# Doubling Phase Shifters for Efficient Hybrid Precoder Design in Millimeter-Wave Communication Systems

Xianghao Yu, Jun Zhang, Khaled B. Letaief

*Abstract*—Hybrid precoding is a cost-effective approach to support directional transmissions for millimeter-wave (mmWave) communications, but its precoder design is highly complicated. In this paper, we propose a new hybrid precoder implementation, namely the double phase shifter (DPS) implementation, which enables highly tractable hybrid precoder design. Efficient algorithms are then developed for two popular hybrid precoder structures, i.e., the fully- and partially-connected structures. For the fully-connected one, the RF-only precoding and hybrid precoding problems are formulated as a least absolute shrinkage and selection operator problem and a low-rank matrix approximation problem, respectively. In this way, computationally efficient algorithms are provided to approach the performance of the fully digital one with a small number of radio frequency (RF) chains. On the other hand, the hybrid precoder design in the partially-connected structure is identified as an eigenvalue problem. To enhance the performance of this cost-effective structure, dynamic mapping from RF chains to antennas is further proposed, for which a greedy algorithm and a modified K-means algorithm are developed. Simulation results demonstrate the performance gains of the proposed hybrid precoding algorithms over existing ones. It shows that, with the proposed DPS implementation, the fully-connected structure enjoys both satisfactory performance and low design complexity while the partially-connected one serves as an economic solution with low hardware complexity.

*Keywords*—5G networks, hybrid precoding, low-rank matrix approximation, millimeter-wave communications, multiple-input multiple-output (MIMO), OFDM

## I. INTRODUCTION

The proliferation of smart mobile devices has resulted in an ever-increasing wireless data explosion, which calls for an exponential increase in the capacity of wireless networks. In particular, the upcoming 5G networks require a 1 000X increase in capacity by 2020[3]. The spectrum crunch in current wireless systems stimulates extensive interests on exploiting new spectrum bands for cellular communications, and millimeter-wave (mmWave) bands from 30 GHz to 300 GHz have been demonstrated to be promising candidates in recent experiments[4]. Thanks to the smaller wavelength of mmWave signals, large-scale antenna arrays can be leveraged at both the transmitter and receiver sides, which can provide spatial multiplexing gains with the help of multiple-input multiple-output (MIMO) techniques. On the other hand, the ten-fold increase of the carrier frequency introduces several challenges to mmWave communication systems, especially the high power consumption and cost of hardware components at mmWave bands[5]. In addition, the large available bandwidth at mmWave frequencies induces severe frequency selectivity, for which multicarrier techniques such as orthogonal frequency-division multiplexing (OFDM) shall be utilized. All the above-mentioned design aspects should be taken into consideration when developing practical transceivers for mmWave MIMO systems.

By utilizing a small number of radio frequency (RF) chains to combine a low-dimensional digital baseband precoder and a high-dimensional analog RF precoder, hybrid precoding stands out as a cost-effective transceiver solution for mmWave MIMO systems[6-8]. Compared with conventional MIMO systems, the additional high-dimensional analog RF precoder is

X. H. Yu. Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen 91058, Germany (e-mail: xianghao.yu@fau.de).

J. Zhang. Department of EIE, The Hong Kong Polytechnic University, Hong Kong 999077, China (e-mail: jun-eie.zhang@polyu.edu.hk).

K. B. Letaief. Department of ECE, The Hong Kong University of Science and Technology, Hong Kong 999077, China (e-mail: eekhaled@ust.hk).

the differentiating part. According to the mapping strategies from RF chains to antennas in the analog RF precoder, hybrid precoders can be categorized into the fully- and partially-connected structures[9]. In the fully-connected structure, each antenna is connected to all the RF chains. In contrast, each antenna is connected to one RF chain in the partially-connected structure, with a significant reduction in the hardware complexity.

To effectively reduce the power consumption in the RF domain, analog RF precoders are usually implemented by phase shifters at the expense of sacrificing the ability to adjust the amplitude of the RF signals[6]. Thus, the analog component forms the major challenge in designing hybrid precoders. Given the large dimension of the design space and the unit modulus constraint induced by the phase shifter implementation, an important design aspect of hybrid precoders is the computational complexity. While various attempts have been made to balance the performance and computational complexity, there is no systematic approach to design computationally efficient hybrid precoders with satisfactory performance in the meanwhile. In this paper, we will show the great potential to develop efficient hybrid precoding algorithms by adopting a novel double phase shifter (DPS) hybrid precoder implementation.

### A. Related Works and Motivation

Most existing works on hybrid precoding focused on the fully-connected structure[7,9-20]. The initial efforts started from single-user single-carrier[1] mmWave systems[7,9,11,12]. Then, the investigation was extended to single-user multicarrier[9,13,14] and multiuser single-carrier systems[15-18]. The main differences in these existing works are the approaches in dealing with the unit modulus constraints on the analog RF precoder.

By choosing the analog beamforming vectors from a predefined candidate set, e.g., array response vectors in Refs. [7,13,15,16] and discrete Fourier transform beamformers in Ref. [12], a greedy algorithm called orthogonal matching pursuit (OMP) has been widely used in designing hybrid precoders. Although its computational complexity is relatively low, the performance is not satisfactory and has been improved by several followed-up works. In Ref. [9], it was shown that the unit modulus constraints define a Riemannian manifold, and manifold optimization was introduced to directly tackle them, which helps to approach the performance of the fully digital one with a small number of RF chains. Furthermore, the contribution of each phase shifter to the spectral efficiency was identified in Refs. [11,18], based on which the analog precoder was optimized in a phase shifter-by-phase

shifter fashion. However, these algorithms all involve iterative procedures to optimize the analog RF precoders, which results in high computational complexity. Moreover, there were also some studies on how to achieve the performance of the fully digital precoder with the hybrid structure[19,20], yet requiring a large number of RF chains, which, to some extend, deviates from the motivation of hybrid precoding.

On the other hand, less attention has been paid on hybrid precoding in the partially-connected structure. In Refs. [21,22], codebook-based design of hybrid precoders was presented for single-user narrowband and OFDM systems, respectively. While using codebook enjoys a low complexity, there will be certain performance loss, and how to design the codebook remains to be clarified. By migrating the concept of successive interference cancellation, an iterative hybrid precoding algorithm in the partially-connected structure was proposed in Ref. [23] for single-user single-carrier systems. Since the partially-connected structure employs much fewer phase shifters, there should be some inevitable degradation in the analog precoding gain, which makes it difficult for such structures to achieve a high spectral efficiency, especially when the analog precoder is shared across all the users and subcarriers as in the multiuser multicarrier systems. Hence, how to efficiently use the limited number of phase shifters is an urgent issue to be solved in the partially-connected structure.

As illustrated above, in both the fully- and partially-connected structures, there is no comprehensive way to efficiently design hybrid precoders with satisfactory performance, which motivates us to seek a new hybrid precoding architecture that can relieve us from the current dilemma. Furthermore, it is still unclear how to design hybrid precoders in multiuser multicarrier systems, where a single analog RF precoder is shared by a large number of subcarriers, and multiple users that interfere with each other. In this paper, we propose a novel DPS implementation for hybrid precoding in the general setting of multiuser OFDM mmWave MIMO systems. Although similar implementations were considered in Refs. [20,24], the systematic design approach and algorithmic advantages of this new implementation have not been exploited, which will be illustrated in this paper via effective algorithms for different hybrid precoder structures.

### B. Contributions

Conventionally, a single phase shifter is used to connect an RF chain and an antenna, i.e., the SPS implementation, which introduces the unit modulus constraints and hinders efficient algorithm design. In this paper, to overcome this algorithmic difficulty, we propose a novel hybrid precoder implementation that makes the precoder design more tractable. Our main contributions are summarized as follows.

- We propose a novel hybrid precoder implementation,

---

[1]In this paper, single-carrier systems refer to single-carrier transmissions over flat-fading channels.

i.e., the DPS implementation, which relaxes the unit modulus constraints of the analog RF precoder and thus enables computationally efficient hybrid precoder design. To the best of the authors' knowledge, this is the first attempt to directly adopt the DPS implementation for designing hybrid precoders in multiuser OFDM mmWave MIMO systems.

• For the fully-connected structure, the optimization of the analog RF precoder is formulated as a least absolute shrinkage and selection operator (LASSO) problem, based on which efficient algorithms are developed. Furthermore, the hybrid precoder design is identified as a low-rank matrix approximation problem, which has a closed-form solution. Furthermore, the efficient algorithm for the DPS implementation inspires an effective heuristic hybrid precoder design for the conventional SPS implementation, which outperforms the state-of-the-art algorithms in both computational complexity and spectral efficiency.

• For the partially-connected structure, we identify that the hybrid precoder design is an eigenvalue problem, and provide closed-form solutions for both analog RF and digital baseband precoders. To further improve the system performance, a dynamic partially-connected structure is proposed. Two effective algorithms, i.e., the greedy and modified K-means algorithms, are proposed to dynamically optimize the mapping strategies from RF chains to antennas.

• For both structures, we discover that the hybrid precoder in the multiuser setting will produce residual interuser interference, as it only approximates the fully digital precoder. To this end, we propose to cascade an additional block diagonalization (BD) precoder at the baseband to cancel the interuser interference, which is shown to be effective to further improve the spectral efficiency and multiplexing gain.

• Analytical results on the performance gap between the fully- and partially-connected structures are provided. Furthermore, extensive comparisons are offered via simulations to unravel valuable design insights. In particular, the proposed algorithm helps the fully-connected structure to easily approach the performance of the fully digital precoder with a reasonably small amount of RF chains, which cannot be achieved by the widely used OMP algorithm. On the other hand, for the partially-connected structure, it turns out that the dynamic mapping from RF chains to antennas is crucial to achieve good performance. Furthermore, while the DPS partially-connected structure employs much fewer phase shifters, its performance is comparable to the SPS fully-connected structure with the OMP algorithm, which shows its great potential for practical implementation.

## C. Organization

The remainder of this paper is organized as follows. We introduce the system model and the problem formulation in section II. Then, hybrid precoder design for the fully- and partially-connected structures are demonstrated in section III and section IV, respectively. Simulation results will be presented in section V. Finally, we conclude this paper in section VI.

## D. Notations

The following notations are used throughout this paper. The imaginary unit is denoted as $\iota = \sqrt{-1}$; $\boldsymbol{a}$ and $\boldsymbol{A}$ symbolize a column vector and a matrix, respectively; $\boldsymbol{A}^{\mathrm{T}}$, $\boldsymbol{A}^*$, $\boldsymbol{A}^{\mathrm{H}}$, and $\boldsymbol{A}^{\dagger}$ stand for the transpose, conjugate, conjugate transpose, and pseudo-inverse of matrix $\boldsymbol{A}$; The $i$th row, the $j$th column, and the $(i,j)$th entry in matrix $\boldsymbol{A}$ are denoted as $\boldsymbol{A}(i,:)$, $\boldsymbol{A}(:,j)$, and $\boldsymbol{A}(i,j)$; The determinant, Frobenius norm, and $\ell_p$-norm of matrix $\boldsymbol{A}$ are expressed as $\det(\boldsymbol{A})$, $\|\boldsymbol{A}\|_{\mathrm{F}}$, and $\|\boldsymbol{A}\|_p$; $\lambda_i(\boldsymbol{A})$ denotes the $i$th largest eigenvalue of matrix $\boldsymbol{A}$, and the corresponding eigenvector is noted as $\boldsymbol{\lambda}_i(\boldsymbol{A})$; $\mathrm{tr}(\boldsymbol{A})$ and $\mathrm{vec}(\boldsymbol{A})$ indicate the trace and vectorization of matrix $\boldsymbol{A}$; $\circ$ and $\otimes$ stand for the Hadamard and Kronecker products between two matrices; Expectation and the real part of a complex variable are denoted by $\mathbb{E}[\cdot]$ and $\mathfrak{R}[\cdot]$.

## II. SYSTEM MODEL AND PROBLEM FORMULATION

### A. System Model

Consider the downlink transmission of a multiuser OFDM mmWave MIMO system, as shown in Fig. 1, where the base station (BS) is equipped with $N_{\mathrm{t}}$ antennas and transmits signals to $K$ $N_{\mathrm{r}}$-antenna users over $F$ subcarriers. On each subcarrier, $N_s$ data streams are transmitted to each user. The limitations of the RF chains are given by $K N_s \leqslant N_{\mathrm{RF}}^{\mathrm{t}} < N_{\mathrm{t}}$ and $N_s \leqslant N_{\mathrm{RF}}^{\mathrm{r}} < N_{\mathrm{r}}$, where $N_{\mathrm{RF}}^{\mathrm{t}}$ and $N_{\mathrm{RF}}^{\mathrm{r}}$ are the numbers of RF chains facilitated for the BS and each user, respectively.

