Hybrid Beamforming for 5G and Beyond Millimeter-Wave Systems: A Holistic View

JUN ZHANG [ID] [1] (Senior Member, IEEE), XIANGHAO YU [ID] [2] (Member, IEEE), AND KHALED B. LETAIEF[3,4] (Fellow, IEEE)

(Invited Paper)

[1] Department of Electronic and Information Engineering, Hong Kong Polytechnic University (PolyU), Hong Kong

[2] Institute for Digital Communications, Friedrich–Alexander-University Erlangen–Nurnberg (FAU), 91054 Erlangen, Germany

[3] Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology (HKUST), Hong Kong

[4] Peng Cheng Laboratory, Shenzhen 518066, China

CORRESPONDING AUTHOR: X. YU (e-mail: xianghao.yu@fau.de)

ABSTRACT Millimeter-wave (mm-wave) communication is a key technology for future wireless networks. To combat significant path loss and exploit the abundant mm-wave spectrum, effective beamforming is crucial. Nevertheless, conventional fully digital beamforming techniques are inapplicable, as they demand a separate radio frequency (RF) chain for each antenna element, which is costly and consumes too much energy. Hybrid beamforming is a cost-effective alternative, which can significantly reduce the hardware cost and power consumption by employing a small number of RF chains. This paper presents a holistic view on hybrid beamforming for 5G and beyond mm-wave systems, based on a new taxonomy for different hardware structures. We take a pragmatic approach and compare different proposals from three key aspects: 1) *hardware efficiency*, i.e., the required hardware components; 2) *computational efficiency* of the associated beamforming algorithm; and 3) achievable *spectral efficiency*, a main performance indicator. Through systematic comparisons, the interplay and trade-off among these three design aspects are demonstrated, and promising candidates for hybrid beamforming in future wireless networks are identified.

INDEX TERMS Hybrid beamforming, millimeter-wave communications, 5G and beyond, wireless communications.

I. INTRODUCTION

THE CONTINUED upsurge of mobile data and the eruption of diversified mobile applications are driving the demand for next-generation wireless networks, i.e., the fifth-generation (5G) networks. Compared with the current fourth-generation (4G) Long Term Evolution (LTE) networks [1], 5G needs to achieve orders of magnitude increase in the peak data rate, area spectral efficiency, network energy efficiency, while supporting a roundtrip latency of about 1 ms [2]. Thus, disruptive technologies will be needed, and deploying 5G systems at millimeter wave (mm-wave) bands has been proposed due to the abundant spectrum. Thanks to the small wavelength of the mm-wave signals, large-scale antenna arrays can be deployed, and

recent advances in massive multiple-input multiple-output (MIMO) [3] can be leveraged to provide beamforming gains to combat the increased path loss and synthesize highly directional beams to support mm-wave communications [4].

To deploy mm-wave systems with large-scale antenna arrays, challenges in hardware implementation and algorithm design need to be addressed. A large number of hardware components will be needed to support conventional digital beamforming, including signal mixers, analog-to-digital/digital-to-analog converters (ADCs/DACs), and power amplifiers [5]. This will put prohibitive burdens on cost and power consumption, especially for mobile terminals, and thus is not feasible. Furthermore, the significantly increased dimension of the beamformers brings stringent

requirements on computational efficiency of beamforming algorithms. These challenges have driven the recent efforts in developing hardware efficient transceivers, supported with efficient beamforming algorithms. One initial proposal is *analog beamforming*, supported with a phase shifter network and low-complexity beam steering. It is currently the de-facto approach for indoor mm-wave systems [6]. However, analog beamforming only supports single-stream transmission, and cannot fully exploit the available spatial resource. To further improve performance, *hybrid beamforming* has been proposed as a cost-effective approach to support spatial multiplexing with a limited number of radio frequency (RF) chains, whose potential has been demonstrated in many recent studies [7], [8]. In particular, compared with analog beamforming, hybrid beamforming supports multi-stream transmission with spatial multiplexing, as well as spatial division multiple access. It achieves spectral efficiency comparable to fully digital beamforming with much reduced hardware complexity. Therefore, it has been regarded as a promising candidate for transceiver structures in mm-wave systems.

The concept of hybrid beamforming can be traced back to early 2000s [9], [10], where the point-to-point single-stream transmission in sub-6 GHz systems was investigated as a special case. Almost a decade later, hybrid beamforming was revisited in mm-wave systems [11] and has drawn considerable attention from both academia and industry. By leveraging the sparsity of mm-wave channels, low-complexity algorithms were proposed for point-to-point hybrid beamforming [11], whose achievable spectral efficiency was further improved in [12]. Then, hybrid beamforming was extended to single-user multicarrier [13]–[15] and multiuser single-carrier systems [16]–[19]. The main challenge of hybrid beamforming design is to optimize the system performance under the hardware constraints, e.g., reduced RF chains and the high-dimensional phase shifter-based analog beamformer. Various algorithms were developed to combat this difficulty, e.g., compressive sensing [11], [20], codebook-based design [14], [21], and manifold optimization [15], [22], which have offered effective design methodologies for hybrid beamforming.

Nevertheless, hybrid beamforming is still facing several critical issues that may hinder its practical applicability. Compared with fully digital beamforming, hardware complexity has been significantly reduced, but it is still quite a concern, especially considering the cost and power consumption of mm-wave devices [23]. Thus, hybrid beamforming structures that are more hardware-efficient should be developed. For this aspect, we can learn little from conventional digital beamforming design, which takes a performance-oriented perspective, e.g., to maximize spectral efficiency or minimize transmit power, but largely neglects hardware complexity. Furthermore, digital beamforming problems are typically convex, and powerful tools from convex optimization can be leveraged [24]. However, hybrid beamforming problems are innately non-convex and challenging to design.

To address these design challenges for hybrid beamforming, a holistic approach should be taken. In particular, we need a comprehensive consideration that accounts for the following three decisive aspects: *hardware efficiency*, *computational efficiency*, and *spectral efficiency*. Accordingly, this paper presents key proposals of hybrid beamforming structures, emphasizing the following three desirable properties:

1) *High hardware efficiency (HE)*, i.e., with as few hardware components as possible, which leads to low cost and low power consumption.
2) *High spectral efficiency (SE)*, which should be close to that of the fully digital beamforming.
3) *High computational efficiency (CE)*, i.e., the hybrid beamforming algorithm should be of low complexity.

A special emphasis is placed on the interplay between hardware structures and beamforming algorithm design. Answers to the following key questions will be revealed through the discussion:

- *How many RF chains and phase shifters are needed?*
- *Can hybrid beamforming approach the performance of fully digital beamforming?*
- *How to effectively design hybrid beamforming algorithms?*

Specifically, we first present the state-of-the-art hybrid beamforming structures, as well as their algorithm design. Limitations of these basic structures are identified. Then, we introduce two new analog network implementations, which greatly simplify algorithm design and reduce hardware complexity, respectively. Finally, we propose a flexible mapping strategy for hybrid beamforming, which helps to strike a good balance between the hardware complexity and spectral efficiency. The paper ends with key conclusions and some future research directions.

Notations: The following notations are used throughout this paper. $j = \sqrt{-1}$ is the imaginary unit; $\mathbb{C}$ and $\mathbb{Z}$ denote the sets of complex numbers and integer numbers; $\mathbf{a}$ and $\mathbf{A}$ stand for a column vector and a matrix, respectively; The i-th row, the j-th column, and the (i, j)-th entry in matrix $\mathbf{A}$ are denoted as $\mathbf{A}(i, :)$, $\mathbf{A}(:, j)$, and $\mathbf{A}(i, j)$, respectively; The conjugate, transpose and conjugate transpose of $\mathbf{A}$ are represented by $\mathbf{A}^*$, $\mathbf{A}^T$ and $\mathbf{A}^H$; $\|\mathbf{a}\|_0$ stands for the ℓ_0-norm of vector $\mathbf{a}$; $\text{blkdiag}(\mathbf{A}_1, \ldots, \mathbf{A}_i)$ establishes a block diagonal matrix using $\mathbf{A}_1, \ldots, \mathbf{A}_i$ as its diagonal terms.

II. A PRIMER ON HYBRID BEAMFORMING

A hybrid beamforming transceiver is depicted in Fig. 1. We consider the downlink transmission of a multiuser mm-wave MIMO-OFDM (orthogonal frequency-division multiplexing) system. A base station (BS) leverages an N_t-size antenna array to serve K users over F subcarriers. The BS transmits N_s data streams to each user on each subcarrier. The number
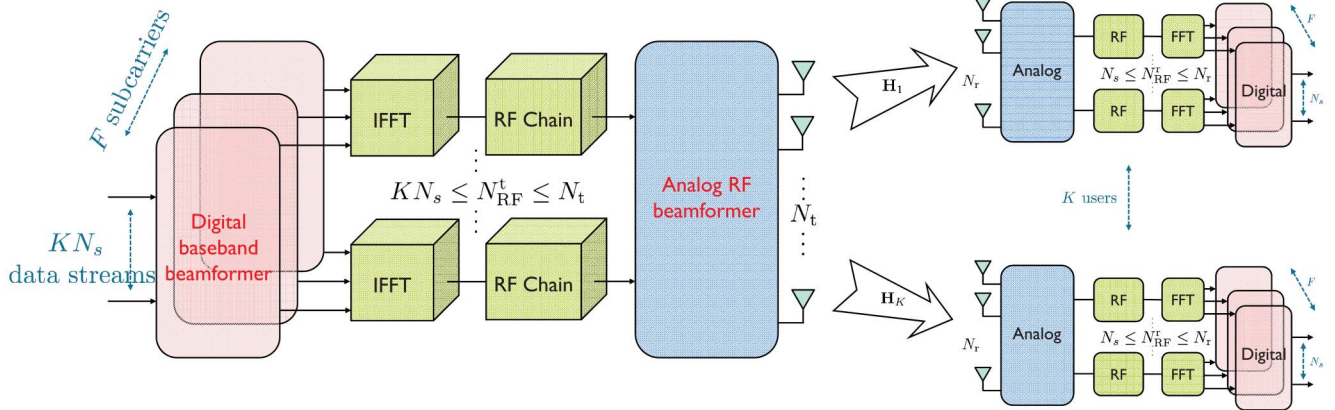
FIGURE 1. Hybrid beamforming in a general multiuser multicarrier system. The base station (BS), on the left hand side of the figure, is equipped with N_t antennas to serve K N_r-antenna users over F subcarriers. In addition, N_S data streams are transmitted to each user on each subcarrier. The available numbers of RF chains are N_{RF}^t and N_{RF}^r for the BS and each user, respectively.

of available RF chains at the BS is N_{RF}^t, which is restricted as $KN_s \leq N_{RF}^t < N_t$.[1]

The hybrid beamformer consists of two components: a digital component and an analog component. The digital part is composed of RF chains, whose structure is common for different proposals to be discussed. Similar to the conventional fully digital beamforming, the digital component in hybrid beamforming can be performed for each user on each subcarrier, denoted as $\mathbf{F}_{BBk,f} \in \mathbb{C}^{N_{RF}^t \times N_s}$. However, this is not the case for the analog component, or *the analog network*, in hybrid beamforming. Since the transmitted signals for all the users are mixed together by the digital beamformers, and analog RF beamforming is a post-IFFT (inverse fast Fourier transform) operation, the analog network $\mathbf{F}_{RF} \in \mathbb{C}^{N_t \times N_{RF}^t}$ is a common component shared by all the users and subcarriers.

