Hybrid Analog-Digital Beamforming Design for SE and EE Maximization in Massive MIMO Networks

Khaled Ardah, Gábor Fodor, *Senior Member, IEEE*, Yuri C. B. Silva, *Member, IEEE*, Walter C. Freitas, Jr., and André L. F. de Almeida, *Senior Member, IEEE*

***Abstract*—Hybrid analog-digital (HAD) beamforming architectures have been proposed to facilitate the practical implementation of massive multiple-input multiple-output (MIMO) systems by reducing the number of employed radio frequency chains. While most prior studies have aimed to maximize spectral efficiency (SE), the present paper proposes a two-stage HAD beamforming design for multi-user MIMO systems that can be used to maximize either the system's overall energy efficiency (EE) or SE. This problem is nonconvex and NP-hard due to the joint optimization between the analog and digital domains and the constant modulus constraints required by the analog domain. To address this problem, we propose a decoupled two-stage design wherein the first stage, the analog beamforming parts are updated, which are then taken into account in the second stage to design the digital beamforming parts to maximize the system's EE or SE. We consider two widely-used HAD beamforming techniques that utilize either fully-connected (FC) or partially-connected (PC) architectures employing variable phase-shifters. Using the most recently available data for the circuitry power consumption of the components, we compare the performance of these two HAD architectures with that of the fully-digital (FD) architecture in terms of the total circuitry power consumption, and achieved SE and EE. We find that there is a certain number of users above which the FC architecture has higher circuitry power consumption than the FD counterpart, in contrast to the PC architecture that always has lower circuitry power consumption. More importantly, our results reveal, contrary to the common opinion, that depending on the circuitry parameters the FD architecture may achieve not only higher SE, but also higher EE than the HAD architectures.**

***Index Terms*—Hybrid analog-digital, MIMO, spectral/energy efficiency maximization.**

Manuscript received April 22, 2019; revised July 3, 2019; accepted July 28, 2019. Date of publication August 5, 2019; date of current version January 15, 2020. This work was supported in part by Ericsson Research, Technical Cooperation Contracts UFC.45 (TIDE5G) and UFC.46 (NAIVE), in part by the Brazilian National Council for Scientific and Technological Development (CNPq), in part by CAPES/PROBRAL Grant 88887.144009/2017-00, and in part by CAPES/PRINT Grant 88887.311965/2018-00. The work of G. Fodor was supported by the Smart and Secure Spectrum Sharing for 5G Long Term Evolution (SMASH) project. The review of this article was coordinated by Dr. Z. Ding. *(Corresponding author: Khaled Ardah.)*

K. Ardah was with the Wireless Telecom Research Group (GTEL), Federal University of Ceará, Fortaleza 60020-181, Brazil. He is now with the Communications Research Laboratory, Technical University of Ilmenau, Ilmenau 98693, Germany (e-mail: khaledardah@gtel.ufc.br).

G. Fodor is with the Ericsson Research and Royal Institute of Technology, Stockholm 114 28, Sweden (e-mail: gabor.fodor@ericsson.com; gaborf@kth.se).

Y. C. B. Silva, W. C. Freitas, Jr., and A. L. F. de Almeida are with the Wireless Telecom Research Group (GTEL), Federal University of Ceará, Fortaleza 60020-181, Brazil (e-mail: yuri@gtel.ufc.br; walter@gtel.ufc.br; andre@gtel.ufc.br).

Digital Object Identifier 10.1109/TVT.2019.2933305

I. Introduction

MASSIVE multiple-input multiple-output (MIMO) communication is considered one of the key techniques for meeting the ambitious goal of 1000-fold increase on area spectral efficiency of 5G systems [1], [2]. However, when the number of antenna elements grows large, the current fully digital (FD) implementation of MIMO processing, which dedicates one radio frequency (RF) chain to each antenna, is prohibitive due to the associated high cost, complexity, and circuitry power consumption. Hybrid analog-digital (HAD) implementation of MIMO processing is seen as a possible solution to realize MIMO systems in practice [3]. With the HAD system, the beamforming matrix is divided into a high-dimensional analog beamforming (ABF) part that is realized, for example, by phase-shifters (PSs) [4], [5] and/or switches [6], [7] and a low-dimensional digital beamforming (DBF) part. In this way, the number of RF chains can be reduced to equal the number of transmitted/received data streams [4], which is, generally, much lower than the number of antenna elements.

Initial works on HAD beamforming design [4]–[7] focused on maximizing the system's spectral efficiency (SE), which measures the number of bits/sec/Hz that can be reliability transmitted. Furthermore, energy efficiency (EE) measures the number of bits/sec/Hz that can be transmitted per Joule has been recognized as an important performance metric for future green 5G networks [2]. In this paper, we consider a multiuser MIMO downlink system model and propose an HAD beamforming design algorithm tackling both EE and SE maximization.

The HAD beamforming design is, generally, a nontrivial task, mainly due to the joint optimization of the analog and digital parts and the nonconvex constraints that arise from the analog part optimization. These involve, for example, constant modulus constraints required by PSs [4] or binary constraints required by switches [7]. Therefore, a sub-optimal approach is widely adopted in practice [4]–[7] by decoupling the optimization of the analog and digital parts and treating them separately. In [4], for instance, the ABF matrix, realized using a network of PSs, is updated from the channel's steering vectors, which naturally admit the constant modulus constraints, using the orthogonal matching pursuit technique. In [7], the ABF matrix, realized using a network of switches, is updated using the cross-entropy machine-learning technique. In [8], we recently proposed a unified analog beamforming design algorithm that is valid for both PS-based or switch-based architectures, which updates the analog matrices such that the equivalent channel's capacity is

maximized. On the other hand, classical beamforming design methods like block diagonalization (BD) and zero-forcing (ZF), for the multiuser scenarios, or maximum ratio transmission (MRT), for the single-user scenarios, can be used directly to update the DBF part considering the resulting equivalent channels [4], [5], [8].

A. Related Works

In the past few years, several EE maximization beamforming design algorithms for HAD systems have been proposed [9]–[12]. In [9], the authors consider a fully-connected architecture and propose an algorithm that jointly optimizes the transmit power and the number of active RF chains using a mixed-integer fractional technique. In [10], the authors consider fully-connected and partially-connected architectures and propose an EE maximization algorithm, in which both the analog and digital parts are updated based on the alternating direction method of multipliers technique. In [11], the authors consider a partially-connected architecture, where the analog part is updated element-wise to minimize the interference-leakage between its sub-blocks and the digital part is then updated to maximize the EE based on the alternating optimization technique. However, both algorithms [10], [11] consider a single-user scenario and are computationally complex, especially with large-scale systems, due to their iterative nature. By contrast, the authors in [12] consider a fully-connected architecture and propose a low-complexity beamforming design to maximize the EE, again for the single-user scenarios, using the singular value decomposition (SVD) technique and a water-filling-like power allocation method. Assuming that the optimal FD beamformers for maximizing the system SE or EE are known a priori, the authors in [13], [14] propose n HAD beamforming design, where the problem is formulated as n Euclidean norm-minimization between the HAD beamformers and the given FD beamformers. Meanwhile, asymptotic SE and EE performance analysis for multi-user HAD multiple-input and single-output (MISO) systems operating with ideal and quantized phase shifters have been investigated in [15], wherein the ABF part is simply updated from the phase angles of users' channels, while the DBF is updated using the classical ZF technique.

Considering the cooperative multi-cell multi-user HAD MISO systems, the authors in [16] propose an optimization problem with the objective of maximizing the system EE by jointly solving the HAD transmit beamforming and the user-to-BS association problems, wherein the former is solved using the Eigen beamforming algorithm and the latter is solved using a Lagrangian approach. Meanwhile, the authors in [17] consider the EE maximization problem for multi-user HAD MISO system and propose an iterative approach, where the ABF and the DBF parts are updated iteratively using convex quadratically constrained quadratic program problem formulations. HAD beamforming systems have been also investigated for energy-harvesting design in [18], where the authors consider an analog-only transmitter with multi antenna array, single RF-chain and a single-antenna user to investigate the channel estimation problem and the optimal average harvested energy at the receiver under phase shifter impairments and channel estimation errors.

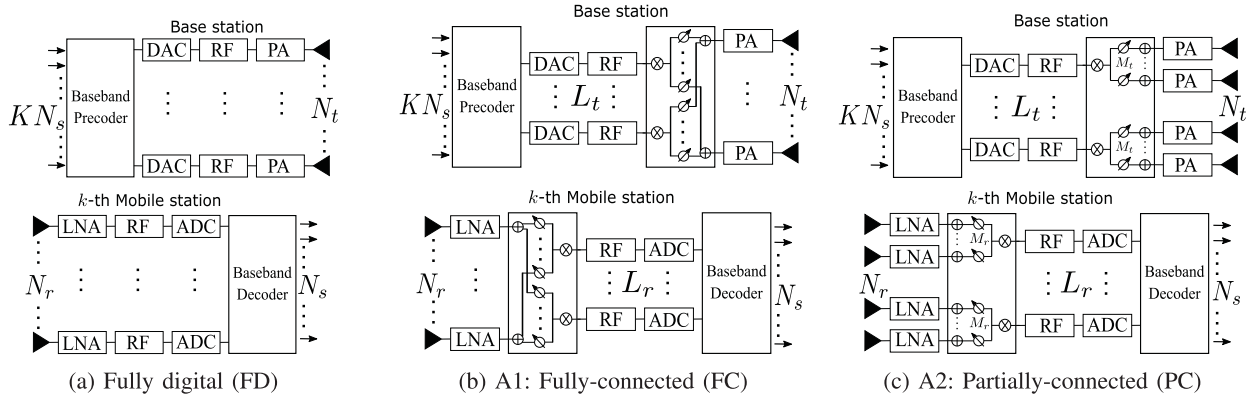
Furthermore, the works of [19], [20] approach the EE maximization problem by formulating the analog/digital beamforming design problem to minimize the total transmit power subject to some quality-of-service constraints; in [19] for single-cell scenarios and in [20] for multicell scenarios. Considering more practical systems, the authors in [21] propose an EE maximization algorithm that takes into account the non-ideal settings of the power amplifiers. It is worth noting that if the analog and digital parts are decoupled, as with all the above proposals, the conventional EE maximization algorithms proposed for FD MIMO systems can be readily used to optimize the digital part, using, for example, the algorithm proposed in [22].

B. Contributions

Unlike the works discussed above, the present paper considers a multiuser MIMO downlink system and proposes a two-stage HAD beamforming design approach tackling both SE and EE maximization, while taking into account the hardware constraints and realistic circuitry power consumption. The main contributions of this paper are summarized as follows.

- At first, we show that, based on the most recently available data for the circuitry energy consumption of PSs and other circuitry components, there is a certain number of users threshold above which the FD beamforming architecture has lower circuitry power consumption and higher EE than the fully-connected and partially-connected hybrid analog-digital beamforming architectures.
- We propose, differently from our work in [8], an iterative ABF design algorithm, called ARAB, with the objective of maximizing the EE based on an alternating optimization technique. The ARAB algorithm is guaranteed to converge monotonically to a local stationary point, but not necessarily to the global optimum.
- We propose a transmit DBF design algorithm, where the beamforming directions are first updated using the well-known BD approach [23], [24] and the power allocation vector, unlike [8], is updated with the objective of maximizing the EE, for which a new power allocation algorithm is proposed.
- We consider both the fully-connected [4] and the partially-connected [5] architectures, where the analog part is realized using a network of PSs, and compare their performance in terms of the achieved SE and EE. Unlike [25]–[27], our analysis is done when the HAD beamforming matrices are designed using both the EE and SE maximization approaches, thus, we provide more insights into their true EE and SE.

We show that, based on the most recently available data for the energy consumption of PSs and other circuitry components, the fully-connected architecture based on the high-resolution PSs actually has higher circuitry power consumption than the FD counterpart if the number of users exceeds a certain threshold given by (9). Thus, in such scenarios, the FD structures are shown to achieve higher EE. By contrast, the partially-connected architecture always has lower circuitry power consumption than the



(a) Fully digital (FD) (b) A1: Fully-connected (FC) (c) A2: Partially-connected (PC)

Fig. 1. Block diagrams of the considered beamforming architectures of the BS and the k-th MS.

