

Energy-Efficient Throughput Maximization in mmWave MU-Massive-MIMO-OFDM: Genetic Algorithm based Resource Allocation

Asil Koc, Farhan Bishe, Tho Le-Ngoc

Department of Electrical and Computer Engineering, McGill University, Montreal, QC, Canada

Email: asil.koc@mail.mcgill.ca, farhan.bishe@mail.mcgill.ca, tho.le-ngooc@mcgill.ca

Abstract—This paper develops a new genetic algorithm based resource allocation (GA-RA) technique for energy-efficient throughput maximization in multi-user massive multiple-input multiple-output (MU-mMIMO) systems using orthogonal frequency division multiplexing (OFDM) based transmission. We employ a hybrid precoding (HP) architecture with three stages: (i) radio frequency (RF) beamformer, (ii) baseband (BB) precoder, (iii) resource allocation (RA) block. First, a single RF beamformer block is built for all subcarriers via the slow time-varying angle-of-departure (AoD) information. For enhancing the energy efficiency, the RF beamformer aims to reduce the hardware cost/complexity and total power consumption via a low number of RF chains. Afterwards, the reduced-size effective channel state information (CSI) is utilized in the design of a distinct BB precoder and RA block for each subcarrier. The BB precoder is developed via regularized zero-forcing technique. Finally, the RA block is built via the proposed GA-RA technique for throughput maximization by allocating the power and subcarrier resources. The illustrative results show that the throughput performance in the MU-mMIMO-OFDM systems is greatly enhanced via the proposed GA-RA technique compared to both equal RA (EQ-RA) and particle swarm optimization based RA (PSO-RA). Moreover, the performance gain ratio increases with the increasing number of subcarriers, particularly for low transmission powers.

Index Terms—Massive MIMO, hybrid precoding, OFDM, genetic algorithm, resource allocation, multi-user, beamforming.

I. INTRODUCTION

MILLIMETER wave (mmWave) communication has attracted considerable attention for the fifth-generation (5G) and beyond wireless communication systems by addressing bandwidth limitation issue [1]. Bandwidth demands have arisen in accordance with the intense growth in user demands in wireless networks and emerging advanced technologies requiring high-throughput data communications. Hence, the mmWave communications is an enabling technology by providing a considerable large bandwidth (i.e., 30 GHz - 300 GHz). Moreover, transmitting information on such high frequencies favors implementing large-dimensional antenna arrays in relatively smaller dimensions. Nevertheless, making a shift to mmWave communications leads to some challenging issues, notably higher atmospheric absorption causing severe path loss due to its short wavelength [1]. To alleviate the path loss effect in mmWave communication, it is required to form high-gain directional beams, which can be done by

This work was supported in part by Huawei Technologies Canada and in part by the Natural Sciences and Engineering Research Council of Canada.

implementing large antenna arrays, introducing the massive multiple-input multiple-output (mMIMO) technology [2].

Hybrid precoding (HP) is a promising technique for the multi-user mMIMO (MU-mMIMO) systems [3]. The two-stage HP architecture interconnects the radio frequency (RF) and baseband (BB) stages via a limited number of power-hungry RF chains [4]–[9]. Thus, HP remarkably enhances the energy-efficiency compared to the single-stage fully digital precoding (FDP). On the other hand, the wideband signal transmission is considered to achieve extremely high data rates, especially in the mmWave communications. To address the frequency selectivity issue, the orthogonal frequency division multiplexing (OFDM) technique is widely considered for the wideband signal transmission [10]. In [10]–[12], OFDM-based HP scheme is investigated for mMIMO systems, where one RF beamformer is designed for all subcarriers, while the number of BB precoders should be the same as the number of subcarriers.

The recent advances in artificial intelligence (AI) enable the development of advanced communication systems [13]. In the field of AI, the genetic algorithm (GA) is a evolutionary population-based optimization technique to solve the complex real-world problems under given constraints [14], [15].

In this paper, we propose a novel genetic algorithm based resource allocation (GA-RA) technique for energy-efficient throughput maximization in the MU-mMIMO-OFDM system. Also, the angular-based hybrid precoding (AB-HP) technique proposed in [5] is adopted for the OFDM-based transmission. Thus, the system model is divided into three stages: (i) RF beamformer, (ii) BB precoder, (iii) resource allocation (RA). First, the RF beamformer is designed via the slow time-varying angle-of-departure (AoD) information. Next, the BB precoder is built for each subcarrier using the reduced-size effective channel state information (CSI) for the corresponding subcarrier. Finally, the proposed GA-RA technique is applied for the throughput maximization by efficiently allocating the power and subcarrier resources among the users. The numerical results present that the proposed GA-RA considerably improves the throughput (i.e., sum-rate)¹ performance compared to the equal RA (EQ-RA) scheme. Also, when the number of subcarriers increases, the performance gain improves, especially for the low transmit powers.

¹We use the terms of throughput and sum-rate interchangeably.

The rest of this paper is organized as follows. The system model is introduced in Section II. The AB-HP scheme is expressed in Section III. Then, Section IV provides the RA problem formulation for the throughput maximization. The proposed GA-RA technique is introduced in Section V. After presenting the illustrative results in Section VI, the paper is concluded in Section VII.

