IJICTR
International Journal of Information & Communication Technology Research

Volume 16- Number 1 – 2024 (1 -10)

A Dynamic Hybrid Precoding Structure for mmWave Massive MIMO Systems

Amirreza Moradi
Department of Electrical Engineering
K. N. Toosi University of Technology
Tehran, Iran
Amirreza_moradi@email.kntu.ac.ir

K. Mohamed-pour *
Department of Electrical Engineering
K. N. Toosi University of Technology
Tehran, Iran
kmpour@kntu.ac.ir

Nasim Jafari Farsani
Department of Electrical Engineering
K. N. Toosi University of Technology
Tehran, Iran
Jafari_nasim@email.kntu.ac.ir

Received: 5 July 2023 – Revised: 23 September 2023 - Accepted: 19 November 2023

Abstract—**Due to the high-power consumption and complexity of fully digital baseband precoding, its implementation in massive millimeter-wave multiple-input multiple-output (MIMO) systems is not cost-efficient and practical; for this reason, hybrid precoding has attracted a lot of attention in recent years. Most hybrid precoding techniques concentrate on the fully-connected structure, although they require lots of phase shifters, which is high energy-consuming. On the contrary, the partially-connected structure has low power consumption, nevertheless, suffers from a severe decrease in spectral efficiency (SE). To enhance SE, this paper proposed a dynamic hybrid precoding structure where a switch network is able to provide dynamic connections from phase shifters to radio frequency (RF) chains. To determine the digital precoder and the states of switch, a novel alternating minimization algorithm is proposed, which leverages closed-form solutions at each iteration to efficiently converge to an optimal solution. Furthermore, the phase shifter matrix is optimized through an iterative solution. The simulation results show that in terms of SE, the proposed algorithm with a dynamic structure achieves higher performance than the partial structure. Also, since the proposed structure reduces the number of phase shifters, it can guarantee better energy efficiency (EE) than the fully connected structure.**

Keywords: Massive MIMO, Spectral efficiency, Hybrid precoding, Millimeter-wave communication, Energy efficiency

Article type: Research Article

Publisher: ICT Research Institute

I. INTRODUCTION

Growing demand for high-speed capacity and flexible network architecture has drawn researchers into promising new technologies such as Massive MIMO to enhance the SE and available bandwidth for the current fifth-generation mobile communication (5G) systems [1, 2]. The past few years have seen extensive studies on various areas of advanced technologies, including relay cooperation, diversity multiplexing, and cognitive radio, which focused on increasing SE. However, excessive demand has put a burden on obtaining proper gain [3], [4], [5]. Accordingly, the best solution is to study other radio frequency bands to prevent the increase of crowded and overused frequency bands. Millimeter-wave (mmWave) MIMO systems are an important technology to meet the high throughput capacity requirements of 5G wireless communication systems [6]. Gigabit-per-second data rates of mmWaves allow them to be implemented in current 5G wireless systems [7]. The main feature of mmWave communications compared to the frequency bands currently used is its tremendous increase in carrier frequency, in which the carrier frequency will experience a tenfold elevation

* Corresponding Author

compared to current wireless systems, meaning that mmWave signals will undergo a gradual increase in the loss of free space path. Therefore, it is difficult to ensure a sufficient signal-to-noise ratio (SNR) for a conventional MIMO systems [8], [11], [12]. The two major advantages of millimeter wavelengths over UHF (Ultra-High Frequency) and microwave bands are the very wide available bandwidth and the very short wavelengths that make it possible to implement tens or hundreds of antenna elements in a reasonable physical space. Therefore, by combining MIMO and mmWave technologies, higher data rates and better SE can be achieved, and lower latency can be expected [13], [14].

Currently, the best performance in mm-wave systems belongs to fully digital precoding [13]. In this case, each antenna requires a separate RF chain, which with a very high number of antennas leads to high energy consumption, cost, and hardware complexity. In fully analog precoding, the antennas are classified and assigned the same RF chain to the antennas of each category. By doing this, the energy consumption, cost, and hardware complexity will be improved [16] ,[17]. However, since it only applies to a single data stream scenario, it is not optimal for improving SE. To address this issue, a concept called hybrid precoding has been introduced, where the precoding process is divided into two stages, analog precoding and digital precoding [18] ,[19]. This process can be done in the digital field by baseband signal processing, and in the analog field, usually by analog phase shifters in a situation where all the matrix elements have an equal norm. Hybrid precoding for Massive MIMO precoding is a practical solution. Massive MIMO precoding and mmWave precoding both face the problem of reducing the number of RF chains. In addition, mmWave communications can be considered as a specific scenario of Massive MIMO operating at high frequencies [18].

Based on how each RF chain is connected to the antennas, the hybrid precoding structures are classified into two categories: Fully connected structure and partially connected structure [14]. In the fully connected structure, every separate RF chain is connected to all antennas, which can achieve excellent precoding performance [19]. However, since the number of phase shifters required is equal to the product of the number of RF chains and antennas, a large number of phase shifters are still needed. Therefore, equipping the system with a huge number of phase shifters leads to high hardware cost, complexity, and power consumption.

Several optimization algorithms are introduced to compute digital and analog precoders, among which one of the most important and widely-used algorithms is the Orthogonal Matching Pursuit (OMP) algorithm [16]. In a two-step procedure with respect to the analog precoding matrix $\boldsymbol{F}_{RF}$ and the digital precoding matrix $\boldsymbol{F}_{BB}$ the authors of [20] propose to maximize the channel mutual information

In the partially connected structure, every RF chain is linked to only a subset of antennas, and each antenna element only connects to one phase shifter. compared to the fully connected structure, reduces the number of phase shifters and hardware cost, complexity, and power consumption. However, the drastic reduction of its SE performance will be disappointing. Based on the successive interference cancellation (SIC), authors in [21] proposed an iterative hybrid precoding algorithm. A hybrid precoding codebook construction algorithm was proposed in [22]. In this method, the digital precoding matrix is reconstructed by using the idea of compressive sampling matching pursuit (CoSaMP) algorithm. Afterward, by combining the dictionary learning algorithm, the digital precoding matrix and analog codebook are updated. Ultimately, the final digital precoding codebook was obtained through iterative optimization. Unfortunately, the certain performance loss is the consequence of using a low-design complexity codebook.

In last recent years, to reduce the power consumption and hardware complexity of phase shifters, the on-off binary states switches have been proposed for hybrid precoding; the authors In [23] proposed three low-complexity algorithms that use switches to select antenna subsets for each RF chain and fixed phase shifters to control the phases of signals in the RF circuit. For designing a switch-based hybrid precoder, a new method was proposed in [24], in which the joint optimization of analog and digital precoders decoupled by bounding the problem to a rank-constrained subspace. Then the solution approximates by majorization (SHD-QRQU) and norm maximization (SHDNM). Sadly, the binary states of the switches limit the SE of switch-based hybrid precoding structures, leading to poor performance. Therefore, a dynamic structure is proposed to achieve the appropriate SE with low power consumption, hardware cost, and complexity.

In summary, the major contributions of this paper are as follows:

- A dynamic structure for hybrid precoding is proposed, which, by using an analog switch network, each RF chain is able to be dynamically connected to all antennas. Also, a unit modulus diagonal analog phase shifter precoding matrix is used since each antenna element only connects to one phase shifter.
- Switch and digital precoding matrix is derived using the Alternating Minimization algorithm. Also, in general cases, the digital precoding matrix is optimized by the least square method. Moreover, we calculate the phase shifter matrix through an iterative solution.
- In the end, we compare the proposed hybrid precoding algorithm with the OMP algorithm, SIC-based precoder, and fully digital precoder regarding SE and EE. Based on the simulation results, the proposed hybrid precoding based on the dynamic structure can achieve proper SE with a huge reduction in power consumption. The remaining of this manuscript is organized as follows. The hybrid precoding system model and the

mmWave channel assumption are introduced in

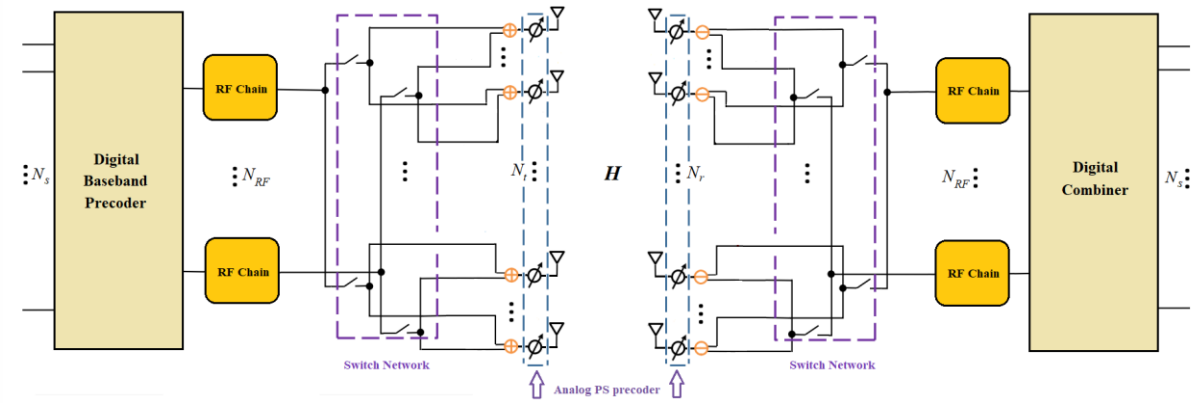

Figure 2. Hybrid precoder design with dynamic structure

Section II; in Section III, first, the hybrid precoding problem of the proposed structure is described, and next, the hybrid precoding algorithm and the EE performance are elaborately discussed. Simulation results are evaluated in Section IV. Finally, in Section V, the conclusion is summarized.