FD and the fully-connected architectures, as given by (10). However, due to the severe degradation on its degrees-of-freedom, we found that it can still achieve not only lower SE but also lower EE than both architectures in some multiuser scenarios. Further, we show that the EE maximization approach is a more appropriate beamforming design, as it provides a better trade-off between maximizing the SE and EE and minimizing the transmit power.

C. Paper Organization and Notation

The rest of this paper is organized as follows. Section II presents the system and power consumption models. Section III formulates the EE-Maximization problem. In Sections IV and V, we present our proposed ABF and DBF design algorithms, respectively, for EE and SE maximization. In Section VI, the computational complexity of the proposed algorithms is analyzed. Next, we show simulation results in Section VII and finally conclude the paper in Section VIII.

Notation: Scalars are denoted by single letters in italic type, while matrices/vectors are denoted by boldface letters. The operations $(\cdot)^{\mathrm{H}}$, $(\cdot)^{\mathrm{T}}$, $(\cdot)^{-1}$, $\|\cdot\|$, $\log(\cdot)$, and $\det(\cdot)$ denote the complex conjugate transpose, the transpose, the inverse, the standard Euclidean norm, the logarithm of base 2, and the determinate function, respectively. $\mathbb{E}(\cdot)$ denotes the statistical expectation. $\mathrm{Bdiag}\{\cdot\}$ denotes the block-diagonal operator of a given vector/matrix. $[\mathbf{A}]_{[i,:]}$ selects the i-th row, while $[\mathbf{A}]_{[:,i]}$ selects the i-th column from matrix $\mathbf{A}$. Finally, $[\mathbf{a}]_{[i]}$ selects the i-th element of vector $\mathbf{a}$.

II. System Model

We consider a multiuser MIMO downlink system consisting of a single base station (BS), equipped with N_t antennas, and K mobile stations (MSs), each equipped with N_r antennas receiving N_s data streams. The BS has $L_t = KN_s \leq N_t$ RF chains and each MS has $L_r = N_s \leq N_r$ RF chains. At the BS, a DBF matrix $\mathbf{B} = [\mathbf{B}_1, \ldots, \mathbf{B}_K] \in \mathbb{C}^{L_t \times KN_s}$ processes KN_s data streams to produce L_t outputs, which are upconverted and mapped via an ABF matrix $\mathbf{F} \in \mathbb{C}^{N_t \times L_t}$ to the N_t antenna elements for transmission. Here, $\mathbf{B}_k \in \mathbb{C}^{L_t \times N_s}$ denotes the k-th MS transmit DBF matrix, $k \in \{1, \ldots, K\}$. The structure at the k-th MS is similar. An ABF matrix $\mathbf{W}_k \in \mathbb{C}^{N_r \times L_r}$ combines the RF signals from the N_r antennas to create L_r outputs, which are downconverted and further combined using a DBF matrix $\mathbf{D}_k \in \mathbb{C}^{L_r \times N_s}$. Therefore, the total transmit and receive beamforming matrices of MS k are given respectively as

$$\mathbf{T}_k = \mathbf{F}\mathbf{B}_k \in \mathbb{C}^{N_t \times N_s} \text{ and } \mathbf{R}_k = \mathbf{W}_k\mathbf{D}_k \in \mathbb{C}^{N_r \times N_s}. \quad (1)$$

The ABF parts are subject to specific constraints depending on the hardware used to implement them. Note that several ABF architectures can be found on the literature, see [4]–[8]. However, in this paper, we focus on the two well-known architectures: A1) fully-connected [4] and A2) partially-connected [5]. Nonetheless, due to our decoupled beamforming design structure, any other ABF architecture, see [6], [8], can be readily used as well. Fig. 1 shows the considered HAD beamforming architectures in this paper along with the classical FD beamforming architecture.

In the fully-connected architecture A1 [4], each RF chain is connected to all antenna elements using a network of PSs. Therefore, the ABF matrices $\mathbf{F}$ and $\mathbf{W}_k$ of user k are given as

$$\mathbf{F} = [\mathbf{f}_1, \ldots, \mathbf{f}_{L_t}] \text{ and } \mathbf{W}_k = [\mathbf{w}_{k,1}, \ldots, \mathbf{w}_{k,L_r}], \quad (2)$$

where $\mathbf{f}_j \in \mathbb{C}^{N_t}, \mathbf{w}_{k,j} \in \mathbb{C}^{N_r}$, and $\left|[\mathbf{f}_j]_{[i]}\right| = \left|[\mathbf{w}_{k,j}]_{[i]}\right| = 1, \forall j, i, k$. Meanwhile, in the partially-connected architecture A2 [5], each RF chain is connected to only a subset of the antenna elements using a network of PSs. The ABF matrices $\mathbf{F}$ and $\mathbf{W}_k$ of user k are given as

$$\mathbf{F} = \mathrm{Bdiag}\{\mathbf{f}_1, \ldots, \mathbf{f}_{L_t}\}, \mathbf{W}_k = \mathrm{Bdiag}\{\mathbf{w}_{k,1}, \ldots, \mathbf{w}_{k,L_r}\}, \quad (3)$$

where $\mathbf{f}_j \in \mathbb{C}^{M_t}, \mathbf{w}_{k,j} \in \mathbb{C}^{M_r}$, $\left|[\mathbf{f}_j]_{[i]}\right| = \left|[\mathbf{w}_{k,j}]_{[i]}\right| = 1, \forall j, i, k$, assuming that $M_t = N_t/L_t$ and $M_r = N_r/L_r$.

A. Circuitry Power Models and Analysis

In the following, we first introduce the circuitry power models of the considered beamforming architectures shown in Fig. 1. Later on, we investigate the number of users threshold, above which the FD architecture has lower power consumption than the fully-connected (A1) and the partially-connected (A2) architectures. To make our analysis applicable to other developed

algorithms in the literature, we consider the well-known and widely used circuitry power model from [9], [25]. Let P_c^X denotes the total circuitry power consumption, in Watts, by the BS and K MSs when using the architecture X, $X \in \{\text{FD}, \text{A1}, \text{A2}\}$, i.e.,

$$P_c^X = P_{c,t}^X + K \cdot P_{c,r}^X, \tag{4}$$

where $P_{c,t}^X$ and $P_{c,r}^X$ denote the circuitry power consumption of architecture X at the BS and each MS, respectively. Following the circuitry power consumption model from [25], the $P_{c,t}^X$ and $P_{c,r}^X$ are given as

$$P_{c,r}^X = \begin{cases} N_r(P_{\text{LNA}} + P_{\text{RF}} + P_{\text{ADC}}) + P_{\text{BB}} & \text{FD} \\ N_r(P_{\text{LNA}} + L_r P_{\text{PS}}) + L_r(P_{\text{RF}} + P_{\text{ADC}}) + P_{\text{BB}} & \text{A1} \\ N_r(P_{\text{LNA}} + P_{\text{PS}}) + L_r(P_{\text{RF}} + P_{\text{ADC}}) + P_{\text{BB}} & \text{A2} \end{cases} \tag{5}$$

$$P_{c,t}^X = \begin{cases} N_t(P_{\text{PA}} + P_{\text{RF}} + P_{\text{DAC}}) + P_{\text{BB}} & \text{FD} \\ N_t(P_{\text{PA}} + L_t P_{\text{PS}}) + L_t(P_{\text{RF}} + P_{\text{DAC}}) + P_{\text{BB}} & \text{A1} \\ N_t(P_{\text{PA}} + P_{\text{PS}}) + L_t(P_{\text{RF}} + P_{\text{DAC}}) + P_{\text{BB}} & \text{A2} \end{cases} \tag{6}$$

where P_{LNA}, P_{PS}, P_{RF}, P_{ADC}, P_{BB}, P_{PA}, and P_{DAC} denote respectively the power consumption by a low-noise-amplifier, PS, RF chain, analog-to-digital converter, baseband amplifier, power-amplifier, and digital-to-analog converter. The power consumption for each of the above components can be written with respect to the reference power $P_{\text{ref}} = 0.02$ W as [9]

$$\begin{aligned} P_{\text{LNA}} &= P_{\text{ref}} \\ P_{\text{PA}} &= 7P_{\text{ref}} \\ P_{\text{ADC}} &= 10P_{\text{ref}} \\ P_{\text{DAC}} &= 5.5P_{\text{ref}} \\ P_{\text{BB}} &= 10P_{\text{ref}} \\ P_{\text{RF}} &= 2P_{\text{ref}} \\ P_{\text{PS}} &= 1.5P_{\text{ref}}. \end{aligned}$$

Let us define the circuitry power consumption ratio between the FD architecture and the HAD architectures A1 and A2 as

$$\alpha^{\{\text{A1,A2}\}} = \frac{P_c^{\{\text{A1,A2}\}}}{P_c^{\text{FD}}}. \tag{7}$$

Substituting the circuitry power consumption values provided above into (5) and (6) and using the assumption that $L_r = N_s$ and $L_t = KN_s = KL_r$, α^{A1} and α^{A2} can be written, after straightforward simplifications, as

$$\alpha^{\text{A1}} = \frac{Kx_1 + c_1}{Ky + b}, \alpha^{\text{A2}} = \frac{Kx_2 + c_2}{Ky + b}, \tag{8}$$

where $x_1 = L_r(1.5N_t + 1.5N_r + 19.5) + N_r + 10$, $x_2 = 2.5N_r + 19.5L_r + 10$, $c_1 = 7N_t + 10$, $c_2 = 8.5N_t + 10$, $y = 13N_r + 10$, and $b = 14.5N_t + 10$. From (8), we are interested in the number of users threshold above which the FD architecture consumes less circuitry power. First, note that $b > c_2 > c_1$. Thus, investigating (8) for the value of K such that $\alpha^{\text{A1}} \geq 1$, gives us

$$\begin{aligned} \alpha^{\text{A1}} \geq 1 \text{ if } K &\geq \left\lceil \frac{b - c_1}{x_1 - y} \right\rceil \\ &\geq \left\lceil \frac{7.5N_t}{L_r(1.5N_t + 1.5N_r + 19.5) - 12N_r} \right\rceil, \end{aligned} \tag{9}$$

where $\lceil \cdot \rceil$ denotes the ceiling function. Meanwhile, investigating (8) for the value of K such that $\alpha^{\text{A2}} \geq 1$, gives us

$$\alpha^{\text{A2}} \geq 1 \text{ if } K \geq \left\lceil \frac{b - c_2}{x_2 - y} \right\rceil \geq \left\lceil \frac{6N_t}{19.5L_r - 10.5N_r} \right\rceil. \tag{10}$$

From (10), since $L_r \leq N_r$, then to have a meaningful K value, i.e., $K \geq 1$, the only option is to have $L_r = N_r$. This implies that for any $N_r > L_r$, which is the natural case in any HAD beamforming architecture, we always have $\alpha^{\text{A2}} < 1$, i.e., the partially-connected architecture A2 always has lower circuitry power consumption than the FD architecture. Thus, we can simplify (10) as $K \geq \lceil \frac{6N_t}{9L_r} \rceil$. For example, if we assume $N_r = L_r = 1$ and $N_t = 8$, then if $K \geq 6$, we have $\alpha^{\text{A2}} \geq 1$. However, if we assume $N_t = 16$, then if $K \geq 11$ we have $\alpha^{\text{A2}} \geq 1$. Please refer to Fig. 2, where we show the circuitry power consumption ratio α versus N_t and K for different N_r and L_r values. From Fig. 2, it is clear that, for a given system setup, when the number of users exceeds a certain threshold, given by (9) or (10), the FD architecture consumes less circuitry power than the HAD architectures A1 and A2.