II. SYSTEM MODEL

We consider an OFDM-based downlink transmission via C subcarriers for the MU-mMIMO systems as shown in Fig. 1. Here, K single-antenna users located in G groups are served by a BS having a uniform rectangular array (URA) with $M = M_x \times M_y$ antennas². The BS utilizes N_{RF} RF chains for interconnecting the BB-stage and RF-stage. According to the multi-carrier transmission, the data signal $\mathbf{d}[i]$ for each subcarrier $i = 1, \dots, C$ is respectively passed through a power allocation block $\mathbf{P}[i] = \text{diag}(\sqrt{p_{1,i}}, \dots, \sqrt{p_{K,i}}) \in \mathbb{R}^{K \times K}$ and a digital BB precoder $\mathbf{B}[i] = [\mathbf{b}_{1,i}, \dots, \mathbf{b}_{K,i}] \in \mathbb{C}^{N_{RF} \times K}$, where $p_{k,i}$ is the allocated power for the k^{th} user at subcarrier i . After applying the inverse fast Fourier transforms (IFFT) and adding the cyclic prefixes (CP), the analog RF beamformer $\mathbf{F} \in \mathbb{C}^{M \times N_{RF}}$ developed via low-cost phase-shifters is employed, which is identical for all subcarriers³ [10]–[12]. Then, the transmitted data signal at subcarrier i is written as:

$$\mathbf{s}[i] = \mathbf{F}\mathbf{B}[i]\mathbf{P}[i]\mathbf{d}[i] \in \mathbb{C}^M. \quad (1)$$

Considering the URA structure [5] and the 3D geometry-based mmWave channel model [17], the channel vector for the k^{th} user at subcarrier $i = 1, \dots, C$ is defined as:

$$\mathbf{h}_k^T[i] = \sum_{l=1}^L \tau_{k_l,i}^{-\eta} z_{k_l,i} \phi^T(\gamma_{x,k_l,i}, \gamma_{y,k_l,i}) e^{-j\frac{2\pi l i}{C}} = \mathbf{z}_k^T[i] \boldsymbol{\Phi}_k[i], \quad (2)$$

where L is the number of paths, $\tau_{k_l,i}$ and $z_{k_l,i} \sim \mathcal{CN}(0, \frac{1}{L})$ are respectively the distance and complex path gain of l^{th} path at subcarrier i , η is the path loss exponent, $\phi(\cdot, \cdot) \in \mathbb{C}^M$ is the phase response vector, $\gamma_{x,k_l,i} = \sin(\theta_{k_l,i}) \cos(\psi_{k_l,i})$ and $\gamma_{y,k_l,i} = \sin(\theta_{k_l,i}) \sin(\psi_{k_l,i})$ are the coefficients reflecting the elevation AoD (EAoD) and azimuth AoD (AAoD) for the corresponding path. Here, $\theta_{k_l,i} \in [\theta_k - \delta_k^\theta, \theta_k + \delta_k^\theta]$ is the EAoD with mean θ_k and spread δ_k^θ , $\psi_{k_l,i} \in [\psi_k - \delta_k^\psi, \psi_k + \delta_k^\psi]$ is the AAoD with mean ψ_k and spread δ_k^ψ . Then, the phase response vector is modeled as [5]:

$$\begin{aligned} \phi(\gamma_x, \gamma_y) &= [1, e^{-j2\pi d\gamma_x}, \dots, e^{-j2\pi d(M_x-1)\gamma_x}]^T \\ &\otimes [1, e^{-j2\pi d\gamma_y}, \dots, e^{-j2\pi d(M_y-1)\gamma_y}]^T \in \mathbb{C}^M, \end{aligned} \quad (3)$$

²Based on the URA structure, M_x and M_y denote the number of antennas along x -axis and y -axis, respectively. There are two main reasons for utilizing the URA structure: (i) it packs a larger number of antennas in a two-dimensional (2D) grid under the physical-limited space requirements in practical applications, (ii) it enables three-dimensional (3D) beamforming by employing both azimuth and elevation domains [2], [5]–[7], [12], [16].

³Since the analog RF beamformer is designed via the AoD information, it is reasonable to assume that each subcarrier experiences a similar AoD support (i.e., mean azimuth/elevation AoD and their spread).

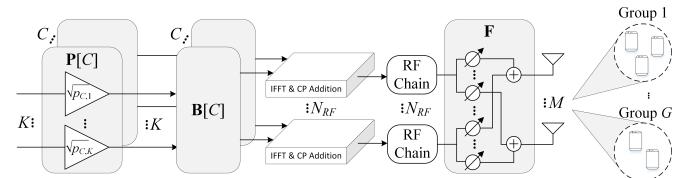


Fig. 1: Hybrid precoding and genetic algorithm based resource allocation (GA-RA) in MU-mMIMO-OFDM systems.

where d is the antenna element spacing normalized by the wavelength. As indicated in (2), the instantaneous channel vector is represented via the fast time-varying path gain vector $\mathbf{z}_k = [\tau_{k_1}^{-\eta} z_{k_1}, \dots, \tau_{k_L}^{-\eta} z_{k_L}]^T \in \mathbb{C}^L$ and slow time-varying phase response matrix $\boldsymbol{\Phi}_k \in \mathbb{C}^{L \times M}$ as a function of AoD information.

By using (1) and (2), the received signal at the k^{th} user at subcarrier i is given by:

$$\begin{aligned} r_{k,i} &= \mathbf{h}_k^T[i] \mathbf{s}[i] + w_{k,i} = \mathbf{h}_k^T[i] \mathbf{F} \mathbf{B}[i] \mathbf{P}[i] \mathbf{d}[i] + w_{k,i} \\ &= \underbrace{\sqrt{p_{k,i}} \mathbf{h}_k^T[i] \mathbf{F} \mathbf{b}_{k,i} d_{k,i}}_{\text{Desired Signal}} + \underbrace{\sum_{t \neq k}^K \sqrt{p_{t,i}} \mathbf{h}_k^T[i] \mathbf{F} \mathbf{b}_{t,i} d_{t,i}}_{\text{Inter-User Interference}} + w_{k,i}, \end{aligned} \quad (4)$$

where $w_{k,i} \sim \mathcal{CN}(0, \sigma^2)$ is the complex circularly symmetric Gaussian noise. Then, we derive the instantaneous sum-rate (i.e., throughput) expression at subcarrier i as follows:

$$R_{\text{sum},i}(\mathbf{F}, \mathbf{B}[i], \mathbf{P}[i]) = \sum_{k=1}^K \log_2 \left(1 + \frac{p_{k,i} |\mathbf{h}_k^T[i] \mathbf{F} \mathbf{b}_{k,i}|^2}{\sum_{t \neq k}^K p_{t,i} |\mathbf{h}_k^T[i] \mathbf{F} \mathbf{b}_{t,i}|^2 + \sigma^2} \right). \quad (5)$$