The following notation is used throughout this paper: $\mathbb{C}$ is the set of complex numbers; $\Re\{\cdot\}$ shows the real part of a complex variable; $\boldsymbol{X}$ is a matrix; $\boldsymbol{x}$ is a vector. The conjugate transpose of $\boldsymbol{X}$ is denoted by $\boldsymbol{X}^{\mathrm{H}}$; $\boldsymbol{X}_{i,j}$ is the entry in the ith row and jth column of $\boldsymbol{X}$; $\|\boldsymbol{X}\|_F$ is the Frobenius norm of $\boldsymbol{X}$; $|\boldsymbol{X}|$ is the determinant of $\boldsymbol{X}$; $(:)$ denotes the trace/Frobenius product; $\boldsymbol{X}^{-1}$ and $\boldsymbol{X}^{\dagger}$express the inverse and the Moore-Penrose pseudo inverse of matrix $\boldsymbol{X}$; Expectation is indicated by $\mathbb{E}[\cdot]$; $\mathrm{tr}(\boldsymbol{X})$ is the trace of matrix $\boldsymbol{X}$; $\mathrm{diag}(\boldsymbol{X})$ return the main diagonal of matrix $\boldsymbol{X}$ as a column vector; $\mathrm{Diag}(\boldsymbol{x})$ creates a diagonal matrix from vector $\boldsymbol{x}$; $\oslash$ stands for element-wise division; $\mathbb{1}(\cdot)$ represents the indicator function; $\mathbf{1}_{M\times N}$ denotes an $M\times N$ matrix of ones. $\mathrm{XN}(x,y)$ denotes the complex Gaussian distribution with x and y as the mean and covariance, respectively; $\boldsymbol{I}_N$ is the $N \times N$ identity matrix.

II. System Model

A Hybrid precoder design with a dynamic structure for a Single-user mmWave Massive MIMO system demonstrates in Fig.1. The base station is equipped with N_t transmitting antennas, and on the receiver side, we suppose a single user has N_r receiving antennas. To transmit N_s data streams, antennas are stacked as a uniform plane array (UPA). The transmitter and receiver are equipped with multiple RF chains to ensure transferring multiple data streams in a parallel mode. The number of RF chains is shown as N_{RF}, and it satisfies $N_s \le N_{RF} \le min\{N_t.N_r\}$ constraints.

Mathematically, the transmitted signal $N_t \times 1$ vector is written as $\boldsymbol{x} = \boldsymbol{F}_{PS}\boldsymbol{F}_S\boldsymbol{F}_{BB}\boldsymbol{s}$. Also, $\boldsymbol{s} = [s_1,\ldots,s_{N_s}]^T$ denotes the $N_S\times 1$ transmitted signal vector, and the average transmitted power of the Gaussian data symbol can be expressed as $\mathbb{E}\left[\boldsymbol{s}\boldsymbol{s}^{\mathrm{H}}=\frac{\boldsymbol{I}_{N_S}}{N_S}\right]$.

The hybrid precoding structure contains three processing stages: digital (baseband) precoder $\boldsymbol{F}_{BB} \in \mathrm{X}^{N_{RF}\times N_S}$, an analog switch precoding matrix $\boldsymbol{F}_S \in \mathrm{X}^{N_t\times N_{RF}}$ followed by an analog phase shifter precoding matrix $\boldsymbol{F}_{PS} \in \mathrm{X}^{N_t\times N_t}$. The following expression is denoted as the received signal after the decoding processes.

$$y = \sqrt{\rho}\,(\boldsymbol{W}_{PS}\boldsymbol{W}_S\boldsymbol{W}_{BB})^{\mathrm{H}}\,\boldsymbol{H}\,\boldsymbol{F}_{PS}\boldsymbol{F}_S\boldsymbol{F}_{BB}\boldsymbol{s} + (\boldsymbol{W}_{PS}\boldsymbol{W}_S\boldsymbol{W}_{BB})^{\mathrm{H}}\,\mathbf{n} \tag{1}$$

where ρ and $\boldsymbol{H}$ denote the average received power and $N_r \times N_t$ scattering channel matrix, respectively. The receiver side decoding process is similar to the precoding at the transmitter. Therefore, we have a low dimensional digital (baseband) combiner $\boldsymbol{W}_{BB} \in \mathrm{X}^{N_{RF}\times N_S}$, an analog switch combining matrix $\boldsymbol{W}_S \in \mathrm{X}^{N_r\times N_{RF}}$ followed by an analog phase shifter combining matrix $\boldsymbol{W}_{PS} \in \mathrm{X}^{N_r\times N_r}$. Noise vector $\mathbf{n}$, which has the dimension of $N_r \times 1$ is assumed to be an independent and identically distributed random variable (i.i.d) with zero mean and average power σ_n^2, i.e., $\mathbf{n} \sim \mathrm{XN}(0, \sigma_n^2)$.

According to (1), the SE obtained by a hybrid precoding system can be expressed as

$$\mathrm{R} = \log_2 \left|\boldsymbol{I}_{N_S} + SINR\right| \tag{2}$$

Where SINR denotes the signal-to-noise-plus-interference ratio and recasts as in

$$\mathrm{SINR} = \frac{\rho}{\sigma_n^2 N_S} \cdot \left|\frac{(\boldsymbol{W}_{PS}\boldsymbol{W}_S\boldsymbol{W}_{BB})^{\mathrm{H}}\,\boldsymbol{H}\,\boldsymbol{F}_{PS}\boldsymbol{F}_S\boldsymbol{F}_{BB}\,(\boldsymbol{F}_{PS}\boldsymbol{F}_S\boldsymbol{F}_{BB})^{\mathrm{H}}\,\boldsymbol{H}^{\mathrm{H}}\,(\boldsymbol{W}_{PS}\boldsymbol{W}_S\boldsymbol{W}_{BB})}{(\boldsymbol{W}_{PS}\boldsymbol{W}_S\boldsymbol{W}_{BB})^{\mathrm{H}}\,(\boldsymbol{W}_{PS}\boldsymbol{W}_S\boldsymbol{W}_{BB})}\right| \tag{3}$$

Let $\boldsymbol{F} = \boldsymbol{F}_{PS}\boldsymbol{F}_S\boldsymbol{F}_{BB}$ and $\boldsymbol{W} = \boldsymbol{W}_{PS}\boldsymbol{W}_S\boldsymbol{W}_{BB}$; then, we can rewrite (3) by (4).

$$\mathrm{SINR} = \frac{\rho}{\sigma_n^2 N_S} \cdot \left|\left(\boldsymbol{W}^{\mathrm{H}}\boldsymbol{W}\right)^{-1}\boldsymbol{W}^{\mathrm{H}}\boldsymbol{H}\boldsymbol{F}(\boldsymbol{F})^{\mathrm{H}}\,\boldsymbol{H}^{\mathrm{H}}\,\boldsymbol{W}\right| \tag{4}$$

Then, we substitute (4) and the definition of noise covariance matrix $\boldsymbol{R}_n = \sigma_n^2\boldsymbol{W}^{\mathrm{H}}\boldsymbol{W}$ into (2). We can rewrite the SE expression with regard to precoding matrices as in (5).

$$R= \log_2 \left| \boldsymbol{I}_{N_S} + \frac{\rho}{N_S} \boldsymbol{R}_n^{-1} \boldsymbol{W}^{H} \boldsymbol{H} \boldsymbol{F} \boldsymbol{F}^{H} \boldsymbol{H}^{H} \boldsymbol{W} \right| \quad (5)$$

In this paper, we suppose that the channel state information (CSI) is perfectly-known for both transceiver sides. Designing the unconstrained optimum digital precoders (combiners), $\boldsymbol{F}_{opt}$ $(\boldsymbol{W}_{opt})$, is usually defined by using the channel singular value decomposition (SVD) [25], [26].

$$\boldsymbol{F}_{opt} = \boldsymbol{V}_{:,1:N_S}$$
$$\boldsymbol{W}_{opt} = \boldsymbol{U}_{:,1:N_S} \quad (6)$$

Where $\boldsymbol{U}$ and $\boldsymbol{V}^{H}$ are left and right singular matrices.