To better understand the impact of circuitry power consumption P_c on the relationship between SE and EE, Fig. 3 shows SE versus EE for different P_c values. From Fig. 3, we can see that when the circuitry power consumption is not taken into account

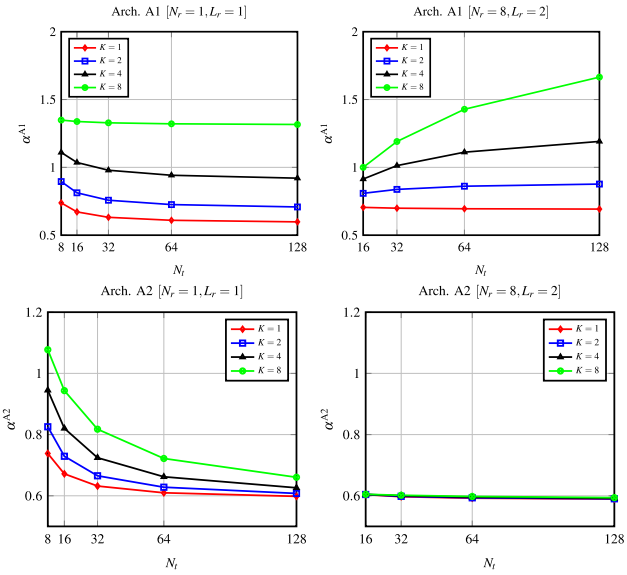


Fig. 2. Circuitry power ratio $\alpha^{\{\text{A1,A2}\}}$ vs. number of transmit antennas N_t and number of users K.

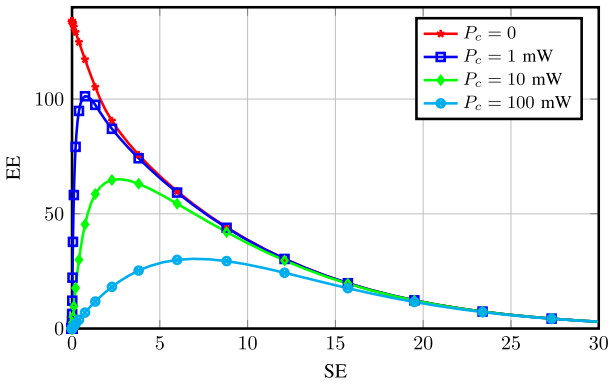


Fig. 3. SE vs. EE for different P_c values considering a system with $[K, N_t, N_r, N_s] = [2, 64, 4, 2]$. The beamforming matrices are updated using the FD BD approach [8], [23] for a range of signal-to-noise ratio (SNR) values, where the SE and EE are given by (12) and (13), respectively.

($P_c = 0$), there is always an inverse relationship between the system's SE and EE. However, for the nonzero P_c scenarios, we can observe that the maximum EE decreases with an increasing P_c value, where the EE increases in the low SE region and decreases in the high SE region.

From the above results, we can conclude that architecture A1 is less energy efficient than the FD architecture when the number of users exceeds a certain threshold, given by (9), which depends on the number of transmit and receive antennas and RF chains that are employed by each user. In contrast, A2 seems a promising architecture to increase the EE, mainly due to its very low circuitry power consumption compared with both the FD and the A1 architectures.

III. Problem Formulation

We consider a narrow-band block-fading propagation channel, where the received signal $\mathbf{y}_k$ at the k-th MS is given as

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{T}_k \mathbf{s}_k + \sum_{j \neq k} \mathbf{H}_k \mathbf{T}_j \mathbf{s}_j + \mathbf{n}_k \in \mathbb{C}^{N_r}, \tag{11}$$

where $\mathbf{H}_k \in \mathbb{C}^{N_r \times N_t}$ is the MIMO channel matrix between the BS and the k-th MS, such that $\mathbb{E}[\|\mathbf{H}_k\|_F^2] = N_r N_t$, $\mathbf{s}_k \in \mathbb{C}^{N_s}$ is the transmitted data vector with $\mathbb{E}[\mathbf{s}_k \mathbf{s}_k^{\mathrm{H}}] = \mathbf{I}_{N_s}$, and $\mathbf{n}_k \in \mathbb{C}^{N_r}$ is the additive white Gaussian noise with variance σ^2.

Throughout this paper, we assume that the BS and each MS has perfect channel state information (CSI). Although perfect CSI cannot be acquired in practice, several high-resolution CSI estimation methods have been proposed in the literature, e.g., the CSI estimation based on the least-square methods proposed in [28]. However, such methods are impractical under massive MIMO setups, since they would entail large training overhead. To avoid such a large overhead, recent CSI estimation solutions [29]–[31] exploit the sparse (or low rank) structure in the angular domain of massive MIMO channels, which appears due to the small number of scatterers compared to the number of antennas, especially in millimeter-wave bands. Exploiting this sparse structure, compressed sensing (CS) tools [32] can be used to estimate the MIMO channel, where the problem can be turned into estimating the parameters of dominant channel paths, namely the angles-of-departure, the angles-of-arrival, and the complex path gains. Using such methods, the pilot overhead can be significantly reduced, while still achieving a high resolution CSI estimation. However, the CSI estimation problem is out of the scope of this paper and we refer to [28]–[31] for more details.

Assuming Gaussian signaling and single-user detection, where the interference is treated as additional noise, the SE of MS k can be written as [4]

$$r_k = \log\det(\mathbf{I}_{N_s} + \mathbf{R}_k^{\mathrm{H}} \mathbf{H}_k \mathbf{T}_k \mathbf{T}_k^{\mathrm{H}} \mathbf{H}_k^{\mathrm{H}} \mathbf{R}_k \mathbf{\Psi}_k^{-1}), \tag{12}$$

where $\mathbf{\Psi}_k = \mathbf{R}_k^{\mathrm{H}} \big(\sum_{j \neq k} \mathbf{H}_k \mathbf{T}_j \mathbf{T}_j^{\mathrm{H}} \mathbf{H}_k^{\mathrm{H}} + \sigma^2 \mathbf{I}_{N_r} \big) \mathbf{R}_k$ denotes the residual inter-user interference plus noise. The system's EE is then defined as

$$\tau = \frac{\sum_k r_k}{P_{\text{tot}}} = \frac{\sum_k r_k}{\sum_k \|\mathbf{F}\mathbf{B}_k\|_{\mathrm{F}}^2 + P_{\mathrm{c}}^X}, \tag{13}$$

where $P_{\text{tot}} = \sum_k \|\mathbf{F}\mathbf{B}_k\|_{\mathrm{F}}^2 + P_{\mathrm{c}}^X$ denotes the total power consumption and $X \in \{\mathrm{A1}, \mathrm{A2}\}$. In this paper, our objective is to design the DBF and the ABF parts of $\mathbf{T}_k$ and $\mathbf{R}_k, \forall k$, such that the system's EE given by (13) is maximized, i.e., we consider the following optimization problem

$$\begin{aligned} \max_{\{\mathbf{F}, \mathbf{W}_k, \mathbf{B}_k, \mathbf{D}_k, \forall k\}} \quad & \tau = \frac{\sum_k r_k}{\sum_k \|\mathbf{F}\mathbf{B}_k\|_{\mathrm{F}}^2 + P_{\mathrm{c}}^X} \\ \text{s.t.} \quad & \mathbf{F} \in \mathscr{F}, \\ & \mathbf{W}_k \in \mathscr{W}, \forall k, \\ & \sum_k \|\mathbf{F}\mathbf{B}_k\|_{\mathrm{F}}^2 \leq P_{\max}, \end{aligned} \tag{14}$$

where $P_{\max}$ is the maximum allowed transmit power, $\mathscr{F}$ and $\mathscr{W}$ denote the sets with all possible analog beamformers satisfying the constraints associated with the considered ABF architecture: A1 or A2.

Problem (14) is a fractional optimization problem, which is nonconvex and NP-hard [33]. The major difficulty comes from the joint optimization of the DBF and ABF parts and the nonconvex constraints in $\mathbf{F} \in \mathscr{F}$ and $\mathbf{W}_k \in \mathscr{W}$. In the following, we relax the joint optimization and decouple the optimization of the DBF and the ABF parts by treating them separately.

IV. Analog Beamforming Design

In this section, we design the ABF parts $\mathbf{F}$ and $\mathbf{W}_k, \forall k$, without considering the DBF parts $\mathbf{B}_k$ and $\mathbf{D}_k, \forall k$. By removing $\mathbf{B}_k$ and $\mathbf{D}_k, \forall k$ from problem (14) and noting that $\|\mathbf{F}\|_{\mathrm{F}}^2$ is a constant, since we impose the constraint of $\mathbf{F} \in \mathscr{F}$ i.e., $|[\mathbf{F}]_{[i,j]}| = 1, \forall i, j$, then the constraint function $\sum_k \|\mathbf{F}\|_{\mathrm{F}}^2 \leq P_{\max}$ becomes inactive and can be removed. Further, the term $\|\mathbf{F}\|_{\mathrm{F}}^2 + P_{\mathrm{c}}^X$ is also a constant and can thus be removed, since multiplying the objective function by a constant does not change the obtained solutions. Therefore, after removing the irrelevant

constant terms, problem (14) reduces to the following fully-analog SE maximization problem

$$\begin{aligned} \max_{\{\mathbf{F},\mathbf{W}_k,\forall k\}} \quad & \sum_k \tilde{r}_k \\ \text{s.t.} \quad & \mathbf{F} \in \mathscr{F}, \\ & \mathbf{W}_k \in \mathscr{W}, \forall k, \end{aligned} \tag{15}$$

where $\tilde{r}_k = \log\det(\mathbf{I}_{L_r} + \mathbf{W}_k^{\mathrm{H}}\mathbf{H}_k\mathbf{F}\mathbf{F}^{\mathrm{H}}\mathbf{H}_k^{\mathrm{H}}\mathbf{W}_k\tilde{\mathbf{\Psi}}_k^{-1})$ and $\tilde{\mathbf{\Psi}}_k = \mathbf{W}_k^{\mathrm{H}}\big(\sum_{j\neq k}\mathbf{H}_k\mathbf{F}\mathbf{F}^{\mathrm{H}}\mathbf{H}_k^{\mathrm{H}} + \sigma^2\mathbf{I}_{N_r}\big)\mathbf{W}_k$. Problem (15) is still nonconvex and NP-hard. Note that if we neglect the constant modulus constraints, i.e., $\mathbf{F} \in \mathscr{F}$ and $\mathbf{W}_k \in \mathscr{W}, \forall k$, a solution to problem (15) can be obtained by using its relationship to the weighted mean-square-error minimization problem, as shown in [34], or alternatively using one of the proposed algorithms in [24]. However, the constant modulus constraints make all the aforementioned solutions unsuitable, since a new set of constant modulus constraints must be satisfied, and thus a new solution approach is required.

We assume that the transmit DBF matrices $\mathbf{B}_k, \forall k$, are updated afterwards using the BD method [23], [35] from the resulting equivalent channels $\mathbf{W}_k^{\mathrm{H}}\mathbf{H}_k\mathbf{F}, \forall k$. This implies that for any given ABF parts $\mathbf{F}$ and $\mathbf{W}_k, \forall k$, the terms $\mathbf{H}_k\mathbf{F}\mathbf{B}_j\mathbf{B}_j^{\mathrm{H}}\mathbf{F}^{\mathrm{H}}\mathbf{H}_k^{\mathrm{H}} = \mathbf{0}, \forall j \neq k$ are always satisfied. Therefore, the inter-user interference plus noise term $\tilde{\mathbf{\Psi}}_k$ can be simplified and written as $\tilde{\mathbf{\Psi}}_k = \sigma^2\mathbf{W}_k^{\mathrm{H}}\mathbf{W}_k$. Note that with architecture A2, we have $\mathbf{W}_k^{\mathrm{H}}\mathbf{W}_k = N_r\mathbf{I}_{L_r}$ due to the block-diagonal structure of $\mathbf{W}_k$. However, with architecture A1, we have $\mathbf{W}_k^{\mathrm{H}}\mathbf{W}_k \approx N_r\mathbf{I}_{L_r}$ with high probability for large N_r [36]. Please note that, in the millimeter wave communication, large N_r is practically feasible, thanks to their very short wavelength, where a large number of antennas can be easily installed in a small physical area [3]. To this end, problem (15) can be simplified and written as

$$\begin{aligned} \max_{\mathbf{F},\mathbf{W}} \quad & \log\det(\mathbf{I}_{KL_r} + \mathbf{W}^{\mathrm{H}}\mathbf{H}\mathbf{F}\mathbf{F}^{\mathrm{H}}\mathbf{H}^{\mathrm{H}}\mathbf{W}), \\ \text{s.t.} \quad & \mathbf{F} \in \mathscr{F}, \\ & \mathbf{W} \in \mathscr{W}, \end{aligned} \tag{16}$$

where $\mathbf{H} = [\mathbf{H}_1^{\mathrm{T}}, \ldots, \mathbf{H}_K^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{KN_r\times N_t}$ and $\mathbf{W} = \mathrm{Bdiag}\{\mathbf{W}_1, \ldots, \mathbf{W}_K\} \in \mathbb{C}^{KN_r\times KL_r}$. Problem (16) is still nonconvex and NP-hard, due to the joint optimization between $\mathbf{F}$ and $\mathbf{W}$ and their constant modulus constraints. In the following, we relax the joint optimization by decoupling the optimization variables and update them in two stages: in the first stage, we update $\mathbf{F}$ for fixed $\mathbf{W}$, while in the second stage, we update $\mathbf{W}$ for fixed $\mathbf{F}$.