The received signal for the $k$th user on the $f$th subcarrier is given by

$$\boldsymbol{y}_{k,f} = \boldsymbol{W}_{\mathrm{BB}k,f}^{\mathrm{H}} \boldsymbol{W}_{\mathrm{RF}k}^{\mathrm{H}} \left( \boldsymbol{H}_{k,f} \sum_{k=1}^{K} \boldsymbol{F}_{\mathrm{RF}} \boldsymbol{F}_{\mathrm{BB}k,f} \boldsymbol{s}_{k,f} + \boldsymbol{n}_{k,f} \right), \tag{1}$$

where the subscript $k,f$ represents the $k$th user on the $f$th subcarrier, and $\boldsymbol{s}_{k,f} \in \mathbb{C}^{N_s}$ is the transmitted symbol vector such that $\mathbb{E}[\boldsymbol{s}_{k,f} \boldsymbol{s}_{k,f}^{\mathrm{H}}] = \frac{1}{K N_s F} \boldsymbol{I}_{N_s}$. The digital baseband precoders and combiners are symbolized by $\boldsymbol{F}_{\mathrm{BB}k,f} \in \mathbb{C}^{N_{\mathrm{RF}}^{\mathrm{t}} \times N_s}$ and $\boldsymbol{W}_{\mathrm{BB}k,f} \in \mathbb{C}^{N_{\mathrm{RF}}^{\mathrm{r}} \times N_s}$, respectively. Because the transmitted signals for all the users are mixed together via the digital baseband precoder, and the analog RF precoder is a post-IFFT (inverse fast Fourier transform) operation, the analog RF precoder is shared by all the users and subcarriers, denoted as $\boldsymbol{F}_{\mathrm{RF}} \in \mathbb{C}^{N_{\mathrm{t}} \times N_{\mathrm{RF}}^{\mathrm{t}}}$. Similarly, the analog RF combiner is subcarrier-independent for each user $k$, denoted as $\boldsymbol{W}_{\mathrm{RF}k} \in \mathbb{C}^{N_{\mathrm{r}} \times N_{\mathrm{RF}}^{\mathrm{r}}}$. Furthermore, the additive noise at the users
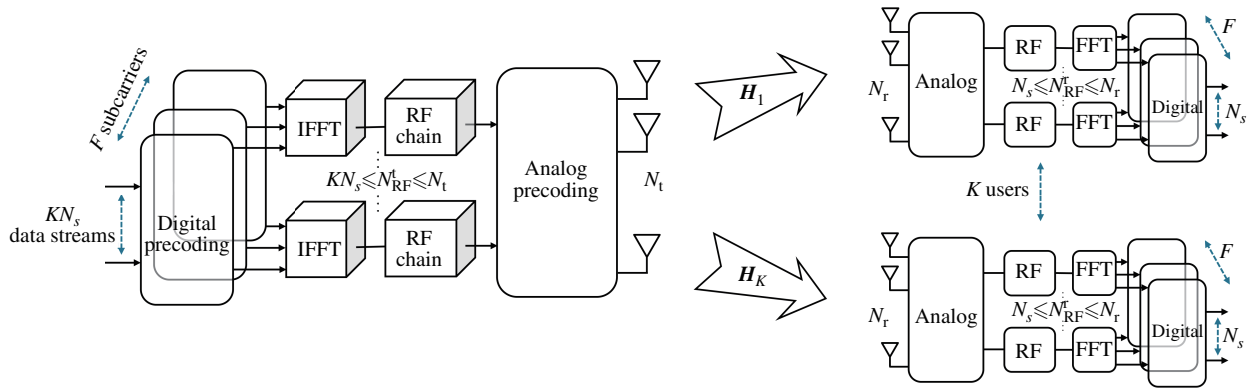
**Figure 1**    A multiuser OFDM mmWave MIMO systems with hybrid precoding

is represented by $\boldsymbol{n}_{k,f} \in \mathbb{C}^{N_r}$, whose elements are independent and identically distributed according to the complex Gaussian distribution $\mathscr{CN}(0, \sigma_n^2)$. The achievable sum rate on the $f$th subcarrier when transmitted symbols follow a Gaussian distribution is given by[7,25]

$$
R_f = \sum_{k=1}^{K} \log\det\left( \boldsymbol{I}_{N_s} + \tfrac{1}{KN_sF} \boldsymbol{W}_{k,f}^{\mathrm{H}} \boldsymbol{H}_{k,f} \boldsymbol{F}_{k,f} \boldsymbol{F}_{k,f}^{\mathrm{H}} \boldsymbol{H}_{k,f}^{\mathrm{H}} \boldsymbol{W}_{k,f} \boldsymbol{\Omega}_{k,f}^{-1} \right),
$$

(2)

where $\boldsymbol{F}_{k,f} = \boldsymbol{F}_{\mathrm{RF}} \boldsymbol{F}_{\mathrm{BB}k,f}$ and $\boldsymbol{W}_{k,f} = \boldsymbol{W}_{\mathrm{RF}k} \boldsymbol{W}_{\mathrm{BB}k,f}$ are the precoder and combiner matrices, and $\boldsymbol{\Omega}_{k,f} = \boldsymbol{W}_{k,f}^{\mathrm{H}} [\tfrac{\rho_k}{KN_sF} \boldsymbol{H}_{k,f} (\sum_{j\neq k} \boldsymbol{F}_{j,f} \boldsymbol{F}_{j,f}^{\mathrm{H}}) \boldsymbol{H}_{k,f}^{\mathrm{H}} + \sigma_n^2] \boldsymbol{W}_{k,f}$ stands for the interference plus noise matrix.

The mmWave MIMO channel between the BS and the $k$th user on the $f$th subcarrier, denoted as $\boldsymbol{H}_{k,f}$, can be characterized by the Saleh-Valenzuela model as[7,9,13]. Although this specific channel model will be used in the simulation, our precoder design approaches are compatible for other general channel models.

## B. New Hybrid Precoder Implementation

According to the mapping strategies from RF chains to antennas, the hybrid precoder structures can be classified into the fully- and partially-connected ones [9, Fig. 1]. The fully-connected structure fully exploits the degrees of freedom (DoFs) in the RF domain with a natural mapping strategy, i.e., to connect each RF chain to all the antennas. On the contrary, in the partially-connected structure, each antenna element is connected to only one RF chain. These two different mapping strategies (structures) correspond to different constraints in the hybrid precoder design problem, which will be illustrated in detail later in sections III and IV.

As mentioned before, the analog RF precoder is practically implemented by phase shifters. Conventionally, in either the fully- or partially-connected structure[9], each connection from a certain RF chain to one of its connected antenna elements is implemented by a single phase shifter, as shown in

Fig. 2(a), which is referred to the *SPS implementation* in this paper. This mapping strategy implies that each non-zero element in the analog precoding and combining matrices should have unit modulus, i.e., $|\boldsymbol{F}_{\mathrm{RF}}(i,j)| = |\boldsymbol{W}_{\mathrm{RF}}(i,j)| = 1$. This is intrinsically a non-convex constraint and difficult to tackle, which forms the main design challenge. Although there exist some approaches that can directly deal with this non-convex constraint[9,11], the design complexity is still unacceptable in mmWave systems with much shorter coherent time compared to current sub-6 GHz systems. As a matter of fact, the main obstacle is that we can only adjust the phase but not the amplitude of the RF signals. This motivates us to consider an alternative hybrid precoder implementation which can adjust the amplitude of the RF signals, yet still realized by phase shifters.

In this paper, we propose a new implementation as shown in Fig. 2(b), referred as the DPS implementation[1], where the phase shifter network is divided into two groups. For each connection from an RF chain to one of its connected antenna elements, one unique phase shifter in each group will be selected and summed up together to compose the analog precoding gain. With this special implementation, each non-zero element in the analog RF precoding and combining matrices corresponds to a sum of two phase shifters. Note that the summation operation creates the possibility to adjust the amplitude of the RF signals, which should be less than two, i.e., the new constraints for the analog RF precoder and combiner are $|\boldsymbol{F}_{\mathrm{RF}}(i,j)| \leqslant 2$ and $|\boldsymbol{W}_{\mathrm{RF}}(i,j)| \leqslant 2$ for all the non-zero entries. By doubling the number of phase shifters, the new constraints become convex and therefore make it more tractable to develop low-complexity design approaches. We impose these amplitude constraints in this paper, and the actual implementation of the phase shifters can then be easily obtained by factorizing a complex number with amplitude less than two into two unit modulus components, expressed as

$$
a\mathrm{e}^{\iota\theta} = \mathrm{e}^{\iota(\theta+\phi)} + \mathrm{e}^{\iota(\theta-\phi)},
$$

(3)

where $a \in [-2, 2]$ and $\theta \in [0, 2\pi)$ are the amplitude and

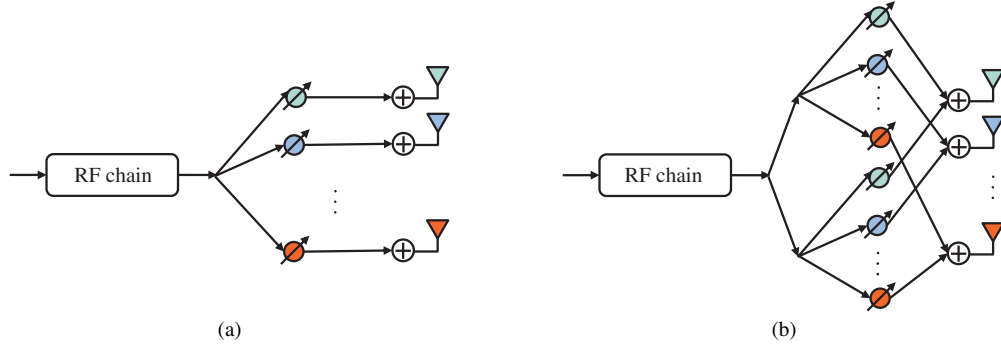(a)                                                                  (b)

**Figure 2** Comparison of two hybrid precoder implementations: a) conventional SPS hybrid precoder implementation; b) DPS hybrid precoder implementation (The main difference is on the number of phase shifters in use to compose each connection from an RF chain to a connected antenna element. In the fully-connected structure, each RF chain can be connected to all $N_t$ antenna elements, and in the partially-connected one, each RF chain is connected to a subset of antennas that do not overlap with each other.)

phase of the non-zero element in $\boldsymbol{F}_{\mathrm{RF}}$ and $\boldsymbol{W}_{\mathrm{RF}}$, and $\phi = \arccos(a/2)$.

**Remark 1** Despite the increased number of phase shifters, as will be shown in this paper, the DPS implementation enjoys unique advantages in both algorithmic and performance aspects. It also provides valuable guidelines for other hybrid precoder design problems. We highlight the benefits of this proposal as follows.

• The DPS implementation greatly simplifies the hybrid precoder design will be greatly simplified when adopting the DPS implementation, , as illustrated in sections III and IV.

• With this new implementation, hybrid precoders can approach the performance of the fully digital one with fewer RF chains than existing works. Thus, this proposal serves as an algorithmically efficient hybrid precoder design for general multiuser multicarrier mmWave systems.

• The DPS fully-connected hybrid precoder structure serves as a performance upper bound for structures that are with lower hardware complexity. It is a tighter upper bound than the fully digital precoder, especially when the number of RF chains is small.

• The precoder design problem becomes a low-rank matrix approximation (eigenvalue) problem for the DPS fully-connected (partially-connected) structure, and theoretical analysis, which is intractable for other structures, becomes possible. It will then help to better understand hybrid precoding systems.

• Thanks to the benefits in both performance and algorithmic perspectives, the proposed DPS implementation would drive the hardware research for this implementation.

## C. Problem Formulation

There exist different problem formulations for hybrid precoding. Some works tried to directly maximize the spectral efficiency based on approximations and bounds in single-user systems[11,18], or based on some extra constraints on the ana-

log precoder to simplify the design in multiuser single-carrier systems[15,17]. However, when it comes to multiuser multicarrier systems, it is highly challenging and intractable to directly optimize the hybrid precoder with the spectral efficiency being the objective function, given that the spectral efficiency of each user on each subcarrier is coupled with each other by the shared analog RF precoder. On the other hand, extensive works showed that minimizing the Euclidean distance[2] between the fully digital precoder and the hybrid precoder is an effective surrogate for maximizing the spectral efficiency in mmWave MIMO systems[1,7,9,12,13,16,22]. In this paper, we adopt this alternative objective as our design goal, whose formulation[3] is given by

$$
\begin{aligned}
\underset{\boldsymbol{F}_{\mathrm{RF}}, \boldsymbol{F}_{\mathrm{BB}}}{\text{minimize}} \quad & \left\| \boldsymbol{F}_{\mathrm{opt}} - \boldsymbol{F}_{\mathrm{RF}} \boldsymbol{F}_{\mathrm{BB}} \right\|_{\mathrm{F}}^{2} \\
\text{subject to} \quad & \begin{cases} \boldsymbol{F}_{\mathrm{RF}} \in \mathscr{A} \\ \left\| \boldsymbol{F}_{\mathrm{RF}} \boldsymbol{F}_{\mathrm{BB}} \right\|_{F}^{2} \leqslant K N_{s} F, \end{cases}
\end{aligned}
\tag{4}
$$

where $\boldsymbol{F}_{\mathrm{opt}} = [\boldsymbol{F}_{\mathrm{opt}_{1,1}}, \cdots, \boldsymbol{F}_{\mathrm{opt}_{k,f}}, \cdots, \boldsymbol{F}_{\mathrm{opt}_{K,F}}] \in \mathbb{C}^{N_t \times K N_s F}$ is the combined fully digital precoder, and $\boldsymbol{F}_{\mathrm{BB}} = [\boldsymbol{F}_{\mathrm{BB}_{1,1}}, \cdots, \boldsymbol{F}_{\mathrm{BB}_{k,f}}, \cdots, \boldsymbol{F}_{\mathrm{BB}_{K,F}}] \in \mathbb{C}^{N_{\mathrm{RF}}^{t} \times K N_s F}$ is the concatenated digital baseband precoder. The second constraint is the transmit power constraint at the BS side. The analog RF precoder $\boldsymbol{F}_{\mathrm{RF}}$ is a common component for all $K$ users and $F$ subcarriers, which is restricted in the candidate set $\mathscr{A} \in \{\mathscr{A}_f, \mathscr{A}_{\mathrm{d}}\}$ induced by the phase shifter implementation. The set $\mathscr{A}$ will be later specified for different hybrid precoder structures. Justifications for the formulation (4) for single-user systems with flat-fading channels were provided in Ref. [7]. Here we provide some intuition for this formula-

---

[2]In this paper, the Euclidean distance between two precoders refers to the Euclidean distance between two points determined by the vectorization of the two precoding matrices.