Due to the characteristics of extreme path loss and high level of correlation between antennas, the traditional channel model is not suited for mmWave Massive MIMO systems. Therefore, the narrowband clustered channel modeling based on the Saleh-Valenzuela model is adopted [27] to describe the propagation environment between the transmitter and receiver. In this model, we assume the channel matrix $\boldsymbol{H}$ contains of N_{cl} clusters, and each cluster is combined of N_{ray} rays. Consequently, the total number of propagation paths would be $L=N_{cl}N_{ray}$; therefore, the channel matrix $\boldsymbol{H}$ can be written as

$$\boldsymbol{H}=\sqrt{\frac{N_t N_r}{L}} \sum_{i=1}^{N_{cl}} \sum_{k=1}^{N_{ray}} \alpha_{ik}\, \mathbf{a}_r\left(\emptyset_{ik}^{r},\theta_{ik}^{r}\right) \mathbf{a}_t^{H}\left(\emptyset_{ik}^{t},\theta_{ik}^{t}\right) \quad (7)$$

Where α_{ik} denotes the complex gain of the kth ray in the ith scattering cluster, The azimuth angle of departure (AOD) and angle of arrival (AOA), and the elevation AOD and AOA associated with the kth ray in ith cluster denoted by $\emptyset_{ik}^{t}$, $\emptyset_{ik}^{r}$, θ_{ik}^{t} and θ_{ik}^{r}, respectively. Moreover, $\mathbf{a}_t\left(\emptyset_{ik}^{t},\theta_{ik}^{t}\right)$ and $\mathbf{a}_r\left(\emptyset_{ik}^{r},\theta_{ik}^{r}\right)$ indicates the array response vector of the transmitter and receiver[28].

The array response vector for an $M \times N$, UPA recast as

$$\boldsymbol{a}_{\mathrm{UPA}}(\emptyset,\theta) = \frac{1}{\sqrt{MN}}\left[1 \ldots e^{j\pi(m\sin\emptyset\sin\theta+n\cos\theta)} \ldots e^{j\pi((M-1)(\sin\emptyset\sin\theta+(N-1)\cos\theta)}\right]^{T}$$

Where $0 \le m \le M-1$ and $0 \le n \le N-1$ and inter-element space is assumed to be half-wavelength.

III. PROBLEM FORMULATION

Based on the proposed structure, the corresponding hybrid precoding problem can be approximately formulated as follows [16]

$$\begin{aligned} \min_{F_{PS},F_S,F_{BB}} & \left\|\boldsymbol{F}_{opt}- \boldsymbol{F}_{PS}\boldsymbol{F}_S\boldsymbol{F}_{BB}\right\|_F^2 \\ \text{s.t.} & [\boldsymbol{F}_S]_{ij} \in \{0,1\}. \quad \forall i,j \\ & \left|f_{PS,m}\right| = 1 \quad \forall m \\ & \|\boldsymbol{F}_{PS}\boldsymbol{F}_S\boldsymbol{F}_{BB}\|_F^2 = N_S \end{aligned} \quad (8)$$

Also, the mathematical formulation of the combining problems is similar to precoding, except that the additional constraint of the transmit power is eliminated in the problem i.e.

$$\begin{aligned} \min_{\boldsymbol{F}_{PS},\boldsymbol{F}_S,\boldsymbol{F}_{BB}} & \left\|\boldsymbol{W}_{opt}- \boldsymbol{W}_{PS}\boldsymbol{W}_S\boldsymbol{W}_{BB}\right\|_F^2 \\ \text{s.t.} & [\boldsymbol{W}_S]_{ij} \in \{0,1\}. \quad \forall i,j \\ & \left|w_{PS,m}\right| = 1 \quad \forall m \end{aligned} \quad (9)$$

Where $\boldsymbol{F}_{PS} = \mathrm{diag}\left\{f_{PS,1}.f_{PS,2}.\ldots.f_{PS,N_t}\right\}$ and $\boldsymbol{W}_{PS} = \mathrm{diag}\left\{w_{PS,1}.w_{PS,2}.\ldots.w_{PS,N_r}\right\}$ are diagonal matrices since each phase shifter only connected to one antenna. To ensure the unit modulus constraints of diagonal elements of $\boldsymbol{F}_{PS}$ and $\boldsymbol{W}_{PS}$, we have

$$\left|f_{PS,i}\right| = 1 \qquad (i=1,2,\ldots,N_t)$$
$$\left|w_{PS,i}\right| = 1 \qquad (i=1,2,\ldots,N_r)$$

Furthermore, Switch and combining precoding matrix have binary constraint $[\boldsymbol{F}_S]_{ij} \in \{0,1\}$ and $[\boldsymbol{W}_S]_{ij} \in \{0,1\}$. Finally, $\|\boldsymbol{F}_{PS}\boldsymbol{F}_S\boldsymbol{F}_{BB}\|_F^2 = N_S$ determines the total transmit power constraint.

To concentrate on the hybrid precoding problem and simplify the analysis, we will assume that the receiver-side problem can be solved similarly to the approach presented in this paper. Thus we will not consider the receiver-side problem in detail. Sadly, the binary constraint of the analog switch network $\boldsymbol{F}_S$ and the non-convex characteristic of the phase shifter matrix $\boldsymbol{F}_{PS}$ will make it extremely difficult to jointly optimize the hybrid precoder over $\boldsymbol{F}_S$, $\boldsymbol{F}_{PS}$, and $\boldsymbol{F}_{BB}$ to obtain the optimal solution of (8). Therefore, In the continuation of this paper, a hybrid precoding algorithm for further analysis of the optimization problem will be described.

A. HYBRID PRECODING STRATEGY

The digital precoder matrix $\boldsymbol{F}_{BB}$ has the dimension of $N_{RF} \times N_S$ and $N_S \le N_{RF}$, which makes $\boldsymbol{F}_{BB}$ a tall matrix [20], [29] proves a semi-orthogonal structure for digital precoding matrix, $\boldsymbol{F}_{BB}$ can achieve near-optimal performance; therefore, by taking a similar approach the semi-orthogonal constraint is specified as

$$\boldsymbol{F}_{BB}^{H}\boldsymbol{F}_{BB} = \alpha^2\, \boldsymbol{F}_{DD}^{H}\, \boldsymbol{F}_{DD} = \alpha^2\, \boldsymbol{I}_{N_S} \quad (10)$$

Where $\boldsymbol{F}_{BB} = \alpha\boldsymbol{F}_{DD}$, and $\boldsymbol{F}_{DD}$ is a semi-unitary matrix. Specifically, first, we derive an upper bound on the objective function, and then, based on this bound, develop an alternating minimization algorithm.

B. Objective upper bound

In this subsection, by using (10), we derive an upper bound of the hybrid precoding problem, which enables us to rewrite the objective function in (8) as

$$\left\|\boldsymbol{F}_{opt}\right\|_F^2 - 2\alpha\Re\mathrm{tr}\left(\boldsymbol{F}_{DD}\,\boldsymbol{F}_{opt}^{H}\boldsymbol{F}_{PS}\boldsymbol{F}_S\right) + \alpha^2\|\boldsymbol{F}_{PS}\boldsymbol{F}_S\boldsymbol{F}_{DD}\|_F^2 \quad (12)$$

The eq. (12) still cannot be optimized directly. To tackle this problem, $\|\boldsymbol{F}_{PS}\boldsymbol{F}_S\boldsymbol{F}_{DD}\|_F^2$ will be upper bounded. As aforementioned, $\boldsymbol{F}_{PS}$ is a unit modulus diagonal matrix, i.e., $\boldsymbol{F}_{PS}\boldsymbol{F}_{PS}^{H} = \boldsymbol{F}_{PS}^{H}\boldsymbol{F}_{PS} = \boldsymbol{I}_{N_S}$. Therefore, we have

$$\|\boldsymbol{F}_{PS}\boldsymbol{F}_S\boldsymbol{F}_{DD}\|_F^2 = tr(\boldsymbol{F}_{DD}^{\mathrm{H}}\boldsymbol{F}_S^{\mathrm{H}}\boldsymbol{F}_{PS}^{\mathrm{H}}\boldsymbol{F}_{PS}\boldsymbol{F}_S\boldsymbol{F}_{DD}) = \mathrm{tr}\left(\begin{bmatrix}\boldsymbol{I}_{N_s} & \\ & 0\end{bmatrix}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{F}_S^{\mathrm{H}}\boldsymbol{F}_S\boldsymbol{V}\right) < \mathrm{tr}(\boldsymbol{V}^{\mathrm{H}}\boldsymbol{F}_S^{\mathrm{H}}\boldsymbol{F}_S\boldsymbol{V}) = \|\boldsymbol{F}_S\|_F^2 \tag{13}$$

Where $\boldsymbol{F}_{DD}\boldsymbol{F}_{DD}^{\mathrm{H}} = \boldsymbol{V}^{\mathrm{H}}\begin{bmatrix}\boldsymbol{I}_{N_s} & \\ & 0\end{bmatrix}\boldsymbol{V}$ is the SVD of $\boldsymbol{F}_{DD}\boldsymbol{F}_{DD}^{\mathrm{H}}$. As a result, the objective function of the problem (8) can be reformulated as

$$\|\boldsymbol{F}_{opt}\|_F^2 - 2\alpha\Re\mathrm{tr}(\boldsymbol{F}_{DD}\boldsymbol{F}_{opt}^{\mathrm{H}}\boldsymbol{F}_{PS}\boldsymbol{F}_S) + \alpha^2\|\boldsymbol{F}_S\|_F^2 \tag{14}$$

