Full-Angle Digital Predistortion of 5G Millimeter-Wave Massive MIMO Transmitters

Chao Yu, *Member, IEEE*, Jianxin Jing, Han Shao, Zhi Hao Jiang, *Member, IEEE*, Pinpin Yan, *Member, IEEE*, Xiao-Wei Zhu, *Member, IEEE*, Wei Hong, *Fellow, IEEE*, and Anding Zhu, *Senior Member, IEEE*

***Abstract*— In this paper, a full-angle digital predistortion (DPD) technique is proposed to linearize fifth-generation (5G) millimeter-wave (mmWave) massive multiple-input-multiple-output (mMIMO) transmitters with low implementation complexity. It is achieved by compensating the differences of power amplifiers (PAs) in different transmitter chains first and then adopting a common digital block to linearize the whole subarray. Based on this operation, all the transmitter chains can be efficiently linearized simultaneously, providing the merits of full-angle linearization including the main beam and sidelobes. To validate the proposed idea, an mmWave full-digital beam-forming transmitter has been developed, which is operated at the center frequency of 24.75–28.5 GHz to meet the 5G candidate frequency bands. Experimental results show that the proposed method can effectively linearize the mmWave mMIMO transmitter in all directions, which provides a promising linearization solution for 5G mMIMO beam-forming systems.**

***Index Terms*— Beam forming, digital predistortion (DPD), millimeter-wave, multiple-input-multiple-output (MIMO), power amplifier (PA).**

I. Introduction

MILLIMETER-WAVE (mmWave) frequency band has been widely accepted as one of the candidates for fifth-generation (5G) communication systems, due to its vastly available spectrum resources to support future large data throughput requirements [1], [2]. To exploit the advantage of these systems, many countries have released candidate frequency bands for 5G wireless system deployment, e.g., 28/39 GHz in the USA and 24/37 GHz in China. However, the signals transmitted at these frequency bands suffer from large path losses with limited link budget [3].

Manuscript received December 11, 2018; revised March 28, 2019; accepted May 11, 2019. Date of publication June 19, 2019; date of current version July 1, 2019. This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 61601117 and Grant 61861136002, in part by the Natural Science Foundation of Jiangsu Province under Grant BK20160698, and in part by Science Foundation Ireland (SFI) through the SFI-NSFC Partnership Program under Grant17/NSFC/4850. *(Corresponding author: Chao Yu.)*

C. Yu, Z. H. Jiang, and W. Hong are with the State Key Laboratory of Millimeter Waves, Southeast University, Nanjing 210096, China, and also with Purple Mountain Laboratories, Nanjing 211111, China (e-mail: chao.yu@seu.edu.cn).

J. Jing, H. Shao, P. Yan, and X.-W. Zhu are with the State Key Laboratory of Millimeter Waves, Southeast University, Nanjing 210096, China.

A. Zhu is with the School of Electrical and Electronic Engineering, University College Dublin, Dublin 4, Ireland (e-mail: anding.zhu@ucd.ie).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMTT.2019.2918450

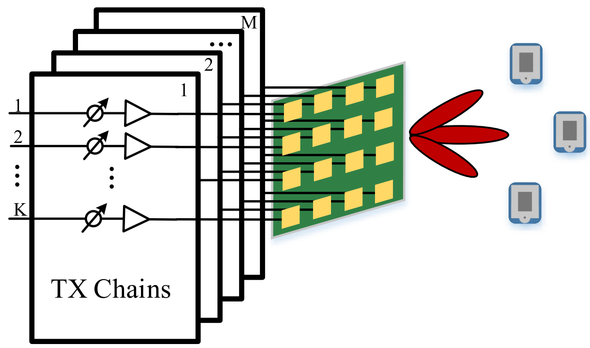


Fig. 1. Scenario of mMIMO beam-forming.

To overcome this issue, beam-forming that uses multiple antennas to form a highly focused beam pointing to a specific direction can be used to increase communication quality and save energy. Another technology, known as multiple-input-multiple-output (MIMO), uses spatial multiplexing coding to split the data stream into multiple channels to increase data capacity. Both technologies require multiple antennas and they are generally called MIMO systems. At mmWave, it is possible to build a large number (e.g., hundreds or thousands) of antenna elements within a small area. Massive MIMO (mMIMO), as shown in Fig. 1, thus has been treated as one of the key technologies in 5G that can provide high link-level gains to overcome path losses and enable super high-speed data transmissions.

