IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 6 June 2022, accepted 8 July 2022, date of publication 14 July 2022, date of current version 21 July 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3190511

RESEARCH ARTICLE

Low-Complexity Hybrid Precoding for Subarray Architecture mmWave MIMO Systems

FAISAL AL-KAMALI[1,2] AND CLAUDE D'AMOURS[1], (Member, IEEE)
[1]School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, ON K1N 6N5, Canada
[2]Department of Electrical Engineering, Ibb University, Ibb, Yemen

Corresponding author: Faisal Al-Kamali (faisalalkamali@gmail.com)

ABSTRACT Hybrid precoding for millimeter wave (mmWave) multiple-input multiple-output (MIMO) systems has attracted much attention in recent years for subarray architecture compared to array architecture because of its low cost and low power consumption. This is due to the small number of required phase shifters in the subarray architecture. In this paper, we investigate the issue of hybrid precoding for the subarray architecture in narrowband mmWave MIMO systems. First, we derive the spectral efficiency of the subarray architecture with hybrid precoding and discuss the problem formulation. Then, we propose two low-complexity hybrid precoding algorithms for the subarray architecture for narrowband mmWave MIMO systems. In the first algorithm, the hybrid precoding matrix is divided into subarrays submatrices and each subarray submatrix is then divided into vectors. The analog precoding of each subarray is determined from the first vector of the subarray submatrix, which is then used to determine the elements of the digital precoder from all vectors in the subarray submatrix (vector by vector) using a simple maximum ratio combining (MRC) method. The proposed algorithm is called vector-by-vector (VBV) hybrid precoding. Finally, to further enhance the system performance, the proposed VBV precoding in the first algorithm is also combined with an iterative solution, and the resulting algorithm is called iterative VBV precoder. Simulation results verify that the proposed precoding algorithms outperform that of the successive interference cancellation-based subarray precoding and has a performance that is close to that obtained by the fully-connected spatially sparse precoding in various system settings, with lower complexity.

INDEX TERMS Hybrid precoding, narrowband millimeter wave MIMO system, subarrays architecture.

I. INTRODUCTION

Millimeter-wave (mmWave) multiple-input multiple-output (MIMO) is considered to be a promising candidate technology in the future wireless communications due to its wider bandwidth and higher spectral efficiency [1]–[4]. Conventional MIMO systems are not suitable for mmWave MIMO systems because the complexity and power consumption of the digital precoder are high, since each antenna must be connected with a radio-frequency (RF) chain [5]–[15]. To alleviate this, hybrid analog and digital precoding is suggested as a suitable solution for mmWave systems [5]–[19]. An analog precoder is implemented using phase shifters [6]–[10]. The main advantages of the hybrid precoding are its lower complexity, lower cost, lower power consumption, and higher energy efficiency as compared to that of fully digital precoding because the required number of the RF chains is low.

In mmWave systems, there are two architectures to connect RF chains to antennas. The first is fully-connected (FC) and is referred to as ''array architecture'' [5]–[19] and the second is sub-connected or ''subarray architecture'' (SA) [20]–[33]. In the FC architecture, each RF chain is connected to antennas, which requires phase shifters for each antenna and results in high cost and power consumption for ultra-massive MIMO systems. However, the spectral efficiency of this approach is high. In the SA architecture, each RF chain is only connected to a subarray of antennas and this technique results in much lower hardware complexity and power consumption than that of the FC architecture, albeit at a decrease in the spectral efficiency of the system.

The associate editor coordinating the review of this manuscript and approving it for publication was Danping He.

Hybrid precoding for FC architecture has been extensively studied in the literature for both narrowband and wideband mmWave systems [5]–[19], [34]. Many of these FC designs achieve near optimum performance compared with a fully digital system [5], [17]. In [5], FC hybrid precoding and combining based on the sparse nature of mmWave channels was studied. The works in [6], [8], [10], [14], and [17] developed low-complexity hybrid precoders and combiners for mmWave MIMO systems, while the works in [13]–[16], and [19] proposed and investigated numerous energy-efficient hybrid precoding and combining schemes. In the ultra-massive MIMO systems, the FC architecture has higher hardware complexity and power consumption than the SA architecture [20]–[33]. As a results of its lower hardware complexity, cost and power consumption, a SA architecture is better suited for practical consideration [20]–[33]. Since each RF chain only connects to a subarray of antennas, the SA architecture changes the constrains of the optimization problem of the hybrid precoding with the FC architecture and the analog precoding must be modeled as a block diagonal matrix [20], [29].

The SA architecture is discussed in the literature, and several hybrid precoding algorithms for SA architecture are presented [20]–[33]. A successive interference cancelation (SIC)-based hybrid precoding for SA was introduced and investigated in [20]. In [20], the authors assumed a diagonal digital precoder with real elements, and they demonstrated that the SA architecture is more energy efficient than the FA architecture. In [23], a hybrid precoder for multi-subarray millimeter-wave communication systems was proposed and investigated. In [22], [27], and [30], many researchers investigated a dynamic subarray architecture that uses switches to adjust the connections between RF chains and subarrays. In the dynamic SA architecture, it is found that the performance of the dynamic SA architecture is better than that of the fixed SA architecture but with higher hardware and computational complexity and power consumption. This is because, the number of the required switches increases linearly with the number of the transmit antennas. To reduce the complexity of the dynamic precoding, [31] proposed a partially dynamic SA structure and assumed a dynamic connection between the subarrays and the RF chains and fixed number of antennas in each subarray. However, the partially dynamic precoder in [31] still necessitates greater computational and hardware complexities, as well as higher power consumption, than fixed precoders. The work in [29] considered the joint design of the precoder and combiner in the subarray architecture. In [26], [28], and [33], an adaptive connection between the antennas and the RF chains in the hybrid precoder was proposed and studied. The adaptive SA precoders provide better performance than that of the fixed SA precoders but with higher hardware and computational complexity. In addition, in the dynamic SA precoders, the antenna location must be updated with the channel fluctuations. So, proposing an efficient, low complexity and power consumption hybrid precoder for the SA architecture is still a very important issue, which is the objective of this paper.

In this paper, we derive the hybrid precoding solution for the SA architecture based on the consideration of the following factors. (1) the division of large antenna arrays into small antenna subarrays, (2) the subarray precoding with RF hardware constraints; and (3) the block-based structure of the precoding matrix. In this work, we only considered narrowband mmWave channels like those in [5], [20], [24], [26], and [29] while the consideration of wideband mmWave channels is also of interest. The extension of this work to wideband mmWave channels will be an important topic for future work.

The main contributions of this paper are:

1. The hybrid precoding optimization problem for the SA architecture mmWave MIMO system is formulated, decomposed into a series of $N_{tRF}N_s$ subproblems, where N_s is the number of data streams and N_{tRF} is the number of RF chains at the transmitter, and then solved.
2. Utilizing the structure of the precoding matrix, a low-complexity algorithm to solve the hybrid precoding problem in the SA architecture, namely, the vector-by-vector (VBV) hybrid precoding, is developed. In the first step, the analog precoding of each subarray is obtained from the first vector of the subarray submatrix. Then, the obtained analog precoding is used to obtain the elements of the digital precoding that are associated with this subarray, i.e., element from each vector, using a simple maximum ratio combining (MRC) method. An efficient iterative solution for each subarray is also combined with the proposed VBV algorithm to enhance the performance of the mmWave MIMO systems, and the resulting algorithm is called iterative VBV hybrid precoding. The complexities of the proposed algorithms are investigated and compared.
3. We carry out extensive simulations for two different scenarios to evaluate the proposed algorithms for the SA architecture. Numerical results show that the proposed algorithms achieve spectral efficiency close to that obtained in [5] and outperforms that obtained in [20] in various system settings.

The main advantages of the proposed VBV hybrid precoding algorithms are that there is no need to make assumption about the digital precoder like that in [20] or about the channel geometry and channel architecture like that in [5], while maintaining low complexity. Compared with the SA dynamic precoders [22], [27], [30], [31], the proposed VBV hybrid precoding algorithms are fixed and there is no need to update the antenna location according to the channel fluctuations and to use switches, which leads to lower computational and hardware complexities. Furthermore, unlike that in [20], the proposed VBV hybrid precoding algorithms are derived for a general case rather than a special case. In [20], a SIC-based precoder is derived for a special case of a fixed

SA architecture, and thus the authors assumed a diagonal digital precoder matrix with real elements and $N_{tRF} = N_s$. Moreover, the matrix inversion is involved in [20] and not involved in the proposed VBV hybrid precoding algorithms, which makes the computational complexities of the proposed VBV hybrid precoding algorithms low.