A. First Stage: Updating $\mathbf{F}$

When $\mathbf{W}$ is fixed, problem (16) simplifies to

$$\begin{aligned} \max_{\mathbf{F}} \quad & \zeta_t = \log\det(\mathbf{I}_{KL_r} + \mathbf{W}^{\mathrm{H}}\mathbf{H}\mathbf{F}\mathbf{F}^{\mathrm{H}}\mathbf{H}^{\mathrm{H}}\mathbf{W}), \\ \text{s.t.} \quad & \mathbf{F} \in \mathscr{F}. \end{aligned} \tag{17}$$

Let $\widetilde{\mathbf{H}} = \mathbf{W}^{\mathrm{H}}\mathbf{H} \in \mathbb{C}^{KL_r\times N_t}$. Then we can write the objective function of problem (17) as [5]

$$\zeta_t = \log\det(\mathbf{E}_j) + \log(1 + \mathbf{f}_j^{\mathrm{H}}\widetilde{\mathbf{H}}^{\mathrm{H}}\mathbf{E}_j^{-1}\widetilde{\mathbf{H}}\mathbf{f}_j), \tag{18}$$

where $\mathbf{E}_j = \mathbf{I}_{KL_r} + \widetilde{\mathbf{H}}\bar{\mathbf{F}}_j\bar{\mathbf{F}}_j^{\mathrm{H}}\widetilde{\mathbf{H}}^{\mathrm{H}} \in \mathbb{C}^{KL_r\times KL_r}$ and $\bar{\mathbf{F}}_j \in \mathbb{C}^{N_t\times L_t-1}$ is a sub-matrix of $\mathbf{F}$ after removing its j-th column $\mathbf{f}_j$. Observing the first term in the right-hand-side (RHS) of (18), i.e., $\log\det(\mathbf{E}_j)$, we can note that it has the same structure as the objective function ζ_t. Thus, it can also be written in a similar method as in (18). In summary, the objective function of problem (17) can be written as a series of $\log(1 + x_j), j = 1, \ldots, L_t$, functions as

$$\zeta_t = \log(1 + x_1) + \cdots + \log(1 + x_{L_t}), \tag{19}$$

where

$$x_j = \mathbf{f}_j^{\mathrm{H}}\widetilde{\mathbf{H}}^{\mathrm{H}}\mathbf{E}_j^{-1}\widetilde{\mathbf{H}}\mathbf{f}_j \in \mathbb{C}, \tag{20}$$

$$\mathbf{E}_j = \mathbf{I}_{KL_r} + \widetilde{\mathbf{H}}\hat{\mathbf{F}}_j\hat{\mathbf{F}}_j^{\mathrm{H}}\widetilde{\mathbf{H}}^{\mathrm{H}} \in \mathbb{C}^{KL_r\times KL_r}, \tag{21}$$

for which $\hat{\mathbf{F}}_j = [\mathbf{f}_1, \ldots, \mathbf{f}_{j-1}]$, i.e., $\hat{\mathbf{F}}_j$ is the matrix that holds the first $i < j$ columns from $\mathbf{F}$. Note that, when $j = 1$, $\hat{\mathbf{F}}_j$ is an empty matrix, which implies that $\mathbf{E}_1 = \mathbf{I}_{KL_r}$.

Writing the objective function ζ_t as given by (19) suggests that problem (17) can be solved sequentially starting from updating the first column $\mathbf{f}_1$ that maximizes $\log(1 + x_1)$ until the last column $\mathbf{f}_{L_t}$ that maximizes $\log(1 + x_{L_t})$. In other words, at the j-th step, problem (17) simplifies to

$$\begin{aligned} \max_{\mathbf{f}_j} \quad & \log(1 + x_j) = \log(1 + \mathbf{f}_j^{\mathrm{H}}\mathbf{G}_j\mathbf{f}_j) \\ \text{s.t.} \quad & \mathbf{f}_j \in \mathscr{F}, \end{aligned} \tag{22}$$

where $\mathbf{G}_j = \widetilde{\mathbf{H}}^{\mathrm{H}}\mathbf{E}_j^{-1}\widetilde{\mathbf{H}} \in \mathbb{C}^{N_t\times N_t}$. Considering the high SNR regime,[1] where $\log(1 + \mathbf{f}_j^{\mathrm{H}}\mathbf{G}_j\mathbf{f}_j) \approx \log(\mathbf{f}_j^{\mathrm{H}}\mathbf{G}_j\mathbf{f}_j)$, we have shown in [8] that the i-th element of $\mathbf{f}_j$, ($[\mathbf{f}_j]_{[i]}$), can be optimally updated from the phase-angle of the product between the i-th row of $\mathbf{G}_j$ ($[\mathbf{G}_j]_{[i,:]}$) and vector $\mathbf{f}_j$, i.e., $[\mathbf{f}_j]_{[i]}$ is updated as

$$[\mathbf{f}_j]_{[i]} = \psi\big([\mathbf{G}_j]_{[i,:]}\mathbf{f}_j\big), \tag{23}$$

where $\psi(z) = \frac{z}{\|z\|}$. Note that with architecture A1, each vector $\mathbf{f}_j$ has N_t nonzero elements, while with architecture A2, each vector $\mathbf{f}_j$ has $\frac{N_t}{L_t}$ nonzero elements.

B. Second Stage: Updating $\mathbf{W}$

When $\mathbf{F}$ is fixed, problem (16) simplifies to

$$\begin{aligned} \max_{\mathbf{W}} \quad & \zeta_r = \log\det(\mathbf{I}_{L_t} + \mathbf{F}^{\mathrm{H}}\mathbf{H}^{\mathrm{H}}\mathbf{W}\mathbf{W}^{\mathrm{H}}\mathbf{H}\mathbf{F}), \\ \text{s.t.} \quad & \mathbf{W} \in \mathscr{W}, \end{aligned} \tag{24}$$

where we have used the property of $\log\det(\mathbf{I} + \mathbf{X}\mathbf{Y}) = \log\det(\mathbf{I} + \mathbf{Y}\mathbf{X})$. Comparing (24) to (17) we can see that both have the same structure, and thus the above formulation can be applied directly to update $\mathbf{W}$. Let $\widehat{\mathbf{H}} = \mathbf{F}^{\mathrm{H}}\mathbf{H}^{\mathrm{H}} \in \mathbb{C}^{L_t\times KL_r}$.

[1] Please note that in the data transmission phase, the assumption of high SNR is feasible due to the use of a large number of antennas. In this case, not only we have large antenna gain, but also the inter-user interference is hugely reduced due to the resulting narrow beams. It is worth pointing out that the assumption of high SNR in a CSI estimation problem is not feasible, since the SNR is normally low before the beamforming [3].

Then, the objective function ζ_r can be written as a series of $\log(1+y_j), j=1,\ldots,KL_r$, functions as

$$\zeta_r = \log(1+y_1)+\cdots+\log(1+y_{KL_r}), \tag{25}$$

where

$$y_j = \mathbf{w}_j^{\mathrm{H}}\widehat{\mathbf{H}}^{\mathrm{H}}\mathbf{S}_j^{-1}\widehat{\mathbf{H}}\mathbf{w}_j \in \mathbb{C}, \tag{26}$$

$$\mathbf{S}_j = \mathbf{I}_{L_t} + \widehat{\mathbf{H}}\hat{\mathbf{W}}_j\hat{\mathbf{W}}_j^{\mathrm{H}}\widehat{\mathbf{H}}^{\mathrm{H}} \in \mathbb{C}^{L_t\times L_t}, \tag{27}$$

for which $\hat{\mathbf{W}}_j$ is formed as $\hat{\mathbf{F}}$ above, i.e., $\hat{\mathbf{W}}_j=[\mathbf{w}_1,\ldots,\mathbf{w}_{j-1}]$ and $\mathbf{S}_1=\mathbf{I}_{L_t}$. Let $\mathbf{C}_j=\widehat{\mathbf{H}}^{\mathrm{H}}\mathbf{S}_j^{-1}\widehat{\mathbf{H}}\in\mathbb{C}^{KL_r\times KL_r}$. Then, the i-th element of $\mathbf{w}_j$, i.e. $[\mathbf{w}_j]_{[i]}$ can be optimally updated as

$$[\mathbf{w}_j]_{[i]} = \psi\big([\mathbf{C}_j]_{[i,:]}\mathbf{w}_j\big). \tag{28}$$

Note that with architecture A1, each vector $\mathbf{w}_j$ has N_r nonzero elements, while with architecture A2, each vector $\mathbf{w}_j$ has $\frac{N_r}{L_r}$ nonzero elements.

C. Proposed ABF Update Algorithm

Algorithm 1 summarizes the solution steps to update the ABF matrices $\mathbf{F}$ and $\mathbf{W}$. In Algorithm 1, we define the function $\Theta(N,M)$ as an $[N\times M]$-matrix initialization function where each nonzero element has unit modulus. Further, to take the different architectures into account, we define the binary matrices $\Xi_t\in\mathbb{Z}^{N_t\times L_t}$ and $\Xi_r\in\mathbb{Z}^{N_r\times KL_r}$, such that the $[i,j]$-th element is equal to one if the i-th antenna is connected to the j-th RF chain and zero otherwise. The convergence proof of Algorithm 1 is given by the following proposition.

Proposition 1: Algorithm 1 convergences monotonically to a local stationary point and the solution is achieved at $\zeta_t^\star=\zeta_r^\star$.

Proof: The first part of the proposition can be proved by noting that the element-wise updates in steps 12 and 25 are optimal [8]. This means that the cost functions $\log(1+x_j^{(\iota+1)})$ and $\log(1+y_i^{(\iota+1)})$ are non-decreasing functions, i.e., we always have

$$\log(1+x_j^{(\iota+1)}) \geq \log(1+x_j^{(\iota)}), \forall j$$
$$\log(1+y_i^{(\iota+1)}) \geq \log(1+y_j^{(\iota)}), \forall i.$$

This implies that $\zeta_t^{(\iota+1)}\geq\zeta_t^{(\iota)}$ and $\zeta_r^{(\iota+1)}\geq\zeta_r^{(\iota)}$, which proves that Algorithm 1 is guaranteed to converge monotonically to, at least, a local stationary point. However, the convergence to the global optimal point cannot be guaranteed, due to the non-convexity of the original problem. The second part of the proposition is straightforward from the property of $\log\det(\mathbf{I}+\mathbf{XY})=\log\det(\mathbf{I}+\mathbf{YX})$. ■

Algorithm 1: Alternating Optimization Method for Updating the Analog Beamforming (ARAB).

