathbf{F}_{BB}\mathbf{x} + \underbrace{\mathbf{W}_{BB}^H \mathbf{\Delta}^H \mathbf{W}_{RF}^H \mathbf{n} + \mathbf{W}_{BB}^H \boldsymbol{\epsilon}}_{\boldsymbol{\eta}}, \quad (8)$$

where $\boldsymbol{\eta}$ is the combined effect of the Gaussian and the quantization noise with $\boldsymbol{\eta} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_\eta)$. Here $\mathbf{R}_\eta \in \mathbb{C}^{L_R \times L_R}$ is the combined noise covariance matrix with,

$$\mathbf{R}_\eta = \sigma_n^2 \mathbf{W}_{BB}^H \mathbf{\Delta}^H \mathbf{W}_{RF}^H \mathbf{W}_{RF} \mathbf{\Delta} \mathbf{W}_{BB} + \mathbf{W}_{BB}^H \mathbf{C}_\epsilon \mathbf{W}_{BB}. \quad (9)$$

III. BIT ALLOCATION AND HYBRID COMBINER DESIGN

A. Problem Formulation

Let us consider a point-to-point MIMO system with the linear quantization model. We define the EE as the ratio of the information rate and the total consumed power as,

$$EE(\mathbf{W}_{\text{RF}}, \mathbf{\Delta}, \mathbf{W}_{\text{BB}}) \triangleq \frac{R(\mathbf{W}_{\text{RF}}, \mathbf{\Delta}, \mathbf{W}_{\text{BB}})}{P(\mathbf{\Delta})} \text{ (bits/Joule)}, \tag{10}$$

where the information rate is defined as,

$$R(\mathbf{W}_{\text{RF}}, \mathbf{\Delta}, \mathbf{W}_{\text{BB}}) \triangleq \log_2 |\mathbf{I}_{L_{\text{R}}} + \frac{\mathbf{R}_\eta^{-1}}{N_{\text{s}}} \mathbf{W}_{\text{BB}}^H \mathbf{\Delta}^H \mathbf{W}_{\text{RF}}^H \mathbf{H} \mathbf{F} \times$$
$$\mathbf{F}^H \mathbf{H}^H \mathbf{W}_{\text{RF}} \mathbf{\Delta} \mathbf{W}_{\text{BB}}| \text{ (bits/s)}, \tag{11}$$

where the A/D hybrid precoder $\mathbf{F} = \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \in \mathbb{C}^{N_{\text{T}} \times N_{\text{s}}}$.

Similar to the power model at the TX in [8], the total consumed power at the RX is expressed as:

$$P(\mathbf{\Delta}) = P_{\text{D}} + N_{\text{R}} P_{\text{R}} + N_{\text{R}} L_{\text{R}} P_{\text{PS}} + P_{\text{CP}} \text{ (W)}, \tag{12}$$

where P_{PS} is the power per phase shifter, P_{R} is the power per antenna, P_{D} is the power associated with the total quantization operation, and following (5) and [14], we have

$$P_{\text{D}} = P_{\text{ADC}} \sum_{i=1}^{L_{\text{R}}} 2^{b_i} = P_{\text{ADC}} \sum_{i=1}^{L_{\text{R}}} \left(\frac{\pi \sqrt{3}}{2(1 - [\mathbf{\Delta}]_{ii}^2)} \right)^{\frac{1}{2}} \text{ (W), } \tag{13}$$

where P_{ADC} is the power consumed per bit in the ADC and P_{CP} is the power required by all circuit components.

Considering the rate and power model in (11) and (12), respectively, we can express the following fractional problem:

$$(\mathcal{P}_1) : \max_{\mathbf{W}_{\text{RF}}, \mathbf{\Delta}, \mathbf{W}_{\text{BB}}} \frac{R(\mathbf{W}_{\text{RF}}, \mathbf{\Delta}, \mathbf{W}_{\text{BB}})}{P(\mathbf{\Delta})}$$
$$\text{subject to } \mathbf{W}_{\text{RF}} \in \mathcal{W}^{N_{\text{R}} \times L_{\text{R}}}, \mathbf{\Delta} \in \mathcal{D}^{L_{\text{R}} \times L_{\text{R}}},$$

where the set $\mathcal{D}$ represents the finite states of the quantizer and is defined as,

$$\mathcal{D} = \left\{ \mathbf{\Delta} \in \mathbb{R}^{L_{\text{R}} \times L_{\text{R}}} \middle| m \le [\mathbf{\Delta}]_{ii} \le M \ \forall \ i = 1, ..., L_{\text{R}} \right\}.$$