Furthermore, as will be revealed in this paper, the analog network is the key differentiating compoent in different hybrid beamforming structures. In particular, the structure of the analog network not only influences the hardware efficiency, but also has a significant impact on both the algorithmic design and achievable spectral efficiency. Hence, our discussion mainly focuses on the analog network. In this section, we first introduce key hardware components, and then introduce a new taxonomy for comparing different hybrid beamforming structures.

A. KEY HARDWARE COMPONENTS

Hardware efficiency is a key consideration when designing hybrid beamforming structures, and we compare different structures by the number of required key components. Note that, given the rapid advances in hardware and diversified choices, it is difficult to make a fair comparison for energy efficiency, which, nevertheless, will be largely determined by hardware efficiency. Therefore, we do not explicitly consider energy efficiency in this paper.

1. The settings for hybrid beamforming at the user side can be defined in a similar way as those at the BS side, which are omitted here to keep the presentation clear and concise.

In the analog RF domain, key hardware components include power amplifiers, phase shifters, and switches. Power amplifiers, as basic components in conventional fully digital beamforming, are needed for each antenna element, and great attention has been drawn on realizing low power amplifiers in integrated circuit (IC) design. In contrast, phase shifters, originally utilized in military radar systems, are the newly-introduced hardware components in hybrid beamforming systems. Hardware suppliers are not yet ready to provide phase shifters for commercial use, and the cost of phase shifters is currently very high, e.g., it can be around a hundred U.S. dollars even with low resolution.[2] It motivates alternative structures to replace phase shifters with other components or to reduce the number of phase shifters. For example, Méndez-Rial *et al.* [25] proposed to replace phase shifters with switches to reduce the hardware complexity. Other proposals will be discussed later in the paper.

As power amplifiers are necessary and cannot be easily replaced, the hardware efficiency of the analog network primarily depends on phase shifters and/or switches. As a matter of fact, switches entail only binary states and therefore outperform phase shifters in terms of implementation complexity, power consumption, and cost. However, limiting to the on-off state will inevitably incur performance loss in spectral efficiency. Later we will show how to combine phase shifters and switches to develop hardware-efficient hybrid beamforming structures with good spectral efficiency.
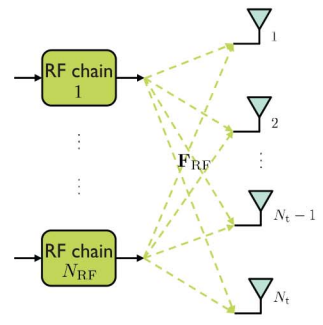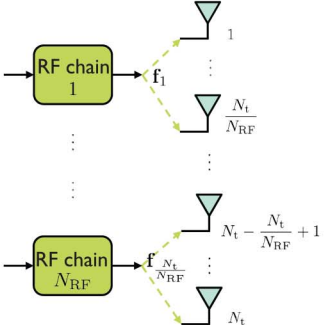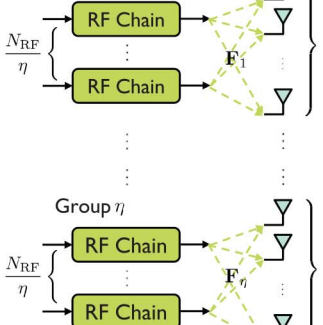
B. A TAXONOMY OF HYBRID BEAMFORMING STRUCTURES

Hybrid beamforming structures differ mainly in the way they use the above-mentioned hardware components to compose the analog network. In particular, the analog network structure is primarily determined by two elements, i.e., the *mapping strategy* and *hardware implementation*, for which different proposals are listed in Tables 1 (a) and (b).
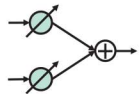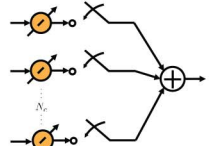
2. http://www.analog.com/en/parametricsearch/10700#

TABLE 1. *Mapping strategies* and *hardware implementations* for hybrid beamforming, with $\mathbf{F}_{\mathbf{RF}}$ as the analog beamforming matrix, and $\mathbf{f}_i$ and $\mathbf{F}_i$ denoting a column vector and a matrix, respectively. To realize a specific analog network structure, one may pick a mapping strategy from (a) to decide how to connect the RF chain-antenna pairs. Then, one should choose a hardware implementation from (b) to realize each RF chain-antenna pair.

(a) Mapping strategies for hybrid beamforming.

	Fully-Connected	Partially-Connected	Group-Connected
Mapping strategy			
Analog beamforming matrix $\mathbf{F}_{\mathrm{RF}}$	Full matrix $\begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1N_{\mathrm{RF}}^{\mathrm{t}}} \\ f_{21} & f_{22} & & \\ \vdots & & \ddots & \\ f_{N_{\mathrm{t}}1} & & & f_{N_{\mathrm{t}}N_{\mathrm{RF}}^{\mathrm{t}}} \end{bmatrix}$	Block diagonal matrix $\begin{bmatrix} \mathbf{f}_1 & & & \\ & \mathbf{f}_2 & & \\ & & \ddots & \\ & & & \mathbf{f}_{\frac{N_{\mathrm{t}}}{N_{\mathrm{RF}}^{\mathrm{t}}}} \end{bmatrix}$	Block diagonal matrix $\begin{bmatrix} \mathbf{F}_1 & & & \\ & \mathbf{F}_2 & & \\ & & \ddots & \\ & & & \mathbf{F}_{\eta} \end{bmatrix}$

(b) Hardware implementations for hybrid beamforming.

	Single Phase Shifter (SPS)	Double Phase Shifter (DPS)	Fixed Phase Shifter (FPS)
Hardware implementation			
Constraint on $\mathbf{F}_{\mathrm{RF}}$	$\lvert\mathbf{F}_{\mathrm{RF}}(i,j)\rvert = 1$	$\lvert\mathbf{F}_{\mathrm{RF}}(i,j)\rvert \leq 2$	$\mathbf{F}_{\mathrm{RF}}(i,j) = \sum_{n=1}^{N_c} b_n e^{j\theta_n},$ $b_n \in \{0,1\}$
Comments	One phase shifter for each RF chain-antenna pair	Two phase shifters for each RF chain-antenna pair	N_c fixed phase shifters shared by all RF chain-antenna pairs

- *The Mapping Strategy:* It determines how the RF chains and antenna elements are connected. As shown in Table 1 (a), there are two basic mapping strategies, namely, the fully- and partially-connected mapping, which will be introduced in Section III. A more flexible mapping strategy, group-connected mapping, will be introduced in Section V. Table 1 (a) further shows the analog beamforming matrix associated with each mapping strategy, which bears a special structure that will affect the beamformer design.

- *The Hardware Implementation:* It specifies the adopted hardware components and the way each RF chain-antenna pair is connected. Among the three implementations shown in Table 1 (b), the single phase shifter (SPS) implementation is the most commonly used one, and the other two, double phase shifter (DPS) and fixed phase shifter (FPS) implementations, are recently proposed and will be introduced in later sections. Different hardware implementations will induce different constraints on $\mathbf{F}_{\mathrm{RF}}$, as shown in the table, which will significantly affect the algorithm design.

As a common example, the SPS fully-connected structure refers to the fully-connected mapping strategy with a single phase shifter connecting each RF chain with a corresponding antenna.

III. BASIC HYBRID BEAMFORMING STRUCTURES

In this section, we present and compare two basic mapping strategies, namely, the *fully-* and *partially-connected* ones. As shown in Table 1 (a), in the fully-connected mapping strategy, every RF chain is connected to all the antenna

elements, while each RF chain is connected to a subset of neighboring antenna elements that do not overlap with each other in the partially-connected mapping strategy. For hardware implementation, we consider the classic SPS implementation, i.e., each connected RF chain-antenna pair is linked via a single phase shifter. Therefore, in terms of hardware efficiency, the SPS fully- and partially-connected ones employ $N_t N_{RF}^t$ and N_t phase shifters, respectively. Through the following comparison of these two basic structures, we shall illustrate their limitations and motivate other proposals.

A. BASIC PRINCIPLES OF HYBRID BEAMFORMING ALGORITHM DESIGN

In this part, we present a basic formulation for hybrid beamforming design, accompanied with some design principles. A common design principle is to approximate the fully digital beamformer subject to the constraint for the analog beamforming matrix [11], [15], whose formulation is correspondingly given by

$$
\begin{aligned}
\underset{\mathbf{F}_{RF}, \mathbf{F}_{BB}}{\text{minimize}} \quad & \left\| \mathbf{F}_{opt} - \mathbf{F}_{RF} \mathbf{F}_{BB} \right\|_F \\
\text{subject to} \quad & \begin{cases} \left\| \mathbf{F}_{RF} \mathbf{F}_{BB} \right\|_F^2 \leq KN_s F \\ \mathbf{F}_{RF} \in \mathcal{A}, \end{cases}
\end{aligned} \tag{1}
$$

where the combined fully digital beamformer is denoted as $\mathbf{F}_{opt} = [\mathbf{F}_{opt_{1,1}}, \ldots, \mathbf{F}_{opt_{k,f}}, \ldots, \mathbf{F}_{opt_{K,F}}] \in \mathbb{C}^{N_t \times KN_s F}$, and $\mathbf{F}_{BB} = [\mathbf{F}_{BB1,1}, \ldots, \mathbf{F}_{BBk,f}, \ldots, \mathbf{F}_{BBK,F}]$ is the concatenated digital beamformer with dimension $N_{RF}^t \times KN_s F$. The first constraint in the formulation is the total transmit power constraint, and the second constraint depends on the adopted hardware implementation, as shown in Table 1. The main merits of this formulation include its general applicability, i.e., it can be applied with any given digital beamformer, and the tractability for algorithm design, to be illustrated below.