The remainder of this paper is organized as follows. The system model of the subarray architecture is introduced in Section II. The spectral efficiency and the problem formulation are discussed in Sec. III. Section IV presents the proposed VBV hybrid precoding algorithms for the subarray architecture. The complexities of the proposed VBV hybrid precoding algorithms are discussed in Sec. V. Simulation results are investigated in Sec. VI. Finally, in Section VII we present our conclusions.

In this paper, we use bold uppercase letters for matrices and bold lowercase letters for vectors, and plaintext letters for scalars. The exponents $(.)^{-1}$, $(.)^*$, $(.)^T$, and $(.)^H$ denote the inversion, conjugate, transpose, and Hermitian transpose and $|.|$ denotes the determinant of a matrix. The functions $\|.\|_1$ and $\|.\|_2$ denote the l_1-norm and l_2-norm of a vector, and $\|.\|_F$ denotes the Frobenius norm of the matrix. $\mathbf{I}_N$ is the $N \times N$ identity matrix.

II. SYSTEM MODEL

Consider a downlink channel with one base station (BS) and one mobile station (MS). The BS has N_t antennas and N_{tRF} RF chains and the MS has N_r antennas and N_{rRF} RF chains. The BS transmit N_s data streams. In this paper, a general case when $N_{tRF} \geq N_s$ is considered. Hybrid precoding in the SA architecture at the BS is depicted in Fig. 1. The antenna array at the BS is divided into L_t subarrays and each subarray is connected with one RF chain. The number of antenna elements in each subarray is equal to $N_{tSA} = N_t/L_t$. The received signal vector at the MS can be modeled as

$$\mathbf{y} = \mathbf{W}^{\mathbf{H}}\sqrt{\rho}\mathbf{H}\mathbf{P}_A\mathbf{P}_D\mathbf{s} + \mathbf{W}^H\mathbf{n} \tag{1}$$

where ρ is the average power of the received signal, $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ is the channel matrix, $\mathbf{P}_A$ is the $N_t \times L_t$ analog precoding matrix, $\mathbf{P}_D$ is the $L_t \times N_s$ digital precoding matrix, $\mathbf{s}$ is the $N_s \times 1$ vector of the transmitted signal such that $\mathbb{E}\left[\mathbf{s}\mathbf{s}^*\right] = \frac{1}{N_s}\mathbf{I}_{N_s}$ and $\mathbf{n}$ is the $N_r \times 1$ vector of independent and identical distribution $\mathcal{CN}\left(0, \sigma_n^2\right)$ additive white noise. $\mathbf{W} = \mathbf{W}_A\mathbf{W}_B$ is the $N_r \times N_s$ hybrid combining matrix, where $\mathbf{W}_A$ is the analog combiner, and $\mathbf{W}_B$ is the baseband combiner. In the SA architecture, the analog precoding matrix can be denoted as

$$\mathbf{P_A} = \begin{bmatrix} \mathbf{p}_{A1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{p}_{A2} & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \vdots & \mathbf{p}_{AL_t} \end{bmatrix} \tag{2}$$

where $\mathbf{p}_{Al}$ is the $N_{tSA} \times 1$ analog precoding vector of the of the lth subarray ($L = 1, 2, \dots, L_t$) whose elements have the same amplitude $1/\sqrt{N_{tSA}}$ but different phases.

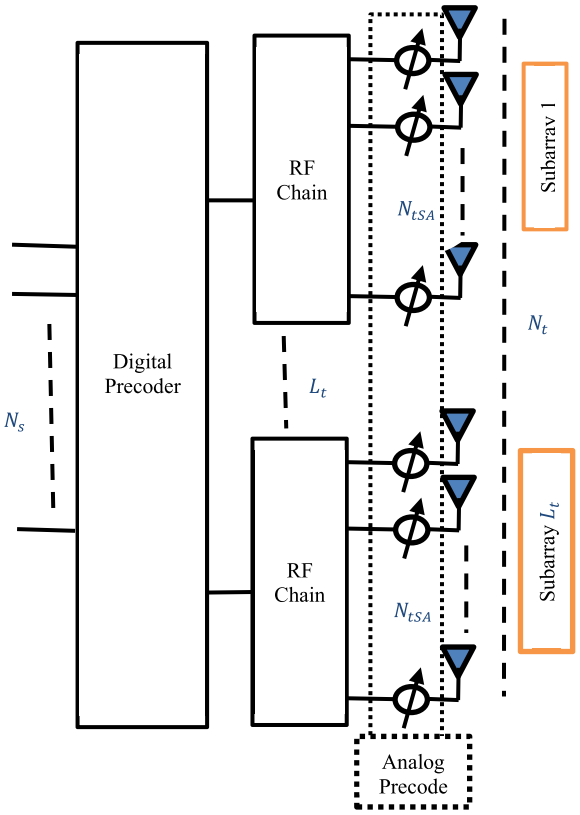

FIGURE 1. Hybrid precoding in the subarray architecture at the BS.

The spectral efficiency that the hybrid precoder and combiner can achieve is given by [5], [17]

$$\boldsymbol{R} = \log_2\left(\left|\mathbf{I}_{N_s} + \frac{\rho}{N_s}\mathbf{R}_n^{-1}\mathbf{W}^H\mathbf{H}\mathbf{P}_A\mathbf{P}_D\mathbf{P}_D^H\mathbf{P}_\mathbf{A}^H\mathbf{H}^H\mathbf{W}\right|\right) \tag{3}$$

where $\mathbf{R}_n$ is the noise covariance. In this paper, we assume a mmWave channel model [5]. $\mathbf{H}$ can be expressed as [5], [17], [20]

$$\begin{aligned}\mathbf{H} = &\sqrt{N_t N_r / N_{cl} N_{ray}} \\ &\times \sum_{i}^{N_{cl}} \sum_{l}^{N_{ray}} \alpha_{il} \mathbf{\Lambda}_r\left(\phi_{il}^r, \theta_{il}^r\right) \mathbf{\Lambda}_t\left(\phi_{il}^t, \theta_{il}^t\right) \\ &\times \mathbf{a}_r\left(\phi_{il}^r, \theta_{il}^r\right) \mathbf{a}_t\left(\phi_{il}^t, \theta_{il}^t\right)^*\end{aligned} \tag{4}$$

where N_{cl} and N_{ray} are number of clusters and paths, respectively. $\alpha_{il} \in \mathbb{C}$ is the lth path's gain in the ith cluster. $\phi_{il}^t(\phi_{il}^r)$ and $\theta_{il}^t(\theta_{il}^r)$ represent the lth path's azimuth (elevation) angles of departure and arrival in the ith cluster, respectively. $\mathbf{\Lambda}_t\left(\phi_{il}^t, \theta_{il}^t\right)$ and $\mathbf{\Lambda}_r(\phi_{il}^r, \theta_{il}^r)$ denote the transmit and receive antenna element gain at their departure and arrival angles. $\mathbf{a}_t\left(\phi_{il}^t, \theta_{il}^t\right)$ and $\mathbf{a}_r\left(\phi_{il}^r, \theta_{il}^r\right)$ are the antenna array response at the transmitter and receiver, respectively. In a uniform planer array, the array response vector can be defined as [5]

$$\begin{aligned}\boldsymbol{a}_{UPA(\phi,\theta)} = &\frac{1}{\sqrt{\mathrm{N_t}}}[1, \dots, e^{jkd(xsin(\phi)sin(\theta)+ycos(\theta))}, \\ &\times \dots, e^{jkd((W-1)sin(\phi)sin(\theta)+(H-1)cos(\theta)}]^{\boldsymbol{T}}\end{aligned} \tag{5}$$

where $k = \frac{2\pi}{\lambda}$, $1 \le x \le (W-1)$ and $1 \le y \le (H-1)$. $d = \frac{\lambda}{2}$, W and H are the inter-antenna spacing, the width and the height of the antenna array. The array size of the transmitter is $N_t = WH$. In this paper, we assume perfect channel state information at the BS and the MS.