1: Input: $\mathbf{H}_k,\forall k$.
2: Output: $\mathbf{F}$ and $\mathbf{W}$.
3: Initialize $\mathbf{F}^{(0)}=\Theta(N_t,L_t)$ and $\mathbf{W}^{(0)}=\Theta(KN_r,KL_r)$
4: Set $\iota=1$
5: **while** not converged **do**

Stage 1: Updating $\mathbf{F}$

6: Compute $\widetilde{\mathbf{H}}^{(\iota)}=(\mathbf{W}^{(\iota)})^{\mathrm{H}}\mathbf{H}$
7: Set $\mathbf{E}_1=\mathbf{I}_{KL_r}$ and $\hat{\mathbf{F}}_j=\emptyset$
8: **for** $j=1$ to L_t **do**
9: Compute $\mathbf{G}_j^{(\iota)}=(\widetilde{\mathbf{H}}^{(\iota)})^{\mathrm{H}}\mathbf{E}_j^{-1}\widetilde{\mathbf{H}}^{(\iota)}$
10: **while** not converged **do**
11: **for** $i=1$ to N_t and $[\Xi_t]_{[i,j]}=1$ **do**
12: Update $[\mathbf{f}_j^{(\iota+1)}]_{[i]}=\psi([\mathbf{G}_j^{(\iota)}]_{[i,:]}\mathbf{f}_j^{(\iota)})$
13: **end for**
14: **end while**
15: Set $\hat{\mathbf{F}}_j=[\hat{\mathbf{F}}_j\cup\mathbf{f}_j^{(\iota+1)}]$
16: Update $\mathbf{E}_j=\mathbf{I}_{KL_r}+\widetilde{\mathbf{H}}^{(\iota)}\hat{\mathbf{F}}_j\hat{\mathbf{F}}_j^{\mathrm{H}}(\widetilde{\mathbf{H}}^{(\iota)})^{\mathrm{H}}$
17: **end for**
18: Set $\mathbf{F}^{(\iota+1)}=\hat{\mathbf{F}}_j$ and go forward to step 19.

Stage 2: Updating $\mathbf{W}$

19: Compute $\widehat{\mathbf{H}}^{(\iota)}=(\mathbf{F}^{(\iota+1)})^{\mathrm{H}}\mathbf{H}^{\mathrm{H}}$
20: Set $\mathbf{S}_1=\mathbf{I}_{L_t}$ and $\hat{\mathbf{W}}_j=\emptyset$
21: **for** $j=1$ to KL_r **do**
22: Compute $\mathbf{C}_j^{(\iota)}=(\widehat{\mathbf{H}}^{(\iota)})^{\mathrm{H}}\mathbf{S}_j^{-1}\widehat{\mathbf{H}}^{(\iota)}$
23: **while** not converged and **do**
24: **for** $i=1$ to N_r and $[\Xi_r]_{[i,j]}=1$ **do**
25: Update $[\mathbf{w}_j^{(\iota+1)}]_{[i]}=\psi([\mathbf{C}_j^{(\iota)}]_{[i,:]}\mathbf{w}_j^{(\iota)})$
26: **end for**
27: **end while**
28: Set $\hat{\mathbf{W}}_j=[\hat{\mathbf{W}}_j\cup\mathbf{w}_j^{(\iota+1)}]$
29: Update $\mathbf{S}_j=\mathbf{I}_{L_t}+\widehat{\mathbf{H}}^{(\iota)}\hat{\mathbf{W}}_j\hat{\mathbf{W}}_j^{\mathrm{H}}(\widehat{\mathbf{H}}^{(\iota)})^{\mathrm{H}}$
30: **end for**
31: Set $\mathbf{W}^{(\iota+1)}=\hat{\mathbf{W}}_j$ and go back to step 6.
32: **end while**

V. Digital Beamforming Design

For given and fixed ABF matrices $\mathbf{F}$ and $\mathbf{W}_k,\forall k$, the main task in this section is to design $\mathbf{B}_k$ and $\mathbf{D}_k,\forall k$, to maximize the system's EE. We assume that each MS k applies the noise whitening filter $\mathbf{Q}_k=(\mathbf{W}_k^{\mathrm{H}}\mathbf{W}_k)^{-\frac{1}{2}}\in\mathbb{C}^{L_r\times L_r}$ at the received signal after the ABF combining. Thus, the received baseband signal at the kth MS is given as

$$\tilde{\mathbf{y}}_k = \tilde{\mathbf{H}}_k\mathbf{B}_k\mathbf{s}_k + \sum_{j\neq k}\tilde{\mathbf{H}}_k\mathbf{B}_j\mathbf{s}_j + \mathbf{Q}_k^{\mathrm{H}}\mathbf{W}_k^{\mathrm{H}}\mathbf{n}_k, \tag{29}$$

where $\tilde{\mathbf{H}}_k=\mathbf{Q}_k^{\mathrm{H}}\mathbf{W}_k^{\mathrm{H}}\mathbf{H}_k\mathbf{F}\in\mathbb{C}^{L_r\times L_t}$. Further, we assume that MS k updates its receive DBF matrix $\mathbf{D}_k$ using the minimum mean-square-error method as [24]

$$\begin{aligned}\mathbf{D}_k &= \arg\min_{\mathbf{D}_k}\mathbb{E}[\|\mathbf{D}_k\tilde{\mathbf{y}}_k-\mathbf{s}_k\|^2]\\ &= (\tilde{\mathbf{H}}_k\mathbf{B}_k\mathbf{B}_k^{\mathrm{H}}\tilde{\mathbf{H}}_k^{\mathrm{H}}+\mathbf{\Phi}_k)^{-1}\tilde{\mathbf{H}}_k\mathbf{B}_k,\end{aligned} \tag{30}$$

where $\mathbf{\Phi}_k = \sum_{j\neq k} \tilde{\mathbf{H}}_k \mathbf{B}_j \mathbf{B}_j^{\mathrm{H}} \tilde{\mathbf{H}}_k^{\mathrm{H}} + \sigma^2 \mathbf{I}_{L_r}$. As a result, the SE function in (12) can be written as

$$r_k = \log\det(\mathbf{I}_{N_s} + \mathbf{B}_k^{\mathrm{H}} \tilde{\mathbf{H}}_k^{\mathrm{H}} \mathbf{\Phi}_k^{-1} \tilde{\mathbf{H}}_k \mathbf{B}_k), \tag{31}$$

and problem (14) simplifies to

$$\begin{aligned} \max_{\{\mathbf{B}_k, \forall k\}} \quad & \tau = \frac{\sum_k r_k}{P_{\mathrm{tot}}}, \\ \text{s.t.} \quad & \sum_k \|\mathbf{F}\mathbf{B}_k\|_{\mathrm{F}}^2 \leq P_{\max}. \end{aligned} \tag{32}$$

According to [33, Theorem 1] on nonlinear fractional programming, problem (32) can be equivalently transformed into a parameterized subtractive form by introducing an auxiliary variable as

$$\begin{aligned} \max_{\{\mathbf{B}_k, \forall k\}} \quad & \sum_k r_k - \mathring{\tau} P_{\mathrm{tot}}, \\ \text{s.t.} \quad & \sum_k \|\mathbf{F}\mathbf{B}_k\|_{\mathrm{F}}^2 \leq P_{\max}. \end{aligned} \tag{33}$$

The existing research on fractional programming problems has shown that solving the above problem is equivalent to looking for a solution to problem (32) such that its objective equals zero, i.e., $\sum_k r_k^* - \mathring{\tau}^* P_{\mathrm{tot}}^* = 0$.

A solution to problem (32) w.r.t the transmit DBF matrices $\mathbf{B}_k, \forall k$, can be found iteratively by investigating its Karush-Kuhn-Tucker conditions [37] as the authors in [11] have followed. However, we note that in massive MIMO settings, the classical solutions, like the BD approach, can provide very similar performance with a much lower complexity. Therefore, differently from [11], we assume that the transmit DBF matrices are given by the well-known BD approach. More precisely, the DBF matrix of MS k is given as

$$\mathbf{B}_k = \mathbf{Z}_k \mathbf{V}_k \mathbf{P}_k, \tag{34}$$

where $\mathbf{Z}_k$ and $\mathbf{V}_k$ define the transmit beamforming direction and $\mathbf{P}_k = \mathrm{diag}\{\sqrt{p_{k,1}}, \ldots, \sqrt{p_{k,L_r}}\}$ is a diagonal matrix holding the power allocations of the L_r data streams (recalling that $L_r = N_s$). In (34), $\mathbf{Z}_k \in \mathbb{C}^{L_t \times L_r}$ holds the nullspace orthonormal vectors of $\breve{\mathbf{H}}_k$, which collects all the users' equivalent channels except user k, i.e.,

$$\breve{\mathbf{H}}_k = [\tilde{\mathbf{H}}_1^{\mathrm{T}}, \ldots, \tilde{\mathbf{H}}_{k-1}^{\mathrm{T}}, \tilde{\mathbf{H}}_{k+1}^{\mathrm{T}}, \ldots, \tilde{\mathbf{H}}_K^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{(K-1)L_r \times L_t}. \tag{35}$$

Meanwhile, $\mathbf{V}_k$ holds the L_r dominant right singular vectors of MS k effective channel

$$\mathbf{H}_k^e = \tilde{\mathbf{H}}_k \mathbf{Z}_k = \mathbf{U}_k \Lambda_k \mathbf{V}_k^{\mathrm{H}} \in \mathbb{C}^{L_r \times L_r}, \tag{36}$$

where $\mathbf{\Lambda}_k = \mathrm{diag}\{\lambda_{k,1}, \ldots, \lambda_{k,L_r}\}$ is the diagonal matrix holding the L_r singular values arranged in a decreasing order, $\mathbf{U}_k \in \mathbb{C}^{L_r \times L_r}$ and $\mathbf{V}_k \in \mathbb{C}^{L_r \times L_r}$ are the left and right singular vectors, respectively.

With the transmit beamforming directions calculated as above, problem (33) simplifies to a power allocation problem

Algorithm 2: Proposed Power Allocation Method (ALG2).

1: Input: $\lambda_{k,\ell}, \forall k, \ell$, $\mathring{\tau}^{(0)} = 0$, $P_{\max}, P_c$
2: **while** not converged **do**
3: update $p_{k,s}^{(\iota)}, \forall k, \ell$ for given $\mathring{\tau}^{(\iota)}$ using (39)
4: update $\mathring{\tau}^{(\iota)} = \sum_k \sum_\ell r_{k,\ell}^{\mathrm{BD}(t)} / \sum_k \sum_\ell p_{k,\ell}^{(\iota)} + P_{\mathrm{c}}$
5: **end while**

that is given as

$$\begin{aligned} \max_{\{p_{k,\ell}\}} \quad & \sum_k \sum_\ell r_{k,\ell}^{\mathrm{BD}} - \mathring{\tau} P_{\mathrm{tot}}, \\ \text{s.t.} \quad & \sum_k \sum_\ell p_{k,\ell} \leq P_{\max}, \end{aligned} \tag{37}$$

where $\ell \in \{1, \ldots, L_r\}$ and $r_{k,\ell}^{\mathrm{BD}}$ is given as

$$r_{k,\ell}^{\mathrm{BD}} = \log(1 + \frac{1}{\sigma^2} \lambda_{k,\ell}^2 p_{k,\ell}). \tag{38}$$

Investigating the Karush-Kuhn-Tucker (KKT) conditions [37] of problem (37), the optimal power allocation $p_{k,\ell}$ of the ℓ-th stream of user k is given as

$$p_{k,\ell} = \max\left[0, \frac{1}{\ln(2)(\tau + \mu)} - \frac{\sigma^2}{\lambda_{k,\ell}^2}\right]^+, \tag{39}$$

where μ is the Lagrangian multiplier associated with the constraint of problem (37), which can be calculated, e.g. using the bi-section method, such that $\mu(\sum_k \sum_\ell p_{k,\ell} - P_{\max}) = 0$.

Algorithm 2 summarizes the BD-based method for the EE-Max approach, which is guaranteed to converge to the optimal solution [12]. Note that, in the case of SE-Max approach, the same algorithm can be used by omitting step 4, thus, reducing Algorithm 2 to the classical water-filling method [38].