The channel's singular value decomposition (SVD) is written as $\mathbf{H} = \mathbf{U}_{\text{H}} \mathbf{\Sigma}_{\text{H}} \mathbf{V}_{\text{H}}^H$, where $\mathbf{U}_{\text{H}} \in \mathbb{C}^{N_{\text{R}} \times N_{\text{R}}}$ and $\mathbf{V}_{\text{H}} \in \mathbb{C}^{N_{\text{T}} \times N_{\text{T}}}$ are unitary matrices, and $\mathbf{\Sigma}_{\text{H}} \in \mathbb{R}^{N_{\text{R}} \times N_{\text{T}}}$ 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. The optimal, fully digital combiner matrix $\mathbf{W}_{\text{opt}}$ consists of the N_{s} columns of the left singular matrix $\mathbf{U}_{\text{H}}$. Our goal, by solving $(\mathcal{P}_1)$, is to obtain the combiner matrices and the bit resolution matrix in an optimal manner. We introduce the novel decomposition of the A/D hybrid combiner to three parts representing the analog combiner matrix, the bit resolution matrix and digital combiner matrix, i.e., $\mathbf{W}_{\text{RF}} \mathbf{\Delta} \mathbf{W}_{\text{BB}}$. So the Euclidean distance $\|\mathbf{W}_{\text{opt}} - \mathbf{W}_{\text{RF}} \mathbf{\Delta} \mathbf{W}_{\text{BB}}\|_F^2$ should be as small as possible for a maximum throughput combiner design. Note that we optimize over the bit resolution matrix with varying resolutions and the choice of combiner matrices at the RX.

Proposition 1. *The maximization of the fractional problem* $(\mathcal{P}_1)$ *is equivalent with the solution of the following problem:*

$$(\mathcal{P}_2) : \min_{\mathbf{W}_{\text{RF}}, \mathbf{\Delta}, \mathbf{W}_{\text{BB}}} \frac{1}{2} \|\mathbf{W}_{\text{opt}} - \mathbf{W}_{\text{RF}} \mathbf{\Delta} \mathbf{W}_{\text{BB}}\|_F^2 + \gamma P(\mathbf{\Delta}),$$
$$\text{subject to } \mathbf{W}_{\text{RF}} \in \mathcal{W}^{N_{\text{R}} \times L_{\text{R}}}, \mathbf{\Delta} \in \mathcal{D}^{L_{\text{R}} \times L_{\text{R}}},$$

where the parameter $\gamma \in \mathbb{R}^+$ *denotes the trade-off between the rate and the power consumption.*

Proof. The main idea to prove the equivalence is first to apply the Dinkelbach approach to transform the fractional problem into an affine one [16]. Afterwards, based on [1], [2], the maximization of the rate R can be expressed as minimization of the Euclidean distance between the computed A/D hybrid combiner and the optimal, fully digital combiner $\mathbf{W}_{\text{opt}}$. The details of this proof are omitted due to space limitations. □

Parameter γ also determines how close is the solution of $(\mathcal{P}_2)$ to $(\mathcal{P}_1)$. In this work, γ is selected after an exhaustive search over all the possible values in the range of [0.001, 0.1] and the value which gives the best result for $(\mathcal{P}_2)$ is selected. Problem $(\mathcal{P}_2)$ is non-convex due to the constraints on the structure of matrix $\mathbf{W}_{\text{RF}}$. Similar non-convex problems have been recently addressed in the literature via alternating direction method of multipliers (ADMM) based solutions [17]–[19].

B. Proposed ADMM Solution

In the following we develop an iterative procedure for solving $(\mathcal{P}_2)$ based on the ADMM approach [17]. This method, is a variant of the standard augmented Lagrangian method that uses partial updates (similar to the Gauss-Seidel method for the solution of linear equations) to solve constrained optimization problems. This method replaces a constrained minimization problem by a series of unconstrained problems and add a penalty term to the objective function. This penalty improves robustness compared to other optimization methods for constrained problems (for example, the dual ascent method) and in particular achieves convergence without the need of specific assumptions for the objective function, i.e., strict convexity and finiteness. The interested reader may refer to [17] for further information.

