IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 6 September 2023, accepted 22 September 2023, date of publication 4 October 2023, date of current version 12 October 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3321681

RESEARCH ARTICLE

Hybrid Precoding/Combining for mmWave MIMO Systems With Hybrid Array Architecture

MOHAMED ALOUZI[1], (Member, IEEE), FAISAL AL-KAMALI[1], CLAUDE D'AMOURS[1], (Member, IEEE), AND FRANCOIS CHAN[1,2], (Senior Member, IEEE)

[1]School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, ON K1N 6N5, Canada
[2]Department of Electrical and Computer Engineering, Royal Military College of Canada, Kingston, ON K7K 7B4, Canada

Corresponding author: Mohamed Alouzi (malou052@uottawa.ca)

ABSTRACT Utilizing hybrid precoding/combining for subarray (SA) architectures in millimeter-wave (mmWave) multi-input multi-output (MIMO) systems offers reduced hardware cost and power consumption when compared with that for full array (FA) architectures, although the spectral efficiency is lower. Hence, this paper introduces a new hybrid array (HA) architecture for mmWave MIMO systems, designed to achieve a balance between spectral efficiency, cost, and power consumption. Initially, the proposed HA architecture partitions the antennas into distinct subarrays. The number of these subarrays equals the count of radio frequency chains at the transmitter/receiver. These subarrays are subsequently organized into separate subsets referred to as groups. Ultimately, within each group, the antennas are connected with a corresponding group of radio frequency chains, employing a connection method similar to that of the FA architecture. Compared to the SA architecture, the proposed HA architecture provides higher spectral efficiency by exploiting spatial diversity through grouping of antennas. Furthermore, the HA architecture achieves cost and power reduction in comparison to the FA architecture by connecting a subset of antennas to each group of radio frequency chains. Two efficient iterative hybrid precoding/combining algorithms are also proposed, studied and compared for HA architecture mmWave MIMO system. During the design derivations, the proposed algorithms consider the block structure inherent in the analog precoding and combining matrices within the HA architecture. This consideration makes it possible to independently optimize hybrid precoding/combining for each group. Results show that the suggested HA hybrid design using Algorithm 2 provides better performance than using Algorithm 1 and both algorithms can achieve higher spectral efficiency than the conventional SA and FA hybrid designs while being less complex.

INDEX TERMS FA, SA, HA, hybrid precoding/combining, mmWave, MIMO.

I. INTRODUCTION

Massive multi-input multi-output (MIMO) and millimeter wave (mmWave) technologies have been an active area of research for fifth-generation (5G) and sixth-generation (6G) wireless broadband networks in recent years due to their ability to significantly increase spectral efficiency and system throughput [1], [2], [3]. Traditional fully digital transceivers are not feasible with these massive MIMO systems because of the high implementation costs and energy consumption of the radio frequency (RF) chains [4], [5], [6], [7]. The solution to this problem is to use hybrid analog/digital transceivers, which allow a flexible trade-off between low cost and power consumption while achieving high spectral efficiency [8], [9], [10], [11], [12], [13]. At the transmitter and receiver, hybrid precoding and combining architectures are used. They are made up of analog precoders/combiners with phase shifters (PSs) in the RF domain and digital precoders/combiners in the baseband domain [4]. The two most common types of hybrid precoding/combining architectures are full array (FA) [4], [5] and subarray (SA) [15], [16], [17], [18], [19], [20], [21]. FA architecture is widely used with hybrid precoding and combining systems. In the FA architecture, PSs are utilized to establish links between each RF chain and individual

The associate editor coordinating the review of this manuscript and approving it for publication was Walid Al-Hussaibi.

antennas. However, with an increasing number of antennas, this results in a linear growth in the number of PSs. In the SA, each RF chain is only connected with a subset of antennas, leading to a reduced requirement for PSs compared to the FA architecture. As a result, the SA architecture is now regarded as a viable mmWave architecture candidate.

In past years, the FA hybrid architecture has received a lot of attention for mmWave systems, and a lot of effort has gone into developing low complexity hybrid precoding/combining algorithms [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14]. In [4], the authors introduced a hybrid design algorithm based on the simultaneous orthogonal matching pursuit (OMP) technique. They demonstrated its capability to achieve performance similar to that of the optimal digital beamformer, although at the expense of increased computational complexity. In [11], a hybrid design with low computational complexity was proposed and studied for mmWave systems. Recently, a new approach to design a hybrid precoder/combiner based on the deep learning technique was suggested in [14]. While many FA hybrid designs can achieve performance similar to the ideal digital beamformer, the FA architecture's high hardware costs, and power consumption persist. This is primarily due to the necessity of connecting each RF chain to every antenna using a PSs. Therefore, a SA structure is considered to be a promising method for practical implementation in mmWave massive MIMO systems. Additionally, hybrid precoding/combining designs using a SA architecture has recently attracted increased interest by balancing spectral efficiency, cost, power consumption, and hardware complexity.

There are two types of SA architectures: fixed SA [15], [16], [17], [18] and dynamic SA [22], [23], [24], [25]. In the fixed SA design, the transmitter/receiver antennas are divided into subarrays, with each subarray connected to a single RF chain. However, in the dynamic SA design, switches are employed to establish connections between RF chains and subarrays. The authors of [15] proposed a hybrid precoding scheme based on successive interference cancellation for fixed SA designs. They considered a diagonal digital precoder consisting of real elements and showed that the fixed SA outperforms the FA in terms of energy efficiency. The work presented in [17] introduced and explored two hybrid algorithms designed for mmWave utilizing SA. The results illustrated that these proposed algorithms, operating for a fixed SA, can achieve a performance level comparable to the FA while maintaining low complexity. As detailed in [18], an overlapped SA architecture employing iterative hybrid precoding can be leveraged to enhance the spectral efficiency of the SA. The results of this study revealed that the overlapped SA architecture can enhance spectral efficiency, introducing a slight increase in hardware complexity compared to the fixed SA architecture. However, this complexity is still below that of the dynamic SA and FA architectures. The dynamic SA architecture achieves superior spectral efficiency over the fixed SA architecture, but at the expense of heightened hardware and computational complexity as well as increased power consumption. This can be attributed to the linear growth in the number of required switches with the increase in the number of antennas. To reduce the this complexity, [26], [27], [28] introduced a partially dynamic SA architecture. In [29], we presented an efficient approach to transform the hybrid precoding from FA architecture to SA architecture.

In recent years, there has been an exploration of hybrid designs using deep learning techniques [30], [31], [32], [33]. In [30], a deep learning framework was introduced to enhance performance and decrease computational time in contrast to conventional methods. The framework employed a convolutional neural network to generate analog precoders and combiners from an imperfect channel matrix. In [32], an unsupervised deep learning approach proposed for designing hybrid beamforming for SA outperformed conventional techniques in simulations. In [33], the researchers presented a deep learning-based wideband hybrid precoding network for Terahertz MIMO systems, designed to address beam squint.

Despite the decreased cost and hardware complexity of hybrid design in the SA architecture compared to the FA architecture, its performance remains lower than that of optimal solution [15], [17]. As a result, the introduction of a novel array architecture, different from both FA and SA architectures, becomes crucial in achieving a balance among performance, hardware cost, and power consumption, and it is the main goal of this study.

To the best of our knowledge, prior to our research in [34], there has been no exploration of the hybrid array (HA) architecture or hybrid precoding/combining for HA systems. These aspects serve as the primary focus of our current paper. In our previous work in [34], we presented several hybrid algorithms designed for mmWave MIMO systems with different array architectures. Building on our previous contributions, this paper further extends the investigation of HA hybrid design in [34]. Specifically, this paper studies two HA algorithms, whereas in [34], only one algorithm was analyzed briefly. Additionally, we provide supplementary analyses and introduce new simulation results to enhance the understanding and applicability of these algorithms.

In the proposed HA architecture, the antennas are divided into groups of subarrays, and analog precoding and combining are implemented as multi-FA precoding/combining with a lower number of antennas than that in the FA architecture. Results showed that the use of HA architecture allows a trade-off among cost, complexity, and power consumption while achieving high spectral efficiency. Our research has demonstrated that the HA architecture offers significant advantages over the FA architecture in terms of cost, complexity, and power consumption reduction. By connecting a subset of antennas to each RF chain group, the HA architecture lowers the number of required PSs, leading to lower hardware cost, and power consumption, while slightly decreasing the spectral efficiency compared to the FA as

shown in the simulation results. The reduction in the number of PSs in the proposed HA depends on the number of groups. For example, the reduction is 50% if the number of groups is 2 and 75% if the number of groups is 4. Compared to the SA architecture, the proposed HA architecture achieves a higher performance with a slight increase in complexity.

The contributions of this paper are summarized as follows:

- A new HA architecture is introduced, along with the derivation and description of a general HA system model for mmWave MIMO communication systems. In the HA architecture, the antennas at both the transmitter and receiver are divided into subarrays whose number matches the number of RF chains. Subsequently, these subarrays are grouped into distinct groups, followed by linking each group's antennas to a corresponding group of RF chains in a manner similar to the connections in the FA.
- The optimization problem of hybrid precoding/ combining in the proposed HA architecture is derived and solved. Our study demonstrates that the optimization problem related to hybrid precoding and combining in the HA architecture can be divided into separate problems for each group, which can be solved independently. This strategy enables us to find hybrid precoding and combining matrices for each group and then combine them to create the complete hybrid precoding and combining matrices for the entire architecture.
- Two iterative algorithms are proposed to determine the hybrid precoding and combining matrices. These algorithms are developed and examined, taking into account the block structure of analog precoding/combining matrices of the HA architecture. The design process avoids any assumption or dependency on the antenna array geometry. To achieve hybrid precoding/combining for the entire system, the algorithms independently calculate the hybrid precoding/combining matrices for each group.
- The complexity of the suggested hybrid design algorithms is studied and discussed to highlight their relative simplicity in comparison to other existing hybrid design algorithms.
- Simulations are conducted to assess the effectiveness of the proposed hybrid design algorithms for mmWave MIMO systems employing HA architecture. In this paper, we compared the performance of our proposed HA hybrid design algorithms with that of the FA and SA hybrid design algorithms in [4], [11], and [17], as well as with that of unconstrained optimal design. According to the results, the proposed HA hybrid design algorithms bridge the gap between SA and FA hybrid designs and provide a valuable trade-off among spectral efficiency, cost, complexity, and power consumption.

The structure for the remainder of the paper is as follow. Section II derives and discusses the HA system model, as well as the hybrid precoding/combining problem formulation. Section III elaborates on the proposed HA hybrid precoding/combining algorithms and derives their mathematical models. Section IV analyzes the complexity of the proposed algorithms. Section V presents simulation results, and Section VI draws conclusions.

II. HYBRID ARRAY ARCHITECTURE SYSTEM MODEL AND PROBLEM FORMULATION

In this section, a system model of the hybrid precoding with HA architecture is discussed for mmWave MIMO system. We consider a scenario with one mobile station equipped with N_r antennas and N_{rRF} RF chains, which is served by a single base station equipped with N_t antennas and N_{tRF} RF chains. N_s distinct data streams are sent out by the base station. The transmitter's and receiver's antennas are partitioned into N_{tSA} and N_{rSA} subarrays, respectively, and then into subsets of subarrays known as groups. Each group of subarrays is connected to a group of RF chains equal to the number of subarrays in this group. N_{tg} groups at the transmitter and N_{rg} groups at the receiver side are assumed. N_{tSAg} represents the number of subarrays within each group at the transmitter, calculated as N_{tSA}/N_{tg}. Similarly, N_{rSAg} represents the number of subarrays within each group at the receiver, calculated as N_{rSA}/N_{rg}. The number of RF chains linking with each group at the transmitter and receiver, respectively, is also equal to N_{tSAg} and N_{rSAg}. In this paper, we assume $N_{tRF} = N_{tSA}$ and $N_{rRF} = N_{rSA}$ and exploring scenarios with different RF chain and subarray numbers in each group is a direction for future research.