Critical Role of the Analog Network: In the hybrid beamforming design problem (1), $\mathcal{A}$ is the feasible set of the analog network, which is distinct for different hybrid beamforming structures. Before we proceed, we would like to emphasize the critical role of the analog network. As we discussed before, the analog network is shared by all the users and subcarriers, so a single analog beamforming matrix should match the channel states of different users on different subcarriers. This is an extremely difficult task, and it is not clear at all how close we can approach the performance of the fully digital beamforming with hybrid beamforming. With such a decisive role on achievable performance, the analog network calls for a delicate design. Moreover, different implementations of the analog network bring different constraints for the analog beamforming matrix, and thus determine the difficulty in beamforming algorithm design. Both of these aspects will be elaborated throughout the discussion in this paper.

As there are two components in a hybrid beamformer, i.e., an analog one and a digital one, *alternating minimization*

(AltMin) serves as a basic design principle [15]. It alternately optimizes the analog and digital parts. It is apparent that the optimization of the digital beamforming matrix $\mathbf{F}_{BB}$ is a least squares problem, which has a closed-form solution. On the other hand, with the SPS implementation, the main difficulty lies in the analog component, for which there is a non-convex unit modulus constraint. In particular, the feasible set $\mathcal{A}$ of the analog network $\mathbf{F}_{RF}$ can be specified by a set of matrices where the amplitude of each non-zero element is forced to be 1, i.e., $|\mathbf{F}_{RF}(i, j)| = 1$ [11]. Design methodologies for the two basic structures are different, as presented in the following two subsections.

B. SPS FULLY-CONNECTED STRUCTURE

Note that when $N_{RF}^t \geq 2KN_s$, the fully digital beamforming can be realized by the SPS fully-connected structure [10], [26], and this case is trivial in terms of algorithm design. For the general case when $N_{RF}^t < 2KN_s$, the orthogonal matching pursuit (OMP) algorithm [11] is the most widely-used algorithm, which treats the analog network design as a sparsity constrained matrix reconstruction problem. In particular, the columns of the analog beamforming matrix $\mathbf{F}_{RF}$ are selected from a candidate set, which typically consists of the array response vectors of mm-wave channels. This codebook-based design inevitably incurs some performance loss when approaching the fully digital beamforming. More recent attention focused on reducing the computational complexity of the OMP algorithm, e.g., by reusing the matrix inversion result in each iteration [20].

In [15], by recognizing that the unit modulus constraints of the analog network define a complex circle Riemannian manifold, a manifold optimization based AltMin (MO-AltMin) algorithm was proposed, which outperforms the OMP algorithm but with increased complexity. In particular, by defining key elements, e.g., inner products and gradients, in the neighborhood of a manifold that is homeomorphic to the Euclidean space, a variety of classic optimization algorithms in the Euclidean space can be transplanted to manifold optimization. For instance, the conjugate gradient method in the Euclidean space was adopted on the complex circle manifold for hybrid beamforming in [12], [15].

As introduced above, the OMP algorithm updates a column of the analog beamforming matrix $\mathbf{F}_{RF}$ at a time while the MO-AltMin algorithm optimizes the whole $\mathbf{F}_{RF}$ matrix in each iteration. To the other extreme, the phase shifters are optimized one by one in [27]. In particular, the contribution of each phase shifter to the spectral efficiency was analytically identified, based on which the analog network was iteratively optimized in a phase shifter-by-phase shifter fashion. This approach also suffers a high complexity since the number of iterations of the algorithm is proportional to the number of phase shifters in use, which is typically a huge number ($N_t N_{RF}^t$) in mm-wave MIMO systems with the SPS fully-connected structure.

TABLE 2. Existing works with the SPS implementation.

Hybrid beamforming structure	Number of phase shifters	Design Approach	Spectral efficiency	Computational complexity
SPS fully-connected	$N_t N_{RF}^t$	OMP [11]	✓✓✓	$\mathcal{O}\left(N_t N_{RF}^t \left(L N_s + N_{RF}^{t\,2} + 2 N_{RF}^t N_s\right)\right)^{\dagger}$
		MO-AltMin [15]	✓✓✓✓	Extremely high
		Element-wise [27]	✓✓✓	$\mathcal{O}\left(N_{iter} N_t^4 N_{RF}^t\right)$
SPS partially-connected	N_t	SIC-based [28]	✓	$\mathcal{O}\left(\frac{N_t^2}{N_{RF}^{t\,2}}\left(N_{RF}^t N_{iter} + N_r\right) + 2 N_{RF}^t N_{iter}\right)$
		SDR-AltMin [15]	✓✓	$\mathcal{O}\left(N_{iter} N_{RF}^{t\,3} N_s^3\right)$

† L is the total number of paths in the channels.
* More ✓ means higher spectral efficiency and N_{iter} denotes the number of iterations involved in the algorithm. The computational complexity is evaluated for single-user single-carrier systems for fair comparison. As there are nested iterations in the MO-AltMin algorithm, its computational complexity is much higher than those of other algorithms.

C. SPS PARTIALLY-CONNECTED STRUCTURE

While most initial efforts on hybrid beamformer design were on the SPS fully-connected structure, the SPS partially-connected one has attracted more recent attention due to its low hardware complexity. In the analog RF domain, the hardware complexity of the SPS partially-connected structure is the same as that of analog beamforming, as the numbers of phase shifters are both equal to the antenna size. In [29], [30], codebook-based design of hybrid beamformers was presented for narrowband and OFDM systems, respectively. Although the codebook-based design enjoys a low complexity, there will be certain performance loss, and it is not clear how much performance gain can be further obtained. Another proposal is based on the concept of successive interference cancellation (SIC) [28]. It decomposes the total achievable rate optimization problem into a series of simple sub-rate optimization problems, each of which only considers the antenna elements connected to one RF chain. However, this approach enforces that the digital beamforming matrix is diagonal, and the number of RF chains should be equal to that of the data streams.

More recently, a semidefinite relaxation based AltMin (SDR-AltMin) algorithm was proposed in [15]. This algorithm effectively designs the hybrid beamformer by offering globally optimal solutions for both subproblems of analog and digital beamformers in each alternating iteration, and thus achieves very good performance. In particular, the hybrid beamformer design problem is decoupled for each RF chain and its connected antenna elements. In this way, each subproblem is reformulated as a non-convex quadratically constrained quadratic programming (QCQP) problem, to which the SDR approach was applied, and the tightness of such an SDR is proved in [15].

The achievable spectral efficiency of existing representative hybrid beamforming algorithms is compared in Fig. 2. As can be observed, the MO-AltMin algorithm achieves the highest spectral efficiency for the SPS fully-connected structure while the SDR-AltMin algorithm outperforms other benchmarks for the SPS partially-connected structure. Furthermore, the numbers of hardware components and computational complexity of corresponding design algorithms
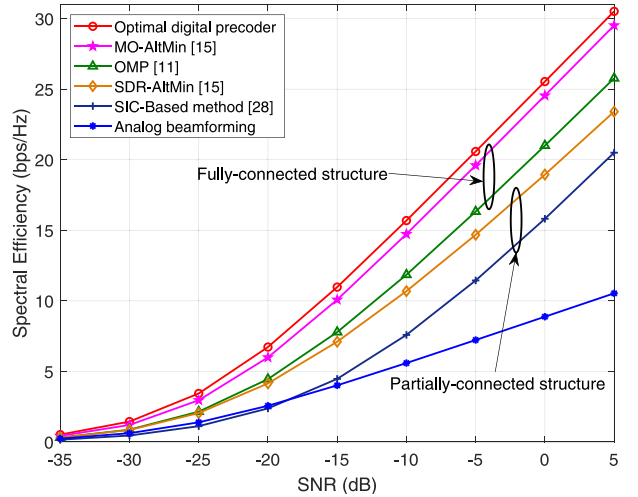

FIGURE 2. Spectral efficiency of different beamforming algorithms for a point-to-point link in a flat-fading channel. The BS and user are equipped with 144 and 36 antennas, respectively, to transmit 3 data streams with 3 RF chains at both the transmitter and receiver sides.

for hybrid beamforming with the SPS implementation are summarized in Table 2. In particular, the partially-connected mapping strategy entails a lower computational complexity thanks to its simpler hardware implementation.

D. LIMITATIONS OF BASIC STRUCTURES

We compare spectral efficiency of the two basic structures in Fig. 3. It shows a clear performance gap between the two structures, with the fully-connected structure providing much higher spectral efficiency than the partially-connected one. Furthermore, the comparison between the MO-AltMin and OMP algorithms demonstrates the importance of efficient algorithms to reach realistic conclusions. In particular, with the MO-AltMin algorithm, the fully-connected structure is shown to approach the performance of the fully digital one with the number of RF chains comparable to the number of data streams, while the OMP algorithm fails to achieve this. These observations demonstrate that the limited number of RF chains in hybrid beamforming is not a
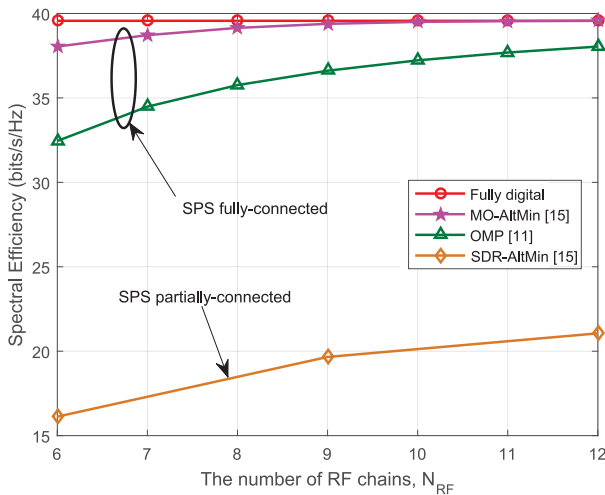
FIGURE 3. Spectral efficiency of different beamforming algorithms for a point-to-point link in a flat-fading channel. The BS and user are equipped with 144 and 36 antennas, respectively, to transmit 6 data streams with SNR as 0 dB. The x-axis stands for the numbers of RF chains at both the transmitter and receiver sides.

FIGURE 4. DPS implementation for hybrid beamforming.

performance bottleneck, but the analog network structure has a decisive effect.

The above comparison reveals several key limitations of the two basic structures.

- *Algorithmic Perspective:* While the SPS fully-connected structure with the MO-AltMin algorithm approaches the performance of the fully digital beamforming, its computational complexity is extremely high [15]. It is not clear how close we can approach fully digital beamforming with more practical algorithms for this structure.
- *Hardware Perspective:* The SPS fully-connected structure has the potential to perform closely to the fully digital one, but still with high hardware complexity in the analog network. The SPS partially-connected structure significantly reduces the number of phase shifters, but with much degraded performance.