We first transform $(\mathcal{P}_2)$ into a form that can be addressed via ADMM. By using the auxiliary variable $\mathbf{Z}$, $(\mathcal{P}_2)$ can be written in the following form:

$$(\mathcal{P}_3) : \min_{\mathbf{Z}, \mathbf{W}_{\text{RF}}, \mathbf{\Delta}, \mathbf{W}_{\text{BB}}} \frac{1}{2} \|\mathbf{W}_{\text{opt}} - \mathbf{Z}\|_F^2 + \mathbb{1}_{\mathcal{W}^{N_{\text{R}} \times L_{\text{R}}}} \{\mathbf{W}_{\text{RF}}\}$$
$$+ \mathbb{1}_{\mathcal{D}^{L_{\text{R}} \times L_{\text{R}}}} \{\mathbf{\Delta}\} + \gamma P(\mathbf{\Delta}),$$
$$\text{subject to } \mathbf{Z} = \mathbf{W}_{\text{RF}} \mathbf{\Delta} \mathbf{W}_{\text{BB}}.$$

Problem $(\mathcal{P}_3)$ formulates the A/D hybrid combiner matrix design as a matrix factorization problem. That is, the overall combiner $\mathbf{Z}$ is sought so that it minimizes the Euclidean distance to the optimal, fully digital combiner $\mathbf{W}_{\text{opt}}$ while supporting decomposition into three factors: the analog combiner matrix $\mathbf{W}_{\text{RF}}$, the matrix $\mathbf{\Delta}$ which is related to the resolution of each ADC and the digital combiner matrix $\mathbf{W}_{\text{BB}}$.

The augmented Lagrangian function of $(\mathcal{P}_3)$ is given by,

$$\mathcal{L}(\mathbf{Z},\mathbf{W}_{\text{RF}},\boldsymbol{\Delta},\mathbf{W}_{\text{BB}},\boldsymbol{\Lambda}) = \frac{1}{2}\|\mathbf{W}_{\text{opt}}-\mathbf{Z}\|_F^2 + \mathbb{1}_{\mathcal{W}^{N_{\text{R}}\times L_{\text{R}}}}\{\mathbf{W}_{\text{RF}}\}$$
$$+ \mathbb{1}_{\mathcal{D}^{L_{\text{R}}\times L_{\text{R}}}}\{\boldsymbol{\Delta}\} + \frac{\alpha}{2}\|\mathbf{Z}+\boldsymbol{\Lambda}/\alpha-\mathbf{W}_{\text{RF}}\boldsymbol{\Delta}\mathbf{W}_{\text{BB}}\|_F^2 + \gamma P(\boldsymbol{\Delta}), \tag{14}$$

where α is a scalar penalty parameter and $\boldsymbol{\Lambda} \in \mathbb{C}^{N_{\text{R}}\times L_{\text{R}}}$ is the Lagrange Multiplier matrix. According to ADMM [17], the solution to $(\mathcal{P}_3)$ is derived by the following iterative steps:

$$(\mathcal{P}_{3\text{A}}): \mathbf{Z}_{(n)} = \arg\min_{\mathbf{Z}} \frac{1}{2}\|(1+\alpha)\mathbf{Z}-\mathbf{W}_{\text{opt}}+\boldsymbol{\Lambda}_{(n-1)}$$
$$- \alpha\mathbf{W}_{\text{RF}(n-1)}\boldsymbol{\Delta}_{(n-1)}\mathbf{W}_{\text{BB}(n-1)}\|_F^2,$$

$$(\mathcal{P}_{3\text{B}}): \mathbf{W}_{\text{RF}(n)} = \arg\min_{\mathbf{W}_{\text{RF}}} \mathbb{1}_{\mathcal{W}^{N_{\text{R}}\times L_{\text{R}}}}\{\mathbf{W}_{\text{RF}}\} + \frac{\alpha}{2}\times$$
$$\left\|\mathbf{Z}_{(n)}+\boldsymbol{\Lambda}_{(n-1)}/\alpha-\mathbf{W}_{\text{RF}}\boldsymbol{\Delta}_{(n-1)}\mathbf{W}_{\text{BB}(n-1)}\right\|_F^2,$$

$$(\mathcal{P}_{3\text{C}}): \boldsymbol{\Delta}_{(n)} = \arg\min_{\boldsymbol{\Delta}}\|\mathbf{y}_{\text{c}}-\boldsymbol{\Psi}\text{vec}(\boldsymbol{\Delta})\|_2^2 + \gamma P(\boldsymbol{\Delta}),$$
$$\text{subject to } \boldsymbol{\Delta} \in \mathcal{D},$$