At the base station, the structures of the proposed HA architecture with hybrid precoding is depicted in Fig. 1. For the purpose of comparison, the structures of FA and SA architectures with hybrid precoding are drawn in Fig. 2. As shown in Fig.1, the transmitter initiates the process with the application of the baseband digital precoder $\mathbf{F}_{\mathrm{D}}$. Following this, the resultant signal undergoes an additional precoding step through the HA analog precoder $\mathbf{F}_{\mathrm{AHA}}$. The expression representing the received signal in the HA architecture can be formulated as follows:

$$\begin{aligned} \mathrm{y} &= \sqrt{\rho}\mathbf{B}_{\mathrm{D}}^{\mathrm{H}}\mathbf{B}_{\mathrm{AHA}}^{\mathrm{H}}\mathbf{H}\mathbf{F}_{\mathrm{AHA}}\mathbf{F}_{\mathrm{D}}\mathrm{s} + \mathrm{n} \\ &= \sqrt{\rho}\mathbf{B}_{\mathrm{HA}}^{\mathrm{H}}\mathbf{H}\mathbf{F}_{\mathrm{HA}}\mathrm{s} + \mathrm{n} \end{aligned} \tag{1}$$

where ρ denotes the average received power, $\mathbf{H}$ is the mmWave channel matrix with a dimension of $N_r \times N_t$, $\mathbf{F}_{\mathrm{AHA}}$ represent the $N_t \times N_{tRF}$ RF precoding matrix and $\mathbf{F}_{\mathrm{D}}$ the $N_{tRF} \times N_s$ baseband precoding matrix, of the HA architecture. $\mathbf{B}_{\mathrm{AHA}}$ and $\mathbf{B}_D$ are the $N_r \times N_{rRF}$ RF combining matrix and the $N_{rRF} \times N_s$ baseband combining matrix, respectively, of the HA architecture. s is the $N_s \times 1$ vector of the transmitted signal such that $\mathbb{E}\left[\mathrm{ss}^*\right] = \frac{1}{N_s}\mathbf{I}_{\mathrm{N_s}}$. n is an $N_r \times 1$ vector representing the additive white Gaussian noise. We assume an independent and identically distributed (i.i.d) complex Gaussian noise $\mathcal{CN}\left(0,\sigma_n^2\right)$. All elements in $\mathbf{P}_{\mathrm{AHA}}$ and $\mathbf{B}_{\mathrm{AHA}}$ have the same amplitudes, which are equal to $1/\sqrt{N_t/N_{tg}}$ and $1/\sqrt{N_r/N_{rg}}$, respectively. Note that, $\mathbf{F}_{\mathrm{HA}}$, and $\mathbf{B}_{\mathrm{HA}}$

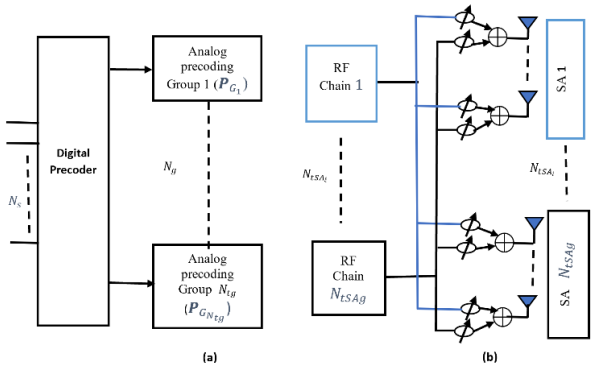


FIGURE 1. The structure of the proposed HA architecture at the base station. (a) Hybrid precoding. (b) Analog precoding in the *ngth* group [34].

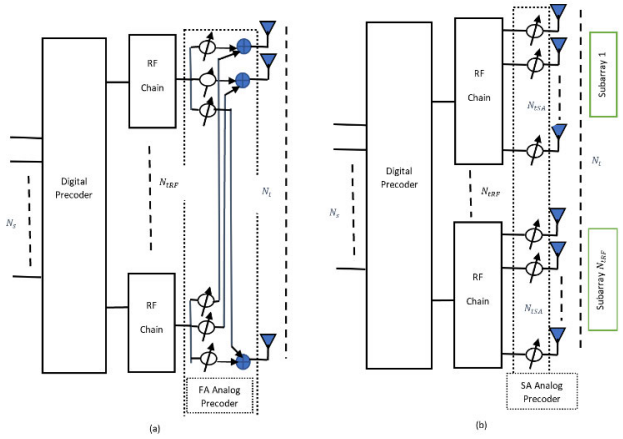


FIGURE 2. A BS with hybrid precoding at different architectures. (a) FA. (b) SA [34].

must satisfy $\|\mathbf{F}_{\mathrm{HA}}\|_F^2 = N_s$ and $\|\mathbf{B}_{\mathrm{HA}}\|_F^2 = N_s$, where $\mathbf{F}_{\mathrm{HA}} = \mathbf{F}_{\mathrm{AHA}}\mathbf{F}_{\mathrm{D}}$ and $\mathbf{B}_{\mathrm{HA}} = \mathbf{B}_{\mathrm{AHA}}\mathbf{B}_{\mathrm{D}}$. This paper assumes the narrowband mmWave channel as detailed in [4], [11], and [17], and $\mathbf{H}$ can be represented as follows:

$$\mathbf{H} = \sqrt{N_t N_r / N_{cl} N_{ray}} \sum_{i}^{N_{cl}} \sum_{l}^{N_{ray}} \alpha_{il} \Gamma_r\left(\phi_{il}^r, \theta_{il}^r\right) \times \Gamma_t\left(\phi_{il}^t, \theta_{il}^t\right) \mathrm{a}_r\left(\phi_{il}^r, \theta_{il}^r\right) \mathrm{a}_t\left(\phi_{il}^t, \theta_{il}^t\right)^* \quad (2)$$

where N_{cl}, N_{ray} and α_{il} are the number of clusters, the number of paths, and the gain of the lth path in the ith cluster, respectively. $\phi_{il}^t(\phi_{il}^r)$ and $\theta_{il}^t(\theta_{il}^r)$ are the azimuth (elevation) angles of departure and arrival of the lth path in the ith cluster. $\Gamma_t\left(\phi_{il}^t, \theta_{il}^t\right)$ and $\Gamma_r(\phi_{il}^r, \theta_{il}^r)$ represent the transmit and receive antenna element gains at their departure and arrival angles. $\mathrm{a}_t\left(\phi_{il}^t, \theta_{il}^t\right)$, and $\mathrm{a}_r\left(\phi_{il}^r, \theta_{il}^r\right)$ denote the antenna array responses at the transmitter and the receiver, respectively. We assume uniform planar array (UPA) with w_1 and w_2 elements on width and height, respectively. Note that $N_t = w_1 w_2$. The antenna array response of the UPA is defined as

$$\mathrm{a}_{UPA(\phi,\theta)} = \frac{1}{\sqrt{N_t}}[1, \ldots, e^{j(\frac{2\pi}{\lambda})d(x\sin(\phi)\sin(\theta)+y\cos(\theta))}, \ldots, e^{j(\frac{2\pi}{\lambda})d((w_1-1)\sin(\phi)\sin(\theta)+(w_2-1)\cos(\theta)}]^T \quad (3)$$

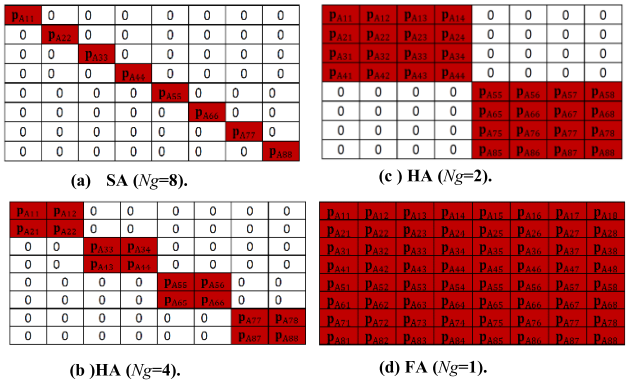


FIGURE 3. The structures of the analog precoding matrix when $N_{tSA} = N_{tRF} = 8$ for different groups. (a) SA Architecture with $N_g = 8$. (b) HA Architecture with $N_g = 4$. (c) HA Architecture with $N_g = 2$. (d) FA Architecture with $N_g = 1$.

where $1 \leq x \leq (w_1-1)$ and $1 \leq y \leq (w_2-1)$. $d = \frac{\lambda}{2}$ is the antennas spacing. In the HA, we can write $\mathbf{F}_{\mathrm{AHA}}$ as

$$\mathbf{F}_{\mathrm{AHA}} = \begin{bmatrix} \mathbf{F}_{\mathrm{G}_1} & 0 & \ldots & 0 \\ 0 & \mathbf{F}_{\mathrm{G}_2} & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & \mathbf{F}_{\mathrm{G}_{\mathrm{N_{tg}}}} \end{bmatrix} \quad (4)$$

The matrix $\mathbf{F}_{\mathrm{G_{ng}}}$ has dimensions $(N_t/N_{tg}) \times (N_{tRF}/N_{tg})$ and represents the matrix of the analog precoding of the $ngth$ group, where $1 \leq ng \leq N_{tg}$. Here, $N_{tg} = 2^n$, with n taking values from 0 to $\log_2 N_{tSA}$. Note that, when $n = 0$, $N_{tg} = 1$, resulting in an FA structure [11], while when $n = \log_2 N_{tSA}$, N_{tg} equals N_{tSA}, yielding a conventional fixed SA structure in [17]. The proposed HA architecture encompasses cases where $1 \leq n \leq (\log_2 N_{tSA}) - 1$.

Additionally, $\mathbf{B}_{\mathrm{AHA}}$ can be also expressed as

$$\mathbf{B}_{\mathrm{AHA}} = \begin{bmatrix} \mathbf{B}_{\mathrm{G}_1} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{\mathrm{G}_2} & \ldots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{B}_{\mathrm{G}_{\mathrm{N_{rg}}}} \end{bmatrix} \quad (5)$$

The matrix $\mathbf{B}_{\mathrm{G_{ng}}}$ has dimensions $(N_r/N_{rg}) \times (N_{rRF}/N_{rg})$ and represents the analog combining matrix of the $ngth$ group, where $1 \leq ng \leq N_{rg}$. The value of N_{rg}, can be calculated using a similar method as N_{tg}, but by replacing N_{tSA} with N_{rSA} in the calculation.

To enhance the visual representation, the structures of the analog precoding matrix when $N_{tSA} = N_{tRF} = 8$ for different groups are drawn in Fig. 3. Fig. 3 illustrates the structure of the analog precoding matrix in the HA, which exhibits a diagonal pattern similar to that of the SA architecture. At the same time, the structure of the analog precoding matrix for each group within the HA matrix is similar to that of the FA architecture. This distinctive combination of structures is why we refer to the proposed architecture as the hybrid array architecture.

In the proposed HA, the spectral efficiency can be written as

$$\mathrm{R} = \left(\left| \mathbf{I}_{\mathrm{N_r}} + \frac{\rho}{N_s} \psi_{\mathrm{k}}^{-1} \mathbf{B}_{\mathrm{D}}^{\mathrm{k^H}} \mathbf{B}_{\mathrm{AHA}}^{k^{\mathrm{H}}} \mathbf{H} \mathbf{F}_{\mathrm{AHA}} \mathbf{F}_{\mathrm{D}} \mathbf{F}_{\mathrm{D}}^{\mathrm{H}} \mathbf{F}_{\mathrm{AHA}}^{\mathrm{H}} \mathbf{H}^{\mathrm{H}} \times \mathbf{B}_{\mathrm{AHA}}^{\mathbf{k}} \mathbf{B}_{\mathrm{D}}^{\mathbf{k}} \right| \right) \tag{6}$$

where $\psi_{\mathbf{k}} = \sigma_{\mathrm{n}}^2 \mathbf{B}_{\mathrm{D}}^{\mathrm{k^H}} \mathbf{B}_{\mathrm{AHA}}^{\mathrm{k^H}} \mathbf{B}_{\mathrm{AHA}}^{\mathrm{k}} \mathbf{B}_{\mathrm{D}}^{\mathrm{k}}$.