The overall sum-rate across all C subcarriers is calculated as $R_{\text{sum}} = \sum_{i=1}^C \mathbb{E}\{R_{\text{sum},i}(\mathbf{F}, \mathbf{B}[i], \mathbf{P}[i])\}$ in the unit of [bps/Hz]. Similarly, the average sum-rate per subcarrier is obtained as $\frac{R_{\text{sum}}}{C}$ in the unit of [bps/Hz/subcarrier]. Hence, we formulate the throughput maximization problem as follows:

$$\begin{aligned} &\max_{\{\mathbf{F}, \mathbf{B}[i], \mathbf{P}[i]\}} R_{\text{sum},i}(\mathbf{F}, \mathbf{B}[i], \mathbf{P}[i]) \\ \text{s.t. } &\mathbb{E}\left\{ \|\mathbf{s}[i]\|_2^2 \right\} = \sum_{k=1}^K p_{k,i} \|\mathbf{F} \mathbf{b}_{k,i}\|_2^2 \leq \frac{P_T}{C}, \quad p_{k,i} \geq 0, \forall k, \\ &|\mathbf{F}|_{m,n} = \frac{1}{\sqrt{M}}, \forall m, n, \end{aligned} \quad (6)$$

where the constraints indicate the transmit power constraint per subcarrier [10] (i.e., $\frac{P_T}{C}$) and the constant modulus (CM) constraint due to the utilization of phase-shifters at the RF-stage. However, it is a non-convex optimization due to the allocated powers entangled with each other (please see (5)) and the CM constraint at the RF beamformer. Thus, we first design the RF beamformer and the BB precoder via AB-HP in Section III, then we propose the GA-RA technique in Section V for optimizing the allocated powers across all subcarriers.

III. HYBRID PRECODING

In this section, we develop the angular-based hybrid precoding (AB-HP) architecture for MU-mMIMO-OFDM systems to reduce the number of RF chains, suppress the inter-user interference and lower the CSI overhead size. After designing the RF beamformer and BB precoder in this section, the proposed GA-RA technique is expressed in Section V.

A. RF Beamformer

The slow-time varying AoD information⁴ is utilized at the RF beamformer for reducing the large CSI overhead size in MU-mMIMO-OFDM systems. Considering the K users located in G groups as shown in Fig. 1, the channel matrix for group g at subcarrier i is given by:

$$\mathbf{H}_g[i] = [\mathbf{h}_1[i], \dots, \mathbf{h}_{K_g}[i]]^T = \mathbf{Z}_g[i] \boldsymbol{\Phi}_g[i] \in \mathbb{C}^{K_g \times M}, \quad (7)$$

where $\mathbf{Z}_g[i] = [\mathbf{z}_1[i], \dots, \mathbf{z}_{K_g}[i]]^T \in \mathbb{C}^{K_g \times L}$, $\boldsymbol{\Phi}_g[i] \in \mathbb{C}^{L \times M}$, K_g is the number of users in group g with $K = \sum_{g=1}^G K_g$. Then, the full-size channel matrix at subcarrier i is defined as $\mathbf{H}[i] = [\mathbf{H}_1^T[i], \dots, \mathbf{H}_G^T[i]]^T \in \mathbb{C}^{K \times M}$. Here, we assume that the users in the same groups experience a similar AoDs as in [5]–[8]. Therefore, G blocks are designed for the RF beamformer as follows:

$$\mathbf{F} = [\mathbf{F}_1, \dots, \mathbf{F}_G] \in \mathbb{C}^{M \times N_{RF}}, \quad (8)$$

where $\mathbf{F}_g \in \mathbb{C}^{M \times N_{RF,g}}$ is the RF beamformer for group g with $N_{RF} = \sum_{g=1}^G N_{RF,g}$. By using (7) and (8), the effective channel matrix seen from the BB-stage is obtained as:

$$\mathbf{H}[i] = \mathbf{H}[i]\mathbf{F} = \begin{bmatrix} \mathbf{H}_1[i]\mathbf{F}_1 & \mathbf{H}_1[i]\mathbf{F}_2 & \cdots & \mathbf{H}_1[i]\mathbf{F}_G \\ \mathbf{H}_2[i]\mathbf{F}_1 & \mathbf{H}_2[i]\mathbf{F}_2 & \cdots & \mathbf{H}_2[i]\mathbf{F}_G \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_G[i]\mathbf{F}_1 & \mathbf{H}_G[i]\mathbf{F}_2 & \cdots & \mathbf{H}_G[i]\mathbf{F}_G \end{bmatrix} \in \mathbb{C}^{K \times N_{RF}}, \quad (9)$$