III. SPECTRAL EFFCIENCY AND PROBLEM FORMULATION

In this section, the spectral efficiency and the problem formulation are discussed. In this paper, we aim to maximize the spectral efficiency of the hybrid precoder at the transmitter side of the SA architecture. So, the combiner analysis at the receiver side is not considered, and the spectral efficiency of the SA architecture with the precoder can only be written as [20]

$$R = \log_2\left(\left|\mathbf{I}_{N_r} + \frac{\rho}{N_s\sigma_n^2}\mathbf{H}\mathbf{P}_A\mathbf{P}_D\mathbf{P}_D^H\mathbf{P}_A^H\mathbf{H}^H\right|\right) \tag{6}$$

The singular value decomposition (SVD) of $\mathbf{H}$ can be defined as [5], [17]

$$\mathbf{H} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H = \begin{bmatrix}\mathbf{U_1} & \mathbf{U_2}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Sigma}_1 & \mathbf{0}\\ \mathbf{0} & \boldsymbol{\Sigma}_2\end{bmatrix}\begin{bmatrix}\mathbf{V}_1^H\\ \mathbf{V}_2^H\end{bmatrix} \tag{7}$$

where $\mathbf{U}$ is an $N_r \times rank(\mathbf{H})$ unitry matrix and $\mathbf{V}$ is an $N_t \times rank(\mathbf{H})$ unitary matrix. $\boldsymbol{\Sigma}$ is an $rank(\mathbf{H}) \times rank(\mathbf{H})$ diagonal matrix containing the singular values of $\mathbf{H}$ in decreasing order. $\boldsymbol{\Sigma_1}$ is an $N_s \times N_s$ diagonal matrix, $\mathbf{V_1}$ is an $N_t \times N_s$ matrix and $\mathbf{U_1}$ is an $N_r \times N_s$ matrix. The solution of the optimal unconstrained hybrid precoding and combining scheme is given by $\mathbf{P}^{opt} = \mathbf{V_1}$ and $\mathbf{W}^{opt} = \mathbf{U}_1^H$ [5], [17]. Given $\mathbf{H} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H$, (6) can be expressed as

$$R = \log_2\left(\left|\mathbf{I}_{N_r} + \frac{\rho}{N_s\sigma_n^2}\boldsymbol{\Sigma}^2\mathbf{V}^H\mathbf{P}_A\mathbf{P}_D\mathbf{P}_D^H\mathbf{P}_A^H\mathbf{V}\right|\right) \tag{8}$$

Then, by separating $\boldsymbol{\Sigma}$ and $\mathbf{V}$ as in (7), (8) can be approximated as [5]

$$\mathrm{R} \approx \log_2\left(\left|\mathbf{I}_{N_s} + \frac{\rho}{N_s\sigma_n^2}\boldsymbol{\Sigma}_\mathbf{1}^2\mathbf{V}_1^H\mathbf{P}_A\mathbf{P}_D\mathbf{P}_D^H\mathbf{P}_A^H\mathbf{V}_\mathbf{1}\right|\right) \tag{9}$$

By using the formula $(\boldsymbol{I} - \boldsymbol{XY})(\boldsymbol{I} - (\boldsymbol{I} + \boldsymbol{X})^{-1}\boldsymbol{X}(\boldsymbol{I} - \boldsymbol{Y}))$ [5], [20], (9) can be rewritten as

$$\begin{aligned}\mathrm{R} = &\log_2\left(\left|\mathbf{I}_{N_s} + \frac{\rho}{N_s\sigma_n^2}\boldsymbol{\Sigma}_\mathbf{1}^\mathbf{2}\right|\right)\\ &+\log_2\left(\left|\mathbf{I}_{N_s} - \left(\mathbf{I}_{N_s} + \frac{\rho}{N_s\sigma_n^2}\boldsymbol{\Sigma}_\mathbf{1}^\mathbf{2}\right)^{-1}\left(\frac{\rho}{N_s\sigma_n^2}\boldsymbol{\Sigma}_\mathbf{1}^\mathbf{2}\right)\right.\right.\\ &\left.\left.\left(\mathbf{I}_{N_s} - \mathbf{V}_1^H\mathbf{P}_A\mathbf{P}_D\mathbf{P}_D^H\mathbf{P}_A^H\mathbf{V}_\mathbf{1}\right)\right|\right)\end{aligned} \tag{10}$$

After employing the high SNR approximation, R can be simplified as [5], [20]

$$R \approx \log_2\left(\left|\mathbf{I}_{N_s} + \frac{\rho}{N_s\sigma_n^2}\boldsymbol{\Sigma}_\mathbf{1}^\mathbf{2}\right|\right) + \log_2\left(\left|\mathbf{V}_1^H\mathbf{P}_A\mathbf{P}_D\mathbf{P}_D^H\mathbf{P}_A^H\mathbf{V}_\mathbf{1}\right|\right) \tag{11}$$

From (11), we observe that maximizing R is equivalent to maximizing $\mathbf{V_1^H}\mathbf{P}_A\mathbf{P}_D\mathbf{P}_D^H\mathbf{P}_A^H\mathbf{V_1} = \left\|\mathbf{V_1^H}\mathbf{P}_A\mathbf{P}_D\right\|_\mathbf{F}^\mathbf{2}$. Define the digital precoding matrix $\mathbf{P_D} = \begin{bmatrix}\mathbf{p}_{D1} & \mathbf{p}_{D2} & \dots & \mathbf{p}_{DL_t}\end{bmatrix}^T$, where $\mathbf{p}_{Dl} = \begin{bmatrix}d_{ln} & d_{l2} & \dots & d_{lN_s}\end{bmatrix}$ is the $1 \times N_s$ vector represents the lth row of $\mathbf{P_D}$ and d_{ln} is the element of $\mathbf{P_D}$ that lies in lth row and nth column, we can reformulate $\left\|\mathbf{V_1^H}\mathbf{P_A}\mathbf{P_D}\right\|_\mathbf{F}^\mathbf{2}$ for all subarray submatrices as

$$\left|\left|\mathbf{V_1^H}\mathbf{P_A}\mathbf{P_D}\right|\right|_\mathbf{F}^\mathbf{2} = \sum_{\mathbf{l=1}}^{\boldsymbol{L_t}}\left|\left|\tilde{\mathbf{V}}_\mathbf{l}^\mathbf{H}\mathbf{p}_{\boldsymbol{Al}}\mathbf{p}_{\boldsymbol{Dl}}\right|\right|_\mathbf{F}^\mathbf{2} \tag{12}$$

where $\mathbf{V_1} = \begin{bmatrix}\tilde{\mathbf{V}}_\mathbf{1}\tilde{\mathbf{V}}_\mathbf{2}\dots\tilde{\mathbf{V}}_{\mathbf{L_t}}\end{bmatrix}^\mathbf{T}$, $\tilde{\mathbf{V}}_\boldsymbol{l}$ is the $N_{tSA} \times N_s$ submatrix of the lth subarray. We can also further decompose (12) for all vectors as

$$\left\|\mathbf{V}_\mathbf{1}^H\mathbf{P}_A\mathbf{P}_D\right\|_F^2 = \sum_{l=1}^{L_t}\sum_{n=1}^{N_s}\left\|\mathbf{v}_{\boldsymbol{ln}}^H(d_{ln}\mathbf{p}_{\boldsymbol{Al}})\right\|_\mathbf{2}^\mathbf{2} \tag{13}$$

where $\tilde{V}_l = \begin{bmatrix}\mathbf{v}_{l1} & \mathbf{v}_{l2} & \dots & \mathbf{v}_{lN_s}\end{bmatrix}$, $\mathbf{v}_{ln}$ is the $N_{tSA} \times 1$ vector of the nth vector in the lth subarray submatrix.

From (12) and (13), we note that maximizing $\mathbf{V}_\mathbf{1}^H\mathbf{P}_A\mathbf{P}_D$ is equivalent to maximizing $\tilde{\boldsymbol{V}}_\boldsymbol{l}^\boldsymbol{H}\mathbf{p}_{\boldsymbol{Al}}\mathbf{p}_{\boldsymbol{Dl}}$ in each subarray submatrix or to maximizing $\mathbf{v}_{\boldsymbol{ln}}^\boldsymbol{H}(d_{ln}\mathbf{p}_{\boldsymbol{Al}})$ in each vector.

IV. THE PROPOSED VBV HYBRID PRECODING SCHEME

In this section, we propose two new algorithms by exploiting the structure of the SA precoding matrix to solve the hybrid precoding problem for narrowband mmWave communication systems.