[3]In this paper, we focus on the precoder design, and the combiner design problem can be formulated in the same way without the transmit power constraint.

tion for general hybrid precoding systems. The fully digital precoder serves as a performance upper bound for the hybrid one, and one ideal design goal is to obtain hybrid precoders that approach the performance of the fully digital one. Therefore, it is intuitive to formulate the design problem as approximating the fully digital precoder with the hybrid one.

With this formulation, the proposed algorithm can be applied with any fully digital precoder. In this paper, we adopt the classical BD precoder as the fully digital one, which is asymptotically optimal in the high signal-to-noise ratio (SNR) regime[26]. We will investigate the hybrid precoder design with the DPS implementation for the fully- and partially-connected structures in sections III and IV, respectively.

## III. HYBRID PRECODING FOR THE FULLY-CONNECTED STRUCTURE

The fully-connected hybrid precoder structure has drawn much research attention in recent years[7,9,11-20], which will be investigated in this section with the new DPS implementation. We will first present an RF-only precoder to demonstrate the advantages of doubling the phase shifters, where the optimization of the analog RF precoder is formulated as a LASSO problem. Afterwards, the hybrid precoder design will be performed via a simple low-rank matrix approximation.

### A. RF-Only Precoding

The main difference between the conventional SPS hybrid precoder implementation and the proposed DPS one is on the analog RF precoder. Therefore, we first present an RF-only precoder design[27], where the analog RF precoder is optimized for a given digital precoder. This problem may arise as a subproblem in hybrid precoder design, as in Refs. [1,11,16], or for situations where the digital precoder has a fixed design, e.g., from a codebook. The investigation of this problem will demonstrate the algorithmic advantage of the DPS implementation. For the fully-connected structure, the feasible set $\mathscr{A}_f$ can be specified as $\mathscr{A}_f = \{ \boldsymbol{A} \| \boldsymbol{A}(i,j) | \leqslant 2 \}$, as each RF chain is connected to all the antenna elements. The optimization of the analog RF precoder design problem is given by

$$
\begin{aligned}
& \underset{\boldsymbol{F}_{\mathrm{RF}}}{\operatorname{minimize}} \quad \left\| \boldsymbol{F}_{\mathrm{opt}} - \boldsymbol{F}_{\mathrm{RF}} \boldsymbol{F}_{\mathrm{BB}} \right\|_{\mathrm{F}}^2 \\
& \text{subject to} \quad \boldsymbol{F}_{\mathrm{RF}} \in \mathscr{A}_f .
\end{aligned} \tag{5}
$$

Note that the power constraint in (4) is temporarily removed. In fact, after designing the analog RF precoder, we can normalize it if the transmit power constraint is not satisfied. It has been shown in [9, Lemma 1] that as long as we can make the Euclidean distance between the fully digital precoder and the hybrid precoder sufficiently small when ignoring the power constraint, the normalization step will also achieve a small

distance to the fully digital precoder. The optimization problem (5) is a convex one and can be solved by solvers such as CVX. Nevertheless, to further reduce the computational complexity, we will exploit the inherent structure of the solution by considering its dual problem.

**Lemma 1** The dual problem of (5) is a LASSO problem, given by

$$
\underset{\boldsymbol{x}}{\operatorname{minimize}} \quad \frac{1}{2} \| \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} \|_2^2 + 2\| \boldsymbol{x} \|_1 . \tag{6}
$$

The parameters $\boldsymbol{A}$ and $\boldsymbol{b}$ are given by

$$
\boldsymbol{A} = \boldsymbol{S}^{\frac{1}{2}} \boldsymbol{U}, \quad \boldsymbol{b} = \boldsymbol{A}\boldsymbol{D}^{\mathrm{H}}\boldsymbol{f}_{\mathrm{opt}} , \tag{7}
$$

where $\boldsymbol{D} = \boldsymbol{F}_{\mathrm{BB}}^{\mathrm{T}} \otimes \boldsymbol{I}_{N_t}$ and $\left( \boldsymbol{D}^{\mathrm{H}}\boldsymbol{D} \right)^{-1} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{U}^{\mathrm{H}}$ is the singular value decomposition (SVD) of $\left( \boldsymbol{D}^{\mathrm{H}}\boldsymbol{D} \right)^{-1}$. The optimal solution of (5) can be written as

$$
\operatorname{vec}(\boldsymbol{F}_{\mathrm{RF}}^{\star}) = \boldsymbol{f}_{\mathrm{RF}}^{\star} = \boldsymbol{A}^{\mathrm{H}}(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}^{\star}) . \tag{8}
$$

**Proof** See Appendix.

Based on Lemma 1, problem (5) is transferred to a LASSO problem. This provides the opportunity to leverage the large body of existing works on low-complexity algorithms to solve the general LASSO problem[28]. Recall that, with the conventional SPS implementation, the analog RF precoder is optimized through high-complexity algorithms such as manifold optimization[9] to achieve good performance. In contrast, doubling the phase shifters equips us with huge potential to significantly reduce the computational complexity when designing the analog RF precoder.

What deserves an additional mention is a special case where we can get a closed-form solution, which will further reduce the computational complexity. It was shown in Refs. [9,11] that a semi-orthogonal structure of the digital baseband precoder, i.e., $\boldsymbol{F}_{\mathrm{BB}}\boldsymbol{F}_{\mathrm{BB}}^{\mathrm{H}} = \boldsymbol{I}_{N_{\mathrm{RF}}^t}$[4], leads to an approximately optimal solution. Therefore, we resort to this special case where the observation matrix $\boldsymbol{A}$ in the LASSO problem (6) is also semi-orthogonal, i.e.,

$$
\begin{aligned}
\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A} &= \left( \boldsymbol{D}^{\mathrm{H}}\boldsymbol{D} \right)^{-1} = \left( (\boldsymbol{F}_{\mathrm{BB}}^{\mathrm{T}} \otimes \boldsymbol{I}_{N_{\mathrm{RF}}^t})^{\mathrm{H}}(\boldsymbol{F}_{\mathrm{BB}}^{\mathrm{T}} \otimes \boldsymbol{I}_{N_{\mathrm{RF}}^t}) \right)^{-1} = \\
& \left( (\boldsymbol{F}_{\mathrm{BB}}^{*} \otimes \boldsymbol{I}_{N_{\mathrm{RF}}^t})(\boldsymbol{F}_{\mathrm{BB}}^{\mathrm{T}} \otimes \boldsymbol{I}_{N_{\mathrm{RF}}^t}) \right)^{-1} = \\
& \left( (\boldsymbol{F}_{\mathrm{BB}}\boldsymbol{F}_{\mathrm{BB}}^{\mathrm{H}})^{\mathrm{T}} \otimes \boldsymbol{I}_{N_{\mathrm{RF}}^t} \right)^{-1} = \boldsymbol{I}_{N_{\mathrm{RF}}^t {}^2} .
\end{aligned} \tag{9}
$$

With the semi-orthogonal observation matrix $\boldsymbol{A}$, the LASSO problem (6) has a closed-form solution called soft-thresholding[28], which is given by

$$
\boldsymbol{x}^{\star} = \exp\{ \imath\angle(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{b}) \} \circ \left( \left| \boldsymbol{A}^{\mathrm{H}}\boldsymbol{b} \right| - 2 \right)^{+} , \tag{10}
$$

---

[4] Note that in mmWave multiuser OFDM systems, $\boldsymbol{F}_{\mathrm{BB}} \in \mathbb{C}^{N_{\mathrm{RF}}^t \times KN_sF}$, where $KN_sF \geqslant N_{\mathrm{RF}}^t$ for practical system parameters, which means $\boldsymbol{F}_{\mathrm{BB}}$ is a fat matrix.

where $\angle(\cdot)$, $|\cdot|$, and $(\boldsymbol{X})^+ = \max\{\boldsymbol{0}, \boldsymbol{X}\}$ are element-wise operations, and the first two extract the phase and amplitude of a complex variable, respectively. Then, substituting Eqs. (9) and (10) to Eq. (8), we obtain the corresponding optimal solution to $\boldsymbol{F}_{\mathrm{RF}}$ in (5) as

$$\boldsymbol{F}_{\mathrm{RF}}^{\star} = \boldsymbol{F}_{\mathrm{opt}}\boldsymbol{F}_{\mathrm{BB}}^{\mathrm{H}} - \exp\left\{\imath\angle\left(\boldsymbol{F}_{\mathrm{opt}}\boldsymbol{F}_{\mathrm{BB}}^{\mathrm{H}}\right)\right\} \circ \left(\left|\boldsymbol{F}_{\mathrm{opt}}\boldsymbol{F}_{\mathrm{BB}}^{\mathrm{H}}\right| - 2\right)^+ . \tag{11}$$

Note that, in order to obtain the optimal analog RF precoder $\boldsymbol{F}_{\mathrm{RF}}^{\star}$ when the digital baseband precoder $\boldsymbol{F}_{\mathrm{BB}}$ is semi-orthogonal, a product of $\boldsymbol{F}_{\mathrm{opt}}$ and $\boldsymbol{F}_{\mathrm{BB}}^{\mathrm{H}}$ is the only required step, which is computationally much more efficient than solving the original problem (5) using an algorithm-embedded solver. This result also suggests that it is beneficial to set the digital baseband precoder as a semi-orthogonal one in the RF-only precoding with the DPS implementation.

### B. Hybrid Precoding

Previously, we demonstrated the benefit of doubling the phase shifters when optimizing the analog part. When the digital baseband precoder can be jointly optimized, the hybrid precoder design problem is further simplified as an unconstrained matrix decomposition problem, i.e.,

$$\underset{\boldsymbol{F}_{\mathrm{RF}}, \boldsymbol{F}_{\mathrm{BB}}}{\operatorname{minimize}} \ \left\|\boldsymbol{F}_{\mathrm{opt}} - \boldsymbol{F}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{BB}}\right\|_{\mathrm{F}}^2 . \tag{12}$$

**Remark 2** The constraint $\boldsymbol{F}_{\mathrm{RF}} \in \mathscr{A}_f$ in (5), i.e., $|\boldsymbol{F}_{\mathrm{RF}}(i,j)| \leqslant 2$, is in fact redundant in hybrid precoding. Once a pair of the unconstrained optimal solution $\{\boldsymbol{F}_{\mathrm{RF}}, \boldsymbol{F}_{\mathrm{BB}}\}$ is obtained, one can always get another pair of optimal solution $\{\frac{\boldsymbol{F}_{\mathrm{RF}}}{\gamma}, \gamma\boldsymbol{F}_{\mathrm{BB}}\}$ with the factor $\gamma = \|\mathrm{vec}(\boldsymbol{F}_{\mathrm{RF}})\|_{\infty}/2$ to satisfy the constraint $\boldsymbol{F}_{\mathrm{RF}} \in \mathscr{A}_f$, which will not affect the objective value. On the other hand, one may consider deploying $n > 2$ phase shifters for each connection from an RF chain to an antenna, and the corresponding constraint would be $|\boldsymbol{F}_{\mathrm{RF}}(i,j)| \leqslant n$. As illustrated above, this constraint is redundant and the factor $\gamma = \|\mathrm{vec}(\boldsymbol{F}_{\mathrm{RF}})\|_{\infty}/n$ can be applied. Therefore, from both performance and algorithmic perspectives, it does not help to further increase the number of phase shifters. Obviously, the minimum number, i.e., two phase shifters, should be adopted due to cost and power consideration. Furthermore, the transmit power constraint $\|\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{BB}}\|_F^2 \leqslant KN_sF$ is automatically satisfied by the optimal solution of the hybrid precoder, which will be elaborated in the following optimization.

While the main focus of this paper is on multiuser multi-carrier systems, some advantages of the proposed DPS implementation in hybrid precoding will be firstly presented in single-carrier systems, as shown in the following result.

**Lemma 2** For single-carrier systems, with the DPS implementation, the fully digital precoder $\boldsymbol{F}_{\mathrm{opt}}$ can be perfectly decomposed into $\boldsymbol{F}_{\mathrm{RF}}$ and $\boldsymbol{F}_{\mathrm{BB}}$ using the minimum number of RF chains, i.e., $N_{\mathrm{RF}}^{\mathrm{t}} = KN_s$ and $N_{\mathrm{RF}}^{\mathrm{r}} = N_s$.

**Proof** The proof can be easily obtained by the rank sufficiency of $\boldsymbol{F}_{\mathrm{RF}}$ and $\boldsymbol{F}_{\mathrm{BB}}$ in the decomposition when $F = 1$, and is omitted due to space limitation.

Lemma 2 shows that, for single-carrier systems with either single-user or multiuser transmissions, the performance of the fully digital precoder can be easily achieved by the hybrid precoder via a simple matrix decomposition. Note that, with the conventional SPS implementation, the number of RF chains should be at least twice that of the data streams in order to realize the fully digital precoder, i.e., $N_{\mathrm{RF}}^{\mathrm{t}} = 2KN_s$ and $N_{\mathrm{RF}}^{\mathrm{r}} = 2N_s$[9,11]. In this case, since the numbers of phase shifters in use are the same, i.e., $2KN_sN_{\mathrm{t}}$ at the BS, for both the SPS and DPS implementations, the proposed DPS implementation, which requires fewer RF chains, is more energy efficient when achieving the fully digital precoder.