where the diagonal block matrix $\mathbf{H}_g[i]\mathbf{F}_g = \mathbf{Z}_g[i]\boldsymbol{\Phi}_g[i]\mathbf{F}_g \in \mathbb{C}^{K_g \times N_{RF,g}}$ is the effective channel matrix and the off-diagonal block matrix $\mathbf{H}_t[i]\mathbf{F}_g = \mathbf{Z}_t[i]\boldsymbol{\Phi}_t[i]\mathbf{F}_g \in \mathbb{C}^{K_t \times N_{RF,g}}$ is the effective interference channel matrix, $\forall t \neq g$. Therefore, the RF beamformer design targets for achieving the following two objectives: (i) maximizing the beamforming gain in the desired direction (i.e., $\text{Span}(\mathbf{F}_g) \subset \text{Span}(\boldsymbol{\Phi}_g[i])$), (ii) successfully suppress the interference among user groups (i.e., $\text{Span}(\mathbf{F}_g) \subset \cup_{t \neq g} \text{Null}(\boldsymbol{\Phi}_t[i])$). As proven in [5], both objectives are accomplished by building the RF beamformer \mathbf{F}_g via the steering vector $\mathbf{e}(\gamma_x, \gamma_y) = \frac{1}{\sqrt{M}} \phi^*(\gamma_x, \gamma_y) \in \mathbb{C}^M$ with (γ_x, γ_y) angle-pairs covering the AoD support of desired user group and excluding the AoD supports of the other user groups (please see (3) for $\phi(\gamma_x, \gamma_y)$). For covering the complete 3D elevation and azimuth angular space with minimum number of angle-pairs, M orthogonal quantized angle-pairs are defined as $\lambda_u^x = -1 + \frac{2u-1}{M_x}$ for $u = 1, \dots, M_x$ and $\lambda_n^y = -1 + \frac{2n-1}{M_y}$

⁴In [16], it is shown that the AoD parameters (i.e., mean and spread) can be efficiently obtained via an offline deep learning and geospatial data-based estimation technique instead of the traditional online channel sounding.

for $n = 1, \dots, M_y$. Considering that $N_{RF,g}$ quantized angle-pairs covers the AoD support of user group g [5, eq. (13)], we build the RF beamformer for group g as follows:

$$\mathbf{F}_g = [\mathbf{e}(\lambda_{u_1}^x, \lambda_{n_1}^y), \dots, \mathbf{e}(\lambda_{u_{N_{RF,g}}}^x, \lambda_{n_{N_{RF,g}}}^y)] \in \mathbb{C}^{M \times N_{RF,g}}. \quad (10)$$

Finally, the complete RF precoder \mathbf{F} satisfying the CM constraint given in (6) is derived by substituting (10) into (8). It is worthwhile to mention that the RF beamformer is a unitary matrix (i.e., $\mathbf{F}^H \mathbf{F} = \mathbf{I}_{N_{RF}}$).

B. BB Precoder

As seen in Fig. 1, we develop C distinct BB precoders for each subcarriers. Here, the main objective is to further suppress the residual inter-user interference. By utilizing the reduced-size effective channel matrix $\mathbf{H}[i]$, the regularized zero-forcing (RZF) technique is applied for each subcarrier. Thus, the BB precoder at subcarrier i is defined as:

$$\mathbf{B}[i] = [(\mathbf{H}[i])^H \mathbf{H}[i] + K\alpha \mathbf{I}_K]^{-1} (\mathbf{H}[i])^H \in \mathbb{C}^{N_{RF} \times K}, \quad (11)$$

where $\alpha = \frac{\sigma^2}{P_T}$ is the regularization parameter [5].

IV. PROBLEM FORMULATION

After deriving the closed-form solutions for RF beamformer \mathbf{F} and BB precoder $\mathbf{B}[i]$ for all subcarriers $i = 1, \dots, C$, the throughput maximization problem given in (6) turns into a resource allocation (RA) problem (i.e., allocating power and subcarrier resources among the users). Thus, for a given \mathbf{F} and $\mathbf{B}[i]$, we formulate the RA optimization problem as follows:

$$\begin{aligned} & \max_{\mathbf{P}[i]} R_{\text{sum},i}(\mathbf{F}, \mathbf{B}[i], \mathbf{P}[i]) \\ \text{s.t. } & \mathbb{E} \left\{ \|\mathbf{s}[i]\|_2^2 \right\} = \sum_{k=1}^K p_{k,i} \mathbf{b}_{k,i}^H \mathbf{b}_{k,i} \leq \frac{P_T}{C}, \quad p_{k,i} \geq 0, \forall k, \end{aligned} \quad (12)$$

where $R_{\text{sum},i}(\mathbf{F}, \mathbf{B}[i], \mathbf{P}[i])$ is the sum-rate at subcarrier i defined in (5). However, it is still a non-convex optimization problem due to the allocated power $p_{k,i}$ interchangeably located in the numerator and denominator in (5). Thus, the traditional optimization algorithms may not be utilized to solve the RA problem.

V. GENETIC ALGORITHM BASED RESOURCE ALLOCATION

Genetic algorithm (GA) is one of the most well-known nature-inspired evolutionary optimization algorithms [15]. It can address the shortcomings of traditional optimization algorithms, which are not able to find the optimal solution for non-convex problems. According to the GA, each solution of the problem is considered a chromosome, where the genes on a chromosome are defined as the problem variables. The chromosomes/genes evolve through generations via three main steps: (i) selection, (ii) crossover (iii) mutation [15], [18].