Remark 1: For the FD approach, $\mathbf{B}_k$ can be computed exactly in the same way by assuming $\mathbf{F} = \mathbf{I}_{N_t}$ and $\mathbf{W}_k = \mathbf{I}_{N_r}, \forall k$. In this case, (35) can be rewritten as

$$\breve{\mathbf{H}}_k = [\mathbf{H}_1^{\mathrm{T}}, \ldots, \mathbf{H}_{k-1}^{\mathrm{T}}, \mathbf{H}_{k+1}^{\mathrm{T}}, \ldots, \mathbf{H}_K^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{(K-1)N_r \times N_t}. \tag{40}$$

Observing the latter equation, we note that the condition $N_t \geq (K-1)N_r + N_s$ should be satisfied in order to have at least N_s vectors in the nullspace of $\breve{\mathbf{H}}_k$. To relax this condition, we resort to a partial BD approach by requiring that the $\mathbf{V}_k$ (i.e., the L_r dominant right singular vectors of the effective channel $\mathbf{H}_k^e$ of MS k) be orthogonal to the dominant L_r left singular vectors of the channels $\mathbf{H}_j, \forall j \neq k$, i.e., the nullspace $\mathbf{Z}_k$ with the FD approach is calculated from

$$\breve{\mathbf{H}}_k = [\bar{\mathbf{H}}_1^{\mathrm{T}}, \ldots, \bar{\mathbf{H}}_{k-1}^{\mathrm{T}}, \bar{\mathbf{H}}_{k+1}^{\mathrm{T}}, \ldots, \bar{\mathbf{H}}_K^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{(K-1)L_r \times N_t}, \tag{41}$$

where $\bar{\mathbf{H}}_j = \mathbf{U}_j^{\mathrm{H}} \mathbf{H}_j \in \mathbb{C}^{L_r \times N_t}$ and $\mathbf{U}_j \in \mathbb{C}^{N_r \times L_r}$ holds the dominant L_r left singular vectors of the channel matrix $\mathbf{H}_j$ of MS j. In this way, each MS beamforming matrix $\mathbf{B}_k$ is orthogonal to the $L_r(K-1)$-dimensional subspace and nulls the most significant part of the interference. Note that in this case we have $\mathbf{Z}_k \in \mathbb{C}^{N_t \times N_t - (K-1)L_r}$ and $\mathbf{H}_k^e = \bar{\mathbf{H}}_k \mathbf{Z}_k \in \mathbb{C}^{L_r \times N_t - (K-1)L_r}$.

TABLE I
COMPUTATIONAL ANALYSIS OF THE ARAB AND ALG. [5]

Notation	Process	Step	ARAB Complexity	Alg. [5] Complexity
c_1	Computing $\widetilde{\mathbf{H}}$	6	$2K^2L_rN_rN_t$	-
c_2	Computing $\mathbf{G}_j$	9	$2N_tK^2L_r^2+2KN_t^2L_r+\frac{2}{3}K^3L_r^3$	$2N_tK^2N_r^2+2KN_t^2N_r+\frac{2}{3}K^3N_r^3$
c_3	Computing $\mathbf{f}_j, \forall j$	10-14	$T_1\cdot(2N_t^3)$	$11N_t^3$
c_4	Computing $\mathbf{E}_j$	16	$4KL_rL_tN_t+2K^2L_r^2N_t$	$4KN_rL_tN_t+2K^2N_r^2N_t$
c_5	Computing $\widehat{\mathbf{H}}$	19	$2L_tN_tKL_r$	-
c_6	Computing $\mathbf{G}_m$	22	$2KN_rL_t^2+2K^2N_r^2L_t+\frac{2}{3}L_t^3$	$2KN_rN_t^2+2K^2N_r^2N_t+\frac{2}{3}N_t^3$
c_7	Computing $\mathbf{w}_j, \forall j$	23-27	$T_2\cdot(2K^2N_r^3)$	$11K^3N_r^3$
c_8	Computing $\mathbf{G}_m$	29	$4L_tL_rK^2N_r+2KL_t^2N_r$	$4N_tL_rK^2N_r+2KN_t^2N_r$
-	Computing $\mathbf{F}$	6 to 17	$T_0\cdot(c_1+L_t(c_2+c_3+c_4))$	$L_t(c_2+c_3+c_4)$
-	Computing $\mathbf{W}$	19 to 30	$T_0\cdot(c_5+KL_r(c_6+c_7+c_8))$	$KL_r(c_6+c_7+c_8)$

TABLE II
COMPUTATIONAL ANALYSIS OF ALG2

Process	Method	Complexity
Computing $\tilde{\mathbf{H}}_k, \bar{\mathbf{H}}_k, \forall k$	HAD	$2L_rK(L_rN_r+N_rN_t+N_tL_t+2/3L_r^2)$
	FD	$K(7N_rN_t^2+4N_t^3+2L_rN_rN_t)$
Computing $\mathbf{Z}_k, \forall k$	HAD	$K(7(K-1)L_rL_t^2+4L_t^3)$
	FD	$K(7(K-1)L_rN_t^2+4N_t^3)$
Computing $\mathbf{V}_k, \forall k$	HAD	$K(2L_r^2L_t+11L_r^3)$
	FD	$K(2L_rN_tN_e+7L_rN_e^2+4N_e^3)$, where $N_e=N_t-(K-1)L_r$

VI. COMPUTATIONAL COMPLEXITY

In this section, we provide the computational complexity analysis of the major steps of the proposed Algorithms 1 and 2. Similarly to the assumptions made in [39], we assume that the computational complexity of the matrix product between $[n\times m]$ and $[m\times r]$ matrices, the SVD of $[n\times m]$ matrix, and the inversion of $[n\times n]$ matrix are given by $2\,nmr, 7nm^2+4m^3$, and $\frac{2}{3}n^3$, respectively. Tables I and II show the detailed computational complexity of Algorithms 1 and 2, respectively. In Table I, T_0 denotes the total number of iterations required by the outer loop steps 5-32, while T_1 (resp. T_2) denotes the total number of iterations required by the inner loop steps 10-14 (resp. 23-27).

In the next section, we show for comparison some simulation results when the ABF matrices $\mathbf{F}$ and $\mathbf{W}$ are updated using the proposed algorithm in [5]. In particular, we follow the same solution steps of [5, Algorithm 2] to update $\mathbf{F}$ that maximizes the function $\hat{\zeta}_t=\log\det(\mathbf{I}_{KL_r}+\mathbf{HF}\mathbf{F}^{\mathrm{H}}\mathbf{H}^{\mathrm{H}})$ and to update $\mathbf{W}$ that maximizes the function $\hat{\zeta}_r=\log\det(\mathbf{I}_{N_t}+\mathbf{H}^{\mathrm{H}}\mathbf{W}^{\mathrm{H}}\mathbf{W}\mathbf{H})$ sequentially from the phase-angles of the largest eigenvectors. Note that the above updates are completely decoupled between $\mathbf{F}$ and $\mathbf{W}$, which is different from our proposed ARAB algorithm that updates one variable while fixing the other. However, we found that the convergence of the coupled version of [5, Algorithm 2] are not monotonic and for some channel realizations it might never converge. Thus, we restrict our comparison to the decoupled updates, which provides a lower-bound comparison to the proposed ARAB algorithm.

Table I shows the detailed computational complexity of the algorithm proposed in [5]. Observing closely the results in Table I, we can see that ARAB has a lower computational complexity than Alg. [5], especially for large N_r and/or N_t, since the former operates on the equivalent channels $\widetilde{\mathbf{H}}\in\mathbb{C}^{KL_r\times N_t}$ and $\widehat{\mathbf{H}}\in\mathbb{C}^{L_t\times KL_r}$ when updating the ABF matrices $\mathbf{F}$ and $\mathbf{W}$, respectively. By contrast, Alg. [5] operates on the true channels $\mathbf{H}\in\mathbb{C}^{KN_r\times N_t}$ and $\mathbf{H}^{\mathrm{H}}\in\mathbb{C}^{N_t\times KL_r}$ when updating the ABF matrices $\mathbf{F}$ and $\mathbf{W}$, respectively, where they have larger dimensions than the $\widetilde{\mathbf{H}}$ and $\widehat{\mathbf{H}}$ counterparts. In Table III, we show computational complexity comparing between the ARAB and Alg. [5] algorithms in terms of the ratio β, which is defined as

$$\beta=\frac{\beta^{\text{ARAB}}}{\beta^{\text{Alg. [5]}}}, \tag{42}$$

where $\beta^X, X\in\{\text{ARAB, Alg. [5]}\}$, denotes the number of flops required by algorithm X when updating the ABF matrices $\mathbf{F}$ and $\mathbf{W}$. From Table III, it is clear that ARAB has significantly lower computational complexity than Alg. [5], especially with a large number of antennas. For instance, when $N_t=64$ and $K=1$, ARAB has approximately 10% of the complexity of Alg. [5], while it increases to around 27% when $K=4$.

TABLE III
COMPUTATIONAL COMPLEXITY COMPARISON BETWEEN ARAB AND ALG. [5]
$[N_r=8, N_s=2, T_0=T_1=T_2=5]$

	$N_t=16$	$N_t=64$
$K=1$	$\beta=0.6689$	$\beta=0.0999$
$K=4$	$\beta=1.0692$	$\beta=0.2692$

VII. NUMERICAL RESULTS

In this section, we show detailed simulation results to evaluate the performance of the proposed algorithm as compared to the reference algorithm in [5] in terms of their achievable SE and EE.

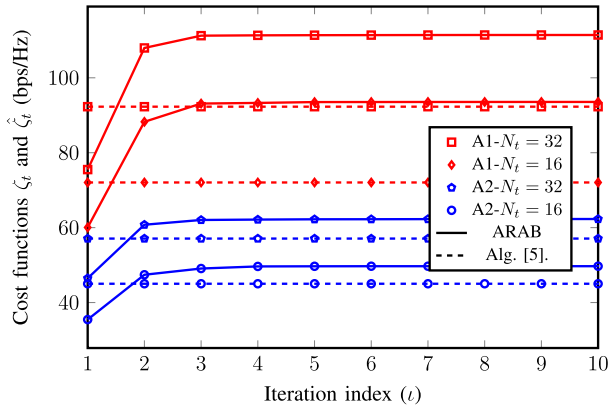


Fig. 4. Convergence behavior of the ARAB algorithm [$N_r = 8$, $N_s = 2$, and $K = 4$].

We assume a geometric channel model with L scatterers, each of which contributes to a single path, where the channel matrix $\mathbf{H}_k \in \mathbb{C}^{N_r \times N_t}$ between the BS and the kth MS is modeled as [3]–[5], [8], [30]

$$\mathbf{H}_k = \frac{1}{\sqrt{L}} \sum_{\ell=1}^{L} a_\ell \mathbf{a}(\theta_\ell)\mathbf{b}^{\mathrm{T}}(\phi_\ell), \tag{43}$$

where L is the total number of channel paths, fixed to $L = 6$ for all the simulation scenarios, for which $a_\ell \sim \mathcal{CN}(0, 1)$, $\theta_\ell \in [0, 2\pi]$, and $\phi_\ell \in [0, 2\pi]$ denote, respectively, the complex path gain, angle of arrival, and angle of departure of the ℓ-th path. Further, $\mathbf{a}(\theta_\ell) \in \mathbb{C}^{N_r \times 1}$ and $\mathbf{b}(\phi_\ell) \in \mathbb{C}^{N_t \times 1}$ denote the array response vectors at MS and BS, respectively. We assume uniform linear arrays with half wavelength between the antenna elements, where the array response vectors $\mathbf{a}(\theta_\ell)$ and $\mathbf{b}(\phi_\ell)$ are given respectively as [3], [4]

$$\mathbf{a}(\theta_\ell) = \left[1, e^{j\pi\cos(\theta_\ell)}, \ldots, e^{j\pi(N_r-1)\cos(\theta_\ell)}\right]^{\mathrm{T}}, \tag{44}$$

$$\mathbf{b}(\phi_\ell) = \left[1, e^{j\pi\cos(\phi_\ell)}, \ldots, e^{j\pi(N_t-1)\cos(\phi_\ell)}\right]^{\mathrm{T}}. \tag{45}$$

Fig. 4 shows the convergence behavior of the proposed ARAB algorithm (solid-lines) when updating the ABF matrix $\mathbf{F}$ in terms of the cost function ζ_t versus the iteration index (ι). Fig. 4 also includes the cost function $\hat{\zeta}_t$ (dashed-lines) when the ABF matrix $\mathbf{F}$ is updated using the reference algorithm from [5]. From Fig. 4, we can observe that ARAB has a monotonic and fast convergence rate, within 4-6 iterations. Obviously, architecture A1 has higher cost function (equivalent channel capacity) than architecture A2, in the expense of a higher energy consumption and computational complexity. Further, ARAB clearly achieves higher cost function than Alg. [5], since it maximizes the cost function iteratively by taking into account both the transmit and the receive ABF matrices, while Alg. [5] completely decouples the transmit and the receive ABF matrices and updates each one separately, which results in a lower cost function.