$$(\mathcal{P}_{3\text{D}}): \mathbf{W}_{\text{BB}(n)} = \arg\min_{\mathbf{W}_{\text{BB}}} \frac{\alpha}{2}\|\mathbf{Z}_{(n)}+\boldsymbol{\Lambda}_{(n-1)}/\alpha$$
$$- \mathbf{W}_{\text{RF}(n)}\boldsymbol{\Delta}_{(n)}\mathbf{W}_{\text{BB}}\|_F^2,$$

$$\boldsymbol{\Lambda}_{(n)} = \boldsymbol{\Lambda}_{(n-1)} + \alpha\left(\mathbf{Z}_{(n)}-\mathbf{W}_{\text{RF}(n)}\boldsymbol{\Delta}_{(n)}\mathbf{W}_{\text{BB}(n)}\right), \tag{15}$$

where n denotes the iteration index, $\mathbf{y}_{\text{c}} = \text{vec}(\mathbf{Z}_{(n)}+\boldsymbol{\Lambda}_{(n-1)}/\alpha)$ and $\boldsymbol{\Psi} = \mathbf{W}_{\text{BB}(n-1)} \otimes \mathbf{W}_{\text{RF}(n)}$ ($\otimes$ is the Khatri-Rao product).

We solve the optimization problems $(\mathcal{P}_{3\text{A}})$-$(\mathcal{P}_{3\text{D}})$ and the solutions are provided in Algorithm 1. The algorithm provides the complete procedure to obtain the optimal analog combiner matrix $\mathbf{W}_{\text{RF}}$, the optimal bit resolution matrix $\boldsymbol{\Delta}$ and the optimal baseband (or digital) combiner matrix $\mathbf{W}_{\text{BB}}$. It starts by initializing the entries of the matrices $\mathbf{Z}$, $\mathbf{W}_{\text{RF}}$, $\boldsymbol{\Delta}$, $\mathbf{W}_{\text{BB}}$ with random values and the entries of the Lagrange multiplier matrix $\boldsymbol{\Lambda}$ with zeros. For iteration index n, $\mathbf{Z}_{(n)}$, $\mathbf{W}_{\text{RF}(n)}$, $\boldsymbol{\Delta}_{(n)}$ and $\mathbf{W}_{\text{BB}(n)}$ are updated at each iteration step using the solutions provided in Steps 4, 7, 8, 10 and 11 of Algorithm 1. In Step 7, $\Pi_{\mathcal{W}}$ is the operator that projects the solution onto the set $\mathcal{W}$. This is computed by solving the following optimization problem [20]:

$$(\mathcal{P}_4): \quad \min_{\mathbf{A}_{\mathcal{W}}}\|\mathbf{A}_{\mathcal{W}}-\mathbf{A}\|_F^2, \text{subject to } \mathbf{A}_{\mathcal{W}} \in \mathcal{W},$$

where $\mathbf{A}$ is an arbitrary matrix and $\mathbf{A}_{\mathcal{W}}$ is its projection onto the set $\mathcal{W}$. The solution to $(\mathcal{P}_4)$ is given by the phase of the complex elements of $\mathbf{A}$. Thus, for $\mathbf{A}_{\mathcal{W}} = \Pi_{\mathcal{W}}\{\mathbf{A}\}$ we have

$$\mathbf{A}_{\mathcal{W}}(x,y) = \begin{cases} 0, & \mathbf{A}(x,y)=0 \\ \frac{\mathbf{A}(x,y)}{|\mathbf{A}(x,y)|}, & \mathbf{A}(x,y)\neq 0 \end{cases}, \tag{16}$$

where $\mathbf{A}_{\mathcal{W}}(x,y)$ and $\mathbf{A}(x,y)$ are the elements at the xth row-yth column of matrices $\mathbf{A}_{\mathcal{W}}$ and $\mathbf{A}$, respectively. Furthermore, as shown in Step 8, the minimization problem in $(\mathcal{P}_{3\text{C}})$ is solved by implementing CVX [21]. A termination criterion related to the maximum permitted number of iterations of the ADMM sequence (N_{max}) is considered. Upon convergence, the number of bits for each ADC is obtained by using (5) and quantized to the nearest integer value.