C. Alternating Minimization

This problem has been studied in [30], [31]. Similarly, the upper bound of (14) is adopted as the surrogate objective function, and the constant term $\|\boldsymbol{F}_{opt}\|_F^2$ is dropped; then, we can reformulate the hybrid precoder design problem as

$$\begin{aligned}\min_{\alpha,\boldsymbol{F}_S,\boldsymbol{F}_{DD}}\ & \alpha^2\|\boldsymbol{F}_S\|_F^2 - 2\alpha\Re\mathrm{tr}(\boldsymbol{F}_{DD}\boldsymbol{F}_{opt}^{\mathrm{H}}\boldsymbol{F}_{PS}\boldsymbol{F}_S)\\ \text{s.t. }& [\boldsymbol{F}_S]_{ij} \in \{0,1\}, \quad \forall i,j\\ & \boldsymbol{F}_{DD}^{\mathrm{H}}\boldsymbol{F}_{DD} = \gamma\boldsymbol{I}_{N_S}\end{aligned} \tag{15}$$

Due to the characteristic of the Alternating Minimization algorithm, we only optimize a subset of the optimization variables, and the other parts should be fixed. By fixing the switch matrix $\boldsymbol{F}_S$ and α, the optimization problem can be recast as

$$\begin{aligned}\max_{\boldsymbol{F}_{DD}}\ & \alpha\Re\mathrm{tr}(\boldsymbol{F}_{DD}\boldsymbol{F}_{opt}^{\mathrm{H}}\boldsymbol{F}_{PS}\boldsymbol{F}_S)\\ \text{s.t. }& \boldsymbol{F}_{DD}^{\mathrm{H}}\boldsymbol{F}_{DD} = \gamma\boldsymbol{I}_{N_S}\end{aligned} \tag{16}$$

Considering the fact in (16) only $\boldsymbol{F}_{DD}$ is variable; it can be reformulated by considering the definition of the dual norm [15],

$$\begin{aligned}\alpha\Re\mathrm{tr}(\boldsymbol{F}_{DD}\boldsymbol{F}_{opt}^{\mathrm{H}}\boldsymbol{F}_{PS}\boldsymbol{F}_S) &\le \left|\mathrm{tr}(\alpha\,\boldsymbol{F}_{DD}\boldsymbol{F}_{opt}^{\mathrm{H}}\boldsymbol{F}_{PS}\boldsymbol{F}_S)\right|\\ &\overset{a}{\le} \|\boldsymbol{F}_{DD}^{\mathrm{H}}\|_\infty\|\alpha\,\boldsymbol{F}_{opt}^{\mathrm{H}}\boldsymbol{F}_{PS}\boldsymbol{F}_S\|_1\\ &\le \|\alpha\,\boldsymbol{F}_{opt}^{\mathrm{H}}\boldsymbol{F}_{PS}\boldsymbol{F}_S\|_1 = \textstyle\sum_{i=1}^{N_S}\sigma_i\,,\end{aligned} \tag{17}$$

Where (a) follows the Holder's inequality. Also, $\|\cdot\|_\infty$ and $\|\cdot\|_1$ stand for the infinite and one Schatten norms, respectively [32]. The equality in (17) holds only if

$$\boldsymbol{F}_{DD} = \boldsymbol{V}\boldsymbol{U}^{\mathrm{H}} \tag{18}$$

Where $\boldsymbol{U}$ and $\boldsymbol{V}$ are the left and right singular matrices of $\alpha\,\boldsymbol{F}_{opt}^{\mathrm{H}}\boldsymbol{F}_{PS}\boldsymbol{F}_S$, respectively. Therefore, in the case when $N_{RF} = N_S$, the digital precoding matrix $\boldsymbol{F}_{BB}$ optimal solution can be obtained by

$$\boldsymbol{F}_{BB}^{opt} = \alpha\boldsymbol{F}_{DD}^{opt} = \alpha\boldsymbol{V}\boldsymbol{U}^{\mathrm{H}} \tag{19}$$

For general cases where ($N_{RF} > N_s$), $\boldsymbol{F}_{BB}$ can be computed by the least square method as

$$\boldsymbol{F}_{BB} = \left(\boldsymbol{F}_S^{\mathrm{H}}\boldsymbol{F}_{PS}^{\mathrm{H}}\boldsymbol{F}_{PS}\boldsymbol{F}_S\right)^{-1}\boldsymbol{F}_S^{\mathrm{H}}\boldsymbol{F}_{PS}^{\mathrm{H}}\boldsymbol{F}_{opt} \tag{20}$$

Usually, the two variables α and $\boldsymbol{F}_S$ can be optimized separately, but to reduce the computational complexity, we propose to update them in parallel. By adding a constant term $\|\Re(\boldsymbol{F}_{PS}^{\mathrm{H}}\boldsymbol{F}_{opt}\boldsymbol{F}_{DD}^{\mathrm{H}})\|_F^2$ to the objective function of (15), the sub-problem of updating $\boldsymbol{F}_S$ and α can be recast as

$$\begin{aligned}\min_{\alpha,\boldsymbol{F}_S}\ & \|\Re(\boldsymbol{F}_{PS}^{\mathrm{H}}\boldsymbol{F}_{opt}\boldsymbol{F}_{DD}^{\mathrm{H}} - \alpha\boldsymbol{F}_S)\|_F^2\\ \text{s.t. }& [\boldsymbol{F}_S]_{ij} \in \{0,1\}, \quad \forall i,j\end{aligned} \tag{21}$$

The optimal solution to this similar problem has been studied in [30]; the problem in Eq. (21) can be equivalently considered by

$$\begin{aligned}\min_{\alpha,\boldsymbol{s}}\ & \|\mathbf{x} - \alpha\mathbf{s}\|_F^2\\ \text{s.t. }& s_i \in \{0,1\}, \quad \forall i\end{aligned} \tag{22}$$

Where $\boldsymbol{x} = \left\|vec\left\{\Re\{\boldsymbol{F}_{PS}^{\mathrm{H}}\boldsymbol{F}_{opt}\boldsymbol{F}_{DD}^{\mathrm{H}}\}\right\}\right\|_F^2$, $\boldsymbol{s} = [s_1,s_2,\ldots,s_k]^{\mathrm{T}} = vec\{\boldsymbol{F}_S\}$, and $k = N_tN_{RF}$ .The entries of $\boldsymbol{x}$ sort in the ascending order as $\tilde{\boldsymbol{x}} = [\tilde{x}_1,\tilde{x}_2,\ldots,\tilde{x}_k]^{\mathrm{T}}$, where $\tilde{x}_1 \le \tilde{x}_2 \le \ldots \le \tilde{x}_k$. To minimize the objective function in (22), if the corresponding $\tilde{x}_j$ is closer to α than 0, then the switch is on, i.e., we have $s_i = 1$, otherwise $s_i = 0$. In that way, the objective function in eq. (22) can be described as

$$\begin{aligned}f(\alpha) &= \|\boldsymbol{x} - \alpha\mathbf{s}\|_2^2\\ &= \begin{cases}\sum_{j=1}^{i}(\tilde{x}_j - \alpha)^2 + \sum_{j=i+1}^{k}\tilde{x}_j^2\,, & \alpha < 0 \quad and\ \frac{\alpha}{2} \in \mathrm{I}_i\\ \sum_{j=1}^{i}\tilde{x}_j^2 + \sum_{j=i+1}^{k}(\tilde{x}_j - \alpha)^2\,, & \alpha > 0 \quad and\ \frac{\alpha}{2} \in \mathrm{I}_i\end{cases}\\ &= \begin{cases}i\alpha^2 - 2\sum_{j=1}^{i}\tilde{x}_j\alpha + \sum_{j=1}^{k}\tilde{x}_j^2\,, & \alpha < 0 \quad and\ \alpha \in \mathrm{P}_i\\ (k-i)\alpha^2 - 2\sum_{j=i+1}^{i}\tilde{x}_j\alpha + \sum_{j=i}^{k}\tilde{x}_j^2\,, & \alpha > 0 \quad and\ \alpha \in \mathrm{P}_i\end{cases}\end{aligned} \tag{23}$$

where $\frac{\alpha}{2} \in I_i = [\tilde{x}_i,\tilde{x}_{i+1}]$, $\alpha \in \mathrm{R}_i = [2\tilde{x}_i,2\tilde{x}_{i+1}]$ i=1,2,…,k

By comparing the values of the quadratic function $f(\gamma)$ in the intervals P_i for all endpoints $2\tilde{x}_i$ and vertexes $\bar{x}_i$, we can reach the optimal solution α_{opt}. It should be noted that since the entries in $\tilde{\boldsymbol{x}}$ are arranged in the ascending form, the optimal solution of α, cannot be the endpoints $2\tilde{x}_i$ of the intervals P_i [30]. Therefore, the optimal solution of α can be written as

$$\alpha_{opt} = \arg\min_{\tilde{x}_i} f(\tilde{x}_i) \tag{24}$$

Where $\bar{x}_i = \begin{cases}\frac{\sum_{j=1}^{i}\tilde{x}_j}{i}, & \bar{x}_i < 0 \quad and\ \bar{x}_\iota \in \mathrm{P}_i\\ \frac{\sum_{j=i+1}^{i}\tilde{x}_j}{k-i}, & \bar{x}_i > 0 \quad and\ \bar{x}_\iota \in \mathrm{P}_i\end{cases}$

And $\tilde{x}_i$ denotes the ith smallest entry in $\boldsymbol{x}$.