When it comes to multiuser multicarrier systems, typically $KN_sF \geqslant N_{\mathrm{t}}$, the rank of $\boldsymbol{F}_{\mathrm{opt}}$ should be $N_{\mathrm{t}}$ (no longer $KN_s$ as single-carrier systems)[5] and thus perfect decomposition can only be achieved when $N_{\mathrm{RF}}^{\mathrm{t}} \geqslant N_{\mathrm{t}}$, which, however, severely deviates from the setting of hybrid precoding. Therefore, the matrix decomposition cannot be perfect for hybrid precoder design due to the rank deficiency, i.e., $N_{\mathrm{RF}}^{\mathrm{t}} = \mathrm{rank}\left(\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{BB}}\right) \ll \mathrm{rank}\left(\boldsymbol{F}_{\mathrm{opt}}\right) = N_{\mathrm{t}}$. Therefore, problem (12) is typically a low-rank matrix approximation problem, with a closed-form solution as

$$(\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{BB}})^{\star} \triangleq \hat{\boldsymbol{F}}_{\mathrm{opt}} = \boldsymbol{U}_1\boldsymbol{S}_1\boldsymbol{V}_1^{\mathrm{H}} . \tag{13}$$

Denote the SVD of $\boldsymbol{F}_{\mathrm{opt}}$ as $\boldsymbol{F}_{\mathrm{opt}} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^{\mathrm{H}}$, where matrices $\boldsymbol{U}_1$ and $\boldsymbol{V}_1$ are the first $N_{\mathrm{RF}}^{\mathrm{t}}$ columns of $\boldsymbol{U}$ and $\boldsymbol{V}$, respectively, and $\boldsymbol{S}_1$ is the diagonal matrix whose diagonal elements are the $N_{\mathrm{RF}}^{\mathrm{t}}$ largest singular values of $\boldsymbol{F}_{\mathrm{opt}}$. This means that the optimal solution of $\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{BB}}$ is simply obtained by extracting the $N_{\mathrm{RF}}^{\mathrm{t}}$ most principle components of $\boldsymbol{F}_{\mathrm{opt}}$. From the optimal solution (13), we observe that

$$\left\|(\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{BB}})^{\star}\right\|_{\mathrm{F}}^2 = \left\|\hat{\boldsymbol{F}}_{\mathrm{opt}}\right\|_{\mathrm{F}}^2 \leqslant \left\|\boldsymbol{F}_{\mathrm{opt}}\right\|_{\mathrm{F}}^2 \leqslant KN_sF, \tag{14}$$

which means the transmit power constraint is satisfied by the optimal solution $(\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{BB}})^{\star}$. Until now we have obtained the optimal solution for the entire hybrid precoder, and our next task is to decompose it into two parts. In fact, a large number of options are available for decomposing $\hat{\boldsymbol{F}}_{\mathrm{opt}}$ into $\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{BB}}$. Nevertheless, we are especially interested in the following one.

**Lemma 3** The matrix $\hat{\boldsymbol{F}}_{\mathrm{opt}}$ can be decomposed into $\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{BB}}$ in the following form:

$$\boldsymbol{F}_{\mathrm{RF}} = \begin{bmatrix} \boldsymbol{I}_{N_{\mathrm{RF}}^{\mathrm{t}}} \\ \hat{\boldsymbol{F}}_{\mathrm{opt},2}\hat{\boldsymbol{F}}_{\mathrm{opt},1}^{\dagger} \end{bmatrix}, \quad \boldsymbol{F}_{\mathrm{BB}} = \hat{\boldsymbol{F}}_{\mathrm{opt},1},$$

---

[5]Without loss of generality, we assume all the precoding matrices in (12) have full rank.

where $\hat{\boldsymbol{F}}_{\mathrm{opt}} = [\hat{\boldsymbol{F}}_{\mathrm{opt},1}\ \hat{\boldsymbol{F}}_{\mathrm{opt},2}]^{\mathrm{T}}$, $\hat{\boldsymbol{F}}_{\mathrm{opt},1}$ and $\hat{\boldsymbol{F}}_{\mathrm{opt},2}$ are the first $N_{\mathrm{RF}}^{\mathrm{t}}$ rows and the $(N_{\mathrm{RF}}^{\mathrm{t}}+1)$th to $N_{\mathrm{t}}$th rows of $\hat{\boldsymbol{F}}_{\mathrm{opt}}$, respectively.

**Proof**      Assume $\boldsymbol{F}_{\mathrm{RF}} = [\boldsymbol{I}_{N_{\mathrm{RF}}^{\mathrm{t}}}\ \boldsymbol{X}]^{\mathrm{T}}$, then the main task to prove Lemma 3 is to find $\boldsymbol{X}$ and $\boldsymbol{F}_{\mathrm{BB}}$ that satisfy $\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{BB}} = \hat{\boldsymbol{F}}_{\mathrm{opt}}$.

First, we have

$$\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{BB}} = \begin{bmatrix} \boldsymbol{I}_{N_{\mathrm{RF}}^{\mathrm{t}}} \\ \boldsymbol{X} \end{bmatrix} \boldsymbol{F}_{\mathrm{BB}} = \begin{bmatrix} \boldsymbol{F}_{\mathrm{BB}} \\ \boldsymbol{X}\boldsymbol{F}_{\mathrm{BB}} \end{bmatrix} = \hat{\boldsymbol{F}}_{\mathrm{opt}} = \begin{bmatrix} \hat{\boldsymbol{F}}_{\mathrm{opt},1} \\ \hat{\boldsymbol{F}}_{\mathrm{opt},2} \end{bmatrix}.$$

Therefore, it is easy to determine that $\boldsymbol{F}_{\mathrm{BB}} = \hat{\boldsymbol{F}}_{\mathrm{opt},1}$. The remaining task is to solve the equation

$$\boldsymbol{X}\hat{\boldsymbol{F}}_{\mathrm{opt},1} = \hat{\boldsymbol{F}}_{\mathrm{opt},2}.$$

Since $\hat{\boldsymbol{F}}_{\mathrm{opt}}$ is with rank $N_{\mathrm{RF}}^{\mathrm{t}}$ and is obtained by the SVD of $\boldsymbol{F}_{\mathrm{opt}}$, the first $N_{\mathrm{RF}}^{\mathrm{t}}$ rows of $\hat{\boldsymbol{F}}_{\mathrm{opt}}$ (the rows of $\hat{\boldsymbol{F}}_{\mathrm{opt},1}$) are linearly independent, and the remaining rows in $\hat{\boldsymbol{F}}_{\mathrm{opt}}$ (the rows of $\hat{\boldsymbol{F}}_{\mathrm{opt},2}$) can be linearly expressed by the rows of $\hat{\boldsymbol{F}}_{\mathrm{opt},1}$. Hence, $\boldsymbol{X} = \hat{\boldsymbol{F}}_{\mathrm{opt},2}\hat{\boldsymbol{F}}_{\mathrm{opt},1}^{\dagger}$ is the solution to the equation, which completes the proof.

The advantage of this decomposition form lies in the pattern of $\boldsymbol{F}_{\mathrm{RF}}$ in Lemma 3. The first $N_{\mathrm{RF}}^{\mathrm{t}}$ rows of $\boldsymbol{F}_{\mathrm{RF}}$ form an identity matrix, which in fact does not need a phase shifter implementation since the zero elements correspond to no connections whereas the diagonal elements refer to direct connections from RF chains to antennas. This means we only need $2N_{\mathrm{RF}}^{\mathrm{t}}(N_{\mathrm{t}} - N_{\mathrm{RF}}^{\mathrm{t}})$ phase shifters in the analog RF precoder, instead of $2N_{\mathrm{RF}}^{\mathrm{t}}N_{\mathrm{t}}$. Although a similar result was presented in Ref. [20], note that the result in Lemma 3 saves $2N_{\mathrm{RF}}^{\mathrm{t}}$ more phase shifters and the method is simpler and more straightforward than the decomposition procedure involving two QR decompositions as in Ref. [20]. Furthermore, the decomposition pattern in Lemma 3 can also be applied to single-carrier systems based on the result in Lemma 2, which will further improve the energy efficiency when achieving the fully digital precoder.

As demonstrated above, by doubling the phase shifters, what we need for the hybrid precoder design is computing a subset of singular values and vectors of the fully digital precoding matrix, i.e., the $N_{\mathrm{RF}}^{\mathrm{t}}$ most principle components of $\boldsymbol{F}_{\mathrm{opt}}$, whose computational complexity is $\mathcal{O}(KN_sFN_tN_{\mathrm{RF}}^{\mathrm{t}})$. Recall that OMP, as the most popular algorithm for the conventional SPS implementation, is with the computational complexity $\mathcal{O}(KN_sFN_tN_{\mathrm{RF}}^{\mathrm{t}}(KL+N_{\mathrm{RF}}^{\mathrm{t}2}))$, which is higher than that of the simple approach we proposed and is related to the channel parameter $L = \sum_{k=1}^{K} N_{\mathrm{cl},k}N_{\mathrm{ray},k}$. In other words, the proposed DPS implementation equips us with precoding algorithms computationally much more efficient than existing ones. Later in section V, its merits on achievable performance will also be demonstrated via simulations.

### C. DPS-Enabled SPS Hybrid Precoding

In this part, inspired by the above hybrid precoder design, we propose an efficient way to design the conventional SPS implementation. In particular, based on the solution for the DPS implementation, we adopt an heuristic way to tackle the unit modulus constraints induced by the SPS implementation.

As shown in Eq. (13), the optimal hybrid precoder $\hat{\boldsymbol{F}}_{\mathrm{opt}}$ can be decomposed by SVD. Therefore, one optimal solution to the hybrid precoder with the DPS implementation is

$$\boldsymbol{F}_{\mathrm{RF}} = \boldsymbol{U}_1, \quad \boldsymbol{F}_{\mathrm{BB}} = \boldsymbol{S}_1\boldsymbol{V}_1^{\mathrm{H}}. \tag{15}$$

Note that the unitary matrix $\boldsymbol{U}_1$ fully extracts the information of the column space of $\hat{\boldsymbol{F}}_{\mathrm{opt}}$, whose basis are the orthonormal columns in $\boldsymbol{F}_{\mathrm{RF}}$.

In contrast, in the SPS implementation, the unit modulus constraints, i.e., $|\boldsymbol{F}_{\mathrm{RF}}(i,j)| = 1$ not only require each column in $\boldsymbol{F}_{\mathrm{RF}}$ to have a constant norm like $\boldsymbol{U}_1$, but also induce an element-wise constraint. Since each element in $\boldsymbol{F}_{\mathrm{RF}}$ can only contain the phase information, we propose to extract the phases of the optimal analog precoder for the DPS implementation to construct the SPS solution, given by

$$\boldsymbol{F}_{\mathrm{RF}} = \exp\{\imath\angle(\boldsymbol{U}_1)\}. \tag{16}$$

Although this step is based on heuristics, it shall be shown in section V that simply extracting the phase information only incurs negligible performance loss. Similar approaches can also be found in Refs. [9,17].

Compared with existing hybrid precoding algorithms with the SPS implementation, e.g., the MO-AltMin[9], OMP[7] algorithms, and the algorithm in Ref. [18], the proposed DPS-enabled design method enjoys much lower computational complexity without any iterative procedure, which makes it a good candidate for low-complexity hybrid precoding with the SPS fully-connected structure.

### D. Interuser Interference Cancellation

While we can perfectly cancel the interuser interference with the fully digital precoder $\boldsymbol{F}_{\mathrm{opt}}$, there will be residual interuser interference when applying the hybrid precoder, which is an approximation of the fully digital one. For the same reason, as hybrid combining is adopted at the receiver side, the interuser interference cannot be canceled by the receiver either. Later in section V, we will see that in multiuser multicarrier systems, interuser interference is a severe problem that will dramatically degrade the hybrid precoding performance, especially at high SNRs.

In this subsection, after designing the hybrid precoder and combiner, we propose to cascade another digital baseband precoder $\boldsymbol{F}_{\mathrm{BD}}$ that is responsible for canceling the residual interuser interference. In particular, with the hybrid precoder

and combiner at hand, we define an effective channel for the $k$th user on the $f$th subcarrier as

$$\hat{H}_{k,f} = W_{\text{BB}k,f}^{\text{H}} W_{\text{RF}k}^{\text{H}} H_{k,f} F_{\text{RF}} F_{\text{BB}f}, \qquad (17)$$

where $F_{\text{BB}f} = \left[ F_{\text{BB}1,f}, \cdots, F_{\text{BB}k,f}, \cdots, F_{\text{BB}K,f} \right] \in \mathbb{C}^{N_{\text{RF}}^{\text{t}} \times KN_s}$ is the composite digital precoder on the $f$th subcarrier. Our goal is to design the precoders $F_{\text{BD}k,f}$, which satisfy the conditions

$$\hat{H}_{j,f} F_{\text{BD}k,f} = \mathbf{0}, \quad k \neq j. \qquad (18)$$

A simple way to achieve the conditions is the BD precoder, and note that the dimension of the effective channel is $N_s \times KN_s$, which is sufficient for BD. More details can be found in Ref. [26]. Therefore, after cascading the BD precoder at the baseband, the overall digital baseband precoder of the $k$th user on the $f$th subcarrier is

$$F_{\text{B}k,f} = F_{\text{BB}k,f} F_{\text{BD}k,f}. \qquad (19)$$

Since now we have obtained an interuser interference free system, we can normalize the precoder to satisfy the maximum transmit power, in order to improve the SNRs of the users. The same approach to cancel the interuser interference will also be used in the partially-connected structure and will not be repeatedly presented in the next section.

## IV. HYBRID PRECODING FOR THE PARTIALLY-CONNECTED STRUCTURE

One of the shortages of the fully-connected structure is the large number of phase shifters. The partially-connected structure, as a more energy efficient and cost-effective structure[9,23], employs notably fewer phase shifters, i.e., $2N_{\text{t}}$ phase shifters with the DPS implementation, which lends itself to practical implementation. Since the DoFs of the analog precoder is greatly reduced, RF-only precoding is far from satisfactory in the partially-connected structure. In this section, we shall first present the hybrid precoding with a fixed mapping from RF chains to antennas. Two algorithms will be then proposed to perform dynamic mapping to further improve the performance.