In this paper, our goal now is to maximize the spectral efficiency in 6, aiming to achieve hybrid precoding/combining in the HA architecture. This requires considering the total transmitted power constraint and the constraints on $\mathbf{F}_{\mathrm{AHA}}$ and $\mathbf{W}_{\mathrm{AHA}}$. Therefore, the optimization problem can be written as

$$\begin{aligned} \max_{\mathbf{F}_{\mathrm{AHA}}, \mathbf{F}_{\mathrm{D}}, \mathbf{B}_{\mathrm{AHA}}, \mathbf{B}_{\mathrm{D}}} \quad & R \\ \text{st. } & \mathbf{F}_{\mathrm{AHA}} \in \boldsymbol{\zeta}_{\mathrm{AHA}} \text{and} \mathbf{B}_{\mathrm{AHA}} \in \boldsymbol{\xi}_{\mathrm{AHA}} \\ & \|\mathbf{F}_{\mathrm{AHA}} \mathbf{F}_{\mathrm{D}}\|_{\mathrm{F}}^2 = N_s \text{ and } \|\mathbf{B}_{\mathrm{AHA}} \mathbf{B}_{\mathrm{D}}\|_{\mathrm{F}}^2 = N_s \end{aligned} \tag{7}$$

where $\boldsymbol{\zeta}_{\mathrm{AHA}}$ includes all possible precoding matrices, and $\boldsymbol{\xi}_{\mathrm{AHA}}$ includes all possible combining matrices. Both $\boldsymbol{\zeta}_{\mathrm{AHA}}$ and $\boldsymbol{\xi}_{\mathrm{AHA}}$ must satisfy the amplitude constraint. A simpler approach to achieve the objective function's maximization in equation (7) is to minimize the Euclidean distance between the optimal digital solution and the hybrid precoding employed in the proposed HA architecture, expressed as

$$\begin{aligned} \left(\mathbf{F}_{\mathrm{AHA}}^{\mathrm{opt}}, \mathbf{F}_{\mathrm{D}}^{\mathrm{opt}}\right) = \underset{\mathbf{F}_{\mathrm{AHA}}, \mathbf{F}_{\mathrm{D}}}{\arg\min} & \left\|\mathbf{F}^{\mathrm{opt}} - \mathbf{F}_{\mathrm{AHA}} \mathbf{F}_{\mathrm{D}}\right\|_{\mathrm{F}}^2, \\ \text{st. } & \mathbf{F}_{\mathrm{AHA}} \in \zeta_{\mathrm{AHA}} \\ & \|\mathbf{F}_{\mathrm{AHA}} \mathbf{F}_{\mathrm{D}}\|_{\mathrm{F}}^2 = N_s \end{aligned} \tag{8}$$

where $\mathbf{F}^{\mathrm{opt}} = \mathbf{V}_1 = \mathbf{V}(:, 1 : N_s)$ is the optimal precoding. $\mathbf{V}$ is the right singular matrix of the singular value decomposition (SVD) of $\mathbf{H}$, denoted by $\mathbf{H} = \mathbf{U}\Sigma\mathbf{V}^{\mathrm{H}}$. Simillar to equation (8), the optimization problem of the combiner in HA can be formulated as follows:

$$\begin{aligned} \left(\mathbf{B}_{\mathrm{AHA}}^{\mathrm{opt}}, \mathbf{B}_{\mathrm{D}}^{\mathrm{opt}}\right) = \underset{\mathbf{B}_{\mathrm{AHA}}, \mathbf{B}_{\mathrm{D}}}{\arg\min} & \left\|\mathbf{B}^{\mathrm{opt}} - \mathbf{B}_{\mathrm{AHA}} \mathbf{B}_{\mathrm{D}}\right\|_{\mathrm{F}}^2 \\ \text{st. } & \mathbf{B}_{\mathrm{AHA}} \in \xi_{\mathrm{AHA}}, \\ & \|\mathbf{B}_{\mathrm{AHA}} \mathbf{B}_{\mathrm{D}}\|_{\mathrm{F}}^2 = N_s \end{aligned} \tag{9}$$

where $\mathbf{B}^{\mathrm{opt}} = \mathbf{U}_1 = \mathbf{U}(:, 1 : N_s)$ is the optimal hybrid combining. The non-convex nature of problems (8) and (9) makes it evident that finding an optimal solution is challenging. By leveraging the block structure inherent in the HA structure, the $\mathbf{F}_{\mathrm{HA}}$ can be rewritten as follows:

$$\begin{aligned} \mathbf{F}_{\mathrm{HA}} &= \mathbf{F}_{\mathrm{AHA}} \mathbf{F}_{\mathrm{D}} \\ &= \begin{bmatrix} \mathbf{F}_{\mathrm{AG}_1} & 0 & \dots & 0 \\ 0 & \mathbf{F}_{\mathrm{AG}_2} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \mathbf{F}_{\mathrm{AG}_{\mathrm{N_{tg}}}} \end{bmatrix} \begin{bmatrix} \mathbf{F}_{\mathrm{DG}_1} \\ \mathbf{F}_{\mathrm{DG}_2} \\ \vdots \\ \mathbf{F}_{\mathrm{DG}_{\mathrm{N_{tg}}}} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{F}_{\mathrm{AG}_1} \mathbf{F}_{\mathrm{DG}_1} \\ \mathbf{F}_{\mathrm{AG}_2} \mathbf{F}_{\mathrm{DG}_2} \\ \vdots \\ \mathbf{F}_{\mathrm{AG}_{\mathrm{N_{tg}}}} \mathbf{F}_{\mathrm{DG}_{\mathrm{N_{tg}}}} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_{\mathrm{HA}_1} \\ \mathbf{F}_{\mathrm{HA}_2} \\ \vdots \\ \mathbf{F}_{\mathrm{HA}_{\mathrm{N_{tg}}}} \end{bmatrix} \end{aligned} \tag{10}$$

where $\mathbf{F}_{\mathrm{DG_{ng}}}$ represents a $(N_{tRF}/N_{tg}) \times N_s$ matrix corresponding to the digital precoder of the *ngth* group, while $\mathbf{F}_{\mathrm{HA_{ng}}}$ denotes the hybrid precoding matrix of the *ngth* group, with dimension $(N_t/N_{tg}) \times N_s$. Moreover, we can decompose the optimal precoding matrix as

$$\mathbf{F}^{\mathrm{opt}} = \begin{bmatrix} \mathbf{F}_{\mathrm{G}_1}^{\mathrm{opt}} \\ \mathbf{F}_{\mathrm{G}_2}^{\mathrm{opt}} \\ \vdots \\ \mathbf{F}_{\mathrm{G}_{\mathrm{N_{tg}}}}^{\mathrm{opt}} \end{bmatrix} \tag{11}$$

In (11), $\mathbf{F}_{\mathrm{G_{ng}}}^{\mathrm{opt}}$ represents the optimal digital precoder of the *ngth* group. Utilizing equations (10) and (11), problem (8) can be reformulated into a set of N_{tg} distinct subproblems, each addressing an individual group independently.

$$\left\|\mathbf{F}^{\mathrm{opt}} - \mathbf{F}_{\mathrm{AHA}} \mathbf{F}_{\mathrm{D}}\right\|_{\mathrm{F}}^2 = \sum_{\mathrm{ng}=1}^{N_{tg}} \left\|\mathbf{F}_{\mathrm{G_{ng}}}^{\mathrm{opt}} - \mathbf{F}_{\mathrm{AG_{ng}}} \mathbf{F}_{\mathrm{DG_{ng}}}\right\|_{\mathrm{F}}^2 \tag{12}$$

To minimize the objective function in (8), the optimization problem can be solved by individually optimizing the N_{tg} subproblems, as follows:

$$\begin{aligned} \left(\mathbf{F}_{\mathrm{AG_{ng}}}^{\mathrm{opt}}, \mathbf{F}_{\mathrm{DG_{ng}}}^{\mathrm{opt}}\right) = \underset{\mathbf{F}_{\mathrm{AG_{ng}}}, \mathbf{F}_{\mathrm{DG_{ng}}}}{\arg\min} & \left\|\mathbf{F}_{\mathrm{G_{ng}}}^{\mathrm{opt}} - \mathbf{F}_{\mathrm{AG_{ng}}} \mathbf{F}_{\mathrm{DG_{ng}}}\right\|_{\mathrm{F}}^2 \\ \text{st. } & \mathbf{F}_{\mathbf{AG_{ng}}} \in \boldsymbol{\zeta}_{\mathbf{AHA}}, \\ & \left\|\mathbf{F}_{\mathbf{AG_{ng}}} \mathbf{F}_{\mathbf{DG_{ng}}}\right\|_{\mathrm{F}}^2 = N_s/N_{tg} \end{aligned} \tag{13}$$

Similarly, $\mathbf{B}_{\mathrm{HA}}$ and $\mathbf{B}^{\mathrm{opt}}$ can be expressed for each group as

$$\mathbf{B}_{\mathrm{HA}} = \begin{bmatrix} \mathbf{B}_{\mathrm{AG}_1} \mathbf{B}_{\mathrm{DG}_1} \\ \mathbf{B}_{\mathrm{AG}_2} \mathbf{B}_{\mathrm{DG}_2} \\ \vdots \\ \mathbf{B}_{\mathrm{AG}_{\mathrm{N_{rg}}}} \mathbf{B}_{\mathrm{DG}_{\mathrm{N_{rg}}}} \end{bmatrix} = \begin{bmatrix} \mathbf{B}_{\mathrm{HA}_1} \\ \mathbf{B}_{\mathrm{HA}_2} \\ \vdots \\ \mathbf{B}_{\mathrm{HA}_{\mathrm{N_{rg}}}} \end{bmatrix} \tag{14}$$

and

$$\mathbf{B}^{\mathrm{opt}} = \begin{bmatrix} \mathbf{B}_{\mathrm{G}_1}^{\mathrm{opt}} \\ \mathbf{B}_{\mathrm{G}_2}^{\mathrm{opt}} \\ \vdots \\ \mathbf{B}_{\mathrm{G}_{\mathrm{N_{rg}}}}^{\mathrm{opt}} \end{bmatrix} \tag{15}$$

where $\mathbf{B}_{\mathrm{DG_{ng}}}$ is a $(N_{rRF}/N_{rg}) \times N_s$ matrix denoting the digital combiner of the *ngth* group, $\mathbf{B}_{\mathrm{HA_{ng}}}$ is a $(N_r/N_{rg}) \times N_s$ matrix representing the hybrid combining of the *ngth* group and $\mathbf{B}_{\mathrm{G_{ng}}}^{\mathrm{opt}}$ is the optimal unconstrained digital combining of the *ngth* group. Similar to (12), the problem (9) can be rewritten as a series of N_{rg} independent subproblems as

$$\left\|\mathbf{B}^{\mathrm{opt}} - \mathbf{B}_{\mathrm{AHA}} \mathbf{B}_{\mathrm{D}}\right\|_{\mathrm{F}}^2 = \sum_{\mathrm{ng}=1}^{N_{rg}} \left\|\mathbf{B}_{\mathrm{G_{ng}}}^{\mathrm{opt}} - \mathbf{B}_{\mathrm{AG_{ng}}} \mathbf{B}_{\mathrm{DG_{ng}}}\right\|_{\mathrm{F}}^2 \tag{16}$$

and the optimization of each subproblem can be written as

$$\left(\mathbf{B}_{\mathrm{AG_{ng}}}^{\mathrm{opt}}, \mathbf{B}_{\mathrm{DG_{ng}}}^{\mathrm{opt}}\right) = \underset{\mathbf{B}_{\mathrm{AG_{ng}}}, \mathbf{B}_{\mathrm{DG_{ng}}}}{\arg\min} \left\|\mathbf{B}_{\mathrm{G_{ng}}}^{\mathrm{opt}} - \mathbf{B}_{\mathrm{AG_{ng}}} \mathbf{B}_{\mathrm{DG_{ng}}}\right\|_{\mathrm{F}}^2$$

$$\text{st. } \mathbf{B}_{\mathbf{AG_{ng}}} \in \boldsymbol{\xi}_{\mathbf{AHA}},$$
$$\left\|\mathbf{B}_{\mathbf{AG_{ng}}}\mathbf{B}_{\mathbf{DG_{ng}}}\right\|_{\mathrm{F}}^{2} = N_s/N_{rg} \quad (17)$$

In the next section, we will iteratively solve subproblems in (13) and (17).

III. THE PROPOSED HA HYBRID DESIGN ALGORITHMS

This section presents two low-complexity hybrid precoding/combining algorithms for the HA architecture. It derives the equations specifically associated with the precoder, as the procedure for deriving the combiner is similar. The optimal solution for unconstrained semi-unitary digital precoding of the *ngth* group is simply given by $\mathbf{F}_{\mathbf{G_{ng}}}^{\mathbf{opt}} = \mathbf{V}_{\mathbf{G_{ng}}}$, where $\mathbf{V}_{\mathbf{G_{ng}}}$ is the first N_s columns of the *ngth* group of $\mathbf{V}$ that is obtained from the SVD of $\mathbf{H}$. Also, it is essential for the HA hybrid precoder of the *ngth* group, denoted as $\mathbf{F}_{\mathbf{AG_{ng}}}\mathbf{F}_{\mathbf{DG_{ng}}}$, to be sufficiently "close" to the optimal precoder $\mathbf{V}_{\mathbf{G_{ng}}}$ of the *ngth* group by utilizing the digital precoder to form linear combinations of RF precoder vectors. Thus, the solution to (8) can be obtained through the solution of the following N_{tg} subproblems:

$$\left(\mathbf{F}_{\mathbf{AG_{ng}}}^{\mathbf{opt}}, \mathbf{F}_{\mathbf{DG_{ng}}}^{\mathbf{opt}}\right) = \underset{\mathbf{F}_{\mathrm{AG_{ng}}},\mathbf{F}_{\mathbf{DG_{ng}}}}{\arg\ min} \left\|\mathbf{F}_{\mathbf{G_{ng}}}^{\mathbf{opt}} - \mathbf{F}_{\mathbf{AG_{ng}}}\mathbf{F}_{\mathbf{DG_{ng}}}\right\|_{F}^{2}$$
$$st.\ \mathbf{F}_{\mathbf{AG_{ng}}} \in \boldsymbol{\zeta}_{\mathbf{AHA}},$$
$$\left\|\mathbf{F}_{\mathbf{AG_{ng}}}\mathbf{F}_{\mathbf{DG_{ng}}}\right\|_{F}^{2} = N_s/N_{tg} \quad (18)$$

Note that the optimization problem in (18) is non-convex, so the optimal solution is difficult. However, in this paper, we will solve this problem by employing two iterative algorithms, as outlined below.