Algorithm 1 Proposed ADMM Solution for the A/D Hybrid Combiner Design

---

1: **Initialize: $\mathbf{Z}$, $\mathbf{W}_{\text{RF}}$, $\boldsymbol{\Delta}$, $\mathbf{W}_{\text{BB}}$** with random values, $\boldsymbol{\Lambda}$ with zeros, $\alpha=1$ and $n=1$
2: **while** $n \leq N_{\text{max}}$ **do**
3:　　$\mathbf{A} = \alpha\mathbf{W}_{\text{RF}(n-1)}\boldsymbol{\Delta}_{(n-1)}\mathbf{W}_{\text{BB}(n-1)}.$
4:　　$\mathbf{Z}_{(n)} = \frac{1}{\alpha+1}\left(\mathbf{W}_{\text{opt}}-\boldsymbol{\Lambda}_{(n-1)}+\mathbf{A}\right).$
5:　　$\mathbf{B} = \boldsymbol{\Lambda}_{(n-1)}+\alpha\mathbf{Z}_{(n)}.$
6:　　$\mathbf{C} = \alpha\boldsymbol{\Delta}_{(n-1)}\mathbf{W}_{\text{BB}(n-1)}\mathbf{W}_{\text{BB}(n-1)}^H\boldsymbol{\Delta}_{(n-1)}^H.$
7:　　$\mathbf{W}_{\text{RF}(n)} = \Pi_{\mathcal{W}}\{\mathbf{B}\mathbf{W}_{\text{BB}(n-1)}^H\boldsymbol{\Delta}_{(n-1)}^H\mathbf{C}^{-1}\}.$
8:　　Update $\boldsymbol{\Delta}_{(n)}$ by solving $(\mathcal{P}_{3\text{C}})$ using CVX [21].
9:　　$\mathbf{D} = \alpha\boldsymbol{\Delta}_{(n)}^H\mathbf{W}_{\text{RF}(n)}^H\mathbf{W}_{\text{RF}(n)}\boldsymbol{\Delta}_{(n)}.$
10:　　$\mathbf{W}_{\text{BB}(n)} = \mathbf{D}^{-1}\boldsymbol{\Delta}_{(n)}^H\mathbf{W}_{\text{RF}(n)}^H\mathbf{B}.$
11:　　$\boldsymbol{\Lambda}_{(n)} = \boldsymbol{\Lambda}_{(n-1)} + \alpha\left(\mathbf{Z}_{(n)}-\mathbf{W}_{\text{RF}(n)}\boldsymbol{\Delta}_{(n)}\mathbf{W}_{\text{BB}(n)}\right).$
12:　　$n \leftarrow n+1$
13: **end while**
14: **return** $\mathbf{W}_{\text{RF}(N_{\text{max}})}$, $\boldsymbol{\Delta}_{(N_{\text{max}})}$, $\mathbf{W}_{\text{BB}(N_{\text{max}})}$

Computational complexity analysis of Algorithm 1: In Algorithm 1, mainly Step 8 involves multiplication by $\boldsymbol{\Psi}$ whose dimensions are $L_{\text{R}}N_{\text{R}} \times N_{\text{s}}L_{\text{R}}$. In general, the solution of $(\mathcal{P}_{3\text{C}})$ can be upper-bounded by $\mathcal{O}((L_{\text{R}}^2 N_{\text{R}}N_{\text{s}})^3)$ which can be improved significantly by exploiting the structure of $\boldsymbol{\Psi}$.

IV. SIMULATION RESULTS

In this section, we evaluate the performance of the proposed ADMM technique using computer simulation results. The results have been averaged over 1,000 Monte-Carlo realizations.

System setup: We set the following parameters, unless specified otherwise, to obtain the desired results: $N_{\text{T}}=32$, $N_{\text{R}}=16$, $L_{\text{R}}=4$, $N_{\text{s}}=4$, $N_{\text{cl}}=2$, $N_{\text{ray}}=4$, $N_{\text{max}}=40$, $m=1$, $M=8$, $\alpha=1$ and $\sigma_{\alpha,i}^2=1$. The azimuth angles of departure and arrival are computed with uniformly distributed mean angles; each cluster follows a Laplacian distribution about the mean angle. The antenna elements in the ULA are spaced by distance $d=\lambda/2$. The signal-to-noise ratio (SNR) is given by the inverse of the noise variance, i.e., $1/\sigma_{\text{n}}^2$. The transmit vector $\mathbf{x}$ is composed of the normalized i.i.d. Gaussian symbols. The values used for the terms in the power model in (12) of Section III are $P_{\text{ADC}}=100$ mW, $P_{\text{CP}}=10$ W, $P_{\text{R}}=100$ mW and $P_{\text{PS}}=10$ mW. Note that to measure the spectral efficiency (SE) performance, we compute the ratio R/B bits/s/Hz where B represents the bandwidth, and for the simulations we set $B=1$ Hz. For simulations, the precoder matrix $\mathbf{F}$ is considered equal to the optimal fully digital precoder matrix [1], [2], i.e., the product of $1/\sqrt{N_{\text{s}}}$ and first N_{s} columns of the right singular matrix $\mathbf{V}_{\text{H}}$.