Then, by having the α_{opt}, the optimal solution of the problem in eq. (22) can be achieved as follows.

$$\boldsymbol{F}_S^{opt} = \begin{cases}\mathbb{1}\left\{\Re\{\boldsymbol{F}_{PS}^{\mathrm{H}}\boldsymbol{F}_{opt}\boldsymbol{F}_{DD}^{\mathrm{H}}\} > \frac{\alpha_{opt}}{2}\mathbf{1}_{N_t\times N_{RF}}\right\}, & \alpha_{opt} > 0\\ \mathbb{1}\left\{\Re\{\boldsymbol{F}_{PS}^{\mathrm{H}}\boldsymbol{F}_{opt}\boldsymbol{F}_{DD}^{\mathrm{H}}\} < \frac{\alpha_{opt}}{2}\mathbf{1}_{N_t\times N_{RF}}\right\}, & \alpha_{opt} < 0\end{cases} \tag{25}$$

Finally, to satisfy the transmit power constraint $\|\boldsymbol{F}_{PS}\boldsymbol{F}_S\boldsymbol{F}_{BB}\|_F^2 = N_s$, we normalize the digital precoder $\boldsymbol{F}_{BB}$ i.e.

$$\boldsymbol{F}_{BB} = \frac{\sqrt{N_s}}{\|\boldsymbol{F}_{PS}\boldsymbol{F}_S\boldsymbol{F}_{BB}\|_F}\boldsymbol{F}_{BB} \tag{26}$$

D. Analog phase shifter precoding matrix optimization

In this subsection, by fixing the optimized analog switch precoding matrix $\boldsymbol{F}_S$ and the digital precoding matrix $\boldsymbol{F}_{BB}$, we optimize the analog phase shifter precoding matrix $\boldsymbol{F}_{PS}$.

$$\min_{F_{PS}} \left\| \boldsymbol{F}_{opt} - \boldsymbol{F}_{PS}\boldsymbol{F}_S\boldsymbol{F}_{BB} \right\|_F^2 \quad \text{s.t.} \left| f_{PS,m} \right| = 1 \quad \forall m \tag{27}$$

Proposition 1. The optimal solution of (27) is given by

$$\boldsymbol{F}_{PS} = \text{Diag}\left(\text{diag}\left(\boldsymbol{F}_{opt}\boldsymbol{F}_{BB}^H\boldsymbol{F}_S^H - \boldsymbol{F}_{PS}\boldsymbol{F}_S\boldsymbol{F}_{BB}\boldsymbol{F}_{BB}^H\boldsymbol{F}_S^H\right)\right) + \boldsymbol{F}_{PS} \tag{28}$$

Proof: The proof is provided in Appendix A. ■

In the proposed method, we considered $\boldsymbol{F}_{PS}$ as a unit modulus diagonal matrix, i.e., $\boldsymbol{F}_{PS} = \text{diag}\left\{f_{PS,1}.f_{PS,2}.\ldots.f_{PS,N_t}\right\}$ in which $\left|f_{PS,i}\right| = 1$; therefore, an element-wise normalization is proposed to satisfy unit modulus constraints.

$$\boldsymbol{F}_{PS} = \boldsymbol{F}_{PS} \oslash (|\boldsymbol{F}_{PS}|) \tag{29}$$

Algorithm 1 states the pseudo-code for the proposed hybrid precoder $\boldsymbol{F} = \boldsymbol{F}_{PS}\boldsymbol{F}_S\boldsymbol{F}_{BB}$ solution.

Algorithm 1 Proposed Hybrid Precoding

Input: $\boldsymbol{F}_{opt}$

Initialization: $\boldsymbol{F}_{PS}^{(0)}$ and $\boldsymbol{F}_{BB}^{(0)} = \alpha^{(0)}\boldsymbol{F}_{DD}^{(0)}$ are generated randomly, $t = 0$

Repeat

1. $t = t + 1$;
2. Fix $\boldsymbol{F}_{BB}^{(t-1)}$ and $\boldsymbol{F}_{PS}^{(t-1)}$, Optimize $\alpha^{(t)}$ and $\boldsymbol{F}_S^{(t)}$ according to (24) and (25), respectively;
3. **if** $N_{RF} = N_S$
4. Fix $\boldsymbol{F}_{PS}^{(t-1)}$, $\alpha^{(t)}$, $\boldsymbol{F}_S^{(t)}$ compute SVD $\alpha^{(t)}\boldsymbol{F}_{opt}^{\mathrm{H}}\boldsymbol{F}_{PS}^{(t-1)}\boldsymbol{F}_S^{(t)} = \boldsymbol{U}^{(t)}\boldsymbol{\Sigma}^{(t)}\boldsymbol{V}^{(t)^{\mathrm{H}}}$;
5. Update $\boldsymbol{F}_{BB}^{(t)}$ with (19);
6. **else**
7. $\boldsymbol{F}_{BB}^{(t)} = \left(\boldsymbol{F}_S^{(t)^{\mathbf{H}}}\boldsymbol{F}_{PS}^{(t-1)^{\mathbf{H}}}\boldsymbol{F}_{PS}^{(t-1)}\boldsymbol{F}_S^{(t)}\right)^{-1}\boldsymbol{F}_S^{(t)^{\mathrm{H}}}\boldsymbol{F}_{PS}^{(t-1)^{\mathrm{H}}}\boldsymbol{F}_{opt}$
8. **end if**
9. Fix $\boldsymbol{F}_{BB}^{(t)}$ and $\boldsymbol{F}_S^{(t)}$, Optimize $\boldsymbol{F}_{PS}^{(t)}$with (28);
10. Element-Wise Normalization: $\boldsymbol{F}_{PS}^{(t)} = \boldsymbol{F}_{PS}^{(t)} \oslash \left(\left|\boldsymbol{F}_{PS}^{(t)}\right|\right)$

Until convergence

11. $\boldsymbol{F}_{BB} = \frac{\sqrt{N_s}}{\|\boldsymbol{F}_{PS}\boldsymbol{F}_S\boldsymbol{F}_{BB}\|_F}\boldsymbol{F}_{BB}$

Output: $\boldsymbol{F}_{BB}$, $\boldsymbol{F}_S$, $\boldsymbol{F}_{PS}$

E. Convergence

For the convergence of the proposed hybrid precoding to the local minimum point, considering that for the case where $N_s = N_{RF}$ or $N_s < N_{RF}$, the digital precoding matrix $\boldsymbol{F}_{BB}$ has a semi-orthogonal structure. In addition, as mentioned before, $\boldsymbol{F}_{PS}$ is a diagonal matrix of unit modulus, and also considering that the switch and digital precoding matrix are the global optimal solution of the problem, therefore, by considering the uniform convergence theorem and considering each precoding matrix as a BCD block, the proposed algorithm converges to a fixed point [33-34].

F. Computational complexity

In this subsection, the computational complexity of the proposed algorithm is examined. The complexity of each section will be as follows:

The computational complexity of the phase shifter matrix is equal to $\mathcal{O}(rN_t^2N_{RF})$, and r is the number of iterations.

The computational complexity of the digital precoding matrix in case $N_s = N_{RF}$ is equal to $\mathcal{O}(nN_sN_tN_{RF})$, and when $N_s < N_{RF}$, it is $\mathcal{O}(rN_t^2N_{RF})$.

The computational complexity of the switch precoding matrix is equal to $\mathcal{O}(nN_tN_{RF}\log N_tN_{RF})$.

Considering that the number of antennas is much more than the number of RF chains, the computational complexity of the proposed algorithm is equal to $\mathcal{O}(rN_t^2N_{RF})$.

The OMP algorithm has the computational complexity of $\mathcal{O}(N_t^2N_{RF}N_s)$; also, the SIC-Based algorithm has the computational complexity of $\mathcal{O}\left(N_t^2(N_{RF}N_s + N_r)\right)$.

Therefore, the computational complexity of the proposed algorithm is more than the OMP algorithm and depends on the setting when comparing the SIC-based algorithm.