### A. Hybrid Precoding With Fixed Mapping

In Refs. [9,23], fixed mapping was considered in the partially-connected structure, i.e., each RF chain is connected to a certain number of antennas in a predetermined manner. To present the hybrid precoder design with fixed mapping clearly, we take one special mapping[9,23,29] as an example in the following, where the $j$th RF chain is connected to the $j$th set of $N_{\text{t}}/N_{\text{RF}}^{\text{t}}$ adjacent antennas. The corresponding constraint on the analog RF precoding matrix can be visualized as a set of

block diagonal matrices $\mathscr{A}_{\text{b}}$, where each block is an $N_{\text{t}}/N_{\text{RF}}^{\text{t}}$ dimension vector, i.e.,

$$F_{\text{RF}} = \begin{bmatrix} p_1 & 0 & \cdots & 0 \\ 0 & p_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & p_{N_{\text{RF}}^{\text{t}}} \end{bmatrix}, \qquad (20)$$

where

$$p_j = \left[ a_{(j-1)\frac{N_t}{N_{\text{RF}}^{\text{t}}}+1}, \cdots, a_{j\frac{N_t}{N_{\text{RF}}^{\text{t}}}} \right]^{\text{T}}.$$

The amplitude of the analog precoding gain for the $i$th connection from RF chains to antennas is denoted as $|a_i| \leqslant 2$. Similar to the hybrid precoding in the DPS fully-connected structure, the constraints $|a_i| \leqslant 2$ are redundant and therefore they are omitted in the following derivation. Furthermore, the transmit power constraint $\|F_{\text{RF}} F_{\text{BB}}\|_F^2 \leqslant KN_s F$ is also automatically satisfied by the optimal solution of the hybrid precoder, which will be shown in the following parts. Thus, the hybrid precoder design problem with fixed mapping can be recast as

$$\begin{aligned} \underset{F_{\text{RF}}, F_{\text{BB}}}{\text{minimize}} \quad & \left\| F_{\text{opt}} - F_{\text{RF}} F_{\text{BB}} \right\|_F^2 \\ \text{subject to} \quad & F_{\text{RF}} \in \mathscr{A}_{\text{b}}. \end{aligned} \qquad (21)$$

Note that there is only one non-zero element in each row of the analog RF precoding matrix $F_{\text{RF}}$. Due to this special structure, different vectors $p_j$ will be multiplied by distinct rows of $F_{\text{BB}}$, which decouples problem (21) into $N_{\text{RF}}^{\text{t}}$ subproblems in an RF chain-by-RF chain sense. The optimization of the hybrid precoder for the $j$th RF chain is given by

$$\mathscr{P}_j: \quad \underset{\{a_i\}, x_j}{\text{minimize}} \sum_{i \in \mathscr{F}_j} \left\| y_i - a_i x_j \right\|_2^2, \qquad (22)$$

where $\mathscr{F}_j = \{ i \in \mathbb{Z} | (j-1)\frac{N_t}{N_{\text{RF}}^{\text{t}}} + 1 \leqslant i \leqslant j\frac{N_t}{N_{\text{RF}}^{\text{t}}}. \}$, $y_i = F_{\text{opt}}^{\text{T}}(i,:)$, and $x_j = F_{\text{BB}}^{\text{T}}(j,:)$.

**Proposition 1**  The optimal solution to the subproblem $\mathscr{P}_j$ is given by the following closed-form expression.

$$x_j^{\star} = \lambda_1 \left( \sum_{i \in \mathscr{F}_j} y_i y_i^{\text{H}} \right), \quad a_i^{\star} = \frac{x_j^{\text{H}} y_i}{\|x_j\|_2^2}. \qquad (23)$$

**Proof**  We check the first order optimality conditions as

$$\frac{\partial}{\partial a_i} f(a_i, x_j) = 0 \Rightarrow -y_i^{\text{H}} x_j + a_i^* \|x_j\|_2^2 = 0 \Rightarrow a_i = \frac{x_j^{\text{H}} y_i}{\|x_j\|_2^2} \qquad (24)$$

$$\frac{\partial}{\partial x_j} f(a_i, x_j) = \mathbf{0} \Rightarrow \sum_{i \in \mathscr{F}_j} -a_i y_i^{\text{H}} + |a_i|^2 x_j^{\text{H}} = \mathbf{0} \Rightarrow$$

$$\sum_{i \in \mathscr{F}_j} |a_i|^2 x_j = \sum_{i \in \mathscr{F}_j} a_i^* y_i, \qquad (25)$$

where $f(a_i, \boldsymbol{x}_j)$ is the objective function in subproblem $\mathscr{P}_j$. Substituting Eq. (24) into Eq. (25), we can get

$$\boldsymbol{x}_j^{\mathrm{H}} \sum_{i \in \mathscr{F}_j} |a_i|^2 = \frac{\boldsymbol{x}_j^{\mathrm{H}}}{\|\boldsymbol{x}_j\|_2^2} \sum_{i \in \mathscr{F}_j} \boldsymbol{y}_i \boldsymbol{y}_i^{\mathrm{H}} \Rightarrow$$

$$\left( \sum_{i \in \mathscr{F}_j} \boldsymbol{y}_i \boldsymbol{y}_i^{\mathrm{H}} \right) \boldsymbol{x}_j = \left( \|\boldsymbol{x}_j\|_2^2 \sum_{i \in \mathscr{F}_j} |a_i|^2 \right) \boldsymbol{x}_j \triangleq \lambda_j \boldsymbol{x}_j, \quad (26)$$

which shows that $\lambda_j$ and $\boldsymbol{x}_j$ are the eigenvalue and eigenvector of $\sum_{i \in \mathscr{F}_j} \boldsymbol{y}_i \boldsymbol{y}_i^{\mathrm{H}}$. Moreover, by substituting Eq. (24) into the objective function in $\mathscr{P}_j$, it can be rewritten as

$$f(a_i, \boldsymbol{x}_j) = \sum_{i \in \mathscr{F}_j} \boldsymbol{y}_i^{\mathrm{H}} \boldsymbol{y}_i - |a_i|^2 \boldsymbol{x}_j^{\mathrm{H}} \boldsymbol{x}_j = \sum_{i \in \mathscr{F}_j} \|\boldsymbol{y}_i\|_2^2 - \lambda_j. \quad (27)$$

Hence, minimizing the objective function is equivalent to taking $\lambda_j$ as the largest eigenvalue of the covariance matrix $\sum_{i \in \mathscr{F}_j} \boldsymbol{y}_i \boldsymbol{y}_i^{\mathrm{H}}$, denoted as $\lambda_1(\sum_{i \in \mathscr{F}_j} \boldsymbol{y}_i \boldsymbol{y}_i^{\mathrm{H}})$.

From Eq. (27), we obtain

$$\|\boldsymbol{F}_{\mathrm{opt}} - \boldsymbol{F}_{\mathrm{RF}} \boldsymbol{F}_{\mathrm{BB}}\|_{\mathrm{F}}^2 =$$

$$\|\boldsymbol{F}_{\mathrm{opt}}\|_{\mathrm{F}}^2 - \sum_{j=1}^{N_{\mathrm{RF}}^{\mathrm{t}}} \lambda_j = \|\boldsymbol{F}_{\mathrm{opt}}\|_{\mathrm{F}}^2 - \sum_{j=1}^{N_{\mathrm{RF}}^{\mathrm{t}}} \left( \|\boldsymbol{x}_j\|_2^2 \sum_{i \in \mathscr{F}_j} |a_i|^2 \right) =$$

$$\|\boldsymbol{F}_{\mathrm{opt}}\|_{\mathrm{F}}^2 - \|\boldsymbol{F}_{\mathrm{RF}} \boldsymbol{F}_{\mathrm{BB}}\|_{\mathrm{F}}^2 \geqslant 0 \Rightarrow$$

$$\|\boldsymbol{F}_{\mathrm{RF}} \boldsymbol{F}_{\mathrm{BB}}\|_{\mathrm{F}}^2 \leqslant \|\boldsymbol{F}_{\mathrm{opt}}\|_{\mathrm{F}}^2 \leqslant K N_s F, \quad (28)$$

which means that the transmit power constraint is naturally satisfied by the optimal solutions. While we fixed the mapping strategy as shown in Eq. (20), the proposed design approach is applicable to an arbitrary mapping strategy.

### B. Hybrid Precoding With Dynamic Mapping

Different from the fully-connected structure that utilizes all the connections from RF chains to antennas, the partially-connected structure will induce non-negligible performance loss[9]. In this section, we propose to improve its performance by optimizing the mapping strategy, i.e., we will dynamically determine for each RF chain which antennas it should be connected. The dynamic mapping problem is given as

$$\underset{\boldsymbol{F}_{\mathrm{RF}}, \boldsymbol{F}_{\mathrm{BB}}}{\text{minimize}} \quad \|\boldsymbol{F}_{\mathrm{opt}} - \boldsymbol{F}_{\mathrm{RF}} \boldsymbol{F}_{\mathrm{BB}}\|_{\mathrm{F}}^2$$
$$\text{subject to} \quad \boldsymbol{F}_{\mathrm{RF}} \in \mathscr{A}_{\mathrm{d}}, \quad (29)$$

where $\mathscr{A}_{\mathrm{d}}$ is a set of matrices for which every row only has one non-zero entry, i.e., $\mathscr{A}_{\mathrm{d}} = \{\boldsymbol{A} \| \|\boldsymbol{A}(i, :)\|_0 = 1\}$, meaning that each antenna can only be connected to one RF chain. As indicated by Eq. (28), once the mapping is fixed, the optimal value of the objective function in (21) is

$$\|\boldsymbol{F}_{\mathrm{opt}}\|_{\mathrm{F}}^2 - \sum_{j=1}^{N_{\mathrm{RF}}^{\mathrm{t}}} \lambda_1 \left( \sum_{i \in \mathscr{D}_j} \boldsymbol{y}_i \boldsymbol{y}_i^{\mathrm{H}} \right). \quad (30)$$

Hence, when we have the freedom to design the mapping strategy from RF chains to antennas, the design target is to seek the mapping that maximizes the sum of the largest eigenvalues, i.e.,

$$\underset{\{\mathscr{D}_j\}_{j=1}^{N_{\mathrm{RF}}^{\mathrm{t}}}}{\text{maximize}} \quad \sum_{j=1}^{N_{\mathrm{RF}}^{\mathrm{t}}} \lambda_1 \left( \sum_{i \in \mathscr{D}_j} \boldsymbol{y}_i \boldsymbol{y}_i^{\mathrm{H}} \right)$$
$$\text{subject to} \quad \begin{cases} \cup_{j=1}^{N_{\mathrm{RF}}^{\mathrm{t}}} \mathscr{D}_j = \{1, \cdots, N_{\mathrm{t}}\} \\ \mathscr{D}_j \cap \mathscr{D}_k = \emptyset, \quad \forall j \neq k, \end{cases} \quad (31)$$

where $\mathscr{D}_j$ is the mapping set containing the antenna indices that are mapped to the $j$th RF chain. The dynamic mapping problem is a combinatorial problem and the optimal solution can be given by exhaustive search with an extremely huge number of possible mapping strategies as $\frac{1}{(N_{\mathrm{RF}}^{\mathrm{t}})!} \sum_{k=0}^{N_{\mathrm{RF}}^{\mathrm{t}}} (-1)^{N_{\mathrm{RF}}^{\mathrm{t}} - k} \binom{N_{\mathrm{RF}}^{\mathrm{t}}}{k} k^{N_{\mathrm{t}}}$, which prevents its practical implementation. Therefore, we first propose a greedy algorithm to solve the problem. The pseudocode of the algorithm is omitted due to its simplicity and space limitation.

In each iteration of the greedy algorithm, we connect the $p^\star$th antenna to the $j^\star$th RF chain, which is the connection with the maximum increment of the largest eigenvalue when this connection is added into the mapping network. Note that the computational complexity of the algorithm is dominated by the calculation of the largest eigenvalue. In the greedy algorithm, the number of times we need to perform the eigenvalue decomposition (EVD) is $\mathscr{O}(N_{\mathrm{t}} N_{\mathrm{RF}}^{\mathrm{t}} (1 + N_{\mathrm{t}})/2)$, which is a quite large number especially when large-scale antenna arrays are leveraged in mmWave MIMO systems. To relieve us from the high computational complexity, we then propose a modified K-means algorithm to solve the dynamic design problem (31).

We reconsider problem (31) as follows. The problem is equivalent to classifying $N_{\mathrm{t}}$ vectors (antennas) into $N_{\mathrm{RF}}^{\mathrm{t}}$ clusters (RF chains). K-means, aiming at partitioning the observation vectors into $K_{\mathrm{cl}}$ clusters, is a prevalent approach for cluster analysis in data mining, where $K_{\mathrm{cl}}$ is a predefined parameter, and turns out to be suitable for problem (31). In the classical K-means algorithm, the objective is to minimize the sum of the Euclidean distances from each observation vector to the centroid of the cluster it belongs to. The distortion function that is to be minimized in the classical K-means algorithm is given by[6]

$$D(\boldsymbol{y}_i, \boldsymbol{x}_j) = \sum_{j=1}^{K_{\mathrm{cl}}} \sum_{i \in \mathscr{K}_j} \|\boldsymbol{y}_i - \boldsymbol{x}_j\|_2^2, \quad (32)$$

---

[6]We present the distortion function in the classified K-means algorithm with a slight abuse of notations $\boldsymbol{y}_i$ and $\boldsymbol{x}_j$ so that the content of the modified one in the following is easier to follow.

where $\boldsymbol{y}_i$ are the observation vectors while $\boldsymbol{x}_j$ is the centroid of the $j$th cluster.