A. ALGORITHM 1

In this subsection, we will apply the iterative FA (IFA) hybrid design algorithm, as previously introduced in [11], to solve (18). However, it is important to note that while [11] proposed the IFA algorithm for the FA architecture, this paper focuses on the HA architecture. Moreover, in this paper, the suggested algorithm computes hybrid precoding for individual groups first, prior to determining hybrid precoding for the whole system, leading to reduced computational complexity. In comparison, the approach outlined in [11] computes hybrid precoding for the entire system directly, which leads to increased computational complexity.

Algorithm 1 outlines the pseudo-code for the suggested HA hybrid algorithm, employing the iterative solution. The algorithm requires two inputs: $\mathbf{F}_{\mathbf{G_{ng}}}^{\mathbf{opt}} \in \mathbb{C}^{(N_t/N_{tg})\mathrm{x}N_S}$ and K. For the case where $N_s < (N_{tRF}/N_{tg})$, K must satisfy the condition: $K \geq (N_{tRF}/N_{tg}) - N_S$ to calculate the $(N_t/N_{tg})\mathrm{x}(N_{tRF}/N_{tg})\mathbf{F}_{\mathbf{AG_{ng}}}$ matrix. In the scenario where $N_s \geq 1$, the algorithm begins by initializing $\mathbf{F}_{\mathbf{AG_{ng}}}$ based on the element-wise normalization of the first N_s columns of $\mathbf{F}_{\mathbf{G_{ng}}}^{\mathbf{opt}}$ when $N_S \leq (N_{tRF}/N_{tg})$ or the first (N_{tRF}/N_{tg}) when $N_s > (N_{tRF}/N_{tg})$, i.e., $\mathbf{F}_{\mathbf{AG_{ng}}} = \mathbf{F}_{\mathbf{G_{ng}}}^{\mathbf{opt}} \oslash \left(\left|\mathbf{F}_{\mathbf{G_{ng}}}^{\mathbf{opt}}\right| \sqrt{(N_t/N_{tg})}\right)$.

Then, step 2 computes the matrix $\mathbf{F}_{\mathbf{DG_{ng}}}$ using the least squares method. Following this, step 3 calculates the residual precoding matrix $\mathbf{F}_{\mathbf{res}}$, and step 4 calculates the proposed analog precoder $\mathbf{F}_{\mathbf{AG_{ng}}}$. Step 5 guarantees that $\mathbf{F}_{\mathbf{AG_{ng}}}$ meets the requirement of having entries with a consistent magnitude. Steps 7 and 8 address the scenario when $N_s < (N_{tRF}/N_{tg})$. In this case, it becomes necessary to complete the $(N_t/N_{tg})\mathrm{x}(N_{tRF}/N_{tg})$ matrix $\mathbf{F}_{\mathbf{AG_{ng}}}$. This is achieved by adding the element-wise normalization of the first singular vector of the residual matrix $\mathbf{F}_{\mathbf{res}}$ to $\mathbf{F}_{\mathbf{AG_{ng}}}$. The algorithm iterates K times to find $\mathbf{F}_{\mathbf{AG_{ng}}}$ and $\mathbf{F}_{\mathbf{DG_{ng}}}$ such that $\left\|\mathbf{F}_{\mathbf{G_{ng}}}^{\mathbf{opt}} - \mathbf{F}_{\mathbf{AG_{ng}}}\mathbf{F}_{\mathbf{DG_{ng}}}\right\|_{F}^{2}$ is minimized. Finally, step 12 ensures that the proposed HA hybrid precoder, $\mathbf{F}_{\mathbf{HA_{ng}}} = \mathbf{F}_{\mathbf{AG_{ng}}}\mathbf{F}_{\mathbf{DG_{ng}}}$, for the *ngth* group satisfies the transmit power constraint. A similar procedure can be applied to calculate the hybrid combiner for the *ngth* group, $\mathbf{B}_{\mathbf{HA_{ng}}}$.

B. ALGORITHM 2

This subsection presents an effective algorithm for solving the problem defined by equation (18). The optimal precoder structure, $\mathbf{V}_{\mathbf{G_{ng}}}$, for the *ngth* group is non-square semi-unitary. This means that when $N_{tg} = 1$, it satisfies the condition $\mathbf{V}_{\mathbf{Gng}}^{\mathbf{H}}\mathbf{V}_{\mathbf{G_{ng}}} = \mathbf{I}_{\mathbf{N_s}}$, and for $N_{tg} > 1$, it approximately satisfies $\mathbf{V}_{\mathbf{Gng}}^{\mathbf{H}}\mathbf{V}_{\mathbf{G_{ng}}} \approx \mathbf{I}_{\mathbf{N_s}}$. To achieve a similar property for the HA hybrid precoder design, we aim to find $\mathbf{F}_{\mathbf{AG_{ng}}}$ and $\mathbf{F}_{\mathbf{DG_{ng}}}$ such that the condition $\mathbf{F}_{\mathbf{DGng}}^{\mathbf{H}}\mathbf{F}_{\mathbf{AGng}}^{\mathbf{H}}\mathbf{F}_{\mathbf{AG_{ng}}}\mathbf{F}_{\mathbf{DG_{ng}}} = \mathbf{I}_{\mathbf{N_s}}$ is met. This particular structure will be utilized to solve the optimization problem, ensuring that the HA hybrid precoder $\mathbf{F}_{\mathbf{AG_{ng}}}\mathbf{F}_{\mathbf{DG_{ng}}}$ closely approximates the optimal precoder $\mathbf{V}_{\mathbf{G_{ng}}}$. By assuming the structure of $\mathbf{F}_{\mathbf{AG_{ng}}}\mathbf{F}_{\mathbf{DG_{ng}}}$ as a semi-unitary matrix, the following optimization problem needs to be solved:

$$\left(\mathbf{F}_{\mathbf{AG_{ng}}}^{\mathbf{opt}}, \mathbf{F}_{\mathbf{DG_{ng}}}^{\mathbf{opt}}\right)$$
$$= \underset{\mathbf{F}_{\mathrm{AG_{ng}}},\mathbf{F}_{\mathbf{DG_{ng}}}}{\arg\min} \left\|\mathbf{F}_{\mathbf{G_{ng}}}^{\mathbf{opt}} - \mathbf{F}_{\mathbf{AG_{ng}}}\mathbf{F}_{\mathbf{DG_{ng}}}\right\|_{\mathrm{F}}^{2}$$
$$st.\ \mathbf{F}_{\mathbf{AG_{ng}}} \in \boldsymbol{\zeta}_{\mathbf{AHA}},$$
$$\mathbf{F}_{\mathbf{AG_{ng}}}^{\mathbf{H}}\mathbf{F}_{\mathbf{AG_{ng}}} = \mathbf{I}_{(N_{tRF}/N_{tg})} \text{ and } \mathbf{F}_{\mathbf{DG_{ng}}}^{\mathbf{H}}\mathbf{F}_{\mathbf{DG_{ng}}} = \mathbf{I}_{\mathbf{N_s}}$$
$$\left\|\mathbf{F}_{\mathbf{AG_{ng}}}\mathbf{F}_{\mathbf{DG_{ng}}}\right\|_{F}^{2} = N_s/N_{tg} \quad (19)$$

The problem presented in (19) is non-convex, making its mathematical solution challenging. So, in this subsection, we will solve (19) iteratively:

$$\left(\mathbf{F}_{\mathbf{AG_{ng}}}^{\mathbf{opt}}, \mathbf{F}_{\mathbf{DG_{ng}}}^{\mathbf{opt}}\right) = \underset{\mathbf{F}_{\mathrm{AG_{ng}}},\mathbf{F}_{\mathbf{DG_{ng}}}}{\arg\min} \left\|\mathbf{F}_{\mathbf{G_{ng}}}^{\mathbf{opt}} - \mathbf{F}_{\mathbf{AG_{ng}}}\mathbf{F}_{\mathbf{DG_{ng}}}\right\|_{\mathrm{F}}^{2} \quad (20)$$

Our aim is to determine the digital precoder $\mathbf{F}_{\mathbf{DG_{ng}}}$ that minimizes the Euclidean distance. We initiate this process by selecting the analog precoder $\mathbf{F}_{\mathbf{AG_{ng}}}$, which consists of the first N_{tRF}/N_{tg} columns of $\mathbf{F}_{\mathbf{G_{ng}}}^{\mathbf{opt}}$ and then normalizing them to ensure that each entry maintains a constant magnitude, i.e., each entry's squared magnitude $\left|\left[\mathbf{F}_{\mathbf{AG_{ng}}}\right]_{\mathrm{I,j}}\right|^2$ is equal

to $1/\left(N_t/N_{tg}\right)$. Next, we seek to find the $\mathbf{F_{DG_{ng}}}$ such that $\mathbf{F_{AG_{ng}}F_{DG_{ng}}}$ closely approximates $\mathbf{F^{opt}_{G_{ng}}}$. So, the following alternative problem must be solved:

$$\left(\mathbf{F^{opt}_{DGng}}\right) = \arg\min_{\mathbf{F_{DGng}}} \left\|\mathbf{F^{opt}_{G_{ng}}} - \mathbf{F_{AG_{ng}}F_{DG_{ng}}}\right\|_F^2 \tag{21}$$

We can expand (21) as follows:

$$\begin{aligned}&\left\|\mathbf{F^{opt}_{G_{ng}}} - \mathbf{F_{AG_{ng}}F_{DG_{ng}}}\right\|_F^2\\ &= \text{tr}\left(\mathbf{F^{opt^H}_{Gng}F^{opt}_{Gng}}\right) - 2\text{tr}\left(\mathbf{F^{opt^H}_{Gng}F_{AG_{ng}}F_{DG_{ng}}}\right) + \left\|\mathbf{F_{AG_{ng}}F_{DG_{ng}}}\right\|_F^2\\ &= 2N_S - 2\text{tr}\left(\mathbf{F^{opt^H}_{Gng}F_{AG_{ng}}F_{DG_{ng}}}\right)\end{aligned} \tag{22}$$

The solution to this problem, which involves maximizing $\mathbf{F^{opt^H}_{Gng}F_{AG_{ng}}F_{DG_{ng}}}$, is addressed using a method known as the orthonormal Procrustes problem. The solution process is as follows:

$$\mathbf{F_{DG_{ng}}} = \mathbf{VU^H} \tag{23}$$

where $\mathbf{F^{opt^H}_{Gng}F_{AG_{ng}}} = \mathbf{U\Sigma V^H}$. Then, while keeping $\mathbf{F_{DG_{ng}}}$ fixed, we proceed to solve the same optimization problem.However, this time, we minimize over $\mathbf{F_{AG_{ng}}}$, which can be achieved as follows [35]:

$$\left(\mathbf{F^{opt}_{AGng}}\right) = \arg\min_{\mathbf{F_{AGng}}} \left\|\mathbf{F^{opt}_{G_{ng}}} - \mathbf{F_{AG_{ng}}F_{DG_{ng}}}\right\|_F^2 \tag{24}$$

Similar to (22), the objective function can be expressed as

$$\left\|\mathbf{F^{opt}_{G_{ng}}} - \mathbf{F_{AG_{ng}}F_{DG_{ng}}}\right\|_F^2 = 2N_S - 2\text{tr}\left(\mathbf{F^{opt^H}_{Gng}F_{AG_{ng}}F_{DG_{ng}}}\right) \tag{25}$$

Under the assumption that $\mathbf{F_{AG_{ng}}}$ is a semi-unitary matrix, we can write the solution to maximize $\mathbf{F^{opt^H}_{Gng}F_{AG_{ng}}F_{DG_{ng}}}$ in (24) as follows:

$$\mathbf{F_{AG_{ng}}} = \mathbf{VU^H} \tag{26}$$

where $\mathbf{F_{DG_{ng}}F_{Gng}}^{\mathbf{opt^H}} = \mathbf{U\Sigma V^H}$. Also, there is another approach to maximize $\mathbf{F^{opt^H}_{Gng}F_{AG_{ng}}F_{DG_{ng}}}$ in (25) as follows:

$$\mathbf{F_{AG_{ng}}} = \mathbf{F^{opt}_{Gng}F^H_{DGng}} \tag{27}$$

Both solutions in (26) and (27) are quite similar due to our assumption regarding $\mathbf{F_{DG_{ng}}}$. Additionally, $\mathbf{F^{opt^H}_{Gng}F^{opt}_{G_{ng}}} = \mathbf{I_{N_s}}$ when $N_{tg} = 1$, and $\mathbf{F^{opt^H}_{Gng}F^{opt}_{G_{ng}}} \approx \mathbf{I_{N_s}}$ when $N_{tg} > 1$ in (24). The primary distinction between the proposed hybrid design in this paper and the design presented in [6] lies in our assumptions about $\mathbf{F_{AG_{ng}}}$ and $\mathbf{F_{DG_{ng}}}$. Specifically, we assume $\mathbf{F_{AG_{ng}}}$ as a non-square semi-unitary matrix, and $\mathbf{F_{DG_{ng}}}$ as either a non-square semi-unitary or square unitary.