G. Energy efficiency

In this subsection, the EE of the proposed dynamic hybrid precoding structure and other relevant works are formulated. According to [20], [34], the EE of a communications system is specified either by the SE and the total power consumption. The EE in mmWave MIMO systems is represented by η, and can be written as

$$\eta = \frac{R}{P_{total}} \tag{30}$$

where R indicates the hybrid precoding achievable SE, and P_{total} denotes the power consumption of each algorithm which can be represented as

$$P_{total} = \begin{cases} P_t + P_{BB} + P_{RF}N_{RF} + P_{PS}N_t + P_SN_{RF}N_t & \text{PRO} \\ P_t + P_{BB} + P_{RF}N_{RF} + P_{PS}N_{RF}N_t & \text{OMP} \\ P_t + P_{BB} + P_{RF}N_{RF} + P_{PS}N_t & \text{SIC} \end{cases} \tag{31}$$

where P_t denotes the transmit power, P_{BB}, P_{RF}, P_{PS}, and P_S indicate the power consumed by the baseband, each RF chain, phase shifter, and switch, respectively. The high number of phase shifters in the OMP algorithm can cause high power consumption.

Moreover, by using lower power consuming switches, the proposed algorithm has a slightly higher power-consumption compared to the SIC-based algorithm.

IV. Simulation Result

In this section, we assess the SE and the EE performance of the proposed hybrid precoding algorithm by numerical simulations, where the optimal fully digital precoding is considered as a benchmark. Furthermore, the OMP and SIC-based algorithms are taken considered as the competitors. The environment for mmWave propagation is assumed to contain of $N_{cl} = 5$ clusters, and each cluster combined of $N_{ray} = 10$, i.e., the total path number equals to $L = 50$. Both the transmitter and receiver are equipped with UPA. For each cluster, the average azimuth and elevation AOD (AOA) are drawn independently from the uniform distribution over $[0,2\pi)$; In clusters, a 10-degree angular spread Laplace distribution conforms to the azimuth and elevation AOD (AOA). Each path is assumed to have the complex gain of $\mathbf{n}\sim\mathcal{CN}(0,1)$. For evaluating the EE simulation, parameters are set as follows:

$P_t = 1\text{W}$, $P_{BB} = 0.2\text{W}$, $P_{RF} = 0.3\text{W}$, $P_{PS} = 0.05\text{W}$, and $P_S = 0.005\text{W}$ [21], [36].

All of the reported results are achieved by calculating the average of over 1000 independent channel realizations.

In Fig. 2, the SE of all schemes versus SNR are shown where $N_t = 12 \times 12$, $N_r = 4 \times 4$, $N_s = 4$, $N_{RF} = 4$. It is obvious that compared to the OMP algorithm, when $\text{SNR} < -5\text{dB}$, there is only 1Bits/s/Hz gap. But the proposed algorithm approaches the OMP scheme as the SNR increase. Also, compared with the SIC algorithm, the proposed algorithm can show approximately 9 bits/s/Hz gain when $\text{SNR} > 0\text{dB}$.

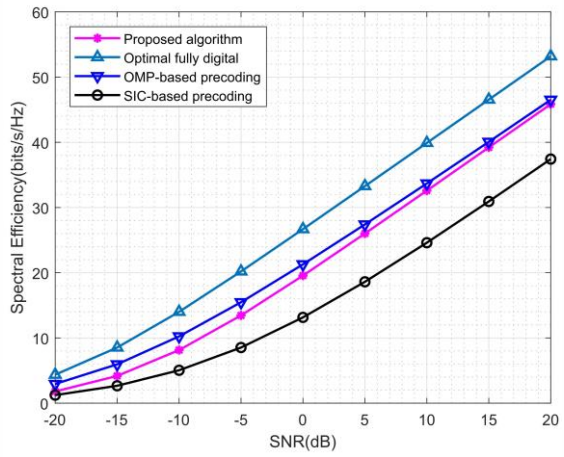

Figure 2. Spectral efficiencies achieved against SNR when $N_t = 12 \times 12$, $N_r = 4 \times 4$, $N_s = 4$, $N_{RF} = 4$

Fig. 3 plots the EE against the SNR. Evaluating meticulously, the proposed algorithm shows a considerable improvement compared to Fully digital and OMP algorithms. Also, compared to the SIC algorithm, the proposed algorithm offers better performance. The EE of our algorithm grows significantly as the SNR increases until when $\text{SNR} = 5\text{dB}$, in which the proposed algorithm has its best performances. The performance of all methods decreases when $\text{SNR} > 5\text{dB}$.

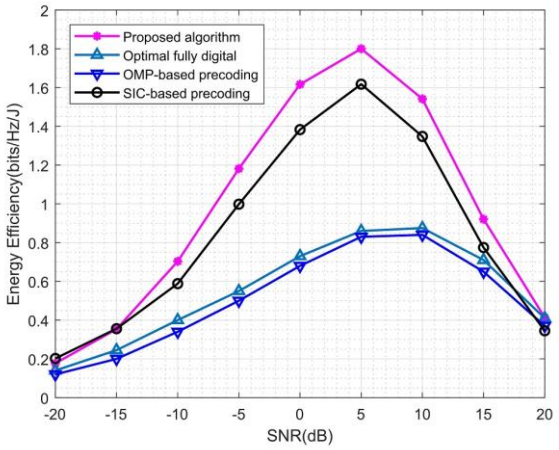

Figure 3. Energy efficiencies achieved against SNR when $N_t = 12 \times 12$, $N_r = 4 \times 4$, $N_s = 4$, $N_{RF} = 4$

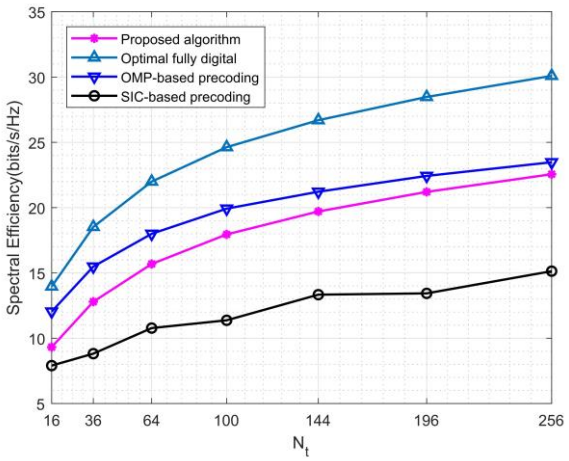

Figure 4. Spectral efficiencies achieved against N_t when $N_r = 4 \times 4$, $N_s = 4$, $N_{RF} = 4$, $\text{SNR} = 0\text{dB}$

Fig.4 and Fig 5 draw the comparisons between SE and the EE performance of the proposed algorithm with other methods as a function of the transmitter antennas N_t when $N_r = 4 \times 4$, $N_s = 4$, $N_{RF} = 4$, $\text{SNR} = 0\text{dB}$. In Fig.4, It's clear by increasing the number of antennas, the SE performance gap between the proposed algorithm and OMP reduced from about 3 bits/s/Hz to only 1 bits/s/Hz. Also, the proposed algorithm provides an approximately 6 bits/s/Hz to 8 bits/s/Hz gain compared to the SIC algorithm when $N_t > 10 \times 10$.

In Fig.5, the EE of various methods decrease with N_t. Compared to the Fully digital and OMP algorithm, the proposed algorithm offers about 1Bits/Hz/J gain when $8 \times 8 < N_t < 12 \times 12$, and 0.5 bits/s/Hz to 0.8 bits/s/Hz when $N_t > 12 \times 12$. Moreover, the difference in the EE of the SIC algorithm compared to the proposed algorithm decreases with the increase in antennas. so that when $8 \times 8 \leq N_t < 12 \times 12$ the difference is around 0.3 bits/s/Hz and when $N_t \geq 12 \times 12$ is around up to 0.2 bits/s/Hz.

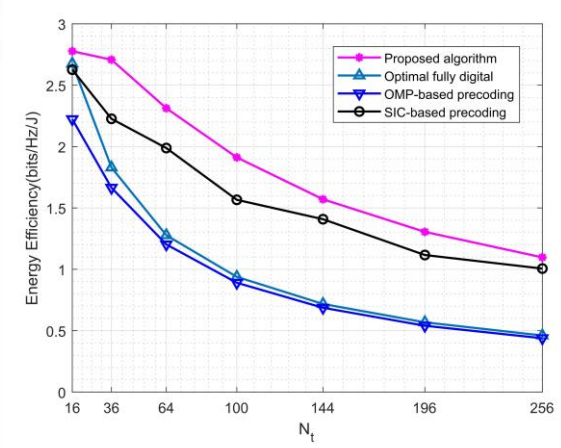

Figure 5. Energy efficiencies achieved against N_t when $N_r = 4 \times 4$, $N_s = 4$, $N_{RF} = 4$, SNR = 0dB

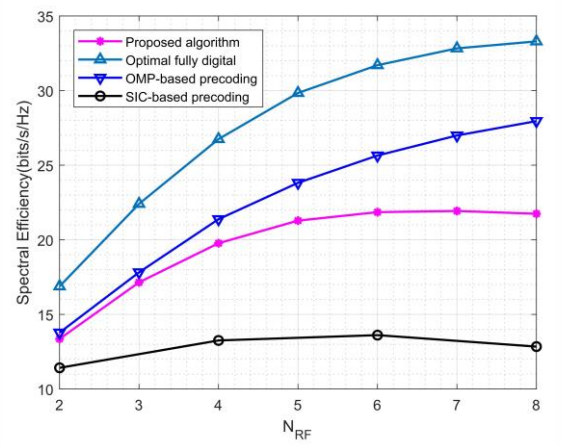

Figure 6. Spectral efficiencies achieved against N_{RF} when $N_t = 12 \times 12$, $N_r = 4 \times 4$, $N_s = 4$, SNR = 0dB

Fig.6 and Fig.7 analyze the SE and the EE versus N_{RF} for all schemes with the parameters set as $N_{RF} = N_s$, $N_t = 12 \times 12$, $N_r = 4 \times 4$ and SNR = 0dB. In Fig.6, The SE of the proposed algorithm increases until when the $N_{RF} = 5$. Despite The OMP algorithm in which its SE enhances as N_{RF} increases, For the proposed algorithm, when $N_{RF} > 5$, the SE almost remains the same value. Also, the proposed algorithm offers between 7 bits/s/Hz to 9 bits/s/Hz gain compared to the SIC algorithm.