Convergence of the proposed ADMM solution: Fig. 2 shows the convergence of the ADMM solution as proposed in Algorithm 1 to obtain the optimal bit resolution at each ADC and corresponding optimal combiner matrices. The proposed solution converges rapidly at around 20 iterations and mean square error (MSE), $\left\|\mathbf{W}_{\text{opt}}-\mathbf{W}_{\text{RF}(N_{\text{max}})}\boldsymbol{\Delta}_{(N_{\text{max}})}\mathbf{W}_{\text{BB}(N_{\text{max}})}\right\|_F^2$, goes as low as -20 dB. A lower number of RX antennas shows
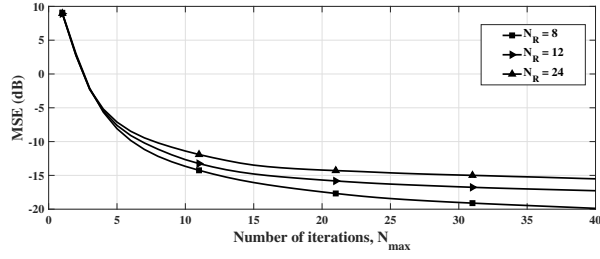
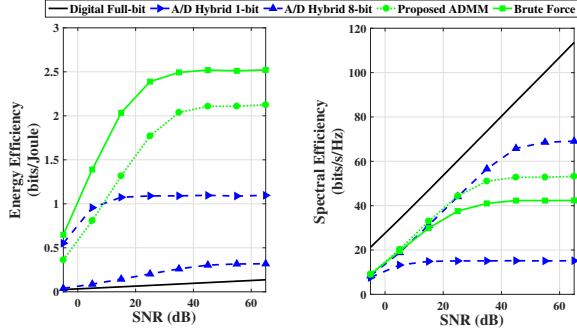Fig. 2. Convergence of the ADMM solution for different N_R at $\gamma = 0.01$.

Fig. 4. EE and SE performance w.r.t. N_R at SNR = 30 dB and $\gamma = 0.01$.

Fig. 3. EE and SE performance w.r.t. SNR at $N_R = 16$ and $\gamma = 0.01$.

Fig. 5. EE and SE performance w.r.t. N_T at SNR = 30 dB and $\gamma = 0.01$.

lower MSE as expected, since fewer parameters are required to be estimated.

Benchmark techniques:

1) Digital combining with full-bit resolution: We consider the conventional fully digital beamforming architecture, where the number of RF chains at the RX is equal to the number of RX antennas, i.e., $L_R = N_R$. The fully digital combining solution may be provided by SVD and waterfilling [22]. In terms of the resolution sampling, we consider full-bit resolution, i.e., $M = 8$-bit, which represents the optimum from the achievable SE perspective.

2) A/D Hybrid combining with 1-bit and 8-bit resolutions: We also consider a A/D hybrid combining architecture with $L_R < N_R$, for two cases of bit resolution: a) 1-bit resolution which usually shows reasonable EE performance, and b) 8-bit resolution which usually shows high SE results.

3) Brute force with A/D hybrid combining: We also implement an exhaustive search approach as an upper bound for EE maximization called brute force (BF), based on [4], which clearly shows the energy-rate performance trade-offs in the simulations. It makes a search over the number of RF chains L_R and all the available bit resolutions, i.e., $b = 1, ..., M$. It then finds the best EE out of all the possible cases and chooses the corresponding optimal resolution for each ADC. This method provides the best possible EE performance, but it is computationally intractable for $L_R > 4$.