We here propose a new genetic algorithm based resource allocation (GA-RA) technique for MU-mMIMO-OFDM systems. As the RA optimization problem defined in (12) aims to find the optimal allocated powers for K users at subcarrier

i , it lies on K -dimensional search space. Thus, each chromosome contains K genes. Since we deal with the continuous values, the continuous GA is adopted instead of the binary GA. Consider the n^{th} chromosome representing the allocated powers of the q^{th} generation at subcarrier i as follows:

$$\mathbf{P}_n^{(q)}[i] = \text{diag} \left(\sqrt{p_{1,i,n}^{(q)}}, \dots, \sqrt{p_{K,i,n}^{(q)}} \right) \in \mathbb{R}^{K \times K}, \quad (13)$$

where $p_{k,i,n}^{(q)}$ is the k^{th} user power at subcarrier i . For satisfying the total transmitted power constraint for each subcarrier given in (12), each chromosome is normalized as:

$$\hat{\mathbf{P}}_n^{(q)}[i] = \frac{\mathbf{P}_n^{(q)}[i]}{\varepsilon_n^{(q)}[i]} = \text{diag} \left(\sqrt{\hat{p}_{1,i,n}^{(q)}}, \dots, \sqrt{\hat{p}_{K,i,n}^{(q)}} \right) \in \mathbb{R}^{K \times K}, \quad (14)$$

where $\varepsilon_n^{(q)}[i] = \sqrt{\frac{P_T/C}{\sum_{k=1}^K \hat{p}_{k,i,n}^{(q)} \mathbf{b}_{k,i}^H \mathbf{b}_{k,i}}}$ and $\hat{p}_{k,i,n}^{(q)} \in [0, 1]$.

The fitness function corresponding to the n^{th} chromosome at subcarrier i is calculated via $R_{\text{sum},i}(\mathbf{F}, \mathbf{B}[i], \mathbf{P}_n^{(q)}[i])$ defined in (5). In the proposed GA-RA, the population size is defined as N_p , which is kept as the same through Q generations. The first generation is initialized randomly, where each allocated power is uniformly distributed as $\hat{p}_{k,i,n}^{(1)} \sim U(0, \kappa)$ with $\kappa \in (0, 1]$. Afterwards, the top ρ_{mating} percent of the population is selected to form the mating pool, and the rest are discarded to free up space for new offspring. In order to generate new chromosomes, we need to choose parents from the mating pool. As new offspring inherits the traits from their parents, if only the best ones are selected the algorithm may be trapped in a local optimum point. To prevent this issue, it is needed to choose parents on a random basis while preserving our desire to choose highly qualified parents [15]. Thus, we employ the tournament selection method [19]. For selecting each parent, ρ_{sub} percent of the mating pool is randomly chosen as a subset. Then, based on the fitness value, the best chromosome of the subset is selected as a parent. Each pair of parents produces a new pair of offspring through the crossover. In the proposed approach, we apply a mixture of two crossover methods: (i) linear crossover [20], (ii) uniform crossover [18]. According to the a given pair of parents as $\hat{\mathbf{P}}_{n_1}^{(q)}[i]$ and $\hat{\mathbf{P}}_{n_2}^{(q)}[i]$ with the randomly chosen n_1 and n_2 indices via the tournament selection method, the linear crossover finds $\hat{\mathbf{P}}_{o_1}^{(q)}[i]$ and $\hat{\mathbf{P}}_{o_2}^{(q)}[i]$ as the new offspring as follows:

$$\begin{aligned} \hat{\mathbf{P}}_{o_1}^{(q)}[i, k] &= \begin{cases} \frac{3\hat{\mathbf{P}}_{n_1}^{(q)}[i, k] - \hat{\mathbf{P}}_{n_2}^{(q)}[i, k]}{2}, & 0 \leq \frac{3\hat{\mathbf{P}}_{n_1}^{(q)}[i, k] - \hat{\mathbf{P}}_{n_2}^{(q)}[i, k]}{2} \leq 1 \\ \hat{\mathbf{P}}_{n_1}^{(q)}[i, k], & \text{otherwise} \end{cases} \\ \hat{\mathbf{P}}_{o_2}^{(q)}[i, k] &= \begin{cases} \frac{3\hat{\mathbf{P}}_{n_2}^{(q)}[i, k] - \hat{\mathbf{P}}_{n_1}^{(q)}[i, k]}{2}, & 0 \leq \frac{3\hat{\mathbf{P}}_{n_2}^{(q)}[i, k] - \hat{\mathbf{P}}_{n_1}^{(q)}[i, k]}{2} \leq 1 \\ \hat{\mathbf{P}}_{n_2}^{(q)}[i, k], & \text{otherwise} \end{cases} \end{aligned} \quad (15)$$

where we abuse the notation $\hat{\mathbf{P}}_{n_1}^{(q)}[i, k]$ to represent the k^{th} diagonal entry of $\hat{\mathbf{P}}_{n_1}^{(q)}[i]$. It is important to note that the linear crossover operation might generate infeasible solutions violating the condition of $\hat{p}_{n_K,i}^{(q)} \in [0, 1]$. To prevent this

Algorithm 1 Proposed GA-RA technique

Input: $\mathbf{F}, \mathbf{B}[i], P_T, N_p, Q, \kappa, \rho_{\text{mating}}, \rho_{\text{sub}}, \rho_{\text{elite}}, \rho_{\text{mut}}$

Output: $\mathbf{P}[1], \mathbf{P}[2], \dots, \mathbf{P}[C]$

```

1: for  $i = 1 : C$  do
2:   for  $n = 1 : N_p$  do
3:     Initialize  $\hat{\mathbf{P}}_n^{(1)}[i]$  with diagonal entries  $\hat{p}_{k,i,n}^{(1)} \sim U(0, \kappa)$ .
4:   end for
5:   for  $q = 1 : Q$  do
6:     Keep top  $\rho_{\text{mating}}$  percent of population in  $q^{th}$  generation as mating pool and discard the rest.
7:     for  $p = 1 : \frac{(1-\rho_{\text{mating}})}{2} N_p$  do
8:       for  $j = 1 : 2$  do
9:         For tournament selection, create a subset by randomly choosing  $\rho_{\text{sub}}$  percent of mating pool.
10:        Select the best subset member as parent  $\hat{\mathbf{P}}_{n_j}^{(q)}[i]$ .
11:      end for
12:      Apply linear crossover via (15) to calculate new offspring  $\hat{\mathbf{P}}_{o_1}^{(q)}[i]$  and  $\hat{\mathbf{P}}_{o_2}^{(q)}[i]$ .
13:      Update  $\hat{\mathbf{P}}_{o_1}^{(q)}[i]$  and  $\hat{\mathbf{P}}_{o_2}^{(q)}[i]$  via uniform crossover by randomly swapping their corresponding genes.
14:    end for
15:    Include all offspring to the population.
16:    Select top  $\rho_{\text{elite}}$  percent of population as elite members, mutate the rest based on mutation rate of  $\rho_{\text{mut}}$ .
17:    Carry elite/mutated population to  $(q+1)^{th}$  generation.
18:  end for
19:  Find  $\mathbf{P}[i]$  via (16).
20: end for
21: return  $\mathbf{P}[1], \mathbf{P}[2], \dots, \mathbf{P}[C]$ 