However, this distortion function cannot be directly adopted to solve the dynamic mapping design problem (31) since the objectives are quite different. In (31), the objective is to maximize the sum of the largest eigenvalues of the covariance matrices of each cluster. Therefore, we propose to modify the distortion function in the K-means algorithm as

$$D'(\boldsymbol{y}_i, \boldsymbol{x}_j) = \sum_{j=1}^{N_{\mathrm{RF}}^{\mathrm{t}}} \frac{\boldsymbol{x}_j^{\mathrm{H}} (\sum_{i \in \mathscr{D}_j} \boldsymbol{y}_i \boldsymbol{y}_i^{\mathrm{H}}) \boldsymbol{x}_j}{\boldsymbol{x}_j \boldsymbol{x}_j^{\mathrm{H}}}. \tag{33}$$

The modified distortion function is the sum of Rayleigh quotients of the covariance matrices of each cluster, whose optimal value is the sum of the largest eigenvalues when we maximize Eq. (33) over $\boldsymbol{x}_j$. The overall clustering problem can be written as

$$\begin{aligned} \underset{\{\mathscr{D}_j, \boldsymbol{x}_j\}_{j=1}^{N_{\mathrm{RF}}^{\mathrm{t}}}}{\text{maximize}} \quad & \sum_{j=1}^{N_{\mathrm{RF}}^{\mathrm{t}}} \frac{\boldsymbol{x}_j^{\mathrm{H}} (\sum_{i \in \mathscr{D}_j} \boldsymbol{y}_i \boldsymbol{y}_i^{\mathrm{H}}) \boldsymbol{x}_j}{\boldsymbol{x}_j \boldsymbol{x}_j^{\mathrm{H}}} \\ \text{subject to} \quad & \begin{cases} \cup_{j=1}^{N_{\mathrm{RF}}^{\mathrm{t}}} \mathscr{D}_j = \{1, \cdots, N_{\mathrm{t}}\} \\ \mathscr{D}_j \cap \mathscr{D}_k = \emptyset, \quad \forall j \neq k. \end{cases} \end{aligned} \tag{34}$$

We propose to adopt alternating maximization to solve this problem, which alternately updates the clustering and centroids when the other one is fixed. This approach results in closed-form solutions for the two update procedures.

In the clustering update, we allocate each vector to the cluster whose centroid has the largest inner product with it, i.e., allocate $\boldsymbol{y}_i$ to the $j^\star$th cluster, where

$$j^\star = \arg\max_j \left| \boldsymbol{y}_i^{\mathrm{H}} \boldsymbol{x}_j \right|^2. \tag{35}$$

In the centroid update, the optimization of the centroids is equivalent to maximizing the Rayleigh quotients for each cluster, whose optimal solution is simply given by the eigenvector corresponding to the largest eigenvalue, i.e.,

$$\boldsymbol{x}_j^\star = \boldsymbol{\lambda}_1 \left( \sum_{i \in \mathscr{D}_j} \boldsymbol{y}_i \boldsymbol{y}_i^{\mathrm{H}} \right). \tag{36}$$

Now we have the modified K-means algorithm, which is summarized as Algorithm 1.

Note that steps 3 and 4 both give the globally optimal solutions to the clustering and centroid. Hence, the algorithm will converge to a stationary point since it is a two block coordinate descent procedure[30]. Because the modified distortion function is not jointly convex with respect to $\{\mathscr{D}_j\}_{j=1}^{N_{\mathrm{RF}}^{\mathrm{t}}}$ and $\{\boldsymbol{x}_j\}_{j=1}^{N_{\mathrm{RF}}^{\mathrm{t}}}$, the modified K-means algorithm converges to a local optimum of problem (34) so the solution is sensitive to the initial centroids selection. For hybrid precoding, the size

---

**Algorithm 1** Modified K-means algorithm for dynamic mapping in the partially-connected structure

---

**Require:** $\{\boldsymbol{y}_i = \boldsymbol{F}_{\mathrm{opt}}^{\mathrm{T}}(i,:)\}_{i=1}^{N_{\mathrm{t}}}$

1: Construct the initial centroids $\{\boldsymbol{x}_j\}_{j=1}^{N_{\mathrm{RF}}^{\mathrm{t}}}$;
2: **repeat**
3:　Fix the centroids, allocate $\{\boldsymbol{y}_i\}_{i=1}^{N_{\mathrm{t}}}$ to the clusters according to Eq. (35);
4:　Optimize the centroids $\{\boldsymbol{x}_j\}_{j=1}^{N_{\mathrm{RF}}^{\mathrm{t}}}$ using Eq. (36) when the clustering is fixed;
5: **until** convergence;
6: Calculate $\boldsymbol{F}_{\mathrm{BB}}^{\mathrm{T}}(j,:) = \boldsymbol{x}_j = \boldsymbol{\lambda}_1 (\sum_{i \in \mathscr{D}_j} \boldsymbol{y}_i \boldsymbol{y}_i^{\mathrm{H}})$ for $j \in \{1, \cdots, N_{\mathrm{RF}}^{\mathrm{t}}\}$;
7: Compute the effective channels $\hat{\boldsymbol{H}}_{k,f} = \boldsymbol{W}_{\mathrm{BB}k,f}^{\mathrm{H}}$ and BD precoders $\boldsymbol{F}_{\mathrm{BD}k,f}$ according to Eqs. (17) and (18);
8: For the digital precoder at the transmit end, normalize $\widehat{\boldsymbol{F}}_{\mathrm{B}} = \frac{\sqrt{KN_sF}}{\|\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{B}}\|_{\mathrm{F}}} \boldsymbol{F}_{\mathrm{B}}$.

---

of the observation set is much larger than the cluster number, i.e., $N_{\mathrm{t}} \gg N_{\mathrm{RF}}^{\mathrm{t}}$. One heuristic rule of thumb to design the initial centroids is to pick $N_{\mathrm{RF}}^{\mathrm{t}}$ observation vectors with small correlations. In our proposed algorithm, we propose to select $N_{\mathrm{RF}}^{\mathrm{t}}/2$ pairs of vectors out of the $N_{\mathrm{t}}$ observation vectors as the initial centroids, which have the $N_{\mathrm{RF}}^{\mathrm{t}}/2$ smallest inner products.

Recall that EVD is the dominant part of the computational complexity in dynamic mapping design. In each alternating iteration in the modified K-means algorithm, $N_{\mathrm{RF}}^{\mathrm{t}}$ times of EVD are needed and therefore the overall times are $\mathscr{O}(NN_{\mathrm{RF}}^{\mathrm{t}})$, where $N$ is the iteration number. For practical settings in section V, the modified K-means algorithm typically converges within 10 iterations, which is much less than $N_{\mathrm{t}}(1+N_{\mathrm{t}})/2$ and thus results in significant complexity reduction compared to the greedy algorithm.

### C. Fully-Connected vs. Partially-Connected Structures

There exist several studies[9,11,15] investigating different design algorithms for the fully- and partially-connected structures, and comparisons between these two structures are provided via simulations. However, to the best of the authors' knowledge, so far there is no analytical quantitative comparisons between different structures. The complicated design approaches to handle the unit modulus constraints induced by the SPS implementation are the main obstacles. With the DPS implementation and its resulting low-complexity design approaches at hand, we are able to fill this gap. Following (13) and (21) in sections III and IV, we obtain that

$$\begin{aligned} f_f^\star &= \sum_{p=N_{\mathrm{RF}}^{\mathrm{t}}+1}^{N_{\mathrm{t}}} \sigma_p^2 \left( \boldsymbol{F}_{\mathrm{opt}} \right), \\ f_p^\star &= \left\| \boldsymbol{F}_{\mathrm{opt}} \right\|_{\mathrm{F}}^2 - \sum_{j=1}^{N_{\mathrm{RF}}^{\mathrm{t}}} \boldsymbol{\lambda}_1 \left( \sum_{i \in \mathscr{D}_j} \boldsymbol{y}_i \boldsymbol{y}_i^{\mathrm{H}} \right), \end{aligned} \tag{37}$$

where $f_f^\star$ and $f_p^\star$ are the optimal values of the objective function in (4) for the fully- and partially-connected structures, respectively, and $\sigma_p(\boldsymbol{F}_{\text{opt}})$ denotes the $p$th largest singular value of $\boldsymbol{F}_{\text{opt}}$. We define the performance gap $\Delta$ between the fully- and partially-connected structures as the different between two objective values, i.e.,

$$
\begin{aligned}
\Delta &\triangleq f_p^\star - f_f^\star = \\
&\|\boldsymbol{F}_{\text{opt}}\|_{\text{F}}^2 - \sum_{j=1}^{N_{\text{RF}}^{\text{t}}} \lambda_1 \left( \sum_{i \in \mathscr{D}_j} \boldsymbol{y}_i \boldsymbol{y}_i^{\text{H}} \right) - \sum_{p=N_{\text{RF}}^{\text{t}}+1}^{N_{\text{t}}} \sigma_p^2 \left( \boldsymbol{F}_{\text{opt}} \right) = \\
&\sum_{p=1}^{N_{\text{RF}}^{\text{t}}} \lambda_p \left( \boldsymbol{F}_{\text{opt}}^{\text{H}} \boldsymbol{F}_{\text{opt}} \right) - \sum_{j=1}^{N_{\text{RF}}^{\text{t}}} \lambda_1 \left( \sum_{i \in \mathscr{D}_j} \boldsymbol{y}_i \boldsymbol{y}_i^{\text{H}} \right) = \\
&\sum_{p=1}^{N_{\text{RF}}^{\text{t}}} \lambda_p \left( \boldsymbol{F}_{\text{opt}}^{\text{H}} \boldsymbol{F}_{\text{opt}} \right) - \sum_{j=1}^{N_{\text{RF}}^{\text{t}}} \lambda_1 \left( \boldsymbol{Y}_j \boldsymbol{Y}_j^{\text{H}} \right),
\end{aligned} \tag{38}
$$

where $\boldsymbol{Y}_j$ is composed of the vectors $\{\boldsymbol{y}_i\}_{i \in \mathscr{D}_j}$ as its columns. Therefore, once the RF-antenna mapping in the partially-connected structure is determined, the performance gap is given by Eq. (38), which provides an analytical comparison of two hybrid precoder structures. This expression indicates that the performance gap depends on the channel realization, as well as the RF chain-antenna mapping strategy.

## V. SIMULATION RESULTS

In this section, we numerically evaluate the performance of the proposed hybrid precoder design for multiuser OFDM mmWave MIMO systems, with $F = 128$ subcarriers are assumed. The channel parameters are given by $N_{\text{cl},k} = 3$ clusters, $N_{\text{ray},k} = 8$ rays, and the average power of each cluster is $\sigma_\alpha^2 = 1$. The AoDs and AoAs follow the Laplacian distribution with uniformly distributed mean angles in $[0, 2\pi)$ and angular spread of 10 degrees. The antenna elements in the USPA are separated by half wavelength, and all simulation results are averaged over 1000 channel realizations.

### A. RF-Only Precoding in the Fully-Connected Structure

First, we test how much performance gain we can get when we double the number of phase shifters via investigating the RF-only precoding in the fully-connected structure. Note that, with the conventional SPS implementation, manifold optimization was shown in Ref. [9] to be an effective method to directly tackle the unit modulus constraints and achieve higher spectral efficiency than other existing works. In this subsection, we adopt Algorithm 1 in Ref. [9] for the SPS implementation as the benchmark to show the advantage of the proposed DPS implementation. For fair comparison, we always keep the same digital baseband precoder $\boldsymbol{F}_{\text{BB}}$ that is
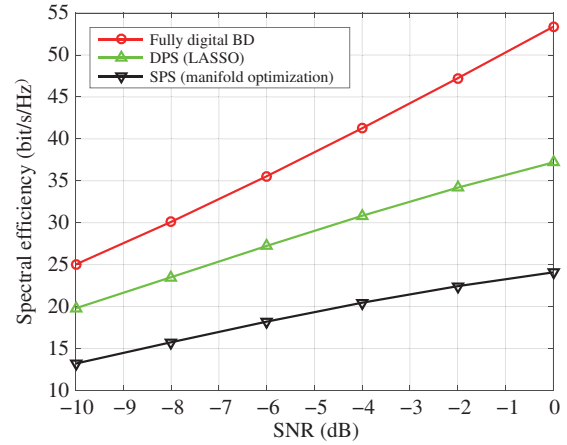


**Figure 3**   Spectral efficiency achieved by different hybrid precoder implementations when $N_{\text{t}} = 64$, $N_{\text{r}} = 9$, $K = 5$, $N_s = 2$, and $N_{\text{RF}}^{\text{t}} = KN_{\text{RF}}^{\text{r}} = KN_s$

semi-orthogonal for both SPS and DPS implementations in the simulation[7].

Fig. 3 shows the spectral efficiency of two different hybrid precoder implementations. First, we see that the proposed DPS implementation outperforms the conventional SPS implementation. This performance gain is obtained from doubling the phase shifters in the analog RF precoder since we adopt the same digital precoder for both implementations. More importantly, the LASSO formulation of the RF-only precoding in section III.A enables low-complexity algorithms and results in a closed-form solution when the digital precoder is semi-orthogonal. In contrast, the algorithm based on manifold optimization involves a complicated iterative procedure and hence leads to high computational complexity[9]. With the conventional SPS implementation, there always exists an obvious trade-off between the design complexity and performance[9,11,17]. On the contrary, with the DPS implementation, by modifying the hardware implementation, we can significantly reduce the design complexity and improve the system performance in the meanwhile.