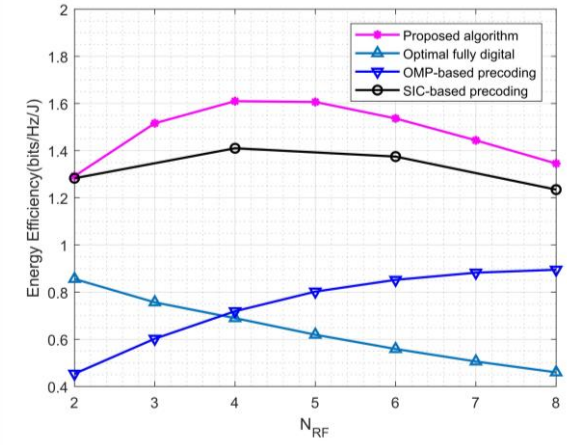

Figure 7. Energy efficiencies achieved against N_{RF} when $N_t = 12 \times 12$, $N_r = 4 \times 4$, $N_s = 4$, SNR = 0dB

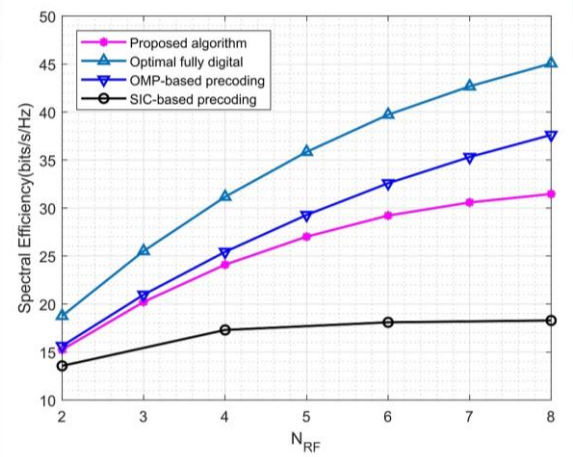

Figure 8. Spectral efficiencies achieved against N_{RF} when $N_t = 12 \times 12$, $N_r = 6 \times 6$, $N_s = 4$, SNR = 0dB

In Fig.7, The EE performance difference between the proposed algorithm and the OMP algorithm decreases as the N_{RF} increases; this reduction accelerates due to the fall of the EE in the proposed algorithm when $N_{RF} > 5$. Also, the SIC algorithm provides lower performance than the proposed algorithm, so that when $N_{RF} = 4$, the difference is around 0.2 bits/s/Hz, and when $N_{RF} = 8$, the difference is about 0.1 bits/s/Hz.

The SE and EE versus N_{RF} for all schemes with the parameters set as $N_{RF} = N_s$, $N_t = 12 \times 12$, $N_r = 6 \times 6$, SNR = 0dB are shown in Fig.8 and Fig.9. In Fig 8 despite, when $N_r = 4 \times 4$, the SE of the proposed algorithm continues elevating by increasing the number of antennas in the receiver, which results in decreasing the performance gap between OMP and the proposed algorithm. Moreover, the SIC algorithm performance slightly increases, but still, there is a huge gap between its performance and the proposed algorithm.

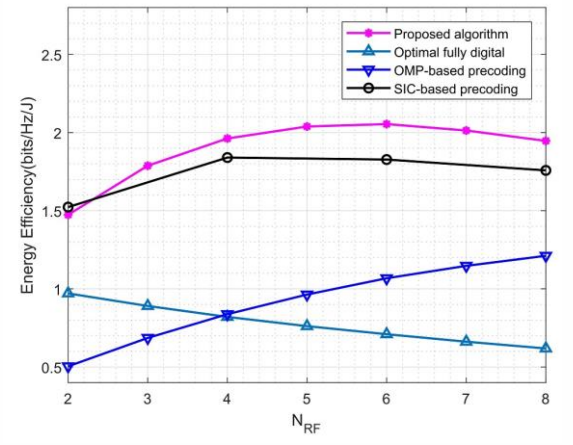

Figure 9. Energy efficiencies achieved against N_{RF} when $N_t = 12 \times 12$, $N_r = 6 \times 6$, $N_s = 4$, SNR = 0dB

Fig.9 determines the EE against N_{RF} for all schemes with the parameters are $N_{RF} = N_s$, $N_t = 12 \times 12$, $N_r = 6 \times 6$, and SNR = 0dB. By increasing the number of antennas in the receiver, the gap between the proposed algorithm and the SIC algorithm is almost preserved.

V. CONCLUSION

This paper concentrates on designing a hybrid precoder for dynamic structures. A hybrid precoding algorithm for a single-user massive mmWave MIMO communication system is introduced. The digital, analog switch, and phase shifter precoding matrices are optimized via an alternative and iterative solution. Simulation results have indicated that in terms of SE, the proposed algorithm can offer a reasonable compromise to the Fully-connected structure OMP precoding with significantly fewer phase shifters. Moreover, compared to the existing SIC scheme for the partially-connected architecture, the proposed algorithm offers much higher SE with the same number of phase shifters. Considering the equal number of phase shifters in the proposed algorithm, a suitable EE improvement is provided by the proposed algorithm.

APPENDIX A
PROOF OF PROPOSITION 1

The objective function in (8) can be recast as follows.

$$\left\|\boldsymbol{F}_{opt} - \boldsymbol{F}_{PS}\boldsymbol{F}_{S}\boldsymbol{F}_{BB}\right\|_F^2$$

Let $A = \boldsymbol{F}_{opt} - \boldsymbol{F}_{PS}\boldsymbol{F}_{S}\boldsymbol{F}_{BB}$.

$$\begin{aligned}
\emptyset &= \|A\|_F^2 = A : A \\
d\emptyset &= 2A : d\,\boldsymbol{F}_{PS}\boldsymbol{F}_{S}\boldsymbol{F}_{BB} \\
&= 2A\boldsymbol{F}_{BB}^{\mathrm{H}}\boldsymbol{F}_{S}^{\mathrm{H}} : d\,\boldsymbol{F}_{PS} \\
&= 2A\boldsymbol{F}_{BB}^{\mathrm{H}}\boldsymbol{F}_{S}^{\mathrm{H}} : \mathrm{Diag}\left(d\,\boldsymbol{f}_{PS}\right) \\
&= 2\mathrm{diag}(A\boldsymbol{F}_{BB}^{\mathrm{H}}\boldsymbol{F}_{S}^{\mathrm{H}}) : d\,\boldsymbol{F}_{PS} \\
\frac{\partial\emptyset}{\partial\,\boldsymbol{F}_{PS}} &= 2\mathrm{diag}(A\boldsymbol{F}_{BB}^{\mathrm{H}}\boldsymbol{F}_{S}^{\mathrm{H}}) \\
&= \mathrm{diag}\left(\boldsymbol{F}_{opt}\boldsymbol{F}_{BB}^{\mathrm{H}}\boldsymbol{F}_{S}^{\mathrm{H}} - \boldsymbol{F}_{PS}\boldsymbol{F}_{S}\boldsymbol{F}_{BB}\boldsymbol{F}_{BB}^{\mathrm{H}}\boldsymbol{F}_{S}^{\mathrm{H}}\right)
\end{aligned}$$

By setting the gradient to zero

$$\mathrm{Diag}\left(\boldsymbol{F}_{opt}\boldsymbol{F}_{BB}^{\mathrm{H}}\boldsymbol{F}_{S}^{\mathrm{H}}\right) = \mathrm{Diag}\left(\boldsymbol{F}_{PS}\boldsymbol{F}_{S}\boldsymbol{F}_{BB}\boldsymbol{F}_{BB}^{\mathrm{H}}\boldsymbol{F}_{S}^{\mathrm{H}}\right) \quad (32)$$

By adding $\boldsymbol{F}_{PS}$ to both sides of (32), eq. (28) will be determined.