```

from happening, infeasible genes are replaced by the original parent's genes. Subsequently, the uniform crossover is applied by randomly swapping the genes of $\hat{\mathbf{P}}_{o_1}^{(q)}[i]$ with the corresponding genes of $\hat{\mathbf{P}}_{o_2}^{(q)}[i]$. The crossover process is applied until reaching the initial population size of N_p . Afterwards, the top ρ_{elite} percent of them are selected as elite members, implying that they do not undergo mutation to keep the good solutions. The rest of the population are mutated by adding a Gaussian noise distributed as $\mathcal{N}(0, 10^{-4})$ to the randomly selected genes based on the mutation rate of ρ_{mut} . All mutated genes are bounded in the interval of $[0, 1]$. Here, we advocate the exploitative behavior of the GA-RA technique by introducing crossover, whereas the mutation encourages the exploratory behavior of the GA-RA technique for preventing it from converging to a local optimum solution. The whole procedure is repeated through Q generations. Finally, the best solution of the last generation is reported as the power allocation block at subcarrier i as follows:

$$\mathbf{P}[i] = \varepsilon_{best}^{(Q)}[i] \hat{\mathbf{P}}_{best}^{(Q)}[i] \quad (16)$$

The summary of the proposed GA-RA technique is expressed in detail in Algorithm 1.

TABLE I: Simulation parameters.

# of antennas [21]	$M = 16 \times 16 = 256$
Cell radius [21]	100m
BS height [21] User height [21]	10m 1.5m-2.5m
User-BS horizontal distance	10m-90m
# of groups	$G = 3$
# of users in each group	$K_g = \frac{K}{G}$
Mean EAoD Mean AAoD	$\theta_g = 60^\circ$ $\psi_g = 21^\circ + 120^\circ(g-1)$
EAoD spread AAoD spread	$\delta_g^\theta = 15^\circ$ $\delta_g^\psi = 11^\circ$
Path loss exponent [7]	$\eta = 3.76$
Noise PSD [7]	-174 dBm/Hz
Channel bandwidth [7]	10 kHz
# of paths	$L = 10$
Antenna spacing (in wavelength)	$d = 0.5$
# of network realizations	1000

VI. ILLUSTRATIVE RESULTS

This section illustrates Monte-Carlo simulation results to evaluate the sum-rate and energy-efficiency performance of the proposed GA-RA technique in MU-mMIMO-OFDM systems. Table I presents the simulation parameters considering the 3D microcell scenario [21]. We have considered a BS equipped with a square URA with $M = 256$ antenna elements. For the proposed GA-RA technique, we assume the population size as $N_p = 100$ and the number of generations as $Q = 10$, unless otherwise stated. Also, other parameters are chosen as $\kappa = \frac{1}{20}$, $\rho_{\text{mating}} = 40\%$, $\rho_{\text{sub}} = 10\%$, $\rho_{\text{elite}} = 80\%$, $\rho_{\text{mut}} = 60\%$.

The energy-efficiency of the MU-mMIMO-OFDM systems is evaluated by taking the ratio of the overall sum-rate R_{sum} and the total power consumption P_{total} as [5], [6], [9]:

$$\beta = \frac{R_{\text{sum}}}{P_{\text{total}}} = \frac{\sum_{i=1}^C \mathbb{E} \{ R_{\text{sum},i} (\mathbf{F}, \mathbf{B}[i], \mathbf{P}[i]) \}}{P_T + N_{RF} P_{RF} + N_{PS} P_{PS}} \text{ [bps/Hz/W]}, \quad (17)$$

where P_T is the transmission power, P_{RF} (P_{PS}) is the power consumption per each RF chain (phase-shifter), N_{RF} (N_{PS}) is the number of RF chains (phase-shifters). We here assume that $P_{RF} = 250$ mW and $P_{PS} = 1$ mW as in [9]. It is important to highlight that the FDP architecture has $N_{RF} = M = 256$ RF chains and does not require any phase-shifters (i.e., $N_{PS} = 0$). On the other hand, the proposed AB-HP architecture requires $N_{RF} = 36$ RF chains and $N_{PS} = N_{RF} \times M = 9216$ phase-shifters to support $M = 256$ antennas in the considered simulation setup. Thus, the number of RF chains reduces from 256 to 36 (i.e., 87% reduction in hardware cost/complexity and CSI overhead size).

The sum-rate and energy-efficiency are respectively plotted versus the number of subcarriers in Fig. 2(a) and Fig. 2(b), where $P_T = 40$ dBm and $K = 15$ users. Here, the proposed GA-RA technique is compared with equal RA (EQ-RA) and particle swarm optimization based RA (PSO-RA)⁵ [7]. Both GA-RA and PSO-RA techniques consider $Q = 10$ generations (e.g., iterations). Also, the power values are equally distributed among the users for all subcarriers in EQ-RA (i.e., $p_{1,1} = \dots = p_{k,i} = \dots = p_{K,C}$, $\forall k, i$). In the

⁵Even though PSO algorithm is employed for power allocation optimization over single subcarrier transmission in [7], it nevertheless serves as benchmark.