Many mMIMO beam-forming architectures have been proposed in recent years [4], such as the analog beam-forming [5], hybrid analog-digital beam-forming [6], and full-digital beam-forming [7], [8]. Like conventional wireless systems, mMIMO systems also encounter linearity issue due to nonideal components used in the RF chains. Some linearity requirements may be relaxed in MIMO configuration, but the distortions induced by high nonlinearities will degrade the transmission signal quality and limit the system capacity, particularly if we want to maintain high power efficiency. For instance, power amplifiers (PAs) can introduce significant distortion if operated at high-efficiency mode.

Digital predistortion (DPD) can provide high linearity while operating the PA at relatively high efficiency, and it has been widely employed in 3G/4G systems because of its high-linearization performance and low-cost implementation [9], [10]. In past decades, many effective

DPD models and system architectures have been developed [11]–[14]. However, due to the complex system architecture, in mMIMO, the conventional DPD is no longer workable. First, the system needs to deal with not only the nonlinearity induced by the PA in each RF chain but also mutual coupling and crosstalk distortion among the multiple channels [15]–[17]. Second, in 5G systems, especially in small cell dense networks, the transmit power of the base stations become much lower, e.g., at watts or lower level, and in the meantime, the transmit signal bandwidth continues to increase. The existing DPD may consume a large amount of power because multiple blocks and multigigahertz digital signal processing are required. The benefits of the technique may be outweighed by the cost and energy consumption of implementing it. Therefore, new linearization strategies must be developed.

In the literature, there are several MIMO DPD approaches proposed [15]–[19], to deal with crosstalks between antennas and RF chains and these approaches, however, are only applicable for small-scale MIMO, e.g., 2×2, systems, but not for mMIMO systems where a very large number, e.g., hundreds, of channels are involved. Choi and Jeong [20] proposed to employ combined feedback that adds all the PA outputs to form single feedback for linearization of MIMO transmitters. This combined feedback approach was further studied in [21]. Hausmair *et al.* [22] proposed a DPD technique to compensate the PA nonlinearity, antenna crosstalk, and impedance mismatch. This technique can effectively reduce the DPD complexity by replacing the MIMO model to a dual-input PA model by introducing the additional extraction with either S-parameter measurements or the proposed identification procedure. Later, Luo *et al.* [23] further improved the model accuracy by employing the canonical piecewise linear technique. Liu *et al.* [24] proposed a beam-oriented DPD (BO-DPD) technique to achieve linearization of the transmitted signal in hybrid beam-forming mMIMO transmitters. Since only one DPD is used, the system complexity is significantly reduced. However, as only the combined signal at the main beam direction is considered, this approach thus can only linearize the PA at the main beam direction, while at other directions, the nonlinear distortions remain. Generally, the power of the sidelobe of the antenna array is usually only 10 dB lower than that of the main beam. The distortion in the sidelobe, thus, can cause large interference to other users. Therefore, it is desirable to remove the nonlinear distortion in all directions.

In this paper, we propose a full-angle DPD technique to linearize the mmWave mMIMO system. This is achieved by first compensating the PA differences in multiple RF chains with low-complexity tuning boxes and then linearizing all PAs by using a single-shared DPD block. This method will provide full-angle linearization with a simple DPD module. The rest of this paper is organized as follows. Section II will give a detailed analysis of the system requirements and review the existing MIMO linearization architectures. In Section III, the proposed DPD method will be introduced in detail. The demo system and experimental results are given in Section IV and V, respectively, followed by a conclusion in Section VI.

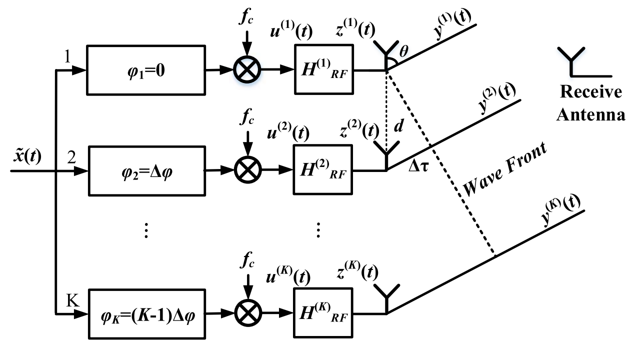


Fig. 2. System diagram for the subarray.