REFERENCES

[1] R. Q. Hu and Y. Qian, "An energy efficient and spectrum efficient wireless heterogeneous network framework for 5G systems," *IEEE Commun. Mag.*, vol. 52, no. 5, 2014.

[2] Q. Li, R. Q. Hu, Y. Qian, and G. Wu, "Cooperative communications for wireless networks: Techniques and applications in LTE-advanced systems," *IEEE Wireless Communications*, vol. 19, no. 2. 2012.

[3] J. Mietzner, R. Schober, L. Lampe, W. Gerstacker, and P. Hoeher, "Multiple-antenna techniques for wireless communications - A comprehensive literature survey," *IEEE Commun. Surv. Tutorials*, vol. 11, no. 2, 2009.

[4] A. Taherpour, M. Nasiri-Kenari and A. Jamshidi, "Efficient cooperative spectrum sensing in cognitive radio networks." In 2007 IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications, pp. 1-6. IEEE, 2007.

[5] Jamshidi, Ali -Azizollah, Nasiri-Kenari, M., Zeinalpour, Z., Taherpour, Abbas. (2008). Space-frequency coded cooperation in OFDM multiple-access wireless networks. Communications, IET. 1. 1152 - 1160.

[6] T. Li and F. Zhao, "A Spectral Efficient Hybrid Precoding Algorithm for MmWave MIMO Systems," in *Procedia Computer Science*, 2020, vol. 174.

[7] J. G. Andrews, T. Bai, M. N. Kulkarni, A. Alkhateeb, A. K. Gupta, and R. W. Heath, "Modeling and Analyzing Millimeter Wave Cellular Systems," *IEEE Trans. Commun.*, vol. 65, no. 1, 2017.

[8] P. B. Papazian and G. A. Hufford, "Study of the local multipoint distribution service radio channel," *IEEE Trans. Broadcast.*, vol. 43, no. 2, 1997.

[9] Z. Cai, Y. Duan, and A. G. Bourgeois, "Delay efficient opportunistic routing in asynchronous multi-channel cognitive radio networks," *J. Comb. Optim.*, vol. 29, no. 4, 2015.

[10] J. Li, X. Guo, L. Guo, S. Ji, M. Han, and Z. Cai, "Optimal routing with scheduling and channel assignment in multi-power multi-radio wireless sensor networks," *Ad Hoc Networks*, vol. 31, 2015.

[11] R. Chataut and R. Akl, "Massive MIMO systems for 5G and beyond networks—overview, recent trends, challenges, and future research direction," *Sensors (Switzerland)*, vol. 20, no. 10. 2020.

[12] S. Sun, T. S. Rappaport, R. W. Heath, A. Nix, and S. Rangan, "MIMO for millimeter-wave wireless communications: Beamforming, spatial multiplexing, or both?," *IEEE Commun. Mag.*, vol. 52, no. 12, 2014.

[13] I. Khan *et al.*, "An efficient precoding algorithm for mmWave massive MIMO systems," *Symmetry (Basel).*, vol. 11, no. 9, 2019.

[14] P. Xia, R. W. Heath, and N. Gonzalez-Prelcic, "Robust Analog Precoding Designs for Millimeter Wave MIMO Transceivers with Frequency and Time Division Duplexing," *IEEE Trans. Commun.*, vol. 64, no. 11, 2016.

[15] V. Venkateswaran and A. J. Van Der Veen, "Analog beamforming in MIMO communications with phase shift networks and online channel estimation," *IEEE Trans. Signal Process.*, vol. 58, no. 8, 2010.

[16] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wirel. Commun.*, vol. 13, no. 3, 2014.

[17] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proc. IEEE*, vol. 102, no. 3, 2014.

[18] X. Liu *et al.*, "Hybrid Precoding for Massive mmWave MIMO Systems," *IEEE Access*, vol. 7, 2019.

[19] X. Song, T. Kuhne, and G. Caire, "Fully-Connected vs. Sub-Connected Hybrid Precoding Architectures for mmWave MU-MIMO," in *IEEE International Conference on Communications*, 2019, vol. 2019-May.

[20] F. Sohrabi and W. Yu, "Hybrid Digital and Analog Beamforming Design for Large-Scale Antenna Arrays," *IEEE J. Sel. Top. Signal Process.*, vol. 10, no. 3, 2016.

[21] 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, 2016.

[22] G. Liu, H. Deng, K. Yang, Z. Zhu, J. Liu, and H. Dong, "A new design of codebook for hybrid precoding in millimeter-wave massive mimo systems," *Symmetry (Basel).*, vol. 13, no. 5, 2021.

[23] E. E. Bahingayi and K. Lee, "Hybrid Combining Based on Constant Phase Shifters and Active/Inactive Switches," *IEEE Trans. Veh. Technol.*, vol. 6.

[24] H. Nosrati, E. Aboutanios, D. Smith, and X. Wang, "Switch-based hybrid precoding in mmwave massive MIMO systems," in *European Signal Processing Conference*, 2019, vol. 2019-Septe.

[25] W. Ni and X. Dong, "Hybrid block diagonalization for massive multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 64, no. 1, 2016.

[26] M. Alouzi, F. Chan, and C. D'Amours, "Low Complexity Hybrid Precoding and Combining for Millimeter Wave Systems," *IEEE Access*, vol. 9, 2021.

[27] A. A. M. Saleh and R. A. Valenzuela, "A Statistical Model for Indoor Multipath Propagation," *IEEE J. Sel. Areas Commun.*, vol. 5, no. 2, 1987.

[28] A. Forenza, D. J. Love, and R. W. Heath, "Simplified spatial correlation models for clustered MIMO channels with different array configurations," *IEEE Trans. Veh. Technol.*, vol. 56, no. 4 II, 2007.

[29] X. Yu, J. C. Shen, J. Zhang, and K. B. Letaief, "Alternating Minimization Algorithms for Hybrid Precoding in Millimeter Wave MIMO Systems," in *IEEE Journal on Selected Topics in Signal Processing*, 2016, vol. 10, no. 3.
[30] X. Yu, J. Zhang, and K. B. Letaief, "Hybrid Precoding in Millimeter Wave Systems: How Many Phase Shifters Are Needed?," in *2017 IEEE Global Communications Conference, GLOBECOM 2017 - Proceedings*, 2017, vol. 2018-Janua.
[31] X. Yu, J. Zhang, and K. B. Letaief, "A Hardware-Efficient Analog Network Structure for Hybrid Precoding in Millimeter Wave Systems," in *IEEE Journal on Selected Topics in Signal Processing*, 2018, vol. 12, no. 2.
[32] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K: Cambridge University Press, 1985.
[33] L. Grippo and M. Sciandrone, "On the convergence of the block nonlinear gauss–seidel method under convex constraints," Operat. Res. Lett., vol. 26, no. 3, pp. 127–136, Mar. 2000.
[34] Kim, J., He, Y. & Park, H. Algorithms for nonnegative matrix and tensor factorizations: a unified view based on block coordinate descent framework. J Glob Optim 58, 285– 319 (2014).
[35] R. Mendez-Rial, C. Rusu, N. Gonzalez-Prelcic, A. Alkhateeb, and R. W. Heath, "Hybrid MIMO Architectures for Millimeter Wave Communications: Phase Shifters or Switches?," *IEEE Access*, vol. 4, 2016.
[36] M. Tian, J. Zhang, Y. Zhao, L. Yuan, J. Yang, and G. Gui, "Switch and Inverter Based Hybrid Precoding Algorithm for mmWave Massive MIMO System: Analysis on Sum-Rate and Energy-Efficiency," *IEEE Access*, vol. 7, 2019.

Amirreza Moradi received his B.Sc. degree in Electrical Engineering from Ekbatan Institute of Higher Education, Qazvin, Iran, in 2019 and M.Sc. degree in Electrical Engineering-Telecommunication systems from K.N.Toosi University of Technology, Tehran, Iran, in 2022. His research interests include Massive MIMO, Wireless Communication, and Mobile Communication.

Kamal Mohamed-pour was born in Iran in 1954, and received his B.Sc. and M.Sc. in Electrical Engineering in the field of Communications in Tehran, 1979 and 1987 respectively. After some lecturing in department of Communication in Electrical Engineering faculty of K.N. Toosi University of Technology, continued his study towards Ph.D., and graduated at 1996 in Electrical Engineering at University of Manchester, Manchester, UK. Dr. Mohamed-pour is a full professor in Communication Department of K.N. Toosi University of Technology, and published several books in digital Signal Processing and Communications. His main interest fields of research are Wide-Band Wireless Communication, MIMO-OFDM, and Relay Networks.

Nasim Jafari Farsani received her B.Sc. degree in Electrical Engineering from Zanjan University, Zanjan, Iran, in 2017 and M.Sc. degree in Electrical Engineering Telecommunication Systems from K.N.Toosi University of Technology, Tehran, Iran, in 2023. Her research interests include Intelligent Reflecting Surface, Massive MIMO and Wireless Communication.