In contrast, the design in [6] is limited to situations where $N_{tRF} = N_s$ and it is only applicable in the FA architecture. Moreover, our proposed design is applicable in all scenarios, regardless of whether $N_{tRF} = N_s$ or $N_{tRF} < N_s$. This flexibility allows for wider adaptability to different communication systems. On the other hand, the design in [6] is restricted to scenarios where $N_{tRF} = N_s$, limiting its usability in more diverse setups. Furthermore, our proposed hybrid design is derived for HA architecture, while the design in [6] is exclusively derived for the FA architecture. This distinction highlights the compatibility of our design with the HA architecture, enabling it to address the specific requirements and characteristics of such systems.

Lastly, our proposed algorithm exhibits simplicity by independently computing the hybrid precoding submarices for each group and then using these submarices to obtain hybrid precoding matrix for the whole HA system. In contrast, the approach described in [6] directly calculates the hybrid precoding matrix for the whole system, which leads to increased computational complexity, as discussed in the next section.

The proposed HA precoder algorithm using Algorithm 2 is described as follows. The algorithm takes several inputs, including $\mathbf{F^{opt}_{G_{ng}}}$, K and the initialized $\mathbf{F_{AG_{ng}}}$, where $K \geq 1$. In the general scenario when $N_s \geq 1$, the algorithm initiates by calculating the *ngth* group $\mathbf{F_{DG_{ng}}}$ as indicated in step 2. Following this, in steps 3, the algorithm updates the *ngth* group RF precoder $\mathbf{F_{AG_{ng}}}$. In order to ensure satisfaction with constant-magnitude entries applicable at RF

Algorithm 1 Proposed HA Hybrid Precoding

Input	Input the optimal unconstrained digital precoding $\mathbf{F^{opt}_{Gng}}$ of all groups, and K.
Algorithm	Initialization: $\mathbf{F_{AGng}} = \mathbf{F^{opt}_{Gng}} \oslash \left(\left\|\mathbf{F^{opt}_{Gng}}\right| \sqrt{(N_t/N_{tg})}\right)$. 1: for i= 1 :K do 2: Update: $\mathbf{F_{DGng}} = (\mathbf{F^H_{AGng}F_{AGng}})^{-1}\mathbf{F^H_{AGng}F^{opt}_{Gng}}$ 3: Update the residual: $\mathbf{F_{res}} = \mathbf{F^{opt}_{Gng}} - \mathbf{F_{AGng}F_{DGng}}$ 4: Update: $\mathbf{F^{i+1}_{AGng}} = \mathbf{F_{res}F^H_{DGng}} + \mathbf{F^i_{AGng}}$ 5: Normalize: $\mathbf{F^{i+1}_{AGng}} = \mathbf{F^{i+1}_{AGng}} \oslash \left(\left|\mathbf{F^{i+1}_{AGng}}\right| \sqrt{(N_t/N_{tg})}\right)$ 6: If $i \leq N^{\mathbf{BS}}_{\mathbf{RF}} - N_S$ 7: $\mathbf{F_{res}} = \mathbf{U\Sigma V^H}$ 8:Add a new column to $\mathbf{F^{i+1}_{AGng}}$: $\mathbf{F^{i+1}_{AGng}} = [\mathbf{F^{i+1}_{AGng}}(\mathbf{U})_1 \oslash \left(|(\mathbf{U})_1| \sqrt{(N_t/N_{tg})}\right)]$ 9: end if 10: end for 11: $\mathbf{F_{DGng}} = (\mathbf{F^H_{AGng}F_{AGng}})^{-1}\mathbf{F^H_{AGng}F^{opt}_{Gng}}$ 12: $\mathbf{F_{DGng}} = \sqrt{N_S/N_{tg}}\,\dfrac{\mathbf{F_{DGng}}}{\left\|\mathbf{F_{AGng}F_{DGng}}\right\|_F}$
Output	13. Return $\mathbf{F_{HAng}} = \mathbf{F_{AGng}F_{DGng}}$.

Algorithm 2 Proposed HA Hybrid Precoding

Input	Input the optimal unconstrained digital precoding $\mathbf{P^{opt}_{Gng}}$ of all groups, K and Initialize $\mathbf{P_{AGng}}$
Algorithm	1: for i= 1 :K do 2: Update: $\mathbf{F_{DGng}} = \mathbf{VU^H}$, where$\mathbf{F^{opt^H}_{Gng}F_{AGng}} = \mathbf{U\Sigma V^H}$ 3: Update: $\mathbf{F_{AGng}} = \mathbf{F^{opt}_{Gng}F^H_{DGng}}$ 4: Normalize: $\mathbf{F_{AGng}} = \mathbf{F_{AGng}} \oslash \left(\left|\mathbf{F_{AGng}}\right| \sqrt{(N_t/N_{tg})}\right)$ 5: end for 6: $\mathbf{F_{DGng}} = \mathbf{F^H_{AGng}F^{opt}_{Gng}}$ 7: $\mathbf{F_{DGng}} = \sqrt{N_S/N_{tg}}\,\dfrac{\mathbf{F_{DGng}}}{\left\|\mathbf{F_{AGng}F_{DGng}}\right\|_F}$
Output	8. Return $\mathbf{F_{HAng}} = \mathbf{F_{AGng}F_{DGng}}$.

using analog phase shifters, step 4 is employed to achieve this condition for the proposed RF precoder $\mathbf{F}_{\mathbf{AG_{ng}}}$. After completing the final iteration of the algorithm, the update for $\mathbf{F}_{\mathbf{DG_{ng}}}$ is calculated using a maximal ratio combining (MRC) method instead of the least square method. This impacts the objective function $\left\|\mathbf{F}^{\mathbf{opt}}_{\mathbf{G_{ng}}} - \mathbf{F}_{\mathbf{AG_{ng}}}\mathbf{F}_{\mathbf{DG_{ng}}}\right\|^2_F$. Specifically, the least solution transforms into MRC after the implementation of the semi-unitary analog precoder, represented as $\mathbf{F}^{\mathbf{H}}_{\mathbf{AG_{ng}}}\mathbf{F}_{\mathbf{AG_{ng}}} = \mathbf{I}_{(\mathbf{N_{tRF}}/\mathbf{N_{tg}})}$. Once the algorithm completes K iterations, it yields the proposed analog precoding matrix $\mathbf{F}_{\mathbf{AG_{ng}}}$ and the digital precoder $\mathbf{F}_{\mathbf{DG_{ng}}}$. The algorithm ensures that these matrices minimize the objective function $\left\|\mathbf{F}^{\mathbf{opt}}_{\mathbf{G_{ng}}} - \mathbf{F}_{\mathbf{AG_{ng}}}\mathbf{F}_{\mathbf{DG_{ng}}}\right\|^2_F$. In steps 7 and 8, the transmit power constraint operation is applied for each *ngth* group and the proposed hybrid analog precoder $\mathbf{F}_{\mathbf{HA_{ng}}} = \mathbf{F}_{\mathbf{AG_{ng}}}\mathbf{F}_{\mathbf{DG_{ng}}}$ for the *ngth* group is computed and returned.

Note that the mmWave channel for both the FA and the proposed HA hybrid beamforming techniques is characterized as a sparse low-rank channel. The FA hybrid beamforming technique was specifically developed taking into account the sparsity of the channel, as described in [4]. In our proposed HA hybrid beamforming technique, we adopted the same objective function proposed in [4] and devised an iterative solution to address it.

IV. COMPLEXITY ANALYSIS OF THE PROPOSED ALGORITHM

This section analyzes the complexity of the suggested hybrid design algorithms 1 and 2. For the sake of simplicity, let us use $N = \max\{N_t, N_r\}$, $N_{\mathrm{RF}} = \max\{N_{tRF}, N_{rRF}\}$, $N_{\mathrm{g}} = \max\{N_{tg}, N_{rg}\}$ and K represents the maximum number of iterations for the suggested algorithms as well as for those algorithms in [6], [11], and [17]. The number of antennas for each subarray in the SA architecture is denoted by N_{SA}. In this section, the total number of floating-point operations (flops) for each hybrid design method are evaluated and compared.

Table 1 shows that the complexity of the suggested hybrid design in Algorithms 1 and 2, as well as those in [6], [11], and [17], is lower in comparison to the complexity of the sparse algorithm in [4].

Moreover, the computational complexity of the suggested HA hybrid design in Algorithms 1 and 2, as well as that in [17], is lower compared to those hybrid design algorithms in [6] and [11], especially for a large N_{g}.

When considering hardware complexity, the suggested hybrid designs in Algorithms 1 and 2 result in reduced hardware complexity in contrast to those in [4], [6], and [11] when $N_{\mathrm{g}} > 1$. Specifically, the number of required PSs in the proposed algorithms depends on the number of groups, equal to $(N_{\mathrm{RF}}/N_{\mathrm{g}})N$, whereas the other hybrid designs in [4], [6], and [11] require NN_{RF}. Moreover, suggested hybrid design algorithms for the HA architecture have the same hardware complexity as that in the SA architecture when $N_{\mathrm{RF}} = N_{\mathrm{g}}$. Based on our analysis, we can conclude that the suggested hybrid design algorithms for the HA architecture exhibits reduced computational and hardware complexities compared to those for the FA architecture and, when $N_{\mathrm{RF}} = N_{\mathrm{g}}$, demonstrates similar complexity to those for the SA architecture.

TABLE 1. Complexity of different hybrid design algorithms.

Algorithm	Constraints	PSs Number	Computational Complexity
Sparse Algorithm [4]	RF codebooks	NN_{RF}	$\mathrm{O}(N^2 N_{RF} N_S)$
HD-AM Algorithm [6]	$N_S = N_{\mathrm{RF}}$	NN_{RF}	$\mathrm{O}(NN_S^2 K)$
IFA Algorithm [11]	None	NN_{RF}	$\mathrm{O}(NN_{RF}^2 K)$
SA Algorithm [17]	None	N	$\mathrm{O}(N_{\mathrm{SA}} N_{\mathrm{RF}} N_S K)$
Proposed Algorithm 1	None	$(N_{\mathrm{RF}}/N_{\mathrm{g}})N$	$\mathrm{O}(NN_{RF}^2 K/N_g^2)$
Proposed Algorithm 2	None	$(N_{\mathrm{RF}}/N_{\mathrm{g}})N$	$\mathrm{O}(NN_{\mathrm{RF}} N_S K/N_g^2)$

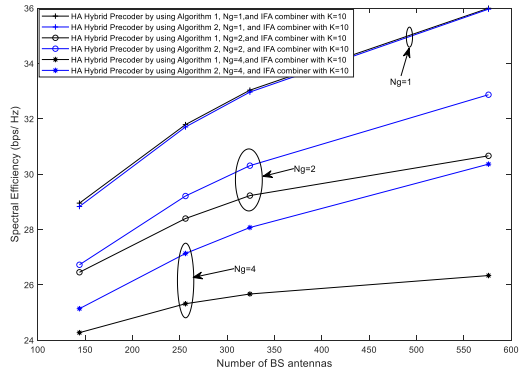
V. SIMULATION RESULTS

In this section, we present simulation results that demonstrate the effectiveness of the suggested Hybrid design algorithms for the HA architecture. These simulations are carried out using the HA architecture shown in Fig. 1. We consider an outdoor scenario with a single BS and a single MS located 100 meters apart, a carrier frequency of 28 GHz, an antenna element spacing of $\lambda/2$, and a path loss exponent of $\mathrm{n} = 3.4$. The mmwave channel model in (2) is used, where $\bar{P_{\alpha,i}} = 1$ holds for all clusters. The angles of arrival (AoAs) and angles of departure (AoDs) for each ray within the clusters follow a random Laplacian distribution. Both the azimuth and elevation angles of the cluster means for both AoA and AoD are assumed to exhibit uniform distributions. In the case of the sparse hybrid design, we utilize the AoD/AoA beamforming codebooks at the BS and MS, respectively. The signal-to-noise ratio (SNR) is defined as $\mathrm{SNR} = \rho/\sigma^2$. in all the depicted plots. We make the assumption of perfect channel estimation and the same total power constraint for all algorithms.

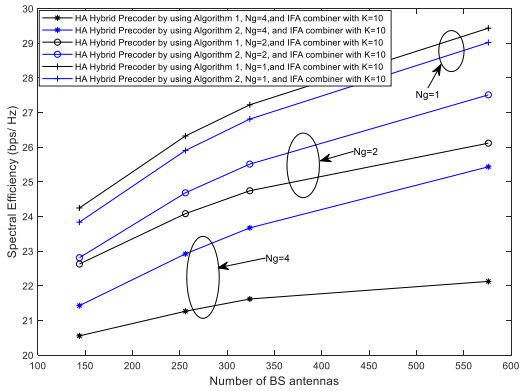
In the next subsection, we will study and compare the performance of the suggested HA algorithms 1 and 2. Based on the obtained results, we will choose the best algorithm and compare it with other existing hybrid designs in the literature in the following subsections.