A. VBV HYBRID PRECODING

In this subsection, we aim to maximize the achievable spectral efficiency of each subarray as illustrated in section III. In the SA architecture, the hybrid precoding matrix $\mathbf{P}$ can be written as

$$\mathbf{P} = \begin{bmatrix}\mathbf{P_1}\\ \mathbf{P_2}\\ \vdots\\ \mathbf{P}_{\boldsymbol{L_t}}\end{bmatrix} = \mathbf{P_A}\mathbf{P_D} = \begin{bmatrix}\mathbf{p}_{A1}\mathbf{p}_{D1}\\ \mathbf{p}_{A2}\mathbf{p}_{D2}\\ \vdots\\ \mathbf{p}_{AL_t}\mathbf{p}_{\boldsymbol{DL_t}}\end{bmatrix} \tag{14}$$

where $\mathbf{P}_l = \mathbf{p}_{Al}\mathbf{p}_{Dl}$ is an $N_{tSA} \times N_S$ hybrid precoding matrix of the lth subarray. We can decompose $\mathbf{P}$ into vectors as follows:

$$\begin{aligned}\mathbf{P} &= \begin{bmatrix}\mathbf{p}_{11} & \mathbf{p}_{12} & \dots & \mathbf{p}_{1N_s}\\ \mathbf{p}_{21} & \mathbf{p}_{22} & \dots & \mathbf{p}_{2N_s}\\ \vdots & \vdots & \vdots & \vdots\\ \mathbf{p}_{\boldsymbol{L_t}\mathbf{1}} & \mathbf{p}_{\boldsymbol{L_t}\mathbf{2}} & \vdots & \mathbf{p}_{L_tN_s}\end{bmatrix}\\ &= \begin{bmatrix}d_{11}\mathbf{p}_{A1} & d_{12}\mathbf{p}_{A1} & \dots & d_{1N_s}\mathbf{p}_{A1}\\ d_{21}\mathbf{p}_{A2} & d_{22}\mathbf{p}_{A2} & \dots & d_{2N_s}\mathbf{p}_{A2}\\ \vdots & \vdots & \vdots & \vdots\\ d_{L_t1}\mathbf{p}_{AL_t} & d_{L_t2}\mathbf{p}_{AL_t} & \vdots & d_{L_tN_s}\mathbf{p}_{AL_t}\end{bmatrix}\end{aligned} \tag{15}$$

where $\mathbf{p}_{\boldsymbol{ln}} = d_{ln}\mathbf{p}_{\boldsymbol{Al}}$ is an $N_{tSA} \times 1$ vector represents the nth vector in the lth subarray. Equation (15) shows that the hybrid precoding matrix $\mathbf{P}$ in the SA architecture consists of L_t subarray submatrices and each subarray submatrix consist

of N_s vectors. The total number of vectors in $\mathbf{P}$ is $L_t N_s$ and the elements in each vector have the same amplitude $|d_{ln}|/\sqrt{N_{tSA}}$. According to the structure of $\mathbf{P}$ in (15), there are two constrains for the design of $\mathbf{P}$ in the SA architecture:

Constrain 1: All elements in $\mathbf{p}_{ln}$ have the same amplitude $|d_{ln}|/\sqrt{N_{tSA}}$ since the amplitude of all elements in $\mathbf{p}_{Al}$ have amplitude $1/\sqrt{N_{tSA}}$.

Constraint 2: The Frobenius norm of $\mathbf{P}$ should satisfy $\|\mathbf{P}_A\mathbf{P}_D\|_F^2 = N_s$ to meet the total transmit power.

As illustrated in Sec. III, the optimization problem in the SA architecture can be written as

$$\begin{aligned}\left(\mathbf{P}_A^{opt}, \mathbf{P}_D^{opt}\right) = \underset{\mathbf{P_A},\mathbf{P_D}}{\arg\min} \quad & \left\|\mathbf{P}^{opt} - \mathbf{P_A}\mathbf{P_D}\right\|_2^F \\ \text{st.} \quad & \mathbf{p}_{ln} \,\varepsilon\, \mathcal{F}_{\boldsymbol{A}}, \\ & \|\mathbf{P}_A\mathbf{P}_D\|_F^2 = N_s \end{aligned} \tag{16}$$

where $\boldsymbol{\mathcal{F}_A}$ includes all possible $N_{tSA} \times 1$ vector satisfying constraint 1. The objective function in (16) can be decomposed into objective functions of each of the $\boldsymbol{L_t}$ subarrays as

$$\left\|\mathbf{P}^{opt} - \mathbf{P}_A\mathbf{P}_D\right\|_{\boldsymbol{F}}^2 = \sum_{\boldsymbol{l=1}}^{\boldsymbol{L_t}} \left\|\mathbf{P}_{\boldsymbol{l}}^{opt} - \mathbf{p}_{Al}\mathbf{p}_{Dl}\right\|_{\boldsymbol{F}}^{\boldsymbol{2}} \tag{17}$$

Moreover, we can further decompose the objective function in (17) into objective functions for the $L_t N_s$ vectors as follows:

$$\left\|\mathbf{P}^{opt} - \mathbf{P}_A\mathbf{P}_D\right\|_{\boldsymbol{F}}^2 = \sum_{\boldsymbol{l=1}}^{\boldsymbol{L_t}} \sum_{\boldsymbol{n=1}}^{\boldsymbol{N_s}} \left\|\mathbf{p}_{ln}^{opt} - d_{ln}\mathbf{p}_{Al}\right\|_{\boldsymbol{2}}^{\boldsymbol{2}} \tag{18}$$

Equation (18) shows that minimizing $\left\|\mathbf{P}^{opt} - \mathbf{P}_A\mathbf{P}_D\right\|_{\mathbf{F}}^{\mathbf{2}}$ is equivalent to separately minimizing the $L_t N_S$ subproblems $\left\|\mathbf{p}_{ln}^{opt} - d_{ln}\mathbf{p}_{Al}\right\|_{\boldsymbol{2}}^{\boldsymbol{2}}$. So, the optimization problem in (16) is equivalent to the following subproblem

$$\begin{aligned}\left(\mathbf{p}_{Al}^{opt}, d_{ln}^{opt}\right) = \underset{\mathbf{p_{Al}},d_{ln}}{\arg\min} \quad & \|\mathbf{v_{ln}} - d_{ln}\mathbf{p}_{Al}\|_2^2 \\ \text{st.} \quad & \mathbf{p}_{ln} \,\varepsilon\, \mathcal{F}_{\boldsymbol{A}}, \\ & \|\mathbf{P}_A\mathbf{P}_D\|_F^2 = N_s \end{aligned} \tag{19}$$

where $\mathbf{p}_{ln}^{opt} = \mathbf{v}_{ln}$ and $\mathbf{v}_{ln}$ can be obtained from $\mathbf{V_1}$ as

$$\mathbf{V}_1 = \begin{bmatrix} \mathbf{v}_{11} & \mathbf{v}_{12} & \cdots & \mathbf{v}_{1N_s} \\ \mathbf{v}_{21} & \mathbf{v}_{22} & \cdots & \mathbf{v}_{2N_s} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{v}_{L_t1} & \mathbf{v}_{L_t2} & \vdots & \mathbf{v}_{L_tN_s} \end{bmatrix} \tag{20}$$

where $\mathbf{v}_{ln}$ is the $N_{tSA} \times 1$ vector containing the optimum elements of the *nth* vector in the *lth* subarray. The unconstrained optimal solution of the problem in (19) can be given as

$$\mathbf{p}_{Al}^{opt} = \frac{1}{\sqrt{N_{tSA}}} e^{j\text{angle}(\mathbf{v_{ln}})} \tag{21}$$

and

$$d_{ln}^{opt} = \frac{\|\mathbf{v_{ln}}\|_1}{\sqrt{N_{tSA}}} \tag{22}$$

Due to constraints 1 and 2, the optimum solutions in (21) and (22) are not suitable for all vectors. However, from the structure of $\mathbf{P}$, it is possible to apply this solution only for one column of $\mathbf{P}$ to find $\mathbf{p}_{Al}^{opt}$ and then $\mathbf{P}_A^{opt}$. So, in this paper, we will apply this solution to find $\mathbf{p}_{Al}^{opt}$ from the first vector of each subarray submatrix and the optimum solution in (21) can be modified as

$$\mathbf{p}_{Al}^{opt} = \frac{\mathbf{1}}{\sqrt{N_{tSA}}} e^{j\text{angle}(\mathbf{v_{l1}})} \tag{23}$$

From this solution, we can find the analog precoding vectors $\mathbf{p}_{Al}^{opt}(l = 1, 2, \ldots, L_t)$ of all subarrays. After that, by fixing the analog precoder $\mathbf{P_A}$ and temporarily ignoring the power constraint $\|\mathbf{P_A}\mathbf{P}_D\|_{\boldsymbol{F}}^{\boldsymbol{2}} = N_s$, we will find the elements of the digital precoding d_{ln}^{opt} from all vectors using a simple MRC (Vector-by-Vector). The proposed VBV hybrid precoding algorithm can be performed by the following three steps.