II. Application Scenarios and Existing Techniques

In this section, we first analyze the application scenarios and system requirements and then discuss the existing mMIMO DPD architectures.

A. System Analysis

In an mMIMO system, antenna arrays are usually divided into subarrays where each subarray transmits one data stream. Although there are many analyses on the MIMO systems in the literature, particularly in baseband signal processing, most of the discussions focus on the main beams only and ignore effects induced by the RF channels. In this section, we intend to give a detailed analysis of the mMIMO system in terms of nonlinear effects in both main beams and sidelobe directions. To simplify the derivation, we use the system shown in Fig. 2 as an example to conduct the system analysis.

To form a beam, the input signal is fed into multiple RF chains with different phase shifts. As shown in Fig. 2, $\widetilde{x}(t)$ is the baseband signal while $\Delta\varphi$ is the phase difference between neighbor chain. ω_c represents the carrier frequency. The inputs of the PAs $u^{(k)}(t)$ can be represented as

$$\begin{cases} u^{(1)}(t) = \text{Re}\{\widetilde{x}(t)e^{j\omega_c t}\} \\ u^{(2)}(t) = \text{Re}\{\widetilde{x}(t)e^{j[\omega_c t+\Delta\varphi]}\} \\ \quad\vdots \\ u^{(K)}(t) = \text{Re}\{\widetilde{x}(t)e^{j[\omega_c t+(K-1)\Delta\varphi]}\} \end{cases} \tag{1}$$

where $\text{Re}\{\cdot\}$ represent the real part of the signal.

The characteristics of the PAs can be represented by a general nonlinear function, e.g., Volterra series, in the time domain, as expressed in the following equation:

$$z^{(k)}(t) = \sum_{p=0}^{P} \int \cdots \int h_p^{(k)}(i_1, i_2, \ldots, i_p) \cdot \prod_{j=1}^{p} u^{(k)}(t - i_j), \quad k = 1, 2, \ldots, K \tag{2}$$

where $u^{(k)}(t)$ and $z^{(k)}(t)$ represents the input and output of the PA, and $h_p^{(k)}(i_1, i_2, \cdots, i_p)$ represents the Volterra kernels. k represents the kth RF chain. P is the nonlinear order and i_j is the time delay. Mutual coupling is not considered here.

For simplification, (2) can be expressed by Volterra operator, as shown in the following:

$$z^{(k)}(t) = \boldsymbol{H}_{\text{RF}}^{(\boldsymbol{k})}[u^{(k)}(t)] \quad k = 1, 2, \ldots, K \tag{3}$$

where $\boldsymbol{H}_{\text{RF}}^{(\boldsymbol{k})}$ represents the Volterra operator for the kth RF chain.

Substituting (1) into (3), we obtain

$$\begin{cases} z^{(1)}(t) = \boldsymbol{H}_{\mathbf{RF}}^{(\mathbf{1})}[\text{Re}\{\widetilde{x}(t)e^{j\omega_c t}\}] \\ z^{(2)}(t) = \boldsymbol{H}_{\mathbf{RF}}^{(\mathbf{2})}[\text{Re}\{\widetilde{x}(t)e^{j[\omega_c t+\Delta\varphi]}\}] \\ \quad\vdots \\ z^{(K)}(t) = \boldsymbol{H}_{\mathbf{RF}}^{(\boldsymbol{K})}[\text{Re}\{\widetilde{x}(t)e^{j[\omega_c t+(K-1)\Delta\varphi]}\}]. \end{cases} \tag{4}$$