Therefore, key innovations in both the hardware and algorithmic aspects are needed before we see the commercial success of hybrid beamforming. From the above discussion, we have already observed that the analog network structure greatly affects the algorithm design. So the key challenge is to design the analog network to reduce hardware complexity, as well as enabling low-complexity beamforming algorithms, which will be addressed in Section IV.

Inevitably, trade-offs need to be made among hardware efficiency, computational efficiency, and spectral efficiency. The two basic structures provide such a trade-off, but in an extreme way. The fully-connected mapping strategy is with too high hardware complexity, as well as algorithm complexity if with the MO-AltMin algorithm, while the partially-connected mapping strategy incurs too much performance degradation. It is thus of practical importance to develop new structures that can achieve more flexible
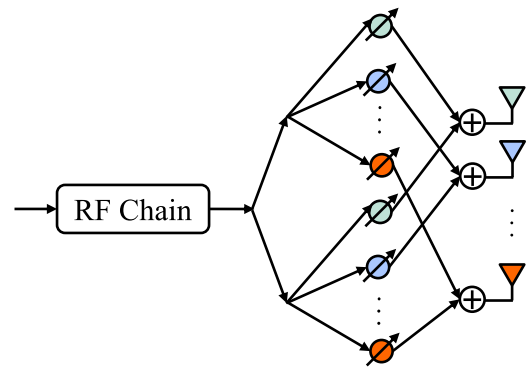
trade-offs. To address this aspect, a flexible mapping strategy will be presented in Section V.

IV. TWO NEW ANALOG NETWORK IMPLEMENTATIONS

In this section, we introduce two recent proposals for the analog network implementation, which improve upon the SPS implementation in different aspects. The first proposal, namely the *double phase shifter* (DPS) implementation, simplifies the algorithm design and improves spectral efficiency, at the cost of more phase shifters. One byproduct of the investigation of this implementation is a convex relaxation approach to develop highly efficient beamforming algorithms. The second proposal, called the *fixed phase shifter* (FPS) implementation, only requires a small number of fixed phase shifters, supplemented with switches, and thus it improves hardware efficiency. As will be shown later, it also does well in computational efficiency and spectral efficiency.

A. DOUBLE PHASE SHIFTER (DPS) IMPLEMENTATION

In this part, we present a new hardware implementation to enable efficient hybrid beamforming algorithms. For the SPS implementation, the unit modulus constraint for the analog network forms the main challenge for algorithm design. The principal obstacle is that we can only adjust the phase but not the amplitude of the RF signals.

To overcome this constraint, the DPS implementation employs two sets of phase shifters, as shown in Fig. 4. Thus, there are $2N_tN_{RF}^t$ and $2N_t$ phase shifters for the DPS fully- and partially-connected structures, respectively. 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 to compose the analog beamforming gain. In this way, each non-zero element in the analog network corresponds to a sum of the outputs of two phase shifters. Correspondingly, the feasible set $\mathcal{A}$ in (1) is specified by a set of matrices where the non-zeros entries have amplitudes less than 2, i.e., $|\mathbf{F}_{RF}(i,j)| = |e^{j\phi} + e^{j\theta}| \leq 2$, where ϕ and θ are two phase shifts from each group, respectively. Thus, the new constraints of the analog beamforming matrix become convex, which makes beamforming algorithm design more tractable.

This new implementation fundamentally changes the algorithm design, and computationally efficient beamforming algorithms have been developed for both the fully- and partially-connected mapping strategies [31].

1) FULLY-CONNECTED MAPPING

For the fully-connected mapping, the hybrid beamforming problem can be specified as

$$\underset{\mathbf{F}_{RF}, \mathbf{F}_{BB}}{\text{minimize}} \quad \left\| \mathbf{F}_{opt} - \mathbf{F}_{RF}\mathbf{F}_{BB} \right\|_F$$
$$\text{subject to} \quad \begin{cases} |\mathbf{F}_{RF}(i,j)| \leq 2 \\ \|\mathbf{F}_{RF}\mathbf{F}_{BB}\|_F^2 \leq KN_sF. \end{cases} \quad (2)$$

It is proved in [31] that the two constraints in (2) are redundant, and the remaining problem turns out to be a low-rank matrix approximation problem, which has been well studied and is with a closed-form solution.

It has been investigated that the fully digital beamforming can be achieved when $N_{RF}^t \geq 2KN_s$ with the SPS fully-connected structure [26]. In other words, $2KN_s$ RF chains and $2KN_sN_t$ phase shifters are enough for achieving fully digital beamforming in single-carrier systems. In contrast, the formulation (2) of the DPS fully-connected structure reveals its optimality in single-carrier systems.

Lemma 1: For single-carrier systems, with the DPS implementation, a fully digital beamformer $\mathbf{F}_{opt}$ can be perfectly decomposed into $\mathbf{F}_{RF}$ and $\mathbf{F}_{BB}$ using the minimum number of RF chains, i.e., $N_{RF}^t = KN_s$ and $N_{RF}^t = N_s$.

Proof: The proof can be easily obtained by the rank sufficiency of $\mathbf{F}_{RF}$ and $\mathbf{F}_{BB}$ in the decomposition when $F = 1$. ∎

This lemma means that KN_s RF chains and $2KN_sN_t$ phase shifters are enough for achieving fully digital beamforming, which reduces the required number of RF chains by half compared to the state-of-the-art with the SPS implementation. This phenomenon clearly demonstrates the superiority of doubling the phase shifters in the analog network for hybrid beamforming.

When it comes to multiuser multicarrier systems, typically $KN_sF \geq N_t$, the rank of $\mathbf{F}_{opt}$ should be N_t (instead of KN_s as single-carrier systems)[3] and thus perfect decomposition can only be achieved when $N_{RF}^t \geq N_t$, which, however, severely deviates from the setting of hybrid beamforming. In this way, the matrix decomposition cannot be perfect for hybrid beamformer design due to the rank deficiency, i.e., $N_{RF}^t = \text{rank}(\mathbf{F}_{RF}\mathbf{F}_{BB}) \ll \text{rank}(\mathbf{F}_{opt}) = N_t$. Therefore, problem (2) is typically a low-rank matrix approximation problem, with a closed-form solution as

$$(\mathbf{F}_{RF}\mathbf{F}_{BB})^\star \triangleq \hat{\mathbf{F}}_{opt} = \mathbf{U}_1\mathbf{S}_1\mathbf{V}_1^H. \quad (3)$$

Denote the SVD of $\mathbf{F}_{opt}$ as $\mathbf{F}_{opt} = \mathbf{U}\mathbf{S}\mathbf{V}^H$, where matrices $\mathbf{U}_1$ and $\mathbf{V}_1$ are the first N_{RF}^t columns of $\mathbf{U}$ and $\mathbf{V}$, respectively, and $\mathbf{S}_1$ is the diagonal matrix whose diagonal elements are

---

3. Without loss of generality, we assume all the beamforming matrices in (2) have full rank.
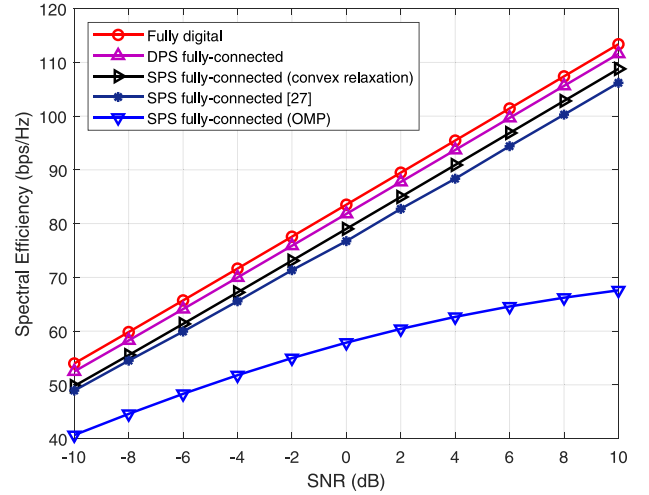

FIGURE 5. Spectral efficiency achieved by different SPS and DPS fully-connected structures. There are 3 mobile users and the BS transmits 3 data streams to each user. The BS and user are equipped with 256 and 16 antennas, respectively.

the N_{RF}^t largest singular values of $\mathbf{F}_{opt}$. This means that the optimal solution of $\mathbf{F}_{RF}\mathbf{F}_{BB}$ is simply obtained by extracting the N_{RF}^t most principle components of $\mathbf{F}_{opt}$.

Convex Relaxation for Efficient Hybrid Beamforming: In addition, inspired by the beamformer design of the DPS fully-connected structure, a convex relaxation approach for the hybrid beamformer design with the SPS fully-connected structure has been developed [31]. Assume that the optimal solution to the low-rank approximation problem (2) is $\hat{\mathbf{F}}_{opt}$, and we propose to extract the phases of the optimal analog network for the DPS implementation to construct the SPS solution, given by

$$\mathbf{F}_{RF} = \exp\{J\angle(\mathbf{U}_1)\}, \quad \mathbf{F}_{BB} = \mathbf{S}_1\mathbf{V}_1^H \quad (4)$$

where $\angle$ extracts the angle information of a complex matrix in an element wise. Note that the unitary matrix $\mathbf{U}_1$ fully extracts the information of the column space of $\hat{\mathbf{F}}_{opt}$, whose basis are the orthonormal columns in $\mathbf{F}_{RF}$. This approach only requires an singular value decomposition (SVD) operation, which leads to a low-complexity beamforming algorithm by extracting phases from the DPS solution.

Fig. 5 shows the spectral efficiency achieved by the DPS fully-connected structure, and that of the SPS fully-connected structure with different algorithms. It shows that the DPS implementation outperforms the SPS implementation, and can achieve a near-optimal performance in terms of spectral efficiency, thanks to the doubling of the phase shifters. In addition, the SPS implementation with the convex relaxation algorithm outperforms the state-of-the-art algorithm in [27], while enjoying much lower computational complexity, which demonstrates the effectiveness of the proposed approach.