Fig. 3 shows the performance of the proposed ADMM solution compared with existing benchmark techniques with respect to (w.r.t.) SNR at $N_R = 16$. The proposed ADMM solution achieves high EE which has performance close to the BF approach and better than the 8-bit hybrid, 1-bit hybrid and full-bit digital baselines. For example, at SNR
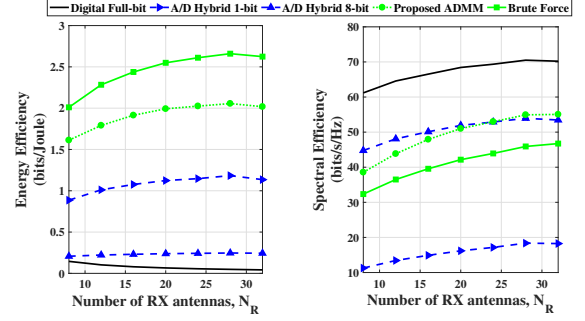
= 20 dB, the proposed ADMM solution outperforms 1-bit hybrid, 8-bit hybrid and full-bit digital baselines by about 0.45 bits/Joule, 1.375 bits/Joule and 1.44 bits/Joule, respectively. It also exhibits better SE than 1-bit hybrid and has similar performance to the 8-bit hybrid baseline.

There is an energy-rate trade-off between the proposed solution and the BF approach as we can achieve better rate with lower EE and vice-versa. Moreover, the proposed solution has lower complexity than the BF approach because the BF involves a search over all the possible bit resolutions while the proposed solution directly optimizes the number of bits to obtain an optimal number of bits at each ADC. We constrain the number of RF chains $L_R = 4$ for the BF approach due to the high complexity order which is $\mathcal{O}(M^{L_R})$. Also note that the proposed approach enables the selection of different resolutions for different ADCs and thus, it offers a better trade-off for EE versus SE than existing approaches which are based on a fixed ADC resolution.

Figs. 4 and 5 show the performance results w.r.t. the number of RX and TX antennas at 30 dB SNR. The proposed ADMM solution again achieves high EE and performs close to the BF approach and better than the 8-bit hybrid, 1-bit hybrid and full-bit digital baselines. For example, at $N_R = 20$, the proposed ADMM solution outperforms 1-bit hybrid, 8-bit hybrid and full-bit digital baselines by about 0.85 bits/Joule, 1.75 bits/Joule and 1.875 bits/Joule, respectively. Also, for $N_T = 20$, the proposed solution outperforms 1-bit hybrid, 8-bit hybrid and full-bit digital baselines by about 1.0 bits/Joule, 1.5 bits/Joule and 1.625 bits/Joule, respectively. The proposed solution also exhibits better SE than 1-bit hybrid and has similar performance to the 8-bit hybrid baseline. Both the
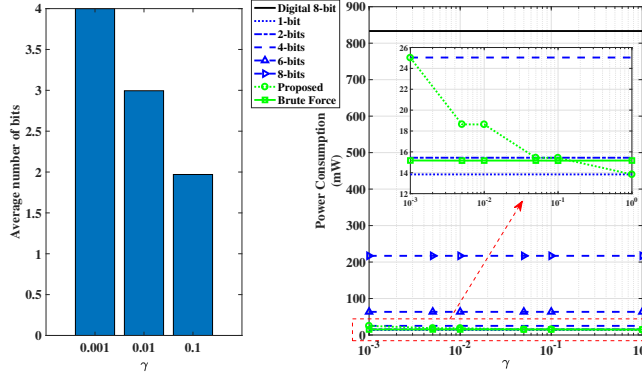
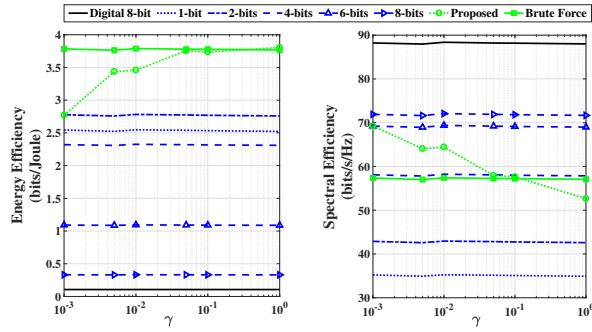Fig. 6. Average number of bits for proposed ADMM and power consumption w.r.t. γ at SNR = 30 dB.

Fig. 7. EE and SE performance w.r.t. γ at SNR = 30 dB.

figures follow the energy-rate trade-off with the BF approach.