### B. Hybrid Precoding in the Fully-Connected Structure

Fig. 4 shows the spectral efficiency achieved by different algorithms with the minimum numbers of RF chains, i.e., $N_{\text{RF}}^{\text{t}} = KN_s$ and $N_{\text{RF}}^{\text{r}} = N_s$. To illustrate the effectiveness of the proposed implementation and algorithms, two hybrid precoding algorithms with the SPS implementation[7,18] are adopted as benchmarks. The widely used OMP algorithm[7] is with the lowest computational complexity among all the available algorithms with the conventional SPS implementation. The algorithm in Ref. [18] achieves the best spectral efficiency in the literature for multiuser OFDM systems with the SPS implementation, which iteratively optimizes the phases in the

---

[7]Note that we cannot adjust the digital precoder in the RF-only precoding so we do not apply the additional BD mentioned in section III.D at this stage.
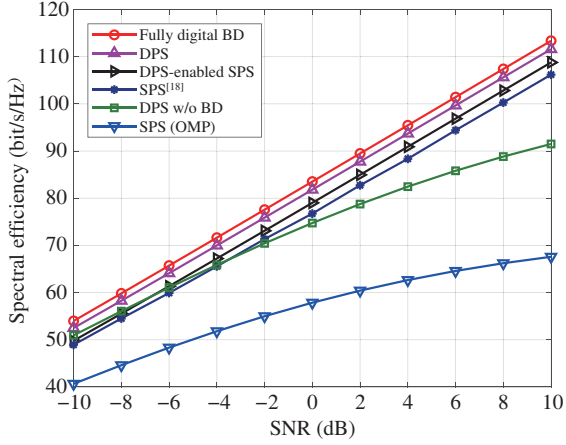
**Figure 4** Spectral efficiency achieved by different hybrid precoding algorithms in the fully-connected structure when $N_t = 256$, $N_r = 16$, $K = 3$, and $N_s = 3$



**Figure 5** Spectral efficiency achieved by different hybrid precoding algorithms in the fully-connected structure for different transmit RF chain numbers $N_{RF}^t$, given $N_t = 256$, $N_r = 16$, $K = 3$, $N_s = 3$, $N_{RF}^r = N_s$, and SNR = 5 dB

analog precoder, therefore with high design complexity.

First, we see that the DPS implementation outperforms the existing algorithm in Ref. [18] with the conventional SPS implementation, and approaches the fully digital one. This means that the proposed algorithm can more accurately approximate the fully digital precoder than existing algorithms, even though the RF chains are limited. While the OMP is already with the lowest computational complexity among all the available algorithms for the conventional SPS implementation, the proposed design approach with simple matrix decomposition enjoys even lower computational complexity and much higher spectral efficiency than the OMP algorithm. This shows that the proposed DPS implementation achieves both good spectral efficiency and computational efficiency.

In Refs. [9,11,16], it has been pointed out that approximating the fully digital precoder with a hybrid structure will lead to a near optimal performance in single-user single-carrier, single-user multicarrier, and multiuser single-carrier mmWave MIMO systems, respectively. Next, we will verify whether this insight still holds in more general multiuser OFDM mmWave MIMO systems.

In Fig. 4, we evaluate the performance of the proposed hybrid precoding without the additional BD operation mentioned in section III.D, i.e., only approximating the fully digital precoder. We discover that, without the BD precoder canceling the interuser interference, there will be residual interuser interference, which results in an obvious performance loss compared to the fully digital one, especially at high SNRs. This phenomenon illustrates that simply approximating the fully digital precoder with the hybrid one is not sufficient in multiuser multicarrier mmWave systems. Although interuser interference also exists in multiuser single-carrier systems, this issue is more prominent in the multicarrier system as the analog precoder is shared by a large number of subcarriers. The comparison in Fig. 4 demonstrates the effectiveness and ne-
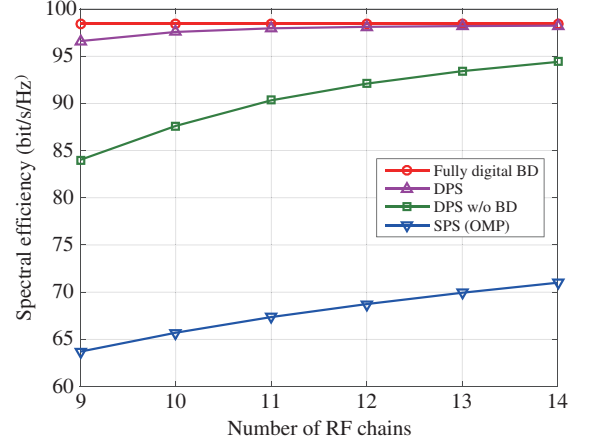
cessity of the additional BD operation proposed in section III.D.

In addition, the performance of the DPS-enabled SPS design is presented. Fig. 4 shows that it only entails little performance loss compared to the DPS one, and outperforms the spectral efficiency of the state-of-the-art algorithm in Ref. [18]. This phenomenon empirically verifies the effectiveness of the phase extraction operation discussed in section III.C. Note that, with the SPS implementation, existing design algorithms involve iterative procedures[7,18]. In contrast, the proposed DPS-enabled one offers a low-complexity design with closed-form solutions. Overall, the proposed DPS-enabled SPS approach provides an effective heuristic way to design SPS hybrid precoders, which enjoy both lower computational complexity and higher achievable spectral efficiency than existing algorithms.

Fig. 5 compares different precoding schemes for different RF chain numbers $N_{RF}^t$ at the BS side while keeping $N_{RF}^r = N_s$ as the minimum number of RF chains at each user. It is shown that the DPS implementation approaches the performance of the fully digital precoder when the number of RF chains is slightly larger than the number of transmitted data streams, which cannot be realized by the existing OMP algorithm. There is no need to further increase the number of RF chains considering the increased power consumption and hardware complexity.

### C. Hybrid Precoding in the Partially-Connected Structure

Fig. 6 shows the performance of the proposed design approaches in the partially-connected structure with the minimum numbers of RF chains, i.e., $N_{RF}^t = KN_s$ and $N_{RF}^r = N_s$. We see that, due to the sharply reduced number of phase shifters, the partially-connected structure does entail

**Table 1**    Comparisons between different hybrid precoder implementations and structures

| Imple-mentation | Structure | Design approach | Hardware complexity (No. of phase shifters) | Computational complexity | Performance |
|---|---|---|---|---|---|
| SPS | Fully-connected | MO-AltMin[9] | $N_{\text{RF}}^{\text{t}}N_{\text{t}}$ | Extremely high | ✓✓✓ |
| | Partially-connected | SDR-AltMin[9] | $N_{\text{t}}$ | High | ✓ |
| DPS | Fully-connected | Matrix decomposition | $2N_{\text{RF}}^{\text{t}}(N_{\text{t}} - N_{\text{RF}}^{\text{t}})$ | $\mathscr{O}(N_{\text{RF}}^{\text{t}}{}^2 N_{\text{t}}F)$ | ✓✓✓✓ |
| | Partially-connected | Modified K-means | $2N_{\text{t}}$ | $\mathscr{O}(NN_{\text{RF}}^{\text{t}}{}^2 N_{\text{t}}F)$ | ✓✓ |

\* For each structure, we pick the design approach with the best performance among existing works. More ✓ means higher spectral efficiency. The comparisons of the computational complexity between the SPS and DPS implementations are based on the running time in the simulation. The computational complexity for the DPS implementation is calculated with the minimum number of RF chains, i.e., $N_{\text{RF}}^{\text{t}} = KN_s$.
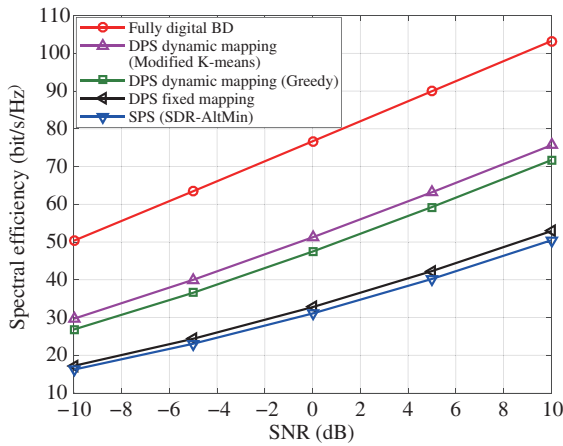


**Figure 6**    Spectral efficiency achieved by different hybrid precoding algorithms in the partially-connected structure when $N_{\text{t}} = 256$, $N_{\text{r}} = 16$, $K = 4$, and $N_s = 2$

non-negligible performance loss compared to the fully digital one. Furthermore, although the computational complexity of the proposed algorithms is significantly reduced compared to the SDR-AltMin algorithm with the SPS implementation in Ref. [9], simply doubling the number of phase shifters with the fixed mapping only has little performance gain over the conventional SPS implementation. However, Fig. 6 demonstrates that dynamic mapping is able to shrink the gap between the fixed mapping and the fully digital precoding by half.

Furthermore, we compare the performance of the proposed greedy and modified K-means algorithms for the dynamic mapping in the partially-connected structure. Fig. 6 illustrates that, although the modified K-means algorithms only converges to a local optimum of (34), it achieves a higher spectral efficiency than the greedy algorithm. Even though the cluster number $N_{\text{RF}}^{\text{t}}$ is much smaller than the number of observation vectors $N_{\text{t}}$ in the modified K-means algorithm, it turns out that the algorithm converges quickly in the simulation[8]. Overall, the modified K-means algorithm steps up as an excellent dynamic mapping design algorithm for hybrid precoding in the

partially-connected structure.

## VI. Conclusions

This paper proposed a new DPS implementation for hybrid precoding, based on which computationally efficient algorithms were developed for different hybrid precoder structures in general multiuser OFDM mmWave systems. The comparisons between different hybrid precoder implementations and structures are listed in Tab. 1, which help reveal important design guidelines as follows.

• For the DPS fully-connected structure, the computational complexity of the hybrid precoder design is significantly reduced, with the performance approaching the fully digital one when the number of RF chains is comparable to that of the data streams. For cost and power consideration, it is unnecessary to further employ more RF chains.

• Dynamic mapping is shown to be critical to improve the spectral efficiency in the DPS partially-connected structure, which is basically a clustering problem. Although there exists some performance loss compared with the fully-connected one, due to the much diminished hardware complexity, the partially-connected structure serves as an economic choice for hybrid precoding.

• Overall, it is beneficial, from both performance and computational complexity points of view, to adopt the DPS implementation in hybrid precoding, in comparison to the conventional SPS implementation. Interestingly, the fully-connected structure enjoys lower computational complexity than the partially-connected one, as dynamic mapping brings additional complexity in order to enhance the performance.

In summary, the proposed DPS implementation enjoys advantages of high spectral efficiency and high computational efficiency, at the cost of more hardware components. We envision that once low-cost high-resolution commercial phase shifters are available, or for cost-insensitive applications, the DPS implementation would be an ideal choice. On the other hand, our investigation on this new implementation also provides valuable guidelines on designing other hybrid precoder structures, as summarized in Remark 1. While this paper

---

[8]The modified K-means algorithm converges within 10 iterations for almost all the channel realizations.

demonstrated the great potential of the proposed DPS implementation, more works will be needed to investigate other problems involving hybrid precoder design, e.g., to consider the hybrid precoder design combined with channel training and feedback. Furthermore, more efforts are needed on quantitative comparisons between different hybrid precoder structures considering realistic hardware components.

## APPENDIX

Since both (5) and (6) are convex problems, it is equivalent to prove that the dual problem of (6) is (5). We introduce a dummy variable $\boldsymbol{y}$ to derive the dual as follows.

$$\underset{\boldsymbol{x},\boldsymbol{y}}{\text{minimize}} \quad \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_2^2 + 2\|\boldsymbol{y}\|_1 \tag{39}$$
$$\text{subject to} \quad \boldsymbol{y}=\boldsymbol{x}.$$

The Lagrangian of (39) is

$$\mathscr{L}(\boldsymbol{x},\boldsymbol{y};\boldsymbol{v}) = \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_2^2 + 2\|\boldsymbol{y}\|_1 + \boldsymbol{v}^{\mathrm{H}}(\boldsymbol{x}-\boldsymbol{y}). \tag{40}$$

Therefore, the dual function is given by

$$\underset{\boldsymbol{v}}{\text{maximize}} \; \underset{\boldsymbol{x},\boldsymbol{y}}{\inf} \; \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_2^2 + \boldsymbol{v}^{\mathrm{H}}\boldsymbol{x} + 2\|\boldsymbol{y}\|_1 - \boldsymbol{v}^{\mathrm{H}}\boldsymbol{y}. \tag{41}$$

By the definition of the conjugate function to the $\ell_1$-norm, we can obtain that

$$\underset{\boldsymbol{y}}{\inf} \, 2\|\boldsymbol{y}\|_1 - \boldsymbol{v}^{\mathrm{H}}\boldsymbol{y} = -\underset{\boldsymbol{y}}{\sup} \, \boldsymbol{v}^{\mathrm{H}}\boldsymbol{y} - 2\|\boldsymbol{y}\|_1 =$$
$$\begin{cases} 0, & \text{if } \|\boldsymbol{v}\|_\infty \leqslant 2, \\ -\infty, & \text{otherwise.} \end{cases} \tag{42}$$

By checking the first order optimality condition of $\boldsymbol{x}$, we can get

$$\frac{\partial}{\partial\boldsymbol{x}}\mathscr{L}(\boldsymbol{x},\boldsymbol{y};\boldsymbol{v}) = \boldsymbol{A}^{\mathrm{H}}(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}) + \boldsymbol{v} = \boldsymbol{0} \Rightarrow$$
$$\boldsymbol{x} = \left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\right)^{-1}\left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{b}-\boldsymbol{v}\right). \tag{43}$$

By substituting (42) and (43) into (41), we can derive the objective function as follows:

$$\underset{\|\boldsymbol{v}\|_\infty \leqslant 2}{\text{maximize}} \; \frac{1}{2}\left\|\boldsymbol{A}\left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\right)^{-1}\left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{b}-\boldsymbol{v}\right)-\boldsymbol{b}\right\|_2^2 +$$
$$\boldsymbol{v}^{\mathrm{H}}\left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\right)^{-1}\left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{b}-\boldsymbol{v}\right). \tag{44}$$

Expand (44) and drop the constant terms, and then we can conclude the dual problem as follows

$$\underset{\|\boldsymbol{v}\|_\infty \leqslant 2}{\text{minimize}} \; \frac{1}{2}\boldsymbol{v}^{\mathrm{H}}\left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\right)^{-1}\boldsymbol{v} - \Re\left\{\boldsymbol{v}^{\mathrm{H}}\left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{b}\right\}. \tag{45}$$

Multiply the objective function in (45) by 2 and add a constant term $\boldsymbol{b}^{\mathrm{H}}\boldsymbol{A}\left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\right)^{-2}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{b}$, we can recast the dual problem as

$$\underset{\boldsymbol{v}}{\text{minimize}} \quad \|\boldsymbol{f}-\boldsymbol{D}\boldsymbol{v}\|_2^2 \tag{46}$$
$$\text{subject to} \quad |\boldsymbol{v}(i)| \leqslant 2.$$

where $\boldsymbol{D}^{\mathrm{H}}\boldsymbol{D} = \left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\right)^{-1}$ and $\left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{b} = \boldsymbol{D}^{\mathrm{H}}\boldsymbol{f}$. By defining $\boldsymbol{f} = \boldsymbol{f}_{\text{opt}}$, $\boldsymbol{v} = \boldsymbol{f}_{\text{RF}}$, $\boldsymbol{D}^{\mathrm{H}}\boldsymbol{D}$, and $\boldsymbol{D} = \boldsymbol{F}_{\text{BB}}^{\mathrm{T}} \otimes \boldsymbol{I}_{N_t}$, and inversely applying the matrix vectorization, it completes the proof of Lemma 1.