1. From the first vector in each subarray submatrix, find the analog precoding vectors as in (23).
2. Obtain the elements of $\mathbf{P_D}$ from all vectors using a simple MRC as follows:
$$d_{ln}^{opt} = (\mathbf{p}_{Al}^{opt})^{\boldsymbol{H}} * \mathbf{v}_{ln} \tag{24}$$
3. Normalize $\mathbf{P}_D$ as $\mathbf{P}_D = \frac{\sqrt{N_s}}{\|\mathbf{P_A}\mathbf{P_D}\|_F}\mathbf{P}_D$.

Algorithm 1 illustrates the pseudo code of the proposed VBV hybrid precoding scheme.

Since the elements of each vector in the hybrid precoding matrix have the same amplitude, it is possible to enhance the performance of **Algorithm 1** by combining it with the iterative solution in [17] and iteratively solving the problem of the hybrid precoding for each subarray.

Algorithm 1 VBV Hybrid Precoding Scheme

```
Input V_1
Decompose V_1 as (20)
For 1 ≤ l ≤ L_t
For 1 ≤ n ≤ N_s
If n = 1
    p_Al^opt = (1/√N_tSA) e^{jangle(v_ln)}
    d_ln^opt = (p_Al^opt)^H * v_ln
else
    d_ln^opt = (p_Al^opt)^H * v_ln
end if
end for
end for
construct P_D and P_A
normalize P_D as P_D = (√N_s / ||P_A P_D||_F) P_D
Return P_A and P_D
```

The vectors $\mathbf{p}_{Al}$ and $\mathbf{p}_{Dl}$ obtained from **algorithm 1** will be used as initial values in the solution as illustrated in the next subsection.

B. ITERATIVE VBV HYBRID PRECODER

In this subsection, the proposed scheme in algorithm 1 is combined with the iterative solution in [17] to further improve the

mmWave communication systems' performance. The optimization problem of the *lth* subarray can be written as

$$\left(\mathbf{p}_{Al}^{opt}, \mathbf{p}_{Dl}^{opt}\right) = \underset{\mathbf{p}_{Al},\mathbf{p}_{Dl}}{\arg\min} \left\|\mathbf{p}_{l}^{opt} - \mathbf{p}_{Al}\mathbf{p}_{Dl}\right\|_2^F \quad \text{st.} \quad \mathbf{p}_{Al} \varepsilon \mathcal{F}_A, \; \|\mathbf{P}_A\mathbf{P}_D\|_F^2 = N_s \tag{25}$$

The iterative solution of this problem can be written as [17]

$$\mathbf{p}_{Al} = \mathbf{P}_{res}\mathbf{p}_{Dl}^H + \mathbf{p}_{Al} \tag{26}$$

where $\mathbf{P}_{res} = \mathbf{P}_l^{opt} - \mathbf{p}_{Al}\mathbf{p}_{Dl}$ is the residual precoding matrix for the *lth* subarray. Equation (26) indicates that the updated $\mathbf{p}_{Al}$ is equal to the $\mathbf{p}_{Al}$ from the previous iteration plus the $\mathbf{P}_{res}\mathbf{p}_{Dl}^H$. After getting $\mathbf{p}_{Al}$ and $\mathbf{p}_{Dl}$ from algorithm 1, it is possible to use these values as an initial value to iteratively solve the optimization problem in (25).

After that, to satisfy the constraint in (25), the obtained $\mathbf{p}_{Al}$ of the *lth* subarray must be normalized as

$$\mathbf{p}_{Al} = \mathbf{p}_{Al}/(|\mathbf{p}_{Al}| * \sqrt{N_{tSA}}) \tag{27}$$

In summary, the pseudo-code of the proposed iterative VBV hybrid precoding can be summarized in **Algorithm 2,** which can be explained as follows. First, algorithm 1 is applied to obtain $\mathbf{p}_{Al}$ and $\mathbf{p}_{Dl}$ of the *lth* subarray. Then, the iterative solution for the *lth* subarray is applied to obtain the optimal $\mathbf{p}_{Al}$ and $\mathbf{p}_{Dl}$ and this operation will be repeated for all subarrays. Finally, $\mathbf{P}_D$ and $\mathbf{P}_A$ are constructed.

Algorithm 2 Iterative VBV Hybrid Precoding scheme

1. Input $\mathbf{V}_1, K$
2. Decompose $\mathbf{V_1}$ as $\mathbf{V}_1 = \left[\tilde{\mathbf{V}}_1\tilde{\mathbf{V}}_2 \ldots \tilde{\mathbf{V}}_{\mathrm{L_t}}\right]^{\mathrm{T}}$
3. For $1 \le l \le L_t$
4. Compute $\mathbf{p}_{Al}^{opt}$ by **Algorithm 1**
5. Compute d_{ln}^{opt} $(n = 1, 2, ..., N_s)$ by **Algorithm 1**
6. Initial $\mathbf{p_{Al}} = \mathbf{p}_{Al}^{opt}$
7. Construct $\mathbf{p}_{Dl} = \left[d_{l1}^{opt} \quad d_{l2}^{opt} \quad \ldots \quad d_{lN_s}^{opt}\right]$
8. For $1 \le k \le K$
9. Compute the residual $\mathbf{P}_{res} = \tilde{\mathbf{V}}_l - \mathbf{p}_{Al}\mathbf{p}_{Dl}$
10. Update $\mathbf{p_{Al}} = \mathbf{P}_{res}\,\mathbf{p}_{Dl}^H + \mathbf{p_{Al}}$
11. Normalize $\mathbf{p}_{Al} = \mathbf{p}_{Al}/(|\mathbf{p}_{Al}| * \sqrt{N_{tSA}})$
12. $\mathbf{p}_{Dl} = \mathbf{p}_{Al}^{\mathbf{H}} * \tilde{\mathbf{V}}_l$
13. end for
14. end for
15. Construct $\mathbf{P}_D$ and $\mathbf{P}_A$
16. Normalize $\mathbf{P}_D$ as $\mathbf{P}_D = \frac{\sqrt{N_s}}{\|\mathbf{P_A}\mathbf{P_D}\|_\mathbf{F}}\mathbf{P}_D$
17. Return $\mathbf{P}_A$ and $\mathbf{P}_D$

Equation (26) satisfies the property of the gradient descent method [17]. Each iteration, from step 8 to 13 in algorithm 2, minimizes the objective function $\left\|\mathbf{P}_l^{opt} - \mathbf{p}_{Al}\mathbf{p}_{Dl}\right\|_F^2$ and, as a result, the convergence of $\mathbf{p}_{Al}$ to a local optimal point is ensured.

The idea of the VBV hybrid precoding algorithms in this paper can also be extended to the combining at the mobile station. Discussion about VBV hybrid combining will be left for future work.

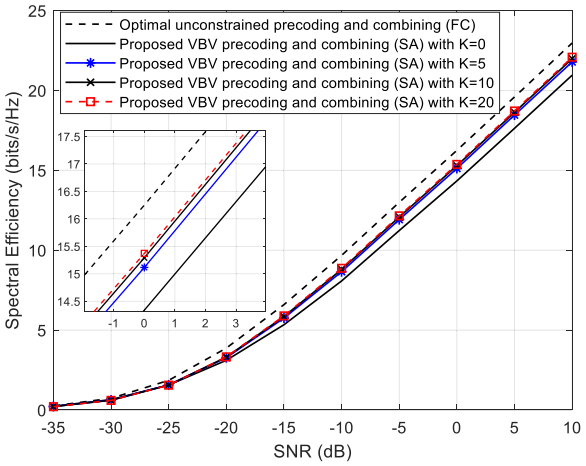

FIGURE 2. Achievable spectral efficiency comparison for $N_t = 64$, $N_r = 16$, $N_{tRF} = 4$ and $N_{rRF} = 4$.

V. COMPLEXITY EVALUATION

In this paper, the complexity of the proposed precoders in algorithms 1 and 2 will be discussed in this section. The proposed precoder in algorithms 1 finds $\mathbf{P}_\mathrm{A}$ from the first vector of all subarrays submatrix and then uses it to find elements of $\mathbf{P}_\mathrm{D}$ from all vectors of each subarray submatrix using a simple MRC. The numbers of multiplications and additions required to find each d_{ln}^{opt} are N_{tSA} and $(N_{tSA}-1)$, respectively, which is very low as compared to that of the sparse hybrid precoding in [5] that uses the least square method to find the $\mathbf{P}_\mathrm{D}$. The complexity of the proposed precoder in algorithms 1 is in the order of $O(N_{tSA}L_tN_s)$ which is very low as compared to that of the sparse hybrid precoding in [5], which requires $O(N_t^2N_{tRF}N_s)$ [17]. When the iterative solution is combined to algorithm 1, the resulting proposed precoder in algorithms 2 requires an additional complexity by about $O(KN_{tSA}N_s)$ and the total complexity is also still lower than that in [5]. The complexities of the proposed precoders are also lower than that of the SIC-based precoder in [20], which requires $O(N_{tSA}^2(L_ts + N_r))$ [20]. Here, s is the number of iterations in [20]. Moreover, the number of the required phase shifters in the BS with the proposed precoders is N_t whereas the required number of phase shifters with the sparse method in [5] is L_tN_t. This indicates that the proposed precoders have lower computational and hardware complexities, lower costs, and lower power consumption than those of [5].