A. COMPARISON OF HA HYBRID PRECODING DESIGNS: ALGORITHM 1 VS. ALGORITHM 2

Figures 4(a) and (b) plot the spectral efficiency of the suggested hybrid precoding algorithms with an IFA combiner

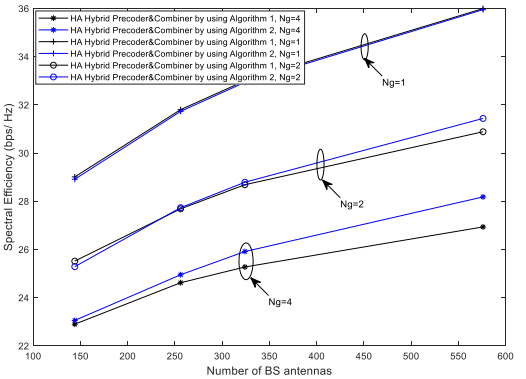


(a) $N_s = 4$

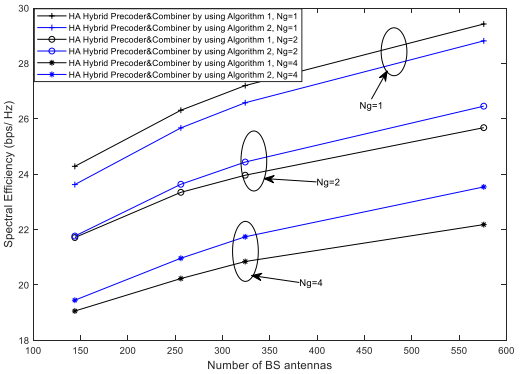


(b) $N_s = 3$

FIGURE 4. Achievable rates using the proposed HA hybrid design with Algorithms 1 and 2 as a precoder and IFA as a combiner for different numbers of BS antennas for $N_r = 16$ with SNR = 0 dB, and $N_{tRF} = N_{rRF} = 4$ and different N_s.



(a) $N_s = 4$



(b) $N_s = 3$

FIGURE 5. Achievable rates using the proposed HA hybrid design with Algorithms 1 and 2 as precoders and combiners for different numbers of BS antennas $N_r = 16$ with SNR = 0 dB, and $N_{tRF} = N_{rRF} = 4$ and different N_s.

for different numbers of BS antennas in a mmWave system when $N_r = 16$, SNR = 0dB, $N_s \in \{3, 4\}$,$N_{tRF} = N_{rRF} = 4$, and $N_g \in \{1, 2, 4\}$. We assume $K = 10$ for the proposed precoding algorithms 1 and 2 as well as for the IFA combining for all cases. As we see in Figs. 4 (a) and (b), the spectral efficiency of the proposed algorithms improves as the number of BS antennas increases. However, both algorithms experience degradation asN_g increases. This is because $N_g$controls the number of PSs, resulting in higher interference between data streams. When $N_g = 1$, the architecture is the same as the FA architecture with the number of PSs equal to $N_{tRF}N_t$, but when $N_g = 4$, we have the SA architecture with the number of PSs equal to N_t. Moreover, the performance of Algorithm 2 outperforms that of Algorithm 1 when $N_g > 1$. We are only interested in the case where $N_g > 1$, as that is where the HA hybrid design exists. The performance gap between the HA hybrid designs using both algorithms increases with the number of groups N_g. The only case where Algorithm 1 outperforms Algorithm 2 is when $N_g = 1$ (FA architecture). This is due to the non-square semi-unitary constraint, i.e., $\mathbf{F}^{\mathbf{H}}_{\mathbf{DGng}}\mathbf{F}^{\mathbf{H}}_{\mathbf{AGng}}\mathbf{F}_{\mathbf{AG_{ng}}}\mathbf{F}_{\mathbf{DG_{ng}}} = \mathbf{I}_{\mathbf{N_s}}$, which is applied in Algorithm 2 but does not work well for the FA architecture. However, this constraint helps to build the block diagonal matrix structure of the proposed HA hybrid precoder, which is why Algorithm 2 outperforms Algorithm 1 when $N_g > 1$. Moreover, in Fig. 4 (a) when $N_g = 1$, the performance gap between Algorithms 1 and 2 is very small compared to the gap in Fig. 4 (b). This is because $N_s = N_{tRF} = N_{rRF} = 4$, which makes $\mathbf{F}_{\mathbf{DG_{ng}}}$ approximately unitary in Algorithm 1, resulting in almost the same performance between the two algorithms.

Figs. 5(a) and (b) show the spectral efficiency achieved by the proposed Algorithms 1 and 2 in a mmWave system with different number of the BS antennas when $K = 10$, $N_r = 64$, SNR = 0dB, $N_s \in \{3, 4\}$, $N_{tRF} = N_{rRF} = 4$, and $N_g \in \{1, 2, 4\}$. In Figs. 5 (a) and (b), we used the proposed algorithms 1 and 2 at both the BS as a precoder and at the MS as a combiner. From these figures, it is clear that the achievable performance is degraded compared to the results obtained in Figs. 4(a) and (b). This degradation is due to the use of the HA hybrid design in both the BS and MS, whereas the IFA combiner was used in Figs. 4 for the MS. Also, the performance of both algorithms decreases as N_g increases.

This is because increasing N_g directly increases the number of PSs, which leads to increased interference between data streams. Fig. 5 also shows that, similar to Fig. 4, the spectral efficiency of the proposed Algorithm 2 is also better than that of the proposed Algorithm 1 when $N_g > 1$, especially when the number of BS antennas is high. The performance gap between the proposed Algorithms is reduced compared to the

results in Figs. 4 (a) and (b) because Algorithm 2 experiences more degradation than Algorithm 1 when the HA is applied in the MS. However, Algorithm 2 still outperforms Algorithm 1 in Figs. 5 when $N_g > 1$, especially with a high number of BS antennas.

In conclusion, although both algorithms yield a good performance, Algorithm 2 outperforms Algorithm 1 in all scenarios when $N_g > 1$, especially for a high number of BS antennas.

The crucial factor in achieving superior performance lies in utilizing the inherent block diagonal structure within the hybrid analog and digital precoding/combining matrices in Algorithm 2. This advantage enables Algorithm 2 to outperform Algorithm 1 when $N_g > 1$. Because we are only interested in the case where $N_g > 1$ where the structure of the HA is applied, we choose Algorithm 2 to collect more results and conduct comparisons against other methods in the literature.

In the subsequent two subsections, we will investigate and analyze Algorithm 2 under two distinct scenarios. In the first scenario, Algorithm 2 will be applied as a precoder at the BS, while the MS will employ the IFA algorithm from [11] as a combiner. In the second scenario, Algorithm 2 will be utilized at both the BS and MS. To ensure a fair evaluation, we will follow the same procedure for the SA hybrid algorithm outlined in [17].

B. FIRST SCENARIO

Fig. 6 displays the achieved spectral efficiency versus the SNR of the proposed HA hybrid design, implemented using Algorithm 2 along with the IFA hybrid combiner. This performance is compared to that of previous hybrid designs in [4], [11], and [17]. For our analysis, we assume a 256 × 64 mmWave system, $N_s \in \{2, 8\}$,$N_{tRF} \in \{4, 16\}$, $N_{rRF} \in \{2, 8\}$, and $N_g \in \{1, 2, 4, 8, 16\}$. K is set to 10 for all algorithms. Fig. 6 shows that the proposed design in this scenario provides superior performance compared to both the SA hybrid design and the FA sparse hybrid design, regardless of N_s and N_g. The performance of the proposed HA hybrid design with Algorithm 2 exhibits deterioration as N_g increases due to the increasing in interference among data streams, which corresponds to a reduction in the number of PSs. The results also indicate that, even with an increase in N_g, the proposed hybrid design consistently outperforms the SA hybrid design in [17]. Furthermore, in scenarios where $N_g = 1$, the performance of the HA hybrid design is comparable to that of the IFA hybrid precoder/combiner.

Fig. 6 shows that the proposed HA hybrid design using Algorithm 2 offers a key advantage by bridging the gap between SA and FA hybrid design approaches. They provide a valuable trade-off between performance, hardware cost, and power consumption, which can be adjusted based on the number of groups utilized. Consequently, the performance of the proposed Algorithm 2 in this scenario is slightly lower than that of the IFA algorithm in [11] but significantly better than that of SA algorithm in [17], while requiring a much lower

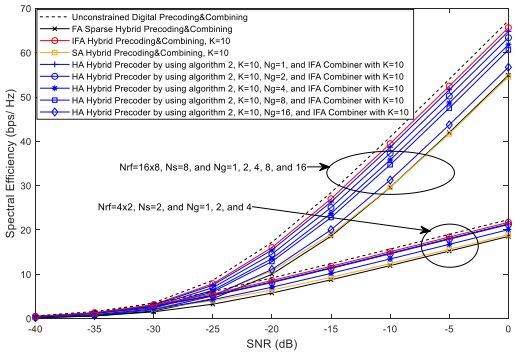


FIGURE 6. Achievable rates using hybrid precoding/combining algorithms for different values of SNR in a 256 × 64 mmWave system, with $N_s \in \{2, 8\}$, $N_{tRF} \in \{4, 16\}$, and $N_{rRF} \in \{2, 8\}$.

hardware cost, and power consumption than that in [11]. For example, in Fig. 6 for SNR = 0 dB, the proposed HA hybrid design demonstrates a spectral efficiency performance loss of approximately 1 dB (HA requires $N_{tRF}N_t$ /2 PSs) compared to the IFA architecture in [11] (which requires $N_{tRF}N_t$ PSs) when employing 2 groups. This indicates that the proposed HA hybrid design reduces the number of PSs by half compared to that of the FA. Conversely, in the same figure for SNR = 0 dB, the suggested Algorithm 2 demonstrates a significant enhancement in performance, approximately 5 dB higher than the SA algorithm presented in [17].

Fig. 7 illustrates the achievable rate versus the SNR of the proposed hybrid design in scenario 1, and compares it with the hybrid design algorithms in [4], [11], and [17]. We consider a 64 × 16 mmWave system with $N_s \in \{2, 4\}$,$N_{tRF} \in \{4, 8\}$, $N_{rRF} \in \{2, 4\}$, and$N_g \in \{1, 2, 4, 8\}$. We use $K = 10$ for all algorithms. It is evident that the proposed HA hybrid design provides better performance than FA algorithm in [4], regardless of the values of N_s and N_g. Moreover, the proposed HA hybrid precoder using Algorithm 2 outperforms the SA algorithm in [17] when $N_g < 8$, and $N_g < 4$ for $N_s = 4$,and 2, respectively; the proposed HA hybrid precoder using Algorithm 2 with $N_g = 4$and8 overlaps with the SA hybrid precoder algorithm in [17] when the number of antennas at the BS and MS is decreased in comparison to the results shown in Fig. 6. As in Fig. 6, the spectral efficiency performance of the proposed Algorithm 2 in Fig. 7 deteriorates as N_g, increases. This corresponds to a reduction in the number of PSs, leading to increased interference among data streams.

In Fig. 8, the spectral efficiency of the proposed hybrid design is evaluated for different number of RF chains and compared with the other algorithms when, $N_{tRF} = N_{rRF}$, is greater than N_s. We assume a 256 × 64 mmWave system, $N_s \in \{2, 4\}$, $N_g \in \{1, 2, 4, 8\}$, and $SNR = 0dB$. $K = 10$ is used for all iterative algorithms. Fig. 8 demonstrates that the performance of the proposed hybrid design in this scenario degrades as N_g increases due to a reduction in the number of PSs, regardless of number of N_s and RF chains. Additionally, it is clear that the proposed HA hybrid design consistently exhibits superior performance compared to the FA algorithm

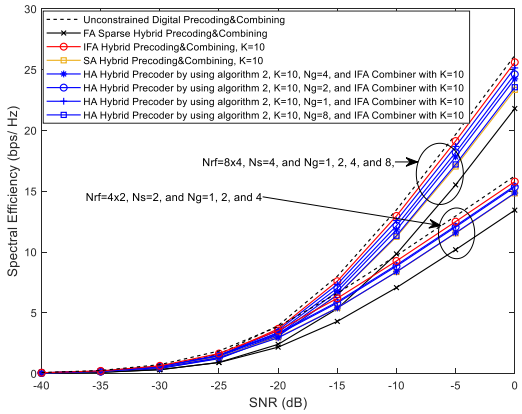


FIGURE 7. Achievable rates using hybrid precoding/combining algorithms for different values of SNR in a 256 × 64 mmWave system.