2) PARTIALLY-CONNECTED MAPPING

On the other hand, similar to the SPS partially-connected structure, the hybrid beamforming design with the DPS

partially-connected mapping can be decoupled in an RF chain-by-RF chain sense. The optimization of the hybrid beamformer for the j-th RF chain is given by

$$\mathcal{P}_j: \quad \underset{\{a_i\}, \mathbf{x}_j}{\text{minimize}} \sum_{i \in \mathcal{F}_j} \left\| \mathbf{y}_i - a_i \mathbf{x}_j \right\|_2^2, \quad (5)$$

where a_i is the non-zero element in $\mathbf{F}_{RF}(i,:)$,[4] $\mathcal{F}_j = \{i \in \mathbb{Z} | (j-1) \frac{N_t}{N_{RF}^t} + 1 \le i \le j \frac{N_t}{N_{RF}^t}.\}$, $\mathbf{y}_i = \mathbf{F}_{opt}^T(i,:)$, and $\mathbf{x}_j = \mathbf{F}_{BB}^T(j,:)$. It is shown in [31] that $\mathcal{P}_j$ is an eigenvalue problem. Thus, the DPS implementation brings great advantages in computational efficiency with the closed-form solutions.

The DPS partially-connected structure employs $2N_t$ phase shifters, which falls in between the numbers of phase shifters in use for the SPS partially-connected structure (N_t) and the DPS fully-connected one ($N_{RF}^t N_t$). To further boost the spectral efficiency with $2N_t$ phase shifters, a dynamic mapping for the DPS partially-connected structure was proposed in [31]. In particular, each RF chain is still connected to a subset of antenna elements, but not necessarily the neighboring ones. In other words, each RF chain is able to select which antenna elements to connect in order to increase the spectral efficiency. For dynamic mapping, the feasible set $\mathcal{A}$ in (1) can be specified as a set of matrices for which every row only has one non-zero entry, i.e., $\mathcal{A} = \{\mathbf{A} | \|\mathbf{A}(i,:)\|_0 = 1\}$, and the dynamic mapping design problem is formulated as [31]

$$\underset{\{\mathcal{D}_j\}_{j=1}^{N_{RF}^t}}{\text{maximize}} \sum_{j=1}^{N_{RF}^t} \lambda_1 \left(\sum_{i \in \mathcal{D}_j} \mathbf{y}_i \mathbf{y}_i^H \right)$$

$$\text{subject to} \begin{cases} \cup_{j=1}^{N_{RF}^t} \mathcal{D}_j = \{1, \dots N_t\} \\ \mathcal{D}_j \cap \mathcal{D}_k = \emptyset, \quad \forall j \ne k, \end{cases} \quad (6)$$

where $\mathcal{D}_j$ is the mapping set containing the antenna indices that are mapped to the j-th RF chain, and $\lambda_1(\cdot)$ denotes the largest eigenvalue of a matrix. The design problem is a combinatorial problem and thus the optimal solution can be given by exhaustive search with an extremely huge number of possible mapping strategies, which prevents its practical implementation. Therefore, a greedy algorithm and a modified K-means algorithm were proposed in [31].

Fig. 6 shows the performance of different design approaches in the DPS partially-connected structure with the minimum numbers of RF chains, i.e., $N_{RF}^t = KN_s$. We see that, due to the sharply reduced number of phase shifters, the partially-connected structure does entail non-negligible performance loss compared to the fully digital one. Furthermore, it shows that simply doubling the number of phase shifters with the fixed mapping only has little performance gain over the conventional SPS implementation [15]. Fig. 6 demonstrates that dynamic mapping is able

---

4. Since each antenna element is connected to one RF chain in the partially-connected mapping strategy, there is only one non-zero element in each row of $\mathbf{F}_{RF}$.
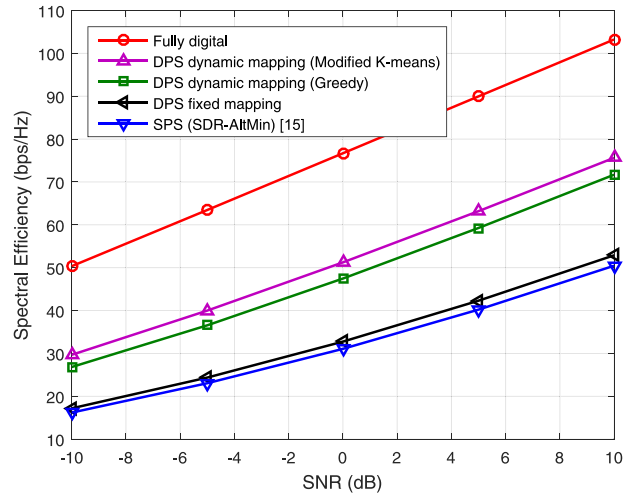
FIGURE 6. Spectral efficiency achieved by different hybrid beamforming algorithms in the DPS partially-connected structure. There are 4 mobile users and the BS transmits 2 data streams to each user. The BS and user are equipped with 256 and 16 antennas, respectively.

to shrink the gap between the fixed mapping and the fully digital beamforming by half.

Considering the increased number of phase shifters, the DPS implementation may not be practical for deployment before low-cost low-power phase shifters are available, but it does provide valuable guidelines to design other hybrid beamforming structures.

1) With computationally efficient and optimal beamforming algorithms, the DPS fully-connected structure can serve as a performance upper bound for structures that are with higher hardware efficiency. It is a tighter upper bound than the fully digital beamforming, especially when the number of RF chains is small.
2) The computationally efficient algorithm for the DPS fully-connected structure has inspired a highly effective algorithm for the SPS fully-connected structure, which enjoys a low computational complexity and outperforms existing algorithms.
3) The algorithmic and performance advantages of the DPS implementation are achieved via passing the same signal through more than one phase shifter, which can inspire similar proposals for improvement, as will be discussed in the next subsection.
4) As the beamforming problem becomes a low-rank matrix approximation (eigenvalue) problem for the DPS fully-connected (partially-connected) structure, theoretical analysis, which is intractable for other structures, becomes possible. It will then help to better understand hybrid beamforming systems.

B. FIXED PHASE SHIFTER (FPS) IMPLEMENTATION

The key weakness of the DPS implementation is the low hardware efficiency. Nevertheless, as discussed above, we can draw valuable lessons for further improvement. The key idea of DPS is to pass the signal out of each RF chain
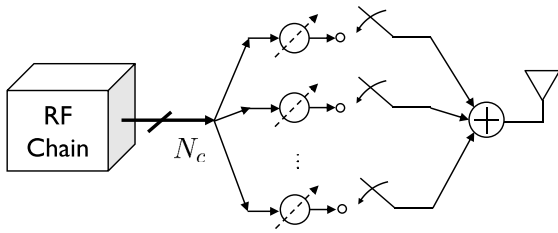
FIGURE 7. The FPS implementation from an RF chain to a connected antenna.

FIGURE 8. Spectral efficiency of different implementations in a 144 × 16 MIMO-OFDM system with 128 subcarriers. The BS transmits 2 data streams to each of the 4 users on each subcarrier, with 8 RF chains at each node. For the left hand side figure, the number of phase shifters is 30, while we change the number of phase shifters for the FPS implementation in the right hand side figure, with SNR being 0 dB.

through more than one phase shifter. Specifically, this will help to overcome the non-convex unit modulus constraint for the analog network, and thus significantly simplifies algorithm design. At the same time, it will provide capability to change the amplitudes of elements of the analog beamforming matrix, which helps to improve the spectral efficiency.

Inspired by these insights, a novel analog network implementation, namely the FPS implementation, has been proposed in [32], which allows each signal to pass multiple phase shifters. A key difference compared with previous proposals is that only a small number of phase shifters, with *quantized* and *fixed* phases, are employed. While existing works on hybrid beamforming commonly assumed a large number of phase shifters with unquantized phases, in practice the phase shifters should be discretized with a coarse quantization, and their number should be reduced to a minimum due to cost and power consideration. Thus, the FPS implementation is very promising for practical systems.

With a small number of fixed phase shifters, the beamformer has limited capability to adapt to the channel states, which will inevitably entail performance loss. To overcome this drawback, a dynamic switch network is cascaded after the fixed phase shifters, as shown in Fig. 7. In particular, a total of N_c multichannel ($N_{\mathrm{RF}}^{\mathrm{t}}$-channel) fixed phase shifters are employed, each of which simultaneously processes the output signals from $N_{\mathrm{RF}}^{\mathrm{t}}$ RF chains in a parallel manner. In this way, these N_c phase shifters generate N_c signals with different phases for the signal of each RF chain. Inspired by the DPS implementation, a subset of these N_c signals are selected and combined to compose the analog beamforming gain from the RF chain to the antenna. As N_c adaptive switches are needed for each RF chain-antenna pair, in total $N_t N_{\mathrm{RF}}^{\mathrm{t}} N_c$ switches are needed for the FPS implementation. The switch network provides dynamic connection from phase shifters to antennas, which is adaptive to channel states. Equipped with a small number of fixed phase shifters and assisted by low-complexity switches, the FPS implementation enjoys hardware complexity comparable to or even lower than the analog beamforming, which needs N_t phase shifters with adaptive phases.

For beamforming algorithm design, different from other implementations, the analog network of the FPS implementation is essentially to determine the states of different switches, with binary variables, whose
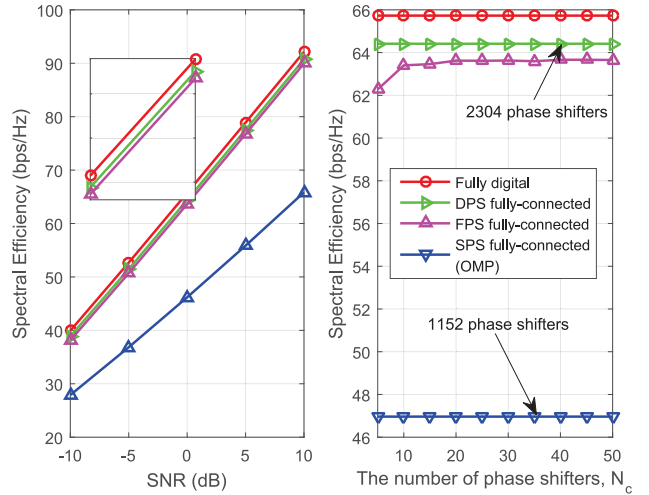
formulation is given by

$$\underset{\mathbf{S},\mathbf{F}_{\mathrm{BB}}}{\text{minimize}} \quad \left\| \mathbf{F}_{\mathrm{opt}} - \mathbf{S}\mathbf{C}\mathbf{F}_{\mathrm{BB}} \right\|_F^2$$
$$\text{subject to} \quad \mathbf{S} \in \{0,1\}^{N_t \times N_c N_{\mathrm{RF}}^{\mathrm{t}}}, \tag{7}$$

where the switch matrix $\mathbf{S}$ is a binary matrix. The matrix $\mathbf{C}$ stands for the phase shift operation carried out by the available fixed phase shifters, given by a block diagonal matrix as

$$\mathbf{C} = \mathrm{blkdiag}\Big(\underbrace{\mathbf{c}, \mathbf{c}, \dots, \mathbf{c}}_{N_{\mathrm{RF}}^{\mathrm{t}}} \Big), \tag{8}$$

where $\mathbf{c} = \frac{1}{\sqrt{N_c}}[e^{j\theta_1}, e^{j\theta_2}, \dots, e^{j\theta_{N_c}}]^T$ is the normalized phase shifter vector containing all N_c fixed phases $\{\theta_i\}_{i=1}^{N_c}$. Note that although there are $N_c N_{\mathrm{RF}}^{\mathrm{t}}$ non-zero parameters in matrix $\mathbf{C}$, only N_c phase shifters are required since the phase shifters are with $N_{\mathrm{RF}}^{\mathrm{t}}$ parallel channels and shared by all RF chain-antenna pairs. To solve this problem, an efficient AltMin algorithm was proposed in [32]. A tight upper bound of the objective function was first derived, based on which closed-form solutions for both the dynamic switch network and the digital baseband beamformer. Note that we may also develop an FPS partially-connected structure to reduce the number of switches, but it has been found to incur significant performance loss. We will explore a more effective approach to achieve hardware-performance trade-offs in Section V.