Furthermore, we investigate the performance over the trade-off parameter γ introduced in ($\mathcal{P}_2$). Fig. 6 shows the bar plot of average of the optimal number of bits selected by the proposed solution for each ADC versus γ. The average optimal number decreases with the increase in γ, for example, it is 4 for $\gamma = 0.001$, 3 for $\gamma = 0.01$ and 2 for $\gamma = 0.1$. Fig. 6 also shows that the power consumption in the proposed case is considerably low and decreases with the increase in the trade-off parameter γ unlike digital 8-bit, several fixed bit hybrid baselines and the BF approach. Fig. 7 shows the EE and SE plots for several solutions w.r.t. γ. It can be observed that the proposed solution achieves higher EE than the fixed bit allocation solutions and achieves comparable EE and SE results to the BF approach. These curves also show that adjusting γ allows the system to vary the energy-rate trade-off.

V. CONCLUSION

This paper proposes an energy efficient mmWave A/D hybrid MIMO system which can vary the ADC bit resolution at the RX. This method uses the decomposition of the A/D hybrid combiner matrix into three parts representing the analog combiner matrix, the bit resolution matrix and the digital combiner matrix. These three matrices are optimized by the novel ADMM solution which outperforms the EE of the full-bit digital, 1-bit hybrid combining and 8-bit hybrid combining baselines. There is an energy-rate trade-off with the BF approach which yields the upper bound for EE maximization.

The proposed approach enables the selection of the optimal resolution for each ADC and thus, it offers better trade-off for data rate versus EE than existing approaches based on fixed ADC resolution. In future work, we will jointly optimize the DAC and ADC bit resolution and hybrid precoder and combiner matrices at the TX and the RX.

REFERENCES

[1] 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.

[2] A. Kaushik et al.,"Sparse hybrid precoding and combining in millimeter wave MIMO systems," *IET Radio Prop. Tech. 5G*, Durham, UK, pp. 1-7, Oct. 2016.

[3] 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.

[4] R. Zi et al., "Energy efficiency optimization of 5G radio frequency chain systems", *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 758-771, Apr. 2016.

[5] 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.

[6] 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.

[7] A. Kaushik et al., "Dynamic RF Chain Selection for Energy Efficient and Low Complexity Hybrid Beamforming in Millimeter Wave MIMO Systems," *IEEE Trans. Green Commun. Netw.*, accepted, July 2019.

[8] A. Kaushik et al., "Energy Efficiency maximization of millimeter wave hybrid MIMO systems with low resolution DACs," *IEEE Int. Conf. Commun. (ICC)*, Shanghai, China, pp. 1-6, May 2019.

[9] J. Mo et al., "Achievable rates of hybrid architectures with few-bit ADC receivers," *VDE Int. ITG Workshop Smart Antennas*, pp. 1-8, 2016.

[10] 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.

[11] A. Kaushik et al.,"Efficient channel estimation in millimeter wave hybrid MIMO systems with low resolution ADCs", *IEEE Europ. Sig. Process.*, Rome, Italy, pp. 1839-1843, Sept. 2018.

[12] T.-C. Zhang et al., "Mixed-ADC massive MIMO detectors: Performance analysis and design optimization," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7738-7752, Nov. 2016.

[13] 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.

[14] O. Orhan et al., "Low power analog-to-digital conversion in millimeter wave systems: Impact of resolution and bandwidth on performance," *Info. Theory Appl. Workshop (ITA)*, San Diego, USA, pp. 191-198, Feb. 2015.

[15] A. Mezghani and J. A. Nossek, "Capacity lower bound of MIMO channels with output quantization and correlated noise," *IEEE Int. Symp. Info. Theory (ISIT)*, Cambridge, USA, Jul. 2012.

[16] W. Dinkelbach, "On nonlinear fractional programming", *Management Science*, vol. 13, no. 7, pp. 492-498, Mar. 1967.

[17] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1-122, 2011.

[18] C. G. Tsinos et al., "Distributed blind hyperspectral unmixing via joint sparsity and low-rank constrained non-negative matrix factorization," *IEEE Trans. Comput. Imag.*, vol. 3, no. 2, pp. 160174, June 2017.

[19] C. G. Tsinos and B. Ottersten, "An efficient algorithm for unit-modulus quadratic programs with application in beamforming for wireless sensor networks," *IEEE Signal Process. Letters*, vol. 25, no. 2, pp. 169-173, Feb. 2018.

[20] D. P. Bertsekas, "Nonlinear programming," 1999.

[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] T. S. Rappaport et al., "Millimeter wave wireless communications," *Prentice-Hall*, Sept. 2014.