Fig. 5 shows the SE versus SNR and the EE versus SNR results while assuming $N_t = 64$, $N_r = 8$, $N_s = 2$, and $K = \{1, 4\}$. For the EE-Max approach, we relax the maximum power threshold $P_{\max}$ so that we make the maximum power constraint

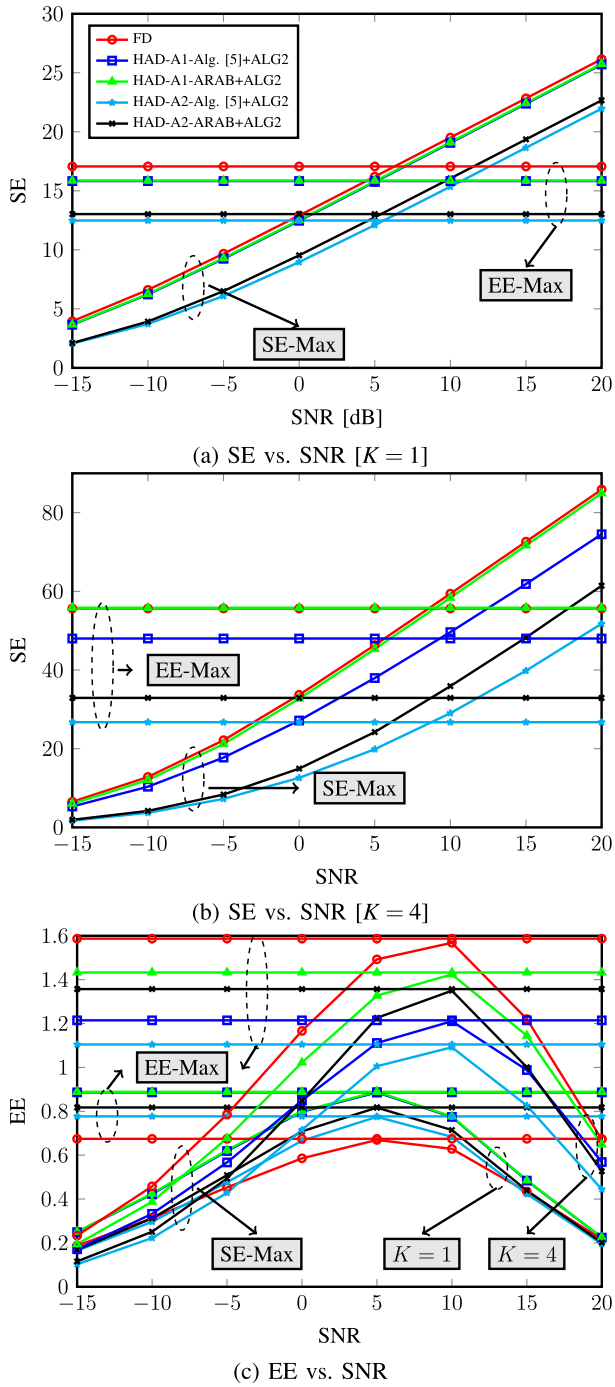


Fig. 5. SE vs. SNR and EE vs. SNR [$N_t = 64$, $N_r = 8$, and $N_s = 2$].

in problem (37) inactive. As a result, neither the SE nor the EE is in function of the SNR level, where the optimal power allocation maximizing the EE is obtained using ALG2.

From Fig. 5a, i.e., when $K = 1$, we can observe that both ARAB and Alg. [5] achieve almost an equal SE performance with both architectures A1 and A2. As expected, the fully-connected architecture A1 achieves a better SE performance than the partially-connected A2 and close to that of the FD architecture counterpart. However, from Fig. 5b, i.e., when K increases to 4, the proposed ARAB algorithm when using the

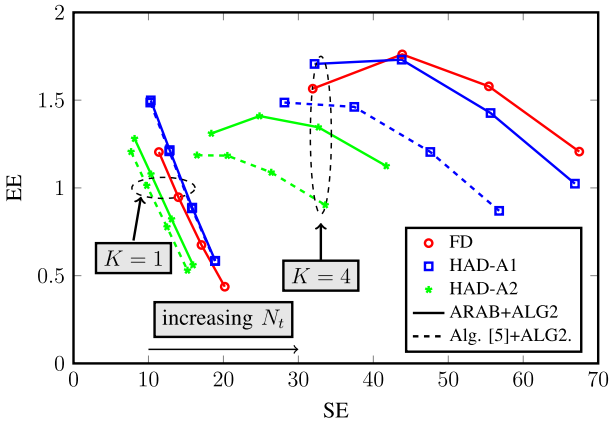


Fig. 6. **EE-Max:** SE vs. EE with varying number of transmit antennas $N_t \in \{16, 32, 64, 128\}$ and number of users $K \in \{1, 4\}$ [$N_r = 8$ and $N_s = 2$].

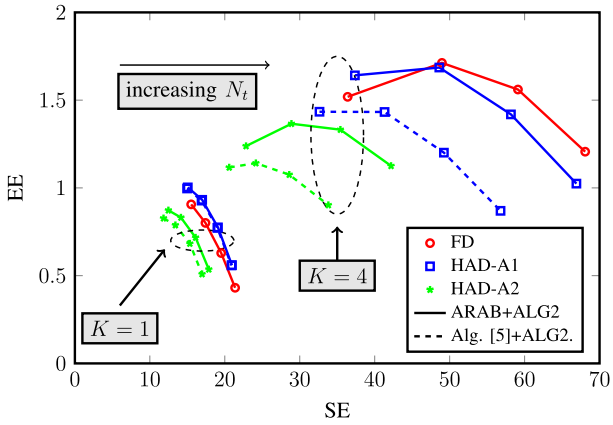


Fig. 7. **SE-Max:** SE vs. EE with varying number of transmit antennas $N_t \in \{16, 32, 64, 128\}$ and number of users $K \in \{1, 4\}$ [$\rho = 10$ dB, $N_r = 8$ and $N_s = 2$].

architecture A1 maintains its very close SE performance to the FD counterpart, unlike the reference Alg. [5]. More interestingly, we can see from Fig. 5c that when $K = 1$, architecture A1 achieves the best EE performance followed by architecture A2, while the FD architecture achieves the worst EE performance.

However, the results are completely different when the number of users increases to $K = 4$, where the FD architecture seems to achieve the best EE performance compared to the both HAD architectures A1 and A2. These results are expected for the following reasons. As Fig. 2 shows, when $K = 4$ and $N_t = 64$, the circuitry power consumption of architecture A1 exceeds that of the FD architecture. Considering that architecture A1 has lower SE than that of the FD architecture, since its major beamforming functionalities are implemented in the analog domain, then it is expected for architecture A1 to have lower EE than FD architecture. On the other hand, Fig. 2 shows that architecture A2 has always lower circuitry power consumption than that of the FD architecture. Therefore, in theory, architecture A2 can be more energy efficient depending on its achievable SE. However, in architecture A2, not only the major beamforming functionalities are implemented in the analog domain, but also each RF chain has significant lack of information, since they are connected with a small number of antennas. This fact causes severe degradation in the system beamforming capabilities and degrees-of-freedom, as compared with the A1 and FD architectures. Therefore, the SE of architecture A2 is significantly degraded, which degrades its EE as well.

To gain more insights about the above observations, Figs. 6, 7, and 8 show the simulation results while varying the number of transmit antennas N_t and the number of users K. Fig. 6 shows SE versus EE results when using the EE-Max approach to update the DBF matrices, while Fig. 7 shows SE versus EE results when using the SE-Max approach to update the DBF matrices. Fig. 8 shows the power allocation, at the convergence, for the EE-Max approach scenarios (note that the SE-Max approach always uses the maximum power). From Figs. 6 and 7, we can see that the above observations from Fig. 5 holds true here, as well, when varying the number of transmit antennas N_t. For instance, when $K = 1$, both HAD architectures A1 and A2 achieve lower SE but higher EE than the FD architecture, where A1 seems to

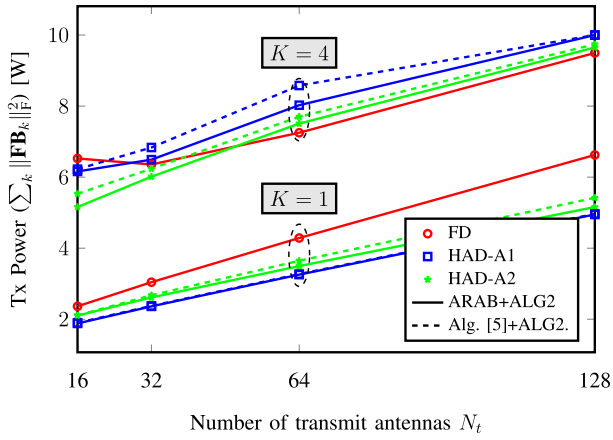


Fig. 8. **EE-Max:** Power allocation with varying number of transmit antennas $N_t \in \{16, 32, 64, 128\}$ and number of users $K \in \{1, 4\}$ [$N_r = 8$ and $N_s = 2$].

outperform A2 in terms of both SE and EE. From Fig. 8 it can also be seen that when $K = 1$, architecture A1 uses the least transmit-power as well. On the other hand, when the number of users increases to $K = 4$, the FD architecture outperforms both the HAD architectures A1 and A2 in terms of both SE and EE, except for the single case when $N_t = 16$, for the same reasons pointed out above.

Further, from Fig. 8, it can be seen that the FD architecture also uses less transmit power than A1 and A2 architectures when $N_t \geq 64$. Note that when the ABF matrices are updated using the proposed ARAB algorithm, both HAD architectures A1 and A2 achieve higher SE and EE than when the ABF matrices are updated using the reference algorithm. Again, this is an expected result, since the ARAB algorithm achieves a higher cost function value than the reference algorithm, i.e., the resulting equivalent channel capacity is larger using the ARAB algorithm than that using the reference algorithm (see Fig 4).

Comparing results from Fig. 6 to Fig. 7, we can see that the above observations hold true as well when the DBF matrices are designed using the SE-Max approach, i.e., when the full transmit

power is used. Note that both approaches achieve comparable SE vs. EE performance, especially when $K = 4$, although the EE-Max approach uses less transmit power compared to the SE-Max approach, as shown in Fig. 8. Obviously, one can increase the transmit power ρ with the SE-Max approach to achieve higher SE at the expense of decreasing the EE. Finally, note that the transmit power increases with N_t and K, which is needed to compensate for increasing circuitry power.

VIII. Conclusion

We have proposed a low-complexity hybrid analog-digital beamforming design for downlink multiuser scenarios to maximize the system's EE. We have shown that, based on the most recently available data for the circuitry components power consumption, the fully-connected architecture based on high-resolution PSs has higher circuitry power consumption than the FD counterpart if the number of users exceeds a certain threshold. In such scenarios, the FD structures are surely more energy-efficient. By contrast, the partially-connected architecture always has lower circuitry power consumption than both architectures: fully-connected and FD. However, due to the severe degradation on its degrees-of-freedom, we found that it can still achieve not only lower SE, but also lower EE than the FD architecture in some multiuser scenarios. Modifying the power consumption model in (4) to account for the computational complexity of beamforming and investigating the impact of low-resolution PSs are left for future works.

Acknowledgment

The authors would like to thank Dr. G. Klang for his continuous support in the TIDE5G and NAIVE projects. Finally, the authors thank the Associate Editor and the anonymous reviewers for their constructive comments.

References

[1] H. Shokri-Ghadikolaei, C. Fischione, G. Fodor, P. Popovski, and M. Zorzi, "Millimeter wave cellular networks: A MAC layer perspective," *IEEE Trans. Commun.*, vol. 63, no. 10, pp. 3437–3458, Oct. 2015.

[2] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3059–3075, Jun. 2015.

[3] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "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, R. W. Heath, S. Abu-Surra, S. Rajagopal, and Z. Pi, "Low complexity precoding for large millimeter wave MIMO systems," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2012, pp. 3724–3729.