C. PERFORMANCE COMPARISON

In Fig. 8, the spectral efficiency of the two presented analog network implementations is evaluated, compared with the fully digital beamforming and the SPS fully-connected structure with the OMP algorithm. As a general multicarrier

multiuser system is considered, the MO-AltMin algorithm is inapplicable due to high complexity. It shows that both the DPS and FPS fully-connected structures achieve performance close to the fully digital one. This is quite an astonishing result, given that a single analog network is shared by all the users and subcarriers, and the number of RF chains is only the same as the number of data streams. The poor performance of the SPS implementation is partly due to the sub-optimality of the beamforming algorithm, as the unit modulus constraint in the analog beamforming matrix makes it difficult to develop high-performance low-complexity algorithms.

Remarkably, the FPS fully-connected structure performs closely to the DPS one, though with much fewer phase shifters. As shown in the figure on the right, around 10 fixed phase shifters are sufficient for the FPS implementation, while the SPS and DPS implementations require 1152 and 2304 phase shifters, respectively. This makes the FPS implementation very attractive for practical deployment. Meanwhile, once low-cost high-resolution commercial phase shifters are available, or for cost-insensitive applications, the DPS implementation would be an ideal choice in terms of both the spectral efficiency and computational efficiency.

V. A FLEXIBLE MAPPING STRATEGY FOR HARDWARE-PERFORMANCE TRADE-OFFS

Among the presented hybrid beamforming structures, the DPS fully-connected structure performs the best in both computational efficiency and spectral efficiency, but with low hardware efficiency. The FPS fully-connected structure achieves a good balance among the three design aspects, but requires a large number of switches. Considering the cost and power consumption of hardware components, especially for mm-wave systems, it is important to further reduce the hardware complexity. Meanwhile, the partially-connected mapping strategy fails to be a good candidate for high hardware efficiency, as it reduces hardware complexity by too much and incurs significant performance loss. Thus, it is highly desirable to have fine granularity when reducing the hardware complexity. In this section, we present a flexible hybrid beamforming mapping strategy, called the *group-connected* mapping, to achieve a better balance between hardware efficiency and spectral efficiency.

As shown in Table 1 (a), with this new mapping strategy, antennas and RF chains are divided into η groups, and signals coming out of each RF chain group are transmitted via its corresponding antenna group. The grouping is flexible, and the numbers of RF chains and antennas in different groups can be different. The mapping strategy within each group is the same as the fully-connected mapping. Thus, the analog beamforming matrix $\mathbf{F}_{RF}$ has the block diagonal structure, with each block corresponding to one RF chain-antenna group. It is easy to observe that conventional fully- and partially-connected mapping strategies are special cases of this flexible one:
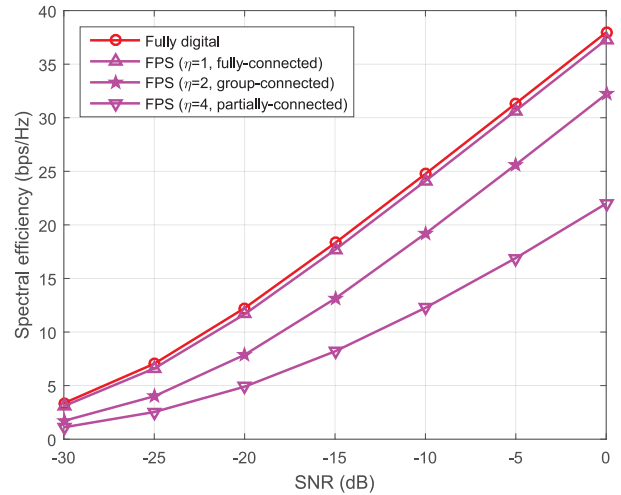

FIGURE 9. Spectral efficiency of different values of η with the FPS group-connected structure. The BS and user are with 256 and 16 antennas. The numbers of RF chains at both the BS and user sides are equal to that of the data streams, which is 4.

- When $\eta = 1$, there is only one RF chain group and one antenna group, and thus we get the fully-connected mapping strategy;
- When $\eta = N_{RF}^{t}$, each RF chain group contains a single RF chain, which is connected to a group of antennas, and thus we get the partially-connected mapping strategy.

By varying the value of η, we can easily obtain hybrid beamforming mapping strategies with different hardware complexities. Moreover, we can apply any of the hardware implementations presented in Table 1 (b) with this group-connected mapping. For the SPS and DPS implementations, the number of phase shifters is $1/\eta$ of the fully-connected one; for the FPS implementation, the number of switches is $1/\eta$ of the one shown in Table 1 (b), while the number of fixed phase shifter keeps the same.

In terms of beamforming algorithm design, due to the block diagonal structure in $\mathbf{F}_{RF}$, we can decouple the design of each block, for which the problem is similar to the conventional fully-connected mapping. Therefore, we can leverage the rich algorithms presented in the previous two sections for different analog network implementations. In other words, this flexible structure does not introduce any additional difficulty in beamforming algorithm design.

In Fig. 9, we ompare spectral efficiency of the FPS group-connected structure with different values of η. Other implementations have the same trend. It shows that varying the value of η helps to effectively balance the hardware complexity and spectral efficiency. To summarize, this new mapping strategy enjoys the following three desirable properties:

1) It provides a flexible way to trade off performance against hardware complexity;
2) It is compatible with different analog network implementations;

TABLE 3. Comparisons of hardware components in the analog network for different hybrid beamforming structures.

Implementation	Mapping strategy	Phase shifter			Other hardware components		
		Number	Type	Power	Hardware	Number	Power
SPS [11], [15]	Fully-connected	$N_{\mathrm{RF}}^{\mathrm{t}} N_{\mathrm{t}}$	Adaptive	50 mW	N/A	N/A	N/A
	Partially-connected	N_{t}					
DPS [31]	Fully-connected	$2 N_{\mathrm{RF}}^{\mathrm{t}} N_{\mathrm{t}}$	Adaptive	50 mW	N/A	N/A	N/A
	Partially-connected	$2 N_{\mathrm{t}}$					
FPS [32]	Fully-connected	$N_c \ll N_{\mathrm{t}}$	Multi-channel Fixed	20 mW	Switch	$N_c N_{\mathrm{RF}}^{\mathrm{t}} N_{\mathrm{t}}$	5 mW
	Group-connected					$\frac{1}{\eta} N_c N_{\mathrm{RF}}^{\mathrm{t}} N_{\mathrm{t}}$	

* The power consumption of (fixed) phase shifters and switches are typical values in [33] and [34].

TABLE 4. Comparisons of computational complexity for different hybrid beamforming algorithms.

| Implementation	Mapping strategy	Design algorithm	Computational complexity	Spectral efficiency
SPS [11], [15]	Fully-connected	MO-AltMin [15]	Extremely high	✓✓✓✓
	Partially-connected	SDR-AltMin [15]	$\mathcal{O}\left(N_{\mathrm{iter}} N_{\mathrm{RF}}^{\mathrm{t}}{}^3 K^3 N_s^3 F^3\right)$	✓
DPS [31]	Fully-connected	Matrix decomposition (3)	$\mathcal{O}\left(N_{\mathrm{RF}}^{\mathrm{t}} N_{\mathrm{t}} K N_s F\right)$	✓✓✓✓✓✓
	Partially-connected	Modified K-means [31]	$\mathcal{O}\left(N_{\mathrm{iter}} N_{\mathrm{RF}}^{\mathrm{t}} K^3 N_s^3 F^3\right)$	✓✓
FPS [32]	Fully-connected	FPS-AltMin [32]	$\mathcal{O}\Big(N_{\mathrm{iter}}\big(K^2 N_s^2 N_{\mathrm{RF}}^{\mathrm{t}} +$	✓✓✓✓
	Group-connected		$N_c N_{\mathrm{RF}}^{\mathrm{t}} N_{\mathrm{t}} \log\left(N_c N_{\mathrm{RF}}^{\mathrm{t}} N_{\mathrm{t}}\right)\big)\Big)$	✓✓✓

3) The hybrid beamformer can be effectively designed by leveraging existing algorithms.

Therefore, this mapping strategy, especially with the FPS implementation, stands out as a promising candidate to support hybrid beamforming in 5G and beyond mm-wave systems. The hardware components in the analog network and design algorithms for different hybrid beamforming structures are compared in Tables 3 and 4, respectively.

VI. CONCLUSION AND FUTURE DIRECTIONS

In this paper, we presented several proposals of hybrid beamforming structures in mm-wave systems, focusing on three key aspects: hardware efficiency, spectral efficiency, and computational efficiency. Through a systematic comparison, important design insights were revealed. In particular, it was shown that hardware implementation significantly affects the algorithm design and achievable spectral efficiency. With a suitable structure, hybrid beamforming can approach the performance of the fully digital one with low hardware complexity. For example, it is sufficient to have RF chains comparable to the number of data streams, and a small number ($\sim$10) of fixed phase shifters are sufficient with the FPS implementation. Furthermore, a flexible structure was proposed to balance hardware efficiency and spectral efficiency. A qualitative comparison of different structures is shown in Fig. 10. Overall, the FPS group-connected structure stands out as a promising candidate for hybrid beamforming in 5G and beyond mm-wave systems. Once low-cost phase shifters are available, the DPS implementation will also be attractive.