VI. SIMULATION RESULTS

In this section, simulation experiments are carried out to evaluate the performance of the proposed VBV hybrid precoding algorithms 1 and 2 at two different scenarios. Note that $K = 0$ denotes Algorithm 1 and $K \neq 0$ denotes Algorithm 2. The proposed VBV precoding algorithms are compared with the optimal unconstrained hybrid precoding, the spatially sparse hybrid precoding in [5] and the SIC-based SA precoding in [20]. In these simulations, we consider one BS and one MS.

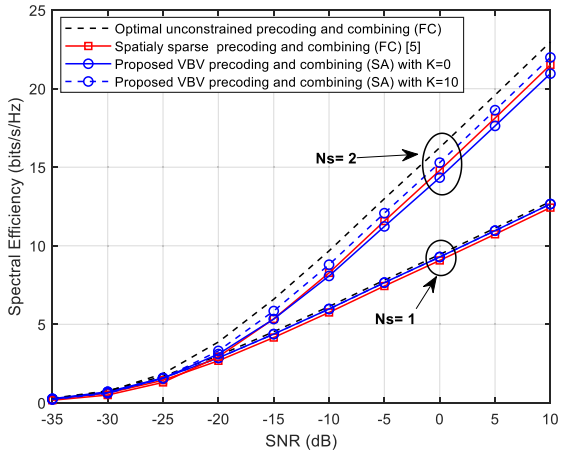

FIGURE 3. Achievable spectral efficiency comparison for $N_t = 64$, $N_r = 16$, $N_{tRF} = 4$ and $N_{rRF} = 4$.

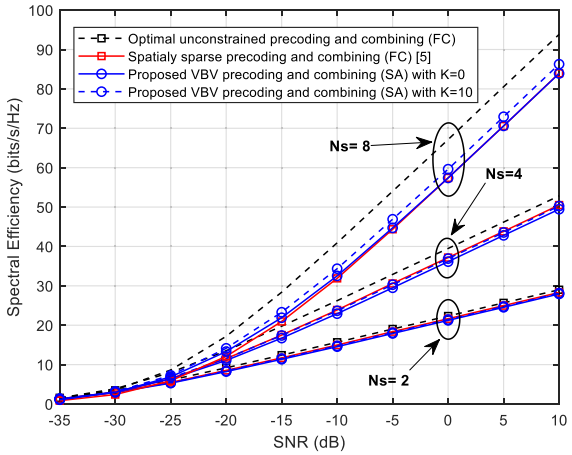

FIGURE 4. Achievable spectral efficiency comparison for $N_t = 256$, $N_r = 64$, $N_{tRF} = 32$ and $N_{rRF} = 8$.

We generate the channel matrix of mmWave MIMO system described in Section II. $N_{cl} = 8$ clusters, $N_{ray} = 8$ rays per cluster and an angular spread of 7.5^o are considered. We assume equal power for all clusters [5]. The transmitter's sector angle is assumed to be 60^o -wide in azimuth domain and 20^o -wide in elevation. Both the transmit and receive antenna arrays are ULAs. $\lambda/2$ is the spacing between antenna elements and the carrier frequency is 28 GHz. Perfect channel state information at the BS and the MS is assumed. The SNR is defined as $SNR = \rho/\sigma_n^2$. All simulation results are averaged over 1000 channel realizations.

A. THE FIRST SCENARIO (PRECODING COMBINING)

In this scenario, the proposed VBV precoder algorithms are studied with an optimal combiner at the receiver side, and the spectral efficiency is studied and calculated using (3). Fig. 2 evaluates the spectral efficiency achieved by the proposed VBV hybrid precoding and combining algorithms with different K compared to the optimal unconstrained precoding and combining scheme for a 64×16 system. $N_{tRF} = 4$ and $N_{rRF} = 4$ are assumed.

It is clearly seen that the performance of Algorithm 2 ($K \neq 0$) improves slowly when K is increased from 10 to 20 and thus $K = 0$ is chosen as a sufficient value for the iterative solution in algorithm 2 and $K = 10$ will be used for the iterative solution in the next figures.

Fig. 3 shows the achievable spectral efficiency performances of the proposed VBV hybrid precoding and combining algorithms 1 ($K = 0$) and 2 ($K = 10$), the sparse precoding and combining, and the optimal unconstrained precoding and combining for a 64×16 system. $N_{tRF} = N_{rRF} = 4$ and $N_s = 1$ and 2 are assumed. We can see from Fig. 3 that the proposed VBV hybrid precoding and combining in algorithm1 provides performance close to that obtained from the sparse hybrid precoding of [5]. Moreover, algorithm 2 provides a little increase in the spectral efficiency compared to that of the sparse hybrid precoding of [5] when $K = 10$. The superiority of the proposed VBV scheme in algorithm 2 as compared with that in [5] is because the proposed VBV precoding and combining algorithm in this scenario uses an optimal combiner at the receiver side, which leads to improved system performance. However, as we will see in the scenario 2, the performances of the proposed hybrid precoding and combining algorithms 1 and 2 approach that in [5] when we compared the precoders only at the transmitter side, but it does not surpass it.

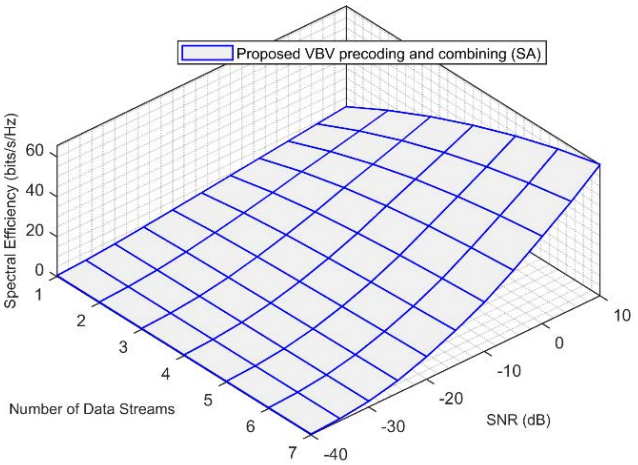

FIGURE 5. Achievable spectral efficiency versus the number of data streams and the SNR in dB for the proposed VBV precoding and combining scheme for $N_t = 256$, $N_r = 64$, $N_{tRF} = 16$ and $N_{rRF} = 4$.

The performances of the proposed VBV hybrid precoding and combining algorithms 1 and 2 are also investigated and tested for a large mmWave MIMO system, as shown in Fig. 4 at different N_s. Fig. 4 illustrates the spectral efficiency achieved by the proposed VBV hybrid precoding and combining algorithms 1 and 2, the sparse hybrid precoding and combining and the optimal unconstrained hybrid precoding and combining for a 256×64 system with $N_{tRF} = 32$, $N_{rRF} = 4$ and $N_s = 2$, 4, and 8. The results in Fig.4 show that the proposed VBV hybrid precoding and combining algorithms 1 ($K = 0$) and 2 ($K = 10$) still provide performances approach that in [5] when $N_s = 2$, and 4 and algorithm 2 outperform that in [5] when $N_s = 8$ because of the optimal combining at the receiver side.

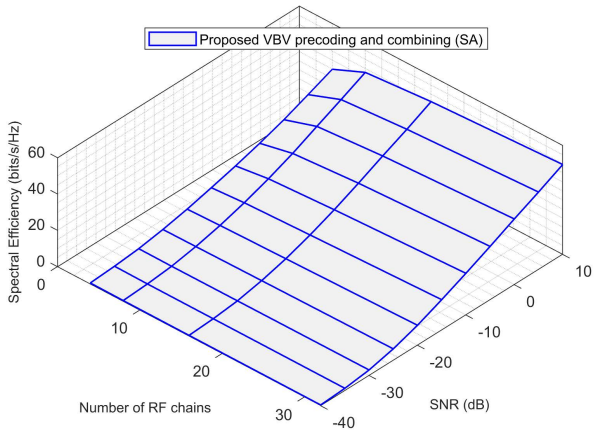

FIGURE 6. Achievable spectral efficiency versus the number of RF chains and the SNR in dB for the proposed VBV precoding and combining scheme for $N_t = 256$, $N_r = 64$, $N_s = 4$ and $N_{rRF} = 4$.