To facilitate the derivation, we employ the baseband representation of the Volterra operator. Equation (4) can be rewritten as

$$\begin{cases} z^{(1)}(t) = \text{Re}\{\boldsymbol{H}^{(\mathbf{1})}[\widetilde{x}(t)]e^{j\omega_c t}\} \\ z^{(2)}(t) = \text{Re}\{\boldsymbol{H}^{(\mathbf{2})}[\widetilde{x}(t)]e^{j[\omega_c t+\Delta\varphi]}\} \\ \quad\vdots \\ z^{(K)}(t) = \text{Re}\{\boldsymbol{H}^{(\boldsymbol{K})}[\widetilde{x}(t)]e^{j[\omega_c t+(K-1)\Delta\varphi]}\} \end{cases} \tag{5}$$

where $\boldsymbol{H}^{(\boldsymbol{k})}, k = 1, 2, \ldots, K$ is the baseband representation of the Volterra operator $\boldsymbol{H}_{\text{RF}}^{(\boldsymbol{k})}$.

The output signals from the PAs will be radiated from the antenna array to the air as shown in Fig. 2 to form a beam in the far-field. At the wavefront plane, the output signals can be expressed as

$$\begin{cases} y^{(1)}(t) = z^{(1)}(t) * \delta(t) \\ y^{(2)}(t) = z^{(2)}(t) * \delta(t - \Delta\tau) \\ \quad\vdots \\ y^{(K)}(t) = z^{(K)}(t) * \delta(t - (K-1)\Delta\tau) \end{cases} \tag{6}$$

where $\delta(t)$ represents Dirac delta function, $\Delta\tau$ is the time delay between neighbor chains, and $*$ represents the convolution operation.

To simplify the derivation, let

$$f^{(k)}(t) = \boldsymbol{H}^{(\boldsymbol{k})}[\widetilde{x}(t)]e^{j\omega_c t}. \tag{7}$$

Based on the Fourier transform, (6) in the frequency domain can be expressed as

$$\begin{cases} Y^{(1)}(\omega) = \text{Re}\{F^{(1)}(\omega)\} \\ Y^{(2)}(\omega) = \text{Re}\{F^{(2)}(\omega)e^{j[\Delta\varphi-\omega\Delta\tau]}\} \\ \quad\vdots \\ Y^{(K)}(\omega) = \text{Re}\{F^{(K)}(\omega)e^{j[(K-1)(\Delta\varphi-\omega\Delta\tau)]}\} \end{cases} \tag{8}$$

where $Y^{(k)}(\omega)$ and $F^{(k)}(\omega)$ are the frequency-domain representation of $y^{(k)}(t)$ and $f^{(k)}(t)$.

The received signal $Y_{\text{receive}}(\omega)$ in the far-field can be represented as the sum of all the signals, that is,

$$Y_{\text{receive}}(\omega) = \sum_{k=1}^{K} \text{Re}\{F^{(k)}(\omega)e^{j(k-1)(\Delta\varphi-\omega\Delta\tau)}\}. \tag{9}$$

From (9), we can see that the received signal is maximized when

$$\Delta\varphi - \omega\Delta\tau = 0. \tag{10}$$

This leads that, in the beam-forming operation, when the receiver location, i.e., the desired wavefront angle θ_m, is known, the main beam will point to the receiver if we set $\Delta\varphi_m$ in the transmitter chain, satisfying

$$\Delta\varphi_m = \omega\Delta\tau_m = \omega d\cos\theta_m/c \tag{11}$$

where $\Delta\varphi_m$ and $\Delta\tau_m$ are the assigned phase and the time delay for the main beam, respectively, θ_m represents the angle of the main beam, d is the distance between neighbor antenna, and c is the speed of the wave. The signal at the main beam $Y_{\text{receive}-m}(\omega)$ can then be obtained

$$Y_{\text{receive}-m}(\omega) = \sum_{k=1}^{K} \text{Re}\{F^{(k)}(\omega)\} = \text{Re}\left\{\sum_{k=1}^{K} F^{(k)}(\omega)\right\}. \tag{12}$$

In the time domain, we can obtain the received signal at the main beam $y_{\text{receive}-m}(t)$

$$\begin{aligned} y_{\text{receive}-m}(t) &= \text{Re}\left\{\sum_{k=1}^{K} f^{(K)}(t)\right\} \\ &= \text{Re}\left\{\sum_{k=1}^{K} \boldsymbol{H}^{(\boldsymbol{k})}[\widetilde{x}(t)]e^{j\omega_c t}\right\}. \end{aligned} \tag{13}$$

The baseband equivalent form can be expressed as

$$\widetilde{y}_{\text{receive}-m}(t) = \sum_{k=1}^{K} \boldsymbol{H}^{(\boldsymbol{k})}[\widetilde{x}(t)]. \tag{14}$$

From (14), we can see that the received signal at the main beam direction is the direct sum of all output signals.