To achieve the full success of hybrid beamforming, more works will be needed, and the followings are some potential future research directions.
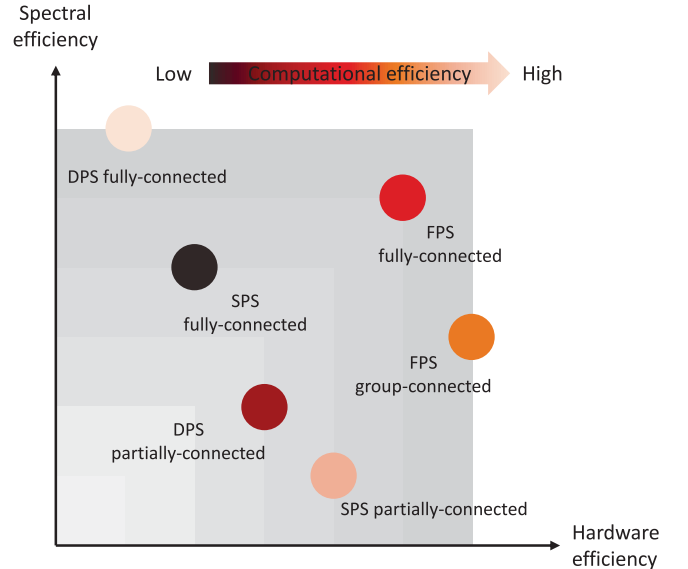
FIGURE 10. Qualitative comparison of different structures, in hardware efficiency, computation efficiency of the typical algorithm, and achievable spectral efficiency.

- *CSI Acquisition for Hybrid Beamforming:* Perfect channel state information (CSI) was assumed in the discussion of this paper, and acquiring large-scale CSI with reduced RF chains is a challenging problem, with some prior studies in [35]–[38]. Different training methods may be needed for different hybrid beamforming structures [25], [39]. The presented results also shed light on hybrid beamforming design during the training stage, which is critical to overcome the low SNR during training. In addition, codebook design for channel

estimation is of another particular interest in mm-wave MIMO systems [40]–[42].

- *Deep Learning for Efficient Hybrid Beamforming:* It is highly desirable to further reduce the computational complexity of hybrid beamforming algorithms. Recently, deep learning has been applied to develop efficient algorithms for large-scale optimization problems in wireless networks [43]–[46], including hybrid beamforming [47]–[49]. While these initial attempts have demonstrated the effectiveness of deep learning-based methods, more investigation will be needed, from both practical and theoretical perspectives.

- *Finite-Precision ADCs:* While the focus in this paper is on the analog network, there are still some gaps to fill in the digital domain. In particular, the quantization effect of ADCs cannot be ignored. How to extend the presented hybrid beamforming structures to systems with low-resolution ADCs deserves delicate investigation, and some previous studies can be found in [50]–[53].

- *Algorithm-Hardware Co-Design:* To effectively design the increasingly complex wireless systems, collaboration among the hardware and algorithm domains will be needed. Hardware-algorithm co-design will play an important role in 5G and beyond systems [54]. The target is to develop hardware-efficient transceiver structures that are also algorithm friendly. The FPS implementation can be regarded as a preliminary attempt of such design approaches in mm-wave systems.

- *Hybrid Beamforming in Networks:* From the network perspective, while mm-wave networks with analog beamforming have been extensively analyzed [55]–[57], the effect of adopting hybrid beamforming has not been fully unraveled. Indeed, hybrid beamforming will result in more intricate signal and interference distributions, which should be carefully investigated.

REFERENCES

[1] A. Ghosh, J. Zhang, J. G. Andrews, and R. Muhamed, *Fundamentals of LTE.* Upper Saddle River, NJ, USA: Prentice-Hall, 2010.

[2] J. G. Andrews *et al.,* "What will 5G be?" *IEEE J. Sel. Areas Commun.,* vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[3] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.,* vol. 52, no. 2, pp. 186–195, Feb. 2014.

[4] T. S. Rappaport *et al.,* "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access,* vol. 1, pp. 335–349, 2013.

[5] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proc. IEEE,* vol. 102, no. 3, pp. 366–385, Mar. 2014.

[6] C. J. Hansen, "WiGiG: Multi-gigabit wireless communications in the 60 GHz band," *IEEE Wireless Commun.,* vol. 18, no. 6, pp. 6–7, Dec. 2011.

[7] A. Alkhateeb, J. Mo, N. González-Prelcic, and R. W. Heath, Jr., "MIMO precoding and combining solutions for millimeter-wave systems," *IEEE Commun. Mag.,* vol. 52, no. 12, pp. 122–131, Dec. 2014.

[8] T. E. Bogale, L. B. Le, A. Haghighat, and L. Vandendorpe, "On the number of RF chains and phase shifters, and scheduling design with hybrid analog-digital beamforming," *IEEE Trans. Wireless Commun.,* vol. 15, no. 5, pp. 3311–3326, May 2016.

[9] A. F. Molisch *et al.,* "Hybrid beamforming for massive MIMO: A survey," *IEEE Commun. Mag.,* vol. 55, no. 9, pp. 134–141, Sep. 2017.

[10] X. Zhang, A. F. Molisch, and S.-Y. Kung, "Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection," *IEEE Trans. Signal Process.,* vol. 53, no. 11, pp. 4091–4103, Nov. 2005.

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

[12] X. Yu, J.-C. Shen, J. Zhang, and K. B. Letaief, "Hybrid precoding design in millimeter wave MIMO systems: An alternating minimization approach," in *Proc. IEEE Glob. Commun. Conf. (GLOBECOM),* San Diego, CA, USA, Dec. 2015, pp. 1–6.

[13] A. Alkhateeb and R. W. Heath, Jr., "Frequency selective hybrid precoding for limited feedback millimeter wave systems," *IEEE Trans. Commun.,* vol. 64, no. 5, pp. 1801–1818, May 2016.

[14] J. Lee and Y. H. Lee, "AF relaying for millimeter wave communication systems with hybrid RF/baseband MIMO processing," in *Proc. IEEE Int. Conf. Commun. (ICC),* Sydney, NSW, Australia, Jun. 2014, pp. 5838–5842.

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

[16] W. Ni and X. Dong, "Hybrid block diagonalization for massive multiuser MIMO systems," *IEEE Trans. Commun.,* vol. 64, no. 1, pp. 201–211, Jan. 2016.

[17] T. E. Bogale and L. B. Le, "Beamforming for multiuser massive MIMO systems: Digital versus hybrid analog-digital," in *Proc. IEEE Glob. Commun. Conf. (GLOBECOM),* Austin, TX, USA, Dec. 2014, pp. 4066–4071.

[18] L. Liang, W. Xu, and X. Dong, "Low-complexity hybrid precoding in massive multiuser MIMO systems," *IEEE Wireless Commun. Lett.,* vol. 3, no. 6, pp. 653–656, Dec. 2014.

[19] F. Sohrabi and W. Yu, "Hybrid analog and digital beamforming for mmWave OFDM large-scale antenna arrays," *IEEE J. Sel. Areas Commun.,* vol. 35, no. 7, pp. 1432–1443, Jul. 2017.

[20] Y.-Y. Lee, C.-H. Wang, and Y.-H. Huang, "A hybrid RF/baseband precoding processor based on parallel-index-selection matrix-inversion-bypass simultaneous orthogonal matching pursuit for millimeter wave MIMO systems," *IEEE Trans. Signal Process.,* vol. 63, no. 2, pp. 305–317, Jan. 2015.

[21] M. Kim and Y. H. Lee, "MSE-based hybrid RF/baseband processing for millimeter-wave communication systems in MIMO interference channels," *IEEE Trans. Veh. Technol.,* vol. 64, no. 6, pp. 2714–2720, Jun. 2015.

[22] T. Lin, J. Cong, Y. Zhu, J. Zhang, and K. B. Letaief, "Hybrid beamforming for millimeter wave systems using the MMSE criterion," *IEEE Trans. Commun.,* vol. 67, no. 5, pp. 3693–3708, May 2019.

[23] S. Han, C.-L. I, Z. Xu, and C. Rowell, "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G," *IEEE Commun. Mag.,* vol. 53, no. 1, pp. 186–194, Jan. 2015.

[24] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization-based beamforming," *IEEE Signal Process. Mag.,* vol. 27, no. 3, pp. 62–75, May 2010.

[25] R. Méndez-Rial, C. Rusu, N. González-Prelcic, A. Alkhateeb, and R. W. Heath, Jr., "Hybrid MIMO architectures for millimeter wave communications: Phase shifters or switches?" *IEEE Access,* vol. 4, pp. 247–267, 2016.

[26] E. Zhang and C. Huang, "On achieving optimal rate of digital precoder by RF-baseband codesign for MIMO systems," in *Proc. 80th IEEE Veh. Technol. Conf. (VTC Fall),* Vancouver, BC, Canada, Sep. 2014, pp. 1–5.

[27] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE J. Sel. Topics Signal Process.,* vol. 10, no. 3, pp. 501–513, Apr. 2016.

[28] X. Gao, L. Dai, S. Han, C.-L. I, and R. W. Heath, Jr., "Energy-efficient hybrid analog and digital precoding for mmwave MIMO systems with large antenna arrays," *IEEE J. Sel. Areas Commun.,* vol. 34, no. 4, pp. 998–1009, Apr. 2016.

[29] J. Singh and S. Ramakrishna, "On the feasibility of beamforming in millimeter wave communication systems with multiple antenna arrays," in *Proc. IEEE Glob. Commun. Conf. (GLOBECOM)*, Dec. 2014, pp. 3802–3808.

[30] C. Kim, T. Kim, and J.-Y. Seol, "Multi-beam transmission diversity with hybrid beamforming for MIMO-OFDM systems," in *Proc. IEEE Glob. Commun. Conf. Workshops (GLOBECOM Wkshps)*, Atlanta, GA, USA, Dec. 2013, pp. 61–65.

[31] X. Yu, J. Zhang, and K. B. Letaief, "Doubling phase shifters for efficient hybrid precoding in millimeter-wave multiuser OFDM systems," *J. Commun. Inf. Netw.*, vol. 4, no. 2, pp. 51–67, Jul. 2019.

[32] X. Yu, J. Zhang, and K. B. Letaief, "A hardware-efficient analog network structure for hybrid precoding in millimeter wave systems," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 2, pp. 282–297, May 2018.

[33] T. S. Rappaport, R. W. Heath, Jr., R. C. Daniels, and J. N. Murdock, *Millimeter Wave Wireless Communications*. New York, NY, USA: Pearson Educ., 2014.