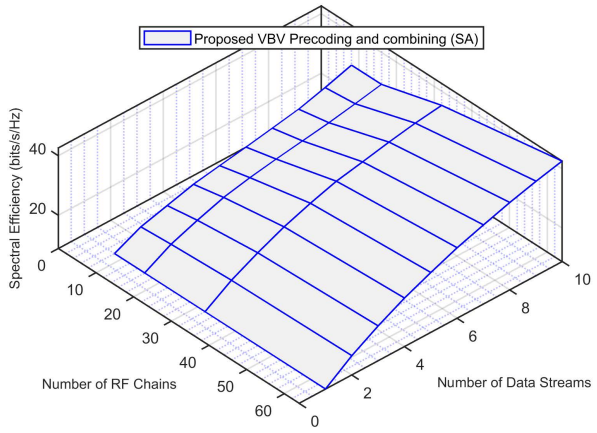

FIGURE 7. Achievable spectral efficiency versus the number of RF chains and the number of data streams for the proposed VBV precoding and combining scheme for $N_t = 256$, $N_r = 64$, $N_{rRF} = 4$, and $SNR = -10(dB)$.

To validate the performance of the proposed VBV hybrid precoding and combining scheme, Fig. 5 shows the spectral efficiency versus the number of data streams and the *SNR* as a three-dimensional plot for a 256×64 system with $K = 0$, $N_{tRF} = 16$ and $N_{rRF} = 4$. It is clear that the spectral efficiency starts to increase rapidly as we increase the number of transmitted streams when $SNR \geq -30\,dB$.

Fig. 6 illustrates the spectral efficiency of the proposed VBV scheme versus the number of RF chains and the *SNR* as a three-dimensional plot for a 256×64 system with $K = 0$, $N_{rRF} = 4$ and $N_s = 4$. From Fig. 6, it can be concluded that as the number of RF increases, the spectral efficiency of the system increases rapidly, even if the number of antennas at the transmitter and receiver, the number of RF chains of the receiver, and data streams are constant, especially when $SNR \geq -30\,dB$. This figure also shows that with the assumed system parameters, the performance of the proposed algorithm improves slowly when $N_{tRF} > 16$ and thus $N_{tRF} = 16$ is sufficient for the proposed VBV precoding and combining in this scenario to provide good performance with lower complexity.

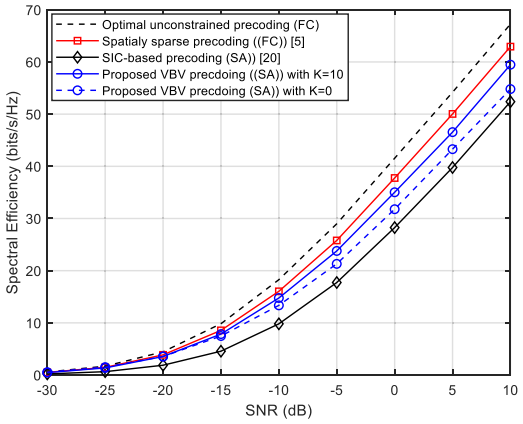

FIGURE 8. Achievable spectral efficiency comparison for $N_t = 64$, $N_r = 16$, and $N_{tRF} = 4$.

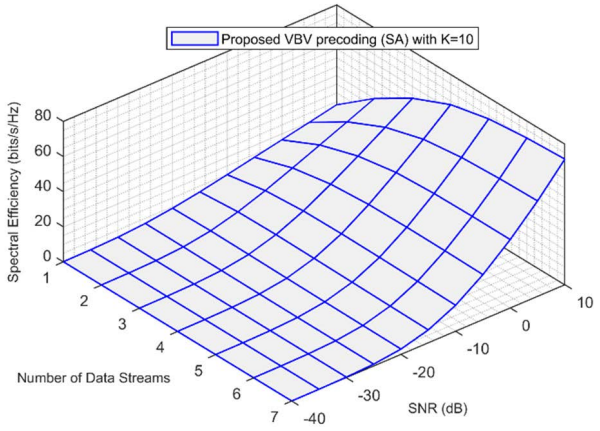

FIGURE 9. Achievable spectral efficiency versus the number of data streams and the SNR in dB for the proposed VBV precoding scheme for $N_t = 64$, $N_r = 16$ and $N_{tRF} = 8$.

The impact of the number of the transmitted data streams and the number of the RF chains on the spectral efficiency performance of the proposed VBV precoding and combining scheme with $K = 0$ is studied and shown in Fig. 7 for a 256×64 system when $SNR = -10\,dB$ and $N_{rRF} = 4$. It is clearly seen that the achievable spectral efficiency has a significant increase with the increase in the numbers of RF chains and data streams. This figure also shows that $N_{tRF} = 16$ is suitable for the proposed VBV precoding and combining scheme when N_s is low and $N_{tRF} = 32$ is suitable when N_s is high.

B. THE SECOND SCENARIO (PRECODING ONLY)

For the sake of fairness, the precoding stage at the transmitter side for the proposed algorithms 1 and 2, the optimal unconstrained precoding, the sparse precoding in [5] and the SIC-based precoding in [20] is only considered in this scenario, and the spectral efficiency is calculated using (6).

Fig. 8 illustrates the spectral efficiencies versus the SNR values of the proposed VBV precoder algorithms 1 ($K = 0$) and 2 ($K = 10$), the optimal unconstrained precoder, the spatially sparse precoder [5], and the SIC-based precoder [20]. In the simulation, we assume a $\mathbf{64 \times 16}$ system with $N_{tRF} = 4$, and $N_s = 4$. From Fig. 8, it can be concluded that as the

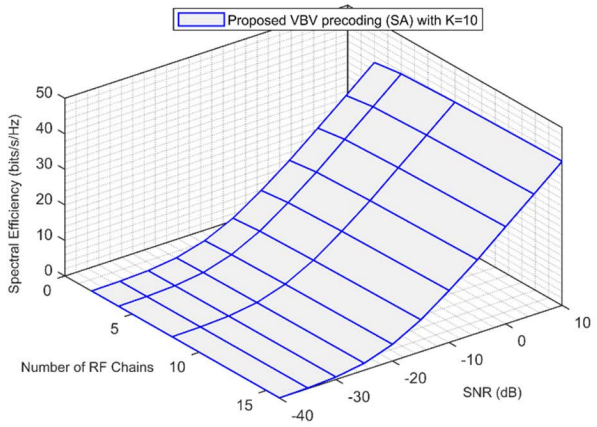

FIGURE 10. Achievable spectral efficiency versus the number of data streams and the SNR in dB for the proposed VBV precoding scheme for $N_t = 64$, $N_r = 16$ and $N_s = 2$.

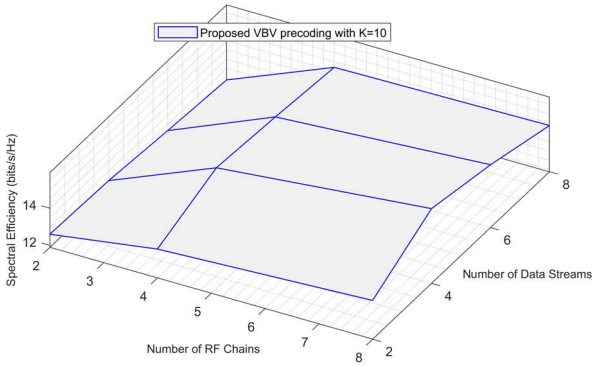

FIGURE 11. Achievable spectral efficiency versus the number of data streams and the SNR in dB for the proposed VBV precoding scheme for $N_t = 64$, $N_r = 16$ and SNR $= -10$dB.

SNR increases, the spectral efficiency of the proposed VBV precoding in algorithms 2 increases rapidly and provides performance near that of [5] with low complexity. Fig 8 also shows that the proposed VBV algorithms 1 and 2 outperform that in [20].

Fig. 9 illustrates the spectral efficiency versus the number of data streams and the SNR. In the simulation, we assume a 64×16 system with $N_{tRF} = 8$ and $K = 10$. From Fig. 9, it can be concluded that as the number of data streams increases, the spectral efficiency of the proposed VBV precoding scheme starts to rapidly increase when $SNR > -25\,dB$, even if the numbers of transmitting antennas and RF chains do not change. On the other hand, the performance of the proposed scheme increases with the increases of the number of the data streams at high SNR, even at a constant SNR.