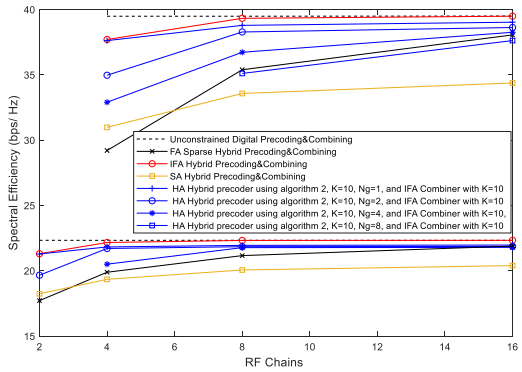


FIGURE 8. Achievable rates using hybrid precoding/combining algorithms for different values of RF chains in a 256 × 64 mmWave system, with $N_s \in \{2, 4\}$, and SNR = 0 dB.

in [4] and the SA algorithm in [17] when $N_s = 2$, regardless of the value of N_g.

The performance gap between the proposed HA hybrid design and the FA sparse hybrid design in [4] reduces as the number of RF chains increases, $N_{tRF} = N_{rRF}$, because both methods take advantage of the increased number of analog precoder columns. When $N_s = 4$, the proposed HA hybrid precoder using Algorithm 2 provides better performance than the SA algorithm in [17], regardless of the value N_g. Also, the proposed HA hybrid design outperforms the hybrid design in [4] when $N_g \in \{1, 2, 4\}$ and is close to it when $N_g = 8$. The performance of the proposed HA hybrid precoder using Algorithm 2 deteriorates as N_g increases, in contrast to the FA sparse hybrid algorithm in [4] and the optimal hybrid algorithm. This degradation is attributed to the increased interference between data streams. Furthermore, when $N_g = 1$, the suggested HA hybrid design in this scenario provides performance close to that in [11] and can be the same if we use Algorithm 1, but we only focus in this subsection on the HA hybrid precoder using Algorithm 2 and $N_g > 1$. When $N_s = N_{tRF} = N_{tRF}$, Fig. 9 plots the spectral efficiency versus the number of RF chains of the proposed HA hybrid precoder using algorithm 2 with IFA hybrid combiner, and compares it with that of the existing hybrid design algorithms in [4], [11], and [17], and the optimal digital design. We assumed a 256 × 64 mmWave systems, $N_s = N_{tRF} = N_{rRF}$, $N_g \in \{1, 2, 4, 8\}$, $K = 10$, and $SNR = 0dB$. Fig. 9 shows that the proposed HA algorithm yields an improvement over the FA hybrid algorithm in [4] and the SA hybrid algorithm in [17], regardless of the values of N_g. When $N_g = 1$, the performance of the proposed HA hybrid design overlaps with that in [11].

To summarize, the constraints on the analog and digital precoding matrices facilitates the creation of a block diagonal structure matrices. This structural advantage contributes to the superiority of the proposed HA hybrid design, which employs Algorithm 2 as a precoder at the BS, along with the IFA hybrid combiner at the MS, compared to the FA hybrid design in [4] and the SA hybrid design in [17].This enhanced performance is particularly evident when $N_g < N_{tRF}$ and for any value of N_s, especially when the number of antennas at the BS antennas is high. Moreover, using the IFA hybrid combiner in the receiver side improves the overall systems performance of HA Architecture. When $N_g = N_{tRF}$, structure of the proposed HA becomes the same as that of the SA; the performance of the proposed HA hybrid precoder using Algorithm 2 might overlap with the performance of the FA algorithm in [4] and the SA algorithm in [17], especially when the number of antennas at the BS is high. It is advisable to consider a value of $K \leq 10$ for the proposed Algorithm 2, as the incremental gain beyond this threshold is negligible. This observation is confirmed by simulation results that are not included in this paper. When $N_g = 1$, the proposed HA structure becomes similar to the FA structure; the performance of the proposed HA hybrid precoder using Algorithm 2 becomes close to that of the IFA hybrid design in [11] and they are identical when Algorithm 1 is used because they are the same algorithm.

C. SECOND SCENARIO

Fig. 10 depicts the spectral efficiency performance achieved by the proposed HA hybrid design using Algorithm 2 at both the BS and MS and compares it with those of the hybrid design algorithms in [4], [11], and [17] as well as with optimal digital design for different SNR values. The results are conducted in a 256 × 64 mmWave system with $N_s \in \{2, 8\}$, $N_{tRF} = N_{rRF} \in \{4, 16\}$, and $N_g \in \{1, 2, 4, 8, 16\}$. A value of $K = 10$ is assumed for the proposed Algorithm 2, as well as for algorithms in [11] and [17]. The results clearly demonstrate that the proposed hybrid design method in this scenario achieves superior performance compared to the SA algorithm in [17], regardless the values of N_s and N_g, and outperforms the FA sparse algorithm in [4] when $N_g \in \{1, 2\}$ for $N_s = 2$ and 8. The performance of the proposed HA design degrades with an increase in N_g. This is attributed to a reduction in the number of PSs, leading to higher interference between data streams. However, when $N_g = 16$, the architecture of HA becomes the same as the SA architecture, and the proposed design using Algorithm 2 outperforms the SA hybrid design algorithm in [17]. Additionally, when $N_g = 1$, the HA architecture is similar to the FA architecture and

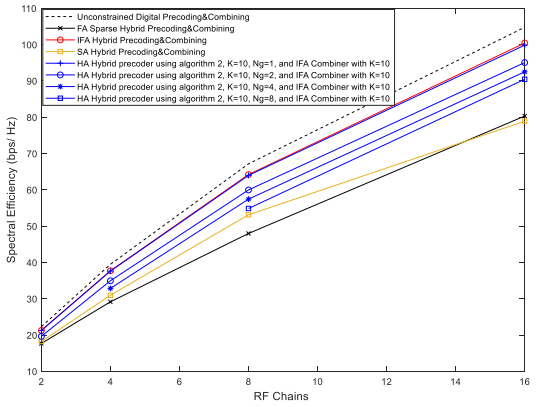


FIGURE 9. Achievable rates using hybrid precoding/combining algorithms for different values of RF chains in a 256 × 64 mmWave system.

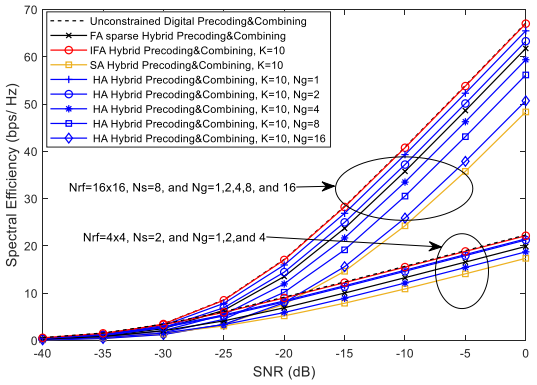


FIGURE 10. Achievable rates using hybrid precoding/combining algorithms for different values of SNR in a 256 × 64 mmWave system., with $N_S \in \{2, 8\}$, *and* $N_{tRF} = N_{rRF} \in \{4, 16\}$.

the proposed hybrid design using Algorithm 2 achieves a performance close to that in [11], which is also comparable to the optimal design.

Fig. 11 evaluates the performance of the proposed HA hybrid precoder/combiner using Algorithm 2, the optimal design algorithm, and the other hybrid design algorithms presented in [4], [11], and [17] at different values of the SNR.

The analysis is carried out in a 64 × 16 mmWave system, with parameters$N_S \in \{2, 4\}$,$N_{tRF} = N_{rRF} \in \{4, 8\}$, and $N_g \in \{1, 2, 4, 8\}$. A constant value of $K = 10$ is assumed for the proposed Algorithm 2, as well as for the iterative algorithms in [11] and [17]. Fig. 11 show the superiority of the proposed hybrid design over [17] when $N_g \in \{1, 2, 4\}$, for $N_S = 4$and $N_S = 2$, respectively. Also, the proposed hybrid design overlaps with that in [17] when $N_g = 8$ for $N_S = 2$and 4, due to the reduction in the number of antennas for both the BS and the MS. Additionally, it outperforms the FA hybrid algorithm in [4] when $N_g \in \{1, 2\}$for $N_S = 2$ and 4.

Similar to the previous scenario, we will now consider a case where $(N_{tRF} = N_{rRF}) > N_S$, and the results of the spectral efficiency at different RF chains will be plotted in Fig. 12. We assume $N_S \in \{2, 4\}$, $N_g \in \{1, 2, 4, 8\}$, $K = 10$, and the $SNR = 0dB$ in a 256 × 64 mmWave systems. The results demonstrate that the performance of the proposed

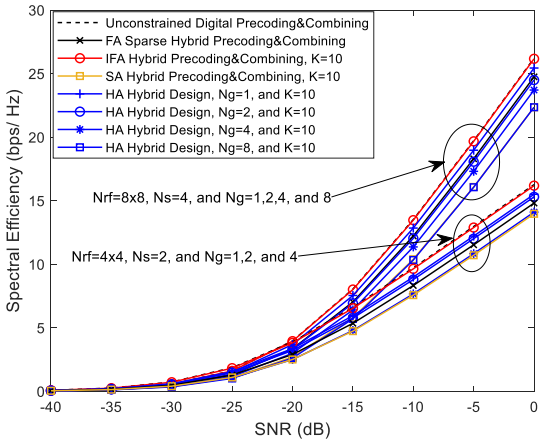


FIGURE 11. Achievable rates using hybrid precoding/combining algorithms for different values of SNR in a 64 × 16 mmWave systems, with $N_S \in \{2, 4\}$, *and* $N_{tRF} = N_{rRF} \in \{4, 8\}$.

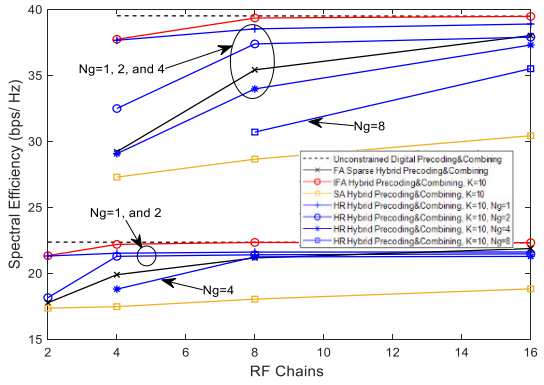


FIGURE 12. Achievable rates using hybrid precoding/combining algorithms for different values of RF chains in a 256 × 64 mmWave system., with SNR = 0 dB and $N_S \in \{2, 4\}$.

hybrid design in this scenario improves with an increase in the number of the RF chains. The performance gap between the proposed scheme and that of FA in [11] is reduced with an increase in the number of RF chains, even with a high value of N_g.

Fig. 13 illustrates the performance of the proposed HA hybrid design in this scenario, along with the hybrid design algorithms in [4], [11], and [17], as well as the optimal digital design algorithm when $N_s = N_{tRF} = N_{rRF}$ in a 256×64 mmWave systems. We assume SNR = 10 dB, $N_g \in \{1, 2, 4, 8\}$, and $K = 10$. The spectral efficiency performance of the proposed design, as well as the hybrid design algorithms in [4], [11], and [17], is clearly observed to improve with an increase in the number of RF chains.

Note that the proposed HA hybrid design using Algorithm 2 effectively bridges the gap between the SA and the FA designs, particularly in scenarios with a high number of RF chains. Results also show that the performance of the proposed HA design using Algorithm 2 becomes better than that of the sparse algorithm in [4], and iterative algorithm in [17] when the number of RF chains equals 16, regardless of the value of N_g.

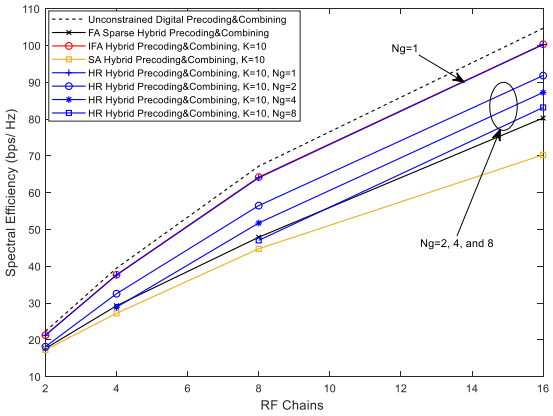


FIGURE 13. Achievable rates using hybrid precoding/combining algorithms for different values of RF chains in a 256 × 64 mmWave system, with SNR = 0 dB and $N_s = N_{tRF} = N_{tRF}$.