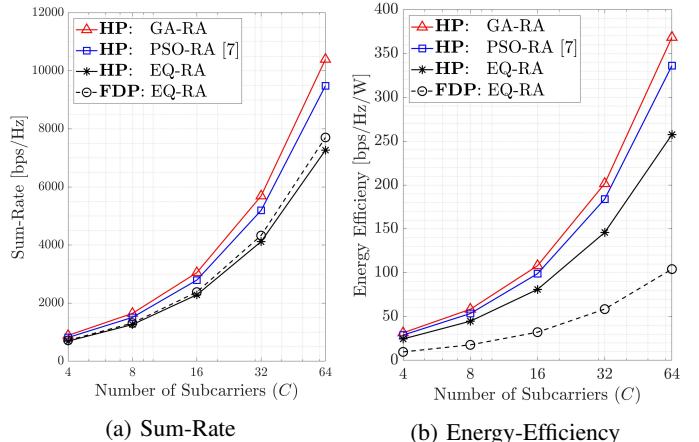


Fig. 2: Performance comparison among HP and FDP schemes ($P_T = 40$ dBm and $K = 15$ users).

EQ-RA scheme, the performance comparison is investigated between FDP and HP schemes. The numerical results in Fig. 2(a) reveal that HP with EQ-RA closely achieves the sum-rate provided by its FDP counterpart. Additionally, Fig. 2(b) illustrates that HP with EQ-RA considerably improves the energy-efficiency compared to the FDP scheme by means of reduced hardware cost/complexity. For instance, when there are $C = 16$ subcarriers, the energy-efficiency increases from 32 bps/Hz/W to 80 bps/Hz/W. In comparison to EQ-RA, the proposed GA-RA technique enhances the sum-rate and energy-efficiency by 28.6%, 33.5%, 43.1% for $C = 4, 16, 64$ subcarriers, respectively. Hence, the performance gap between GA-RA and EQ-RA keeps increasing for larger number of subcarriers. On the other hand, the numerical results demonstrate that the proposed GA-RA technique outperforms PSO-RA by further enhancing both sum-rate and energy-efficiency.

The average sum-rate per subcarrier is plotted versus the total transmit power for $C = 8$ and $C = 64$ subcarriers in Fig. 3, where $K = 3, 9, 15$ downlink users are served. It is observed that the proposed GA-RA technique significantly enhances the system capacity. For example, when the transmit power is $P_T = 40$ dBm, applying the proposed GA-RA technique with $Q = 10$ generations enhances the sum-rate performance for $C = 8$ subcarriers by 19.66%, 29.4%, 30.5% and for $C = 64$ subcarriers by 25.41%, 38%, 43.06% compared with EQ-RA technique for $K = 3, 9, 15$ users, respectively. In addition, the performance gap between GA-RA and EQ-RA increases for the larger number of users. On the other hand, when there are less number of users (e.g., $K = 3$ users), the sum-rate performance is saturated in earlier generations.

Fig. 4(a) compares the average sum-rate per subcarrier performance of GA-RA and EQ-RA techniques versus the number of subcarriers, where the transmit power is chosen as $P_T = 20, 40, 60$ dBm. Here, we observe that the sum-rate improvement provided by the GA-RA technique is almost constant for different subcarrier scenarios. Specifically, the performance of GA-RA is better than EQ-RA by approximately 48.5 bps/Hz/subcarrier at $P_T = 40, 60$ dBm. However,

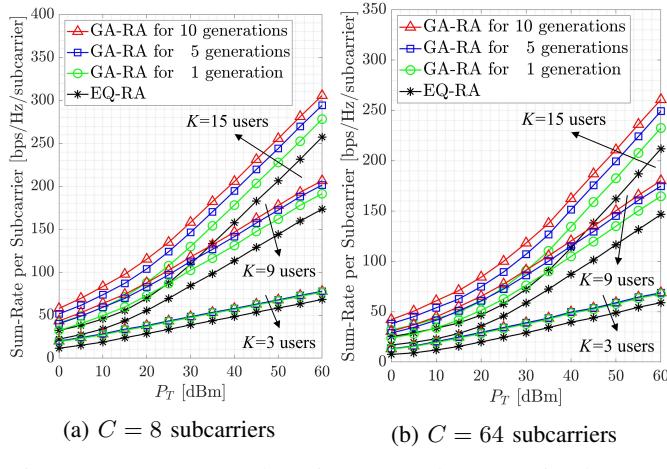


Fig. 3: Sum-rate per subcarrier versus the transmitted power.

at $P_T = 20$ dBm, the performance improvement is 47.6 bps/Hz/subcarrier for $C = 1$ subcarrier and it decays 38.2 bps/Hz/subcarrier for $C = 64$ subcarriers. Moreover, it is shown that by increasing the number of subcarriers, the sum-rate per subcarrier performance degrades since less power is allocated to each subcarrier as discussed in expressed in (12).

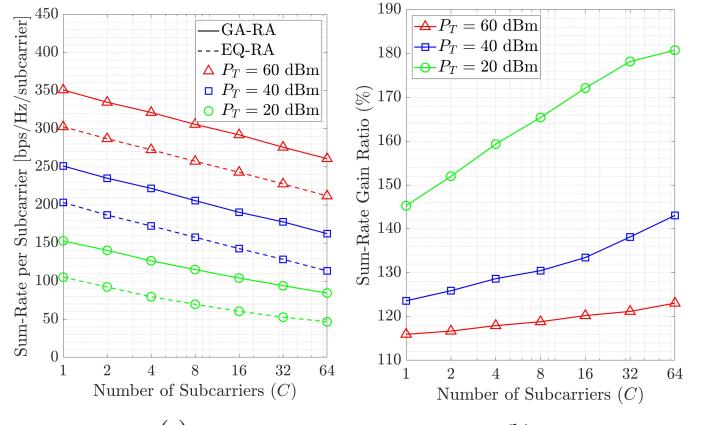
By defining the sum-rate gain ratio as $\frac{R_{sum,GA}^{(Q)}}{R_{sum,EQ}^{(Q)}}$, Fig. 4(b) depicts that the proposed GA-RA technique enhances further the sum-rate gain ratio by increasing the number of subcarriers. The importance of this fact appears is more apparent in the case of low transmitted power.