[5] X. Gao, L. Dai, S. Han, C. L. I, and R. W. Heath, "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.

[6] R. Mendez-Rial, N. Gunzalez-Prelcic, A. Alkhateeb, and J. R. W. Heath, "Hybrid MIMO architectures for millimeter wave communications: Phase shifters or switches?" *IEEE Access*, vol. 4, pp. 247–267, 2016.

[7] X. Gao, L. Dai, Y. Sun, S. Han, and I. Chih-Lin, "Machine learning inspired energy-efficient hybrid precoding for mmwave massive MIMO systems," in *Proc. IEEE Int. Conf. Commun.*, May 2017, pp. 1–6.

[8] K. Ardah, G. Fodor, Y. C. B. Silva, W. C. Freitas, Jr, and F. R. Cavalcanti, "A unifying design of hybrid beamforming architectures employing phase-shifters or switches," *IEEE Trans. Veh. Technol.*, vol. 67, no. 11, pp. 11243–11247, Nov. 2018.

[9] H. Engler, A. Zappone, and E. A. Jorswieck, "EE maximization for massive MIMO with fully connected hybrid beamformers," in *Proc. IEEE 7th Int. Workshop Comput. Advances Multi-Sensor Adaptive Process.*, Dec. 2017, pp. 1–5.

[10] C. G. Tsinos, S. Maleki, S. Chatzinotas, and B. Ottersten, "On the energy-efficiency of hybrid analog–digital transceivers for single- and multi-carrier large antenna array systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1980–1995, Sep. 2017.

[11] S. He, C. Qi, Y. Wu, and Y. Huang, "Energy-efficient transceiver design for hybrid sub-array architecture MIMO systems," *IEEE Access*, vol. 4, pp. 9895–9905, 2016.

[12] F. Zhu, S. He, R. Li, Y. Huang, and X. You, "Energy-efficient hybrid precoding for broadband millimeter wave communication systems," in *Proc. 9th Int. Conf. Wireless Commun. Signal Process.*, Oct. 2017, pp. 1–5.

[13] X. Yu, J. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 485–500, Apr. 2016.

[14] C. Rusu, R. Mèndez-Rial, N. González-Prelcic, and R. W. Heath, "Low complexity hybrid precoding strategies for millimeter wave communication systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8380–8393, Dec. 2016.

[15] W. Tan, D. Xie, J. Xia, W. Tan, L. Fan, and S. Jin, "Spectral and energy efficiency of massive MIMO for hybrid architectures based on phase shifters," *IEEE Access*, vol. 6, pp. 11751–11759, 2018.

[16] I. Akhtar *et al.*, "Energy efficient hybrid precoding for cooperative multi-cell multiuser massive MIMO systems with multiple base station association," in *Proc. 14th Int. Wireless Commun. Mobile Comput. Conf.*, Jun. 2018, pp. 418–423.

[17] V. N. Ha, D. H. N. Nguyen, and J. Frigon, "Energy-efficient hybrid precoding for mmWave multi-user systems," in *Proc. IEEE Int. Conf. Commun.*, May 2018, pp. 1–6.

[18] D. Mishra and H. Johansson, "Efficacy of hybrid energy beamforming with phase shifter impairments and channel estimation errors," *IEEE Signal Process. Lett.*, vol. 26, no. 1, pp. 99–103, Jan. 2019.

[19] W. Ni, P. Chiang, and S. Dey, "Energy efficient hybrid beamforming in massive MU-MIMO systems via eigenmode selection," in *Proc. IEEE Int. Conf. Internet Things, IEEE Green Comput. Commun., IEEE Cyber, Phys., Social Comput., IEEE Smart Data*, Jan. 2017, pp. 400–406.

[20] C. Fang, B. Makki, J. Li, and T. Svensson, "Coordinated hybrid precoding for energy-efficient millimeter wave systems," in *Proc. IEEE 19th Int. Workshop Signal Process. Advances Wireless Commun.*, Jan. 2018, pp. 1–5.

[21] Y. Zhang, Q. Cui, W. Ni, and P. Zhang, "Energy-efficient transmission of hybrid array with non-ideal power amplifiers and circuitry," *IEEE Trans. Wireless Commun.*, vol. 17, no. 6, pp. 3945–3958, Jan. 2018.

[22] H. H. Kha and T. Do-Hong, "Energy efficiency beamformers for K-user MIMO interference channels with interference alignment," in *Proc. 3rd Int. Conf. Inf. Technol., Comput., Elect. Eng.*, Oct. 2016, pp. 425–428.

[23] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.

[24] K. Ardah, Y. C. B. Silva, and F. R. P. Cavalcanti, "Decentralized linear transceiver design in multicell MIMO broadcast channels," *J. Commun. Inf. Syst.*, vol. 32, no. 1, pp. 102–115, Oct. 2017.

[25] S. Buzzi and C. D'Andrea, "Are mmwave low-complexity beamforming structures energy-efficient? Analysis of the downlink MU-MIMO," in *Proc. IEEE Globecom Workshops*, Dec. 2016, pp. 1–6.

[26] J. Du, W. Xu, H. Shen, X. Dong, and C. Zhao, "Hybrid precoding architecture for massive multiuser MIMO with dissipation: Sub-connected or fully connected structures?" *IEEE Trans. Wireless Commun.*, vol. 17, no. 8, pp. 5465–5479, Aug. 2018.

[27] L. N. Ribeiro, S. Schwarz, M. Rupp, and A. L. F. de Almeida, "Energy efficiency of mmWave massive MIMO precoding with low-resolution ACs," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 2, pp. 298–312, May 2018.

[28] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 884–893, Mar. 2006.

[29] R. Zhang, H. Zhao, and J. Zhang, "Distributed compressed sensing aided sparse channel estimation in FDD massive MIMO system," *IEEE Access*, vol. 6, pp. 18383–18397, 2018.
[30] A. Alkhateeb, O. E. Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.
[31] K. Ardah, A. L. F. de Almeida, and M. Haardt, "A gridless CS approach for channel estimation in hybrid massive MIMO systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2019, pp. 4160–4164.
[32] M. Rani, S. B. Dhok, and R. B. Deshmukh, "A systematic review of compressive sensing: Concepts, implementations and applications," *IEEE Access*, vol. 6, pp. 4875–4894, 2018.
[33] C. Yoon and D. H. Cho, "Energy efficient beamforming and power allocation in dynamic TDD based C-RAN system," *IEEE Commun. Lett.*, vol. 19, no. 10, pp. 1806–1809, Oct. 2015.
[34] S. S. Christensen, R. Agarwal, E. de Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.
[35] K. Ardah, Y. C. B. Silva, and F. R. P. Cavalcanti, "Block diagonalization for multicell multiuser MIMO systems with other-cell interference," in *Proc. Simpósio Brasileiro de Telecomunicações e Processamento de Sinais*, Sep. 2016, pp. 1–5.
[36] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 501–513, Apr. 2016.
[37] J. M. Borwein and A. S. Lewis, *Karush-Kuhn-Tucker Theory*. New York, NY, USA: Springer, 2000, pp. 153–177.
[38] D. P. Palomar and J. R. Fonollosa, "Practical algorithms for a family of waterfilling solutions," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 686–695, Feb. 2005.
[39] J. Shin and J. Moon, "Weighted-sum-rate-maximizing linear transceiver filters for the K-user MIMO interference channel," *IEEE Trans. Commun.*, vol. 60, no. 10, pp. 2776–2783, Oct. 2012.

Khaled Ardah received the M.Sc. degree in mobile systems from the Luleå University of Technology, Luleå, Sweden, in 2013, and the Ph.D. degree from the Federal University of Ceará, Fortaleza, Brazil, in 2018. He was a Postdoctoral Researcher with the Wireless Telecom Research Group (GTEL), Fortaleza, Brazil, in 2018. He is currently a Research Fellow with the Communications Research Laboratory (CRL), Technical University of Ilmenau, Ilmenau, Germany. His research interests include sparse signal recovery, compressed sensing, and massive MIMO systems.

Gábor Fodor (SM'08) received the Ph.D. degree in electrical engineering from the Budapest University of Technology and Economics, Budapest, Hungary, in 1998, the Docent degree from the KTH Royal Institute of Technology, Stockholm, Sweden, in 2019, and the D.Sc. degree from the Hungarian Academy of Sciences, Budapest, Hungary, in 2019. He is currently a Master Researcher with Ericsson Research and an Adjunct Professor with the KTH Royal Institute of Technology, Stockholm, Sweden. He was a co-recipient of the IEEE Communications Society Stephen O. Rice Prize in 2018. He is serving as the Chair of the IEEE Communications Society, Full Duplex Emerging Technologies Initiative and as an Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.

Yuri C. B. Silva (S'01–M'09) received the B.Sc. and M.Sc. degrees in electrical engineering from the Federal University of Ceará, Fortaleza, Brazil, in 2002 and 2004, respectively, and the Dr.-Ing. degree in electrical engineering from the Technische Universität Darmstadt, Darmstadt, Germany, in 2008. From 2001 to 2004, he was with the Wireless Telecom Research Group (GTEL), Fortaleza, Brazil. In 2003, he was a Visiting Researcher with Ericsson Research, Stockholm, Sweden. From 2005 to 2008, he was with the Communications Engineering Lab of the Technische Universität Darmstadt and since 2010 he has been a Professor with the Federal University of Ceará, Fortaleza, Brazil. He currently holds a productivity fellowship in technological development and innovation from CNPq. His main research interests include wireless communications systems, multi-antenna processing, interference management, multicast services, and cooperative communications.

Walter C. Freitas, Jr., received the B.S. and M.S. degrees in electrical engineering from the Federal University of Ceará (UFC), Fortaleza, Brazil, and the Ph.D. degree in teleinformatic engineering from UFC, Fortaleza, Brazil, in 2006. During his studies, he was supported by the Brazilian agency FUNCAP and Ericsson. From July–September of 2015 to April–June of 2016, he was a Postdoctoral Researcher with I3S/CNRS Laboratory, University of Nice, Sophia Antipolis, France. In 2005, he was a Senior Researcher of Nokia Technology Institute. He is currently an Assistant Professor with the Department of Teleinformatics Engineering, Federal University of Ceará and a Researcher of Wireless Telecom Research Group (GTEL) one of the most important research groups in telecommunication in Brazil. His main research interest include development to improve the performance of the wireless communication systems, interference avoidance tools, multilinear algebra, and tensor-based signal processing applied to communications.

André L. F. de Almeida (M'08–SM'13) received the B.Sc. and M.Sc. degrees in electrical engineering from the Federal University of Ceará, Fortaleza, Brazil, in 2001 and 2003, respectively, and the double Ph.D. degrees in sciences and teleinformatics engineering from the University of Nice, Sophia-Antipolis, France, and the Federal University of Ceará, Fortaleza, Brazil, in 2007. He is currently an Associate Professor with the Department of Teleinformatics Engineering, Federal University of Ceará. During fall 2002, he was a Visiting Researcher with Ericsson Research Labs, Stockholm, Sweden. From 2007 to 2008, he held a Teaching Position with the University of Nice Sophia-Antipolis, France. In 2012, 2013, and 2018, he was awarded Visiting Professor positions with the University of Nice Sophia-Antipolis, France. His current research interests include tensor decompositions and multilinear algebra with applications to communications and signal processing. Dr. de Almeida was an Associate Editor for several journals, such as the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE SIGNAL PROCESSING LETTERS, and *Wireless Communications and Mobile Computing*, and a Guest Editor for the *EURASIP Journal of Advances in Signal Processing*. He is a member of the Sensor Array and Multichannel (SAM) Technical Committee of the IEEE Signal Processing Society (SPS) and a member of the EURASIP Signal Processing for Multi-Sensor Systems Special Area Team (SPMuS SAT). He was the General Co-Chair of the IEEE CAMSAP'2017, the Technical Co-Chair of the Symposium on "Tensor Methods for Signal Processing and Machine Learning" at GlobalSIP 2018 and 2019, and currently serves as the Technical Co-Chair of IEEE SAM'2020, Hangzhou, China. He is a Research Fellow of the CNPq (The Brazilian National Council for Scientific and Technological Development) and an Affiliate Member of the Brazilian Academy of Sciences.