At other directions, e.g., sidelobe, we denote the angle of them as θ_s, that is,

$$\Delta\varphi_m - \omega\Delta\tau_s = \Delta\varphi_m - \omega d\cos\theta_s/c \neq 0. \tag{15}$$

The received signal $Y_{\text{receive}-s}(\omega)$ becomes

$$\begin{aligned} Y_{\text{receive}-s}(\omega) &= \text{Re}\left\{\sum_{k=1}^{K} F^{(k)}(\omega)e^{j(k-1)(\Delta\varphi_m-\omega\Delta\tau_s)}\right\} \\ &= \text{Re}\left\{\sum_{k=1}^{K} F^{(k)}(\omega)e^{j\omega(k-1)d/c(\cos\theta_m-\cos\theta_s)}\right\}. \end{aligned} \tag{16}$$

Let $a = d/c(\cos\theta_m - \cos\theta_s)$, we can obtain

$$Y_{\text{receive}-s}(\omega) = \text{Re}\left\{\sum_{k=1}^{K} F^{(k)}(\omega)e^{j\omega(k-1)a}\right\}. \tag{17}$$

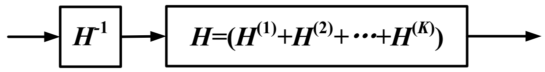


Fig. 3. Single-DPD inverse.

In the time domain, the received signal at sidelobe $y_{\text{receive}-s}(t)$

$$\begin{aligned} y_{\text{receive}-s}(t) &= \text{Re}\left\{\sum_{k=1}^{K} f^{(k)}(t) * \delta[t+(k-1)a]\right\} \\ &= \text{Re}\left\{\sum_{k=1}^{K} f^{(k)}[t+(k-1)a]\right\} \\ &= \text{Re}\left\{\sum_{k=1}^{K} \boldsymbol{H}^{(k)}[\widetilde{x}(t+(k-1)a)]e^{j\omega_c[t+(k-1)a]}\right\} \end{aligned} \quad (18)$$

and the baseband equivalent form can be represented by

$$\widetilde{y}_{\text{receive}-s}(t) = \sum_{k=1}^{K} \boldsymbol{H}^{(k)}[\widetilde{x}(t+(k-1)a)]e^{j\omega_c(k-1)a}. \quad (19)$$

Comparing (19) with (14), we can see that the signals received at the sidelobe direction are not only rotated by $j\omega_c(k-1)a$ but also delayed by $(k-1)a$ before being combined. This phenomenon is similar to the multichannel time delay issue discussed in [25].

B. Existing Techniques

There are mainly two types of DPD for linearizing mMIMO transmitters.

1) Single-DPD Approach: The first mMIMO DPD architecture is based on the single-DPD method. In this system, only the signal located at the main beam is the target for linearization. The multiple RF chains are treated together as one integrated system. In this case, all the PAs are combined and treated as a "new PA," as shown in Fig. 3. The single-input single-output DPD technique can, thus, be employed.

Because only one DPD block is required for each subarray, the implementation complexity can be significantly reduced, especially for hybrid beam-forming structure, in which the number of digital chains is less than the number of PAs and antennas. By utilizing the structure effectively, Liu *et al.* [24] proposed a BO-DPD method to deal with the linearity issue by only linearizing the "virtual" main beam signal instead of the signal captured at the receiver.