## REFERENCES

[1] X. Yu, J. Zhang, K. B. Letaief. Alternating minimization for hybrid precoding in multiuser OFDM mmWave systems [C]//Proc. Asilomar Conf. Signals Syst. Comput., Pacific Grove, CA, USA, 2016: 281-285.

[2] X. Yu, J. Zhang, K. B. Letaief. Partially-connected hybrid precoding in mmWave systems with dynamic phase shifter networks [C]//IEEE Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC), Sapporo, Japan, 2017: 1-5.

[3] J. G. Andrews, S. Buzzi, W. Choi, et al. What will 5G be [J]. IEEE J. Sel. Areas Commun., 2014, 32(6): 1065-1082.

[4] T. S. Rappaport, S. Sun, R. Mayzus, et al. Millimeter wave mobile communications for 5G cellular: It will work! [J]. IEEE Access, 2013(1): 335-349.

[5] T. S. Rappaport, R. W. Heath Jr., R. C. Daniels, et al. Millimeter wave wireless communications [M]. New York, NY, USA: Pearson Education, 2014.

[6] X. Zhang, A. F. Molisch, S. Y. Kung. Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection [J]. IEEE Trans. Signal Process., 2005, 53(11): 4091-4103.

[7] O. El Ayach, S. Rajagopal, S. Abu-Surra, et al. Spatially sparse precoding in millimeter wave MIMO systems [J]. IEEE Trans. Wireless Commun., 2014, 13(3): 1499-1513.

[8] C. Rusu, R. Mèndez-Rial, N. González-Prelcic, et al. Low complexity hybrid precoding strategies for millimeter wave communication systems [J]. IEEE Trans. Wireless Commun., 2016, 15(12): 8380-8393.

[9] X. Yu, J. C. Shen, J. Zhang. Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems [J]. IEEE J. Sel. Topics Signal Process., 2016, 10(3): 485-500.

[10] X. Yu, J. Zhang, K. B. Letaief. A hardware-efficient analog network structure for hybrid precoding in millimeter wave systems [J]. IEEE J. Sel. Topics Signal Process., 2018, 12(2): 282-297.

[11] F. Sohrabi, W. Yu. Hybrid digital and analog beamforming design for large-scale antenna arrays [J]. IEEE J. Sel. Topics Signal Process., 2016, 10(3): 501-513.

[12] M. Kim, Y. H. Lee. MSE-based hybrid RF/baseband processing for millimeter-wave communication systems in MIMO interference channels [J]. IEEE Trans. Veh. Technol., 2015, 64(6): 2714-2720.

[13] J. Lee, Y. H. Lee. AF relaying for millimeter wave communication systems with hybrid RF/baseband MIMO processing [C]//Proc. IEEE Int. Conf. Commun. (ICC), Sydney, NSW, Australia, 2014: 5838-5842.

[14] A. Alkhateeb, R. W. Heath Jr. Frequency selective hybrid precoding for limited feedback millimeter wave systems [J]. IEEE Trans. Commun., 2016, 64(5): 1801-1818.

[15] W. Ni, X. Dong. Hybrid block diagonalization for massive multiuser MIMO systems [J]. IEEE Trans. Commun, 2016, 64(1): 201-211.

[16] T. E. Bogale, L. B. Le. Beamforming for multiuser massive MIMO systems: Digital versus hybrid analog-digital [C]//Proc. IEEE Global Commun. Conf. (GLOBECOM), Austin, TX, USA, 2014: 4066-4071.

[17] L. Liang, W. Xu, X. Dong. Low-complexity hybrid precoding in massive multiuser MIMO systems [J]. IEEE Wireless Commun. Lett., 2014, 3(6): 653-656.

[18] F. Sohrabi, W. Yu. Hybrid analog and digital beamforming for mmWave OFDM large-scale antenna arrays [J]. IEEE J Sel. Areas Commun., 2017, 35(7): 1432-1443.

[19] E. Zhang, C. Huang. On achieving optimal rate of digital precoder by RF-baseband codesign for MIMO systems [C]//Proc. 80th IEEE Veh. Technol. Conf. (VTC Fall), Vancouver, BC, 2014: 1-5.

[20] T. E. Bogale, L. B. Le, A. Haghighat, et al. On the number of RF chains and phase shifters, and scheduling design with hybrid analog-digital beamforming [J]. IEEE Trans. Wireless Commun., 2016, 15(5): 3311-3326.

[21] J. Singh, S. Ramakrishna. On the feasibility of codebook-based beamforming in millimeter wave systems with multiple antenna arrays [J]. IEEE Trans. Wireless Commun., 2015, 14(5): 2670-2683.

[22] C. Kim, T. Kim, J. Y. Seol. Multi-beam transmission diversity with hybrid beamforming for MIMO-OFDM systems [C]//Proc. IEEE Globecom Workshops (GLOBECOM Wkshps), Atlanta, GA, USA, 2013: 61-65.

[23] X. Gao, L. Dai, S. Han, et al. Energy-efficient hybrid analog and digital precoding for mmWave MIMO systems with large antenna arrays [J]. IEEE J. Sel. Areas Commun., 2016, 34(4): 998-1009.

[24] Y. P. Lin. On the quantization of phase shifters for hybrid precoding systems [J]. IEEE Trans. Signal Process., 2017, 65(9): 2237-2246.

[25] Q. Shi, M. Razaviyayn, Z. Q. Luo, et al. An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel [J]. IEEE Trans. Signal Process., 2011, 59(9): 4331-4340.

[26] Q. H. Spencer, A. L. Swindlehurst, M. Haardt. Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels [J]. IEEE Trans. Signal Process., 2004, 52(2): 461-471.

[27] A. Liu, V. Lau. Phase only RF precoding for massive MIMO systems with limited RF chains [J]. IEEE Trans. Signal Process., 2014, 62(17): 4505-4515.

[28] T. Hastie, R. Tibshirani, M. Wainwright. Statistical learning with sparsity: The lasso and generalizations [M]. Boca Raton, FL, USA: CRC Press, 2015.

[29] R. Mèndez-Rial, C. Rusu, N. González-Prelcic, et al. Hybrid MIMO architectures for millimeter wave communications: Phase shifters or switches [J]. IEEE Access, 2016(4): 247-267.

[30] L. Grippo, M. Sciandrone. On the convergence of the block nonlinear gauss-seidel method under convex constraints [J]. Operat. Res. Lett., 2000, 26(3): 127-136.

## About the Authors

**Xianghao Yu** received his Ph.D. degree from the Hong Kong University of Science and Technology in 2018. He is currently a postdoc fellow with Humboldt Fellowship at the Friedrich-Alexander-Universität Erlangen-Nürnberg. He was the recipient of the IEEE GLOBECOM 2017 Best Paper Award and the 2018 IEEE Signal Processing Society Young Author Best Paper Award. He is a co-author of the book Stochastic Geometry Analysis of Multi-Antenna Wireless Networks (Springer, 2019). His research interests include millimeter wave communications, stochastic geometry, and large intelligent surface-assisted communications.

**Jun Zhang** [corresponding author] received his B.Eng. degree in electronic engineering from University of Science and Technology of China in 2004, his M.Phil. degree in information engineering from the Chinese University of Hong Kong in 2006, and his Ph.D. degree in electrical and computer engineering from the University of Texas at Austin in 2009. He is currently an assistant professor in the Department of Electronic and Information Engineering at the Hong Kong Polytechnic University (PolyU), and previously was a research assistant professor in the Department of Electronic and Computer Engineering at the Hong Kong University of Science and Technology (HKUST). His research interests include wireless communications and networking, mobile edge computing and edge learning, distributed learning and optimization, and big data analytics.

Dr. Zhang co-authored the books Fundamentals of LTE (Prentice-Hall, 2010) and Stochastic Geometry Analysis of Multi-Antenna Wireless Networks (Springer, 2019). He is a co-recipient of the 2016 Marconi Prize Paper Award in Wireless Communications, the 2014 Best Paper Award for the EURASIP Journal on Advances in Signal Processing, an IEEE GLOBECOM Best Paper Award in 2017, an IEEE ICC Best Paper Award in 2016, and an IEEE PIMRC Best Paper Award in 2014. Two papers he co-authored received the Young Author Best Paper Award of the IEEE Signal Processing Society in 2016 and 2018, respectively. He also received the 2016 IEEE ComSoc Asia-Pacific Best Young Researcher Award. He is an editor of IEEE Transactions on Wireless Communications and Journal of Communications and Information Networks, and was a guest editor of the special section on "Mobile Edge Computing for Wireless Networks" in IEEE Access. He also served as a MAC track TPC co-chair for IEEE WCNC 2011.

**Khaled B. Letaief** is an internationally recognized leader in wireless communications and networks with research interest in green communications, Internet of things, big data analytics systems, mobile cloud and edge computing, tactile Internet, 5G systems and beyond. In these areas, he has over 600 journal and conference papers and given keynote talks as well as courses all over the world. He also has 15 patents, including 11 US patents.

He served as consultants for different organizations including Huawei, ASTRI, ZTE, Nortel, PricewaterhouseCoopers, and Motorola. He is the founding editor-in-chief of the prestigious IEEE Transactions on Wireless Communications and has served on the editorial board of other premier journals including the IEEE Journal on Selected Areas in Communications-Wireless Series (as editor-in-chief). He has been involved in organizing many flagship international conferences. These include serving as GLOBECOM'18 Executive Vice-Chair, ICCC'12 General Chair, TTM'11 General Co-Chair, ICC'08 Technical Program Co-Chair, and WCNC'07 General Co-Chair.

Professor Letaief has been a dedicated teacher committed to excellence in teaching and scholarship. He received the Mangoon Teaching Award from Purdue University; Teaching Excellence Appreciation Award by the School of Engineering at HKUST (4 times); and the Michael G. Gale Medal for Distinguished Teaching (Highest university-wide teaching award and only one recipient/year is honored for his/her contributions).

He is the recipient of many other distinguished awards and honors including the 2017 IEEE Cognitive Networks Technical Committee Publication Award, 2016 IEEE Signal Processing Society Young Author Best Paper Award, 2016 IEEE Marconi Prize Paper Award in Wireless Communications, 2011 IEEE Wireless Communications Technical Committee Recognition Award, 2011 IEEE Communications Society Harold Sobol Award, 2010 Purdue University Outstanding Electrical and Computer Engineer Award, 2009 IEEE Marconi Prize Award in Wireless Communications, 2007 IEEE Communications Society Joseph LoCicero Publications Exemplary Award, and over 15 IEEE Best Paper Awards. From 1990 to 1993, he was a faculty member at the University of Melbourne, Australia. Since 1993, he has been with the Hong Kong University of Science & Technology (HKUST) where he

is known as one of HKUST's most distinguished professors for his boundless energy, collegial nature, dedication, and excellence in research, education, and service. While at HKUST, he has held many administrative positions, including the Head of the Electronic and Computer Engineering department, Director of the Wireless IC Design Center, founding Director of Huawei Innovation Laboratory, and Director of the Hong Kong Telecom Institute of Information Technology. While at HKUST he has also served as Chair Professor and Dean of Engineering. Under his leadership, the School of Engineering has not only transformed its education and scope and produced very high caliber scholarship, it has also actively pursued knowledge transfer and societal engagement in broad contexts. It has also dazzled in international rankings (rising from # 26 in 2009 to # 14 in the world in 2015 according to QS World University Rankings.) From September 2015 to March 2018, he joined HBKU as Provost to help establish a research-intensive university in Qatar in partnership with strategic partners that include Northwestern University, Carnegie Mellon University, Cornell, and Texas A&M. While at HBKU, he led the University academic strategic planning that emphasizes research, excellence, innovation, and entrepreneurship. He also led the University in

hiring a talented pool of faculty, attracting outstanding students, as well as developing research integrated and interdisciplinary academic programs. Dr. Letaief is well recognized for his dedicated service to professional societies and in particular IEEE where he has served in many leadership positions. These include Treasurer of IEEE Communications Society, IEEE Communications Society Vice-President for Conferences, Chair of IEEE Committee on Wireless Communications, elected member of IEEE Product Services and Publications Board, and IEEE Communications Society Vice-President for Technical Activities. He is currently President of the IEEE Communications Society, the world's leading organization for communications professionals with headquarter in New York City and members in 162 countries.

Dr. Letaief received his B.S. degree with distinction in electrical engineering from Purdue University at West Lafayette, Indiana, USA, in December 1984. He received his M.S. and Ph.D. degrees in electrical engineering from Purdue University, in August 1986, and May 1990, respectively. Dr. Letaief is a fellow of IEEE and a fellow of HKIE. He is also recognized by Thomson Reuters as an ISI Highly Cited Researcher.