[34] A. M. Abbosh, "Broadband fixed phase shifters," *IEEE Microw. Compon. Lett.*, vol. 21, no. 1, pp. 22–24, Jan. 2011.

[35] J. Lee, G.-T. Gil, and Y. H. Lee, "Exploiting spatial sparsity for estimating channels of hybrid MIMO systems in millimeter wave communications," in *Proc. IEEE Glob. Commun. Conf. (GLOBECOM)*, Austin, TX, USA, Dec. 2014, pp. 3326–3331.

[36] A. Alkhateeb, G. Leus, and R. W. Heath, Jr., "Compressed sensing based multi-user millimeter wave systems: How many measurements are needed?" in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Brisbane, QLD, Australia, Apr. 2015, pp. 2909–2913.

[37] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, Jr., "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.

[38] Z. Gao, C. Hu, L. Dai, and Z. Wang, "Channel estimation for millimeter-wave massive MIMO with hybrid precoding over frequency-selective fading channels," *IEEE Commun. Lett.*, vol. 20, no. 6, pp. 1259–1262, Jun. 2016.

[39] K. Venugopal, A. Alkhateeb, N. González-Prelcic, and R. W. Heath, Jr., "Channel estimation for hybrid architecture-based wideband millimeter wave systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1996–2009, Sep. 2017.

[40] J. Song, J. Choi, and D. J. Love, "Common codebook millimeter wave beam design: Designing beams for both sounding and communication with uniform planar arrays," *IEEE Trans. Commun.*, vol. 65, no. 4, pp. 1859–1872, Apr. 2017.

[41] Z. Xiao, P. Xia, and X.-G. Xia, "Codebook design for millimeter-wave channel estimation with hybrid precoding structure," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 141–153, Jan. 2017.

[42] J. Zhang, Y. Huang, Q. Shi, J. Wang, and L. Yang, "Codebook design for beam alignment in millimeter wave communication systems," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 4980–4995, Nov. 2017.

[43] N. C. Luong *et al.*, "Applications of deep reinforcement learning in communications and networking: A survey," *IEEE Commun. Surveys Tuts.*, to be published.

[44] H. Ye, G. Y. Li, and B. Juang, "Power of deep learning for channel estimation and signal detection in OFDM systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 114–117, Feb. 2018.

[45] W. Cui, K. Shen, and W. Yu, "Spatial deep learning for wireless scheduling," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 6, pp. 1248–1261, Jun. 2019.

[46] Y. Shen, Y. Shi, J. Zhang, and K. B. Letaief, "LORM: Learning to optimize for resource allocation in wireless networks with few training samples," *IEEE Trans. Wireless Commun.*, to be published.

[47] X. Gao, L. Dai, Y. Sun, S. Han, and C.-L. I, "Machine learning inspired energy-efficient hybrid precoding for mmWave massive MIMO systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Paris, France, May 2017, pp. 1–6.

[48] Y. Long, Z. Chen, J. Fang, and C. Tellambura, "Data-driven-based analog beam selection for hybrid beamforming under mm-Wave channels," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 2, pp. 340–352, May 2018.

[49] T. Lin and Y. Zhu, "Beamforming design for large-scale antenna arrays using deep learning," *IEEE Wireless Commun. Lett.*, to be published.

[50] J. Mo, A. Alkhateeb, S. Abu-Surra, and R. W. Heath, Jr., "Hybrid architectures with few-bit ADC receivers: Achievable rates and energy-rate tradeoffs," *IEEE Trans. Wireless Commun.*, vol. 16, no. 4, pp. 2274–2287, Apr. 2017.

[51] H. He, C.-K. Wen, and S. Jin, "Bayesian optimal data detector for hybrid mmWave MIMO-OFDM systems with low-resolution ADCs," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 3, pp. 469–483, Jun. 2018.

[52] Y. Wang, W. Xu, H. Zhang, and X. You, "Wideband mmWave channel estimation for hybrid massive MIMO with low-precision ADCs," *IEEE Wireless Commun. Lett.*, vol. 8, no. 1, pp. 285–288, Feb. 2019.

[53] K. Roth and J. A. Nossek, "Achievable rate and energy efficiency of hybrid and digital beamforming receivers with low resolution ADC," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 2056–2068, Sep. 2017.

[54] K. B. Letaief, W. Chen, Y. Shi, J. Zhang, and Y. A. Zhang, "The roadmap to 6G: AI empowered wireless networks," *IEEE Commun. Mag.*, vol. 57, no. 8, pp. 84–90, Aug. 2019.

[55] X. Yu, J. Zhang, M. Haenggi, and K. B. Letaief, "Coverage analysis for millimeter wave networks: The impact of directional antenna arrays," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1498–1512, Jul. 2017.

[56] M. Di Renzo, "Stochastic geometry modeling and analysis of multi-tier millimeter wave cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 9, pp. 5038–5057, Sep. 2015.

[57] T. Bai and R. W. Heath, Jr., "Coverage and rate analysis for millimeter-wave cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 1100–1114, Feb. 2015.

JUN ZHANG (S'06–M'10–SM'15) received the B.Eng. degree in electronic engineering from the University of Science and Technology of China in 2004, the M.Phil. degree in information engineering from the Chinese University of Hong Kong in 2006, and the Ph.D. degree in electrical and computer engineering from the University of Texas at Austin in 2009.

He is an Assistant Professor with the Department of Electronic and Information Engineering, Hong Kong Polytechnic University (PolyU). He has coauthored the books entitled *Fundamentals of LTE* (Prentice-Hall, 2010) and *Stochastic Geometry Analysis of Multi-Antenna Wireless Networks* (Springer, 2019). His research interests include wireless communications and networking, mobile edge computing and edge learning, distributed learning and optimization, and big data analytics. He is a co-recipient of the 2019 IEEE Communications Society and Information Theory Society Joint Paper Award, the 2016 Marconi Prize Paper Award in Wireless Communications (the Best Paper Award of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS), and the 2014 Best Paper Award for the *EURASIP Journal on Advances in Signal Processing*. The two papers coauthored received the Young Author Best Paper Award of the IEEE Signal Processing Society in 2016 and 2018, respectively, and the 2016 IEEE ComSoc Asia–Pacific Best Young Researcher Award. He is an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON COMMUNICATIONS, and the JOURNAL OF COMMUNICATIONS AND INFORMATION NETWORKS.

XIANGHAO YU (S'15–M'19) received the B.Eng. degree in information engineering from Southeast University, Nanjing, China, in 2014, and the Ph.D. degree in electronic and computer engineering from the Hong Kong University of Science and Technology (HKUST), Hong Kong, in 2018. He is currently a Humboldt Post-Doctoral Research Fellow with the Institute for Digital Communications, Friedrich–Alexander-University Erlangen–Nurnberg (FAU), Erlangen, Germany. He has coauthored the book entitled *Stochastic Geometry Analysis of Multi-Antenna Wireless Networks* (Springer, 2019). His research interests include millimeter wave communications, intelligent reflecting surface assisted communications, MIMO systems, mathematical optimization, and stochastic geometry. He was the recipient of the IEEE GLOBECOM 2017 Best Paper Award, the 2018 IEEE Signal Processing Society Young Author Best Paper Award, and the IEEE GLOBECOM 2019 Best Paper Award.

KHALED B. LETAIEF (S'85–M'86–SM'97–F'03) received the B.S. degree (with Distinction) in electrical engineering from Purdue University at West Lafayette, West Lafayette, IN, USA, in December 1984, and the M.S. and Ph.D. degrees in electrical engineering from Purdue University in August 1986 and May 1990, respectively.

From 1990 to 1993, he was a Faculty Member with the University of Melbourne, Australia. He has been with the Hong Kong University of Science and Technology (HKUST) since 1993, where he is one of HKUST's Most Distinguished Professors for his boundless energy, collegial nature, dedication, and excellence in research, education, and service. He has held numerous administrative positions, including Head of the Electronic and Computer Engineering Department and Director of the Hong Kong Telecom Institute of Information Technology. He has also served as the Chair Professor and the Dean of the HKUST School 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 2015 to 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 University, and Texas A&M University. He also has 15 patents, including 11 U.S. patents. He is an internationally recognized Leader in wireless communications and networks with research interest in artificial intelligence, big data analytics systems, mobile cloud and edge computing, tactile Internet, 5G systems and beyond. He has over 620 journal and conference papers and given keynote talks as well as courses all over the world in the above areas.

Dr. Letaief received the Mangoon Teaching Award from Purdue University in 1990; the Teaching Excellence Appreciation Award by the School of Engineering at HKUST (four times); the M. G. Gale Medal for Distinguished Teaching (Highest University-Wide Teaching Award and only one recipient per year is honored for their contributions); the 2019 Distinguished Research Excellence Award by HKUST School of Engineering (highest research award and only one recipient every three years is honored for their contributions); the 2019 IEEE Communications Society and Information Theory Society Joint Paper Award; the 2018 IEEE Signal Processing Society Young Author Best Paper Award; the 2017 IEEE Cognitive Networks Technical Committee Publication Award; the 2016 IEEE Signal Processing Society Young Author Best Paper Award; the 2016 IEEE Marconi Prize Paper Award in Wireless Communications; the 2011 IEEE Wireless Communications Technical Committee Recognition Award; the 2011 IEEE Communications Society Harold Sobol Award; the 2010 Purdue University Outstanding Electrical and Computer Engineer Award; the 2009 IEEE Marconi Prize Award in Wireless Communications; the 2007 IEEE Communications Society Joseph LoCicero Publications Exemplary Award; and over 15 IEEE Best Paper Awards. He is also recognized by Thomson Reuters as an ISI Highly Cited Researcher. He served as consultant for different organizations, including Huawei, ASTRI, ZTE, Nortel, PricewaterhouseCoopers, and Motorola. He is the Founding Editor-in-Chief of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and has served on the Editorial Board for other prestigious journals, including the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS–WIRELESS SERIES as the Editor-in-Chief. He has been involved in organizing a number of flagship international conferences and events. In addition to his active research and professional activities, he has been a dedicated teacher committed to excellence in teaching and scholarship. He is well recognized for his dedicated service to professional societies and in particular IEEE where he has served in many leadership positions. These include the Treasurer of IEEE Communications Society, IEEE Communications Society Vice-President for Conferences, the Chair of IEEE Committee on Wireless Communications, an Elected Member of IEEE Product Services and Publications Board, and IEEE Communications Society Vice-President for Technical Activities. He is currently the President of the IEEE Communications Society, the world's leading organization for communications professionals with headquarters in New York City and members in 162 countries. He is a fellow of HKIE.