VII. CONCLUSIONS

In this work, a novel genetic algorithm based resource allocation (GA-RA) technique has been developed for energy-efficient throughput maximization in the MU-mMIMO-OFDM systems. Furthermore, the angular-based hybrid precoding (AB-HP) scheme has been developed for the OFDM-based downlink transmission to reduce the number of RF chains and the CSI overhead size. The AB-HP scheme first develops a single RF beamformer block for all subcarriers via AoD information, then builds a distinct BB precoder for each subcarrier via the reduced-size effective CSI. Next, the GA-RA technique allocates power and subcarrier resources among the users. The promising numerical results demonstrate that the GA-RA technique greatly enhances the sum-rate (i.e., throughput) performance of MU-mMIMO-OFDM systems compared to the conventional equal RA (EQ-RA). Also, the performance gain ratio with GA-RA increases for the larger number of subcarriers, especially in case of low transmission power.

REFERENCES

- [1] A. N. Uwaechia *et al.*, “A comprehensive survey on millimeter wave communications for fifth-generation wireless networks: Feasibility and challenges,” *IEEE Access*, vol. 8, pp. 62 367–62 414, 2020.
- [2] S. A. Busari *et al.*, “Millimeter-wave massive MIMO communication for future wireless systems: A survey,” *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 836–869, 2nd Quart. 2018.
- [3] I. Ahmed *et al.*, “A survey on hybrid beamforming techniques in 5G: Architecture and system model perspectives,” *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 3060–3097, 4th Quart. 2018.
- [4] A. Koc *et al.*, “Full-duplex mmWave massive MIMO systems: A joint hybrid precoding/combinining and self-interference cancellation design,” *IEEE Open J. Commun. Soc.*, vol. 2, pp. 754–774, 2021.
- [5] A. Koc *et al.*, “3D angular-based hybrid precoding and user grouping for uniform rectangular arrays in massive MU-MIMO systems,” *IEEE Access*, vol. 8, pp. 84 689–84 712, May 2020.
- [6] M. Mahmood *et al.*, “Energy-efficient MU-massive-MIMO hybrid precoder design: Low-resolution phase shifters and digital-to-analog converters for 2D antenna array structures,” *IEEE Open J. Commun. Soc.*, vol. 2, pp. 1842–1861, 2021.
- [7] A. Koc *et al.*, “Swarm intelligence based power allocation in hybrid massive MIMO systems,” in *2021 IEEE Wireless Commun. and Netw. Conf. (WCNC)*, Mar. 2021, pp. 1–7.
- [8] W. Zheng *et al.*, “Sub-connected hybrid precoding architectures in massive MIMO systems,” in *2020 IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2020, pp. 1–6.
- [9] X. Gao *et al.*, “Energy-efficient hybrid analog and digital precoding for mmwave MIMO systems with large antenna arrays,” *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 998–1009, Apr. 2016.
- [10] F. Sohrabi *et al.*, “Hybrid analog and digital beamforming for mmWave OFDM large-scale antenna arrays,” *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1432–1443, July 2017.
- [11] J. Du *et al.*, “Weighted spectral efficiency optimization for hybrid beamforming in multiuser massive MIMO-OFDM systems,” *IEEE Trans. Veh. Technol.*, vol. 68, no. 10, pp. 9698–9712, 2019.
- [12] K. B. Dsouza *et al.*, “Hybrid precoding with partially connected structure for millimeter wave massive MIMO OFDM: A parallel framework and feasibility analysis,” *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8108–8122, 2018.
- [13] K. B. Letaief *et al.*, “The roadmap to 6G: AI empowered wireless networks,” *IEEE Commun. Mag.*, vol. 57, no. 8, pp. 84–90, 2019.
- [14] O. Kramer, “Genetic algorithms,” in *Genetic algorithm essentials*. Springer, 2017, pp. 11–19.
- [15] S. Mirjalili, “Evolutionary algorithms and neural networks,” in *Studies in Computational Intelligence*. Springer, 2019, vol. 780.
- [16] X. Zhu *et al.*, “A deep learning and geospatial data based channel estimation technique for hybrid massive MIMO systems,” *IEEE Access*, vol. 9, pp. 145 115–145 132, 2021.
- [17] X. Cheng *et al.*, “Communicating in the real world: 3D MIMO,” *IEEE Wireless Commun.*, vol. 21, no. 4, pp. 136–144, Aug. 2014.
- [18] R. L. Haupt *et al.*, “Practical genetic algorithms,” 2004.
- [19] B. L. Miller *et al.*, “Genetic algorithms, tournament selection, and the effects of noise,” *Complex systems*, vol. 9, no. 3, pp. 193–212, 1995.
- [20] A. H. Wright, “Genetic algorithms for real parameter optimization,” in *Found. genetic algorithms*. Elsevier, 1991, vol. 1, pp. 205–218.
- [21] 3GPP TR 38.901, “5G: Study on channel model for frequencies from 0.5 to 100 GHz,” Tech. Rep. Ver. 16.1.0, Nov. 2020.

Fig. 4: (a) Sum-rate per subcarrier and (b) sum-rate gain ratio versus number of subcarriers ($K = 15$ users).