In summary, the performance of the suggested hybrid design using Algorithm 2 at the BS for precoding and at the MS for combining proves satisfactory. Its computational and hardware complexities are also lower compared to the hybrid designs outlined in [4] and [11]. Conversely, the proposed HA architecture requires $(\frac{N_t N_{tRF}}{N_g}) + (\frac{N_r N_{rRF}}{N_g})$ PSs, while the FA architecture requires $N_t N_{tRF} + N_r N_{rRF}$ PSs. When N_g equals the number of RF chaines, the structure of the proposed HA design using Algorithm 2 becomes similar to the SA structure and outperforms the SA hybrid design algorithm in [17], especially when the number of antennas is high. When N_g is equal to 1, the structure of the proposed HA design using Algorithm 2 becomes similar to that of the FA architecture and provides performance close to the IFA hybrid design algorithm in [11].

VI. CONCLUSION

In this paper, a new HA architecture is introduced recent and upcoming wireless generations, aiming to achieve a balance between hardware cost, power consumption, and spectral efficiency. The HA architecture divides the antennas into subarrays, followed by grouping them, and subsequently establishes connections to RF chains in a manner similar to that used in the FA architecture. We also proposed two iterative algorithms for single user hybrid precoding/combining for HA architecture. The Numerical results demonstrate that the proposed Algorithms 1 and 2 can enhance the spectral efficiency of the HA architecture while maintaining low complexity. Furthermore, the results indicate that Algorithm 2 yields superior performance compared to Algorithm 1, as well as the currently existing SA algorithm in [17] and FA algorithm in [4]. This performance advantage is particularly noticeable when N_g is set to 2. Furthermore, it has been observed that the proposed Algorithm 2 outperforms the performance of the FA algorithm in [4] and the SA algorithm in [17], particularly in scenarios where N_g is less than N_{tRF} for any N_s, especially when the number of antennas at the BS is high.

When $N_g = N_{tRF}$, the structure of the HA architecture is similar to that of SA and proposed Algorithm 2 outperforms the FA algorithm in [4] and the SA algorithm in [17], regardless of the values of N_g, especially for a large number of $N_s = N_{tRF} = N_{rRF}$. When $N_g = 1$, the structure of the HA architecture is similar to that of FA and proposed Algorithm 2. In this scenario, the proposed Algorithm 2 demonstrates superior performance compared to both the FA algorithm in [4] and the SA algorithm in [17], while achieving performance that is comparable to that of the IFA in [11]. The value for K should be chosen as 10 in both proposed Algorithm 1 and Algorithm 2, as any increment beyond this point would only yield marginal gains. In summary, the proposed HA hybrid design using Algorithms 1 or 2 offers an efficient and practical solution for achieving high spectral efficiency performance while reducing complexity and cost in mmWave MIMO systems. In terms of future work, there is an interest in expanding this research to include multiuser scenarios and the wideband mmWave channel environment.

REFERENCES

[1] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[2] W. Roh, J.-Y. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 106–113, Feb. 2014.

[3] M. Xiao, S. Mumtaz, Y. Huang, L. Dai, Y. Li, M. Matthaiou, G. K. Karagiannidis, E. Björnson, K. Yang, I. Chih-Lin, and A. Ghosh, "Millimeter wave communications for future mobile networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1909–1935, Sep. 2017.

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

[5] A. Alkhateeb, G. Leus, and R. W. Heath, "Limited feedback hybrid precoding for multi-user millimeter wave systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6481–6494, Nov. 2015.

[6] 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.

[7] J. Jin, Y. R. Zheng, W. Chen, and C. Xiao, "Hybrid precoding for millimeter wave MIMO systems: A matrix factorization approach," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 3327–3339, May 2018.

[8] M. Soleimani, R. C. Elliott, W. A. Krzymien, J. Melzer, and P. Mousavi, "Hybrid beamforming for mmWave massive MIMO systems employing DFT-assisted user clustering," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 11646–11658, Oct. 2020.

[9] K. Izadinasab, A. W. Shaban, and O. Damen, "Detection for hybrid beamforming millimeter wave massive MIMO systems," *IEEE Commun. Lett.*, vol. 25, no. 4, pp. 1168–1172, Apr. 2021.

[10] W. Park and J. Choi, "Hybrid precoding and combining strategy for MMSE-based rate balancing in mmWave multiuser MIMO systems," *IEEE Access*, vol. 10, pp. 88043–88057, 2022.

[11] M. Alouzi, F. Chan, and C. D'Amours, "Low complexity hybrid precoding and combining for millimeter wave systems," *IEEE Access*, vol. 9, pp. 95911–95924, 2021.

[12] Q. Wan, J. Fang, Z. Chen, and H. Li, "Hybrid precoding and combining for millimeter wave/sub-THz MIMO-OFDM systems with beam squint effects," *IEEE Trans. Veh. Technol.*, vol. 70, no. 8, pp. 8314–8319, Aug. 2021.

[13] S. Wang, Z. Li, M. He, T. Jiang, R. Ruby, H. Ji, and V. C. M. Leung, "A joint hybrid precoding/combining scheme based on equivalent channel for massive MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 10, pp. 2882–2893, Oct. 2022.

[14] J. Luo, J. Fan, and J. Zhang, "MDL-AltMin: A hybrid precoding scheme for mmWave systems with deep learning and alternate optimization," *IEEE Wireless Commun. Lett.*, vol. 11, no. 9, pp. 1925–1929, Sep. 2022.
[15] X. Gao, L. Dai, S. Han, I. Chih-Lin, 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.
[16] C.-C. Hu and C.-W. Hsu, "Efficient adaptive subarrays in millimeter-wave MIMO systems with hybrid RF/baseband precoding/combining design," *IEEE Syst. J.*, vol. 13, no. 4, pp. 3735–3746, Dec. 2019.
[17] F. Al-Kamali and C. D'amours, "Low-complexity hybrid precoding for subarray architecture mmWave MIMO systems," *IEEE Access*, vol. 10, pp. 74921–74930, 2022.
[18] F. Al-Kamali, C. D'Amours, and F. Chan, "Hybrid precoding for mmWave MIMO systems with overlapped subarray architecture," *IEEE Access*, vol. 10, pp. 130699–130707, 2022.
[19] J. Zhang, Y. Huang, T. Yu, J. Wang, and M. Xiao, "Hybrid precoding for multi-subarray millimeter-wave communication systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 3, pp. 440–443, Jun. 2018.
[20] C.-C. Hu and J.-H. Zhang, "Hybrid precoding design for adaptive sub-connected structures in millimeter-wave MIMO systems," *IEEE Syst. J.*, vol. 13, no. 1, pp. 137–146, Mar. 2019.
[21] Z. Zhang, X. Wu, and D. Liu, "Joint precoding and combining design for hybrid beamforming systems with subconnected structure," *IEEE Syst. J.*, vol. 14, no. 1, pp. 184–195, Mar. 2020.
[22] S. Park, A. Alkhateeb, and R. W. Heath, "Dynamic subarrays for hybrid precoding in wideband mmWave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2907–2920, May 2017.
[23] L. Yan, C. Han, and J. Yuan, "A dynamic array of sub-array architecture for hybrid precoding in the millimeter wave and terahertz bands," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2019, pp. 1–5.
[24] L. Yan, C. Han, and J. Yuan, "A dynamic array-of-subarrays architecture and hybrid precoding algorithms for terahertz wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 9, pp. 2041–2056, Sep. 2020.
[25] F. Yang, J.-B. Wang, M. Cheng, J.-Y. Wang, M. Lin, and J. Cheng, "A partially dynamic subarrays structure for wideband mmWave MIMO systems," *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7578–7592, Dec. 2020.
[26] N. Li, Z. Wei, H. Yang, X. Zhang, and D. Yang, "Hybrid precoding for mmWave massive MIMO systems with partially connected structure," *IEEE Access*, vol. 5, pp. 15142–15151, 2017.
[27] K. B. Dsouza, K. N. R. S. V. Prasad, and V. K. Bhargava, "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, Dec. 2018.
[28] X. Zhao, T. Lin, Y. Zhu, and J. Zhang, "Partially-connected hybrid beamforming for spectral efficiency maximization via a weighted MMSE equivalence," *IEEE Trans. Wireless Commun.*, vol. 20, no. 12, pp. 8218–8232, Dec. 2021.
[29] M. Alouzi, F. Al-Kamali, C. D'Amours, and F. Chan, "Direct conversion of hybrid precoding and combining from full array architecture to sub-array architecture for mmWave MIMO systems," *IEEE Access*, vol. 11, pp. 35457–35468, 2023.
[30] A. M. Elbir and A. K. Papazafeiropoulos, "Hybrid precoding for multiuser millimeter wave massive MIMO systems: A deep learning approach," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 552–563, Jan. 2020.
[31] K. M. Attiah, F. Sohrabi, and W. Yu, "Deep learning for channel sensing and hybrid precoding in TDD massive MIMO OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 12, pp. 10839–10853, Dec. 2022.
[32] H. Hojatian, J. Nadal, J. Frigon, and F. Leduc–Primeau, "Flexible unsupervised learning for massive MIMO subarray hybrid beamforming," in *Proc. IEEE Global Commun. Conf.*, Rio de Janeiro, Brazil, Dec. 2022, pp. 3833–3838.
[33] Q. Yuan, H. Liu, M. Xu, Y. Wu, L. Xiao, and T. Jiang, "Deep learning-based hybrid precoding for terahertz massive MIMO communication with beam squint," *IEEE Commun. Lett.*, vol. 27, no. 1, pp. 175–179, Jan. 2023.
[34] F. Al-Kamali, M. Alouzi, C. D'Amours, and F. Chan, "Architectures for hybrid precoding and combining techniques in massive MIMO systems operating in the mmWave band," in *MIMO Communications—Fundamental Theory, Propagation Channels, and Antenna Systems*. U.K.: IntechOpen, Jun. 2023, doi: 10.5772/intechopen.112113.
[35] P. H. Schönemann, "A generalized solution of the orthogonal Procrustes problem," *Psychometrika*, vol. 31, no. 1, pp. 1–10, Mar. 1966.

MOHAMED ALOUZI (Member, IEEE) received the B.S. degree in electrical and computer engineering from Zawiya University, Libya, in 2009, the M.A.Sc. degree in electrical engineering from the Royal Military College of Canada, Kingston, Canada, in 2017, and the Ph.D. degree in electrical engineering from the University of Ottawa, Ottawa, Canada, in 2023. His research interests include physical layer technologies for wireless communication systems, millimeter wave, 5G and beyond systems, and hybrid beamforming.

FAISAL AL-KAMALI received the B.Sc. degree in electronics and communication engineering from the University of Baghdad, Baghdad, Iraq, in 2001, and the M.Sc. and Ph.D. degrees in communication engineering from Menoufia University, Egypt, in 2008 and 2011, respectively. He joined the Department of Electrical Engineering, University of Ibb, Ibb, Yemen, in 2011, as an Assistant Professor, and became an Associate Professor, in January 2017. He was the Head of the Department of Electrical Engineering, Faculty of Engineering, University of Ibb, from 2013 to 2014, where he was the Vice Dean of Academic Affairs with the Faculty of Engineering, from 2018 to 2020. Since November 2021, he has been a Visiting Researcher with the Faculty of Engineering, School of Electrical Engineering and Computer Science, University of Ottawa, Canada. His research interests include physical layer technologies for wireless communication systems, such as interference cancellation, channel equalization and channel estimation, image transmission over wireless communication systems, hybrid precoding and combining, mmWave systems, and massive MIMO systems. He was a recipient of the prize for supervising the best graduation project with Ibb University, in 2016.

CLAUDE D'AMOURS (Member, IEEE) received the B.A.Sc., M.A.Sc., and Ph.D. degrees in electrical engineering from the University of Ottawa, in 1990, 1992, and 1995, respectively. He was a Systems Engineer with Calian Communications Systems Ltd., and Communications Research Centre, Ottawa, in 1992 and 1995, respectively. Later in 1995, he joined the Department of Electrical and Computer Engineering, Royal Military College, Canada, as an Assistant Professor. He joined the School of Electrical Engineering and Computer Science, University of Ottawa, in 1999, where he was the Vice Dean of Undergraduate Studies with the Faculty of Engineering, from 2007 to 2011, and the Director of the School of Electrical Engineering and Computer Science, from 2013 to 2022. He is currently the Interim Vice Dean of Graduate Studies with the Faculty of Engineering, University of Ottawa. His research interests include physical layer technologies for wireless and optical communication systems, such as multiple access, interference cancellation, modulation, coding, and signal processing.

FRANCOIS CHAN (Senior Member, IEEE) received the B.Eng. degree in electrical engineering from McGill University, Montreal, Canada, and the M.Sc.A. and Ph.D. degrees in electrical engineering from École Polytechnique de Montréal, Canada. He was a Visiting Researcher with the University of California at Irvine, Irvine, in 2002 and 2005. He is currently a Professor with the Department of Electrical and Computer Engineering, Royal Military College of Canada, Kingston, ON, Canada. He is also an Adjunct Professor and a Visiting Professor with the University of Ottawa. His research interests include digital communications, wireless communications, and digital signal processing.