However, since only the sum of the output signals is used as the reference for training the DPD, it cannot guarantee each single PA is linearized. In fact, if the characteristic of the PA at each branch is different, the each individual output will not be linear, though the sum is linear. This leads that the output at other directions remains nonlinear, as shown in Fig. 4. This can be further derived from (14) and (19) and explained as follows. For example, assuming we have two PAs and the normalized baseband equivalent outputs at the wavefront are $\widetilde{y}^{(1)}(t)$ and $\widetilde{y}^{(2)}(t)$, at the main beam direction, the received signal is $\widetilde{y}_m(t) = \widetilde{y}^{(1)}(t) + \widetilde{y}^{(2)}(t)$, while at the sidelobe, the received signal is $\widetilde{y}_s(t) = \widetilde{y}^{(1)}(t)+\widetilde{y}^{(2)}(t)e^{j\omega_c a}$. If we use

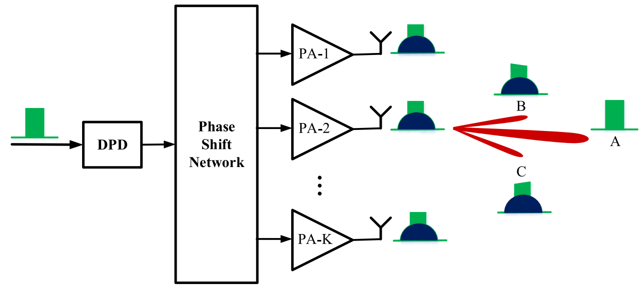


Fig. 4. BO single-DPD linearization approach.

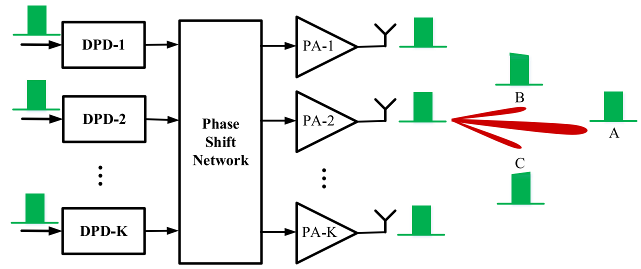


Fig. 5. Multi-DPD linearization approach.

$\widetilde{y}_m(t)$ as the reference signal to model the DPD, the output at the main beam direction can be linearized; however, at the sidelobe, unless $\widetilde{y}_s(t) = \alpha\widetilde{y}_m(t)$, where α is a scaling factor, the signal at the sidelobe cannot be simultaneously linearized. In fact, $\widetilde{y}_s(t) = \alpha\widetilde{y}_m(t)$ only occurs when $\widetilde{y}^{(1)}(t) = \widetilde{y}^{(2)}(t)$. On the other hand, in a real system, $\widetilde{y}^{(1)}(t) \neq \widetilde{y}^{(2)}(t)$, this single-DPD method, therefore, suffers an intrinsic disadvantage that the signal at the sidelobe, which is located at different angles, will not be simultaneously linearized, leading to inevitable degradation of the system performance in the neighbor areas. In [24], an angle broadening technique was proposed via averaging, which can reduce the distortion at neighbor directions but the linearization performance at the main beam is compromised.

This method also encounters problems when mutual coupling occurs. For instance, if there are leakages from one PA output to the other, e.g., the coupled PA outputs will be $\widetilde{y'}^{(1)}(t) = \widetilde{y}^{(1)}(t)+\beta_1\widetilde{y}^{(2)}(t)$ and $\widetilde{y'}^{(2)}(t) = \widetilde{y}^{(2)}(t)+\beta_2\widetilde{y}^{(1)}(t)$, where β_1 and β_2 are the leaking factors. The linearization with $\widetilde{y}^{(1)}(t) + \widetilde{y}^{(2)}(t)$ will not work for $\widetilde{y'}^{(1)}(t) + \widetilde{y'}^{(2)}(t)$ unless $\beta_1 = \beta_2$ or $\widetilde{y}^{(1)}(t) = \widetilde{y}^{(2)}(t)$.

2) Multi-DPD Approach: To alleviate this issue, another architecture is to use multiple DPD blocks, as shown in Fig. 5. In this structure, a separate DPD is assigned to each RF chain individually. The input signal for each chain will be predistorted in baseband before feeding into the PA and each antenna element. After this operation, all the nonlinearity of RF chains will be effectively removed. Therefore, all the beams including the main beam and sidelobe can be linearized simultaneously.