Fig. 10 shows the spectral efficiency against the number of RF chains and the SNR at the same time. In the simulation, we assume a 64×16 system with $N_s = 2$ and $K = 10$. We can easily observe that the performance of the proposed VBV precoder when $SNR > -25\,dB$ starts to increase, even if the number of transmitting antennas and data streams does not change. Moreover, Fig. 10 shows that increasing RF chains lead to an increase in the system performance. It is also noted that the increase in system performance when the number of RF chains is greater than 4 is small, and thus this value is sufficient to use with the proposed algorithm in this scenario and to provide performance outperforms that of the SIC-based precoding in [20] as shown in Fig. 8, with lower complexity.

The impact of the numbers of data streams and RF chains on the spectral efficiency of the proposed VBV precoder is studied at a fixed transmitter antennas and a $SNR = -10\,dB$ using a three-dimensional plot and shown in Figure 11. In the simulation, we assume a 64×16 system and $K = 10$. It is clearly seen that the spectral efficiency of the proposed VBV precoder increases slowly when $N_{tRF} > 4$ and $N_s > 4$ and thus $N_{tRF} = 4$ is sufficient to provide the required performance and maintain low complexity.

VII. CONCLUSION

In this paper, the issue of single user hybrid precoding for SA architecture in narrowband mmWave MIMO systems has been studied. We first showed that the optimization problem of the hybrid precoding for SA architecture can be decomposed into a series of L_t subproblems, where each of them represents one subarray, and then into a series of $L_t N_s$ subproblems, where each of them represents one vector. We showed that the analog precoding vector of each subarray can be determined from the first vector of each subarray hybrid precoding submatrix, and then the obtained analog precoding vector will be used to determine the elements of the digital precoding matrix from each vector in the subarray hybrid precoding submatrix using a simple MRC method. Then, a low-complexity VBV hybrid precoding algorithm is proposed for SA architecture narrowband mmWave MIMO systems for a general case where there is no need to assume equal numbers of data streams and RF chains. Finally, an iterative VBV hybrid precoding algorithm for SA architecture is proposed. Simulation experiments have been performed on two scenarios, and results have shown that the proposed VBV hybrid precoding algorithms achieve better performance than that of the SIC-based subarray precoder in [20] and close performance to that of the FC spatially sparse hybrid precoding in [5]. It is also found that $K = 10$ in Algorithm 2 is sufficient to provide performance close to that in [5] and outperforms that in [2], even at low N_{tRF}. This indicates that the proposed VBV hybrid precoding algorithms are suitable for the SA architecture in narrowband mmWave systems. For future work, it would be interesting to extend this work to a wideband mmWave system and also to consider the combining at the receiver side.

REFERENCES

[1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[2] F. Boccardi, R. W. Heath, Jr., A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[3] 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.

[4] S. A. Busari, S. Mumtaz, S. Al-Rubaye, and J. Rodriguez, "5G millimeter-wave mobile broadband: Performance and challenges," *IEEE Commun. Mag.*, vol. 56, no. 6, pp. 137–143, Jun. 2018.
[5] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, Jr., "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Jan. 2014.
[6] L. Liang, W. Xu, and X. Dong, "Low-complexity hybrid precoding in massive multiuser MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 653–656, Dec. 2014.
[7] S. Han, C.-L. I, Z. Xu, and C. Rowell, "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G,' *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 186–194, Jan. 2015.
[8] R. Méndez-Rial, C. Rusu, N. González-Prelcic, and R. W. Heath, Jr., "Dictionary-free hybrid precoders and combiners for mmWave MIMO systems," in *Proc. IEEE 16th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2015, pp. 151–155.
[9] A. Alkhateeb, G. Leus, and R. W. Heath, Jr., "Limited feedback hybrid precoding for multi-user millimeter wave systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6481–6494, Nov. 2015.
[10] C. Rusu, R. Mèndez-Rial, N. González-Prelcic, and R. W. Heath, Jr., "Low complexity hybrid precoding strategies for millimeter wave communication systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8380–8393, Dec. 2016.
[11] A. Li and C. Masouros, "Hybrid analog-digital millimeter-wave MU-MIMO transmission with virtual path selection," *IEEE Commun. Lett.*, vol. 21, no. 2, pp. 438–441, Feb. 2017.
[12] 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.
[13] 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.
[14] Y. Liu, Q. Feng, Q. Wu, Y. Zhang, M. Jin, and T. Qiu, "Energy-efficient hybrid precoding with low complexity for mmWave massive MIMO systems," *IEEE Access*, vol. 7, pp. 95021–95032, 2019.
[15] Y. Hei, S. Yu, C. Liu, W. Li, and J. Yang, "Energy-efficient hybrid precoding for mmWave MIMO systems with phase modulation array," *IEEE Trans. Green Commun. Netw.*, vol. 4, no. 3, pp. 678–688, Sep. 2020.
[16] T. Ren and Y. Li, "Hybrid precoding design for energy efficient millimeter-wave massive MIMO systems," *IEEE Commun. Lett.*, vol. 24, no. 3, pp. 648–652, Mar. 2020.
[17] 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.
[18] H. M. T. Mustafa and S. J. Lee, "Hybrid beamforming for sum rate maximization in wideband multi-user MIMO relay systems," *IEEE Access*, vol. 9, pp. 167521–167544, 2021.
[19] Y. Zhang, Y. Lian, Y. Liu, Q. Zhang, M. Jin, and T. Qiu, "Energy-efficient multi-antenna hybrid block diagonalization precoding and combining for mmWave massive multi-user MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 70, no. 10, pp. 10461–10476, Oct. 2021.
[20] X. Gao, L. Dai, S. Han, I. Chih-Lin, and R. W. Heath, Jr., "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.
[21] 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.
[22] S. Park, A. Alkhateeb, and R. W. Heath, Jr., "Dynamic subarrays for hybrid precoding in wideband mmWave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2907–2920, May 2017.
[23] 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.
[24] P. Zhao and Z. Wang, "Hybrid precoding for millimeter wave communications with fully connected subarrays," *IEEE Commun. Lett.*, vol. 22, no. 10, pp. 2160–2163, Oct. 2018.
[25] 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.
[26] 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.
[27] 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.
[28] 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.
[29] 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.
[30] 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.
[31] 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.
[32] N. T. Nguyen and K. Lee, "Unequally sub-connected architecture for hybrid beamforming in massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 1127–1140, Feb. 2020.
[33] J.-C. Guo, Q.-Y. Yu, W.-X. Meng, and W. Xiang, "Energy-efficient hybrid precoder with adaptive overlapped subarrays for large-array mmWave systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 1484–1502, Mar. 2020.
[34] R. Magueta, D. Castanheira, A. Silva, R. Dinis, and A. Gameiro, "Hybrid multi-user equalizer for massive MIMO millimeter-wave dynamic subconnected architecture," *IEEE Access*, vol. 7, pp. 79017–79029, 2019.

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 the University of Minufia, Egypt, in 2008 and 2011, respectively. He joined the Electrical Engineering Department, University of Ibb, Ibb, Yemen, as an Assistant Professor, in 2011, where he became an Associate Professor, in January 2017. He worked as the Head of the Electrical Engineering Department, Faculty of Engineering, University of Ibb, from 2013 to 2014, and as the Vice Dean of the Academic Affairs of the Faculty of Engineering, University of Ibb, from 2018 to 2020. Since November 2021, he has been a Visiting Scholar with the Faculty of Engineering, School of Electrical Engineering and Computer Science, University of Ottawa, Canada. His research interest includes 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 mmWave and massive MIMO systems. He was a recipient of the Prize of 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. In 1992, he was employed as a Systems Engineer at Calian Communications Systems Ltd. In 1995, he was employed as a Systems Engineer at the Communications Research Centre, Ottawa. Later in 1995, he joined the Department of Electrical and Computer Engineering, Royal Military College of Canada, as an Assistant Professor. He joined the School of Information Technology and Engineering, which has since been renamed as the School of Electrical Engineering and Computer Science, University of Ottawa, in 1999. He was the Vice Dean of Undergraduate Studies with the Faculty of Engineering, University of Ottawa, from 2007 to 2011, and has been the Director of the School of Electrical Engineering and Computer Science, since 2013. His research interest includes physical layer technologies for wireless and optical communication systems, such as multiple access, interference cancellation, modulation, coding, and signal processing.