To resolve mutual coupling and crosstalk issue, Hausmair *et al.* [22] proposed a dual-input DPD model to reduce the multiple inputs to only two inputs for each DPD block. In this model, one input is the same as the conventional single-input single-output system, and the other is taken from the combination for the other inputs, which

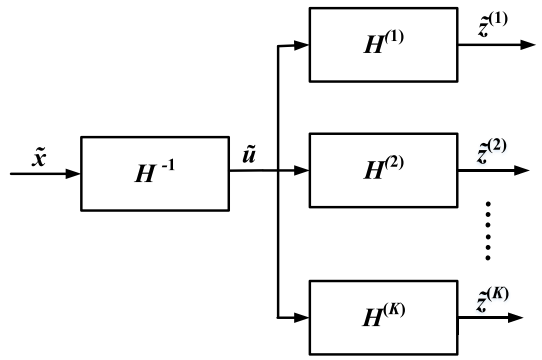


Fig. 6. System diagram in the practical scenario.

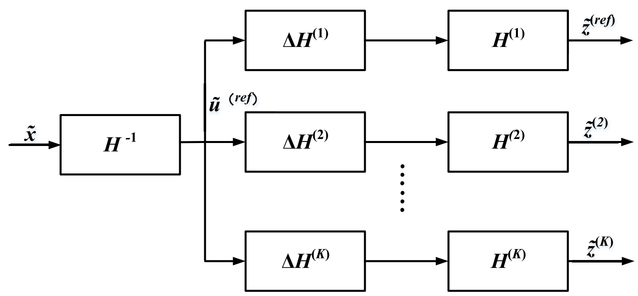


Fig. 7. System diagram of the proposed idea.

is generated based on a weighting factor. This method can effectively deal with the nonlinearity caused by antenna crosstalk. Nevertheless, the multi-DPD approach requires a separate DPD block for each RF chain, which will result that the complexity increases with the number of chains. This approach is also only suitable for full digital beamforming. In analog or hybrid beamformer, multiple RF chains usually share one digital input path, and thus there is no separate digital path available for each PA.

III. Proposed Technique

In this section, we will propose a new linearization structure for mMIMO beam-forming systems that can linearize the signals in all directions.

As analyzed in Section II, from (14) and (19), it can be seen that the signal received at different directions depends on the delay a between neighboring channels and the output signals radiated from the antennas are weighted by $e^{j\omega_c(k-1)a}$ before being combined. This leads that the linearization achieved at one receiver direction cannot be applied at other directions unless the output signals at all PAs are the same.

In a practical system, the characteristics of the PAs in different RF chains are inevitably different from each other, as shown in Fig. 6, where $\boldsymbol{H}^{(1)} \neq \boldsymbol{H}^{(2)} \neq \cdots \neq \boldsymbol{H}^{(K)}$, due to variations in design, fabrication, or configuration.

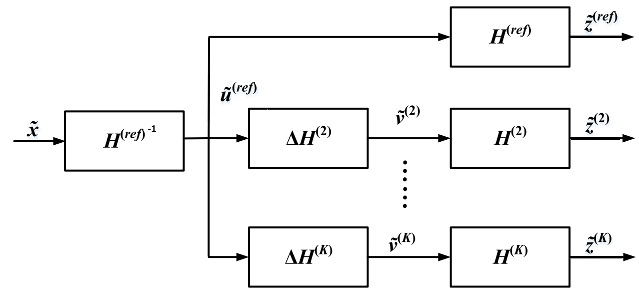


Fig. 8. Practical implementation block diagram of the proposed idea.

A. Full-Angle Linearization Architecture

To resolve this issue, we propose to introduce a tuning box in the RF chain before the PA, as shown in Fig. 7. If the tuning boxes can compensate the differences between the channels, i.e., $\boldsymbol{H}^{(1)}\Delta\boldsymbol{H}^{(1)} = \boldsymbol{H}^{(2)}\Delta\boldsymbol{H}^{(2)} = \cdots = \boldsymbol{H}^{(K)}\Delta\boldsymbol{H}^{(K)}$, a common single DPD can then be adopted to linearize all the PAs simultaneously.

To reduce the system complexity, one of the RF chains can be chosen as the reference chain, namely, we can make

$$\boldsymbol{H}^{(\mathbf{ref})} = \boldsymbol{H}^{(2)}\Delta\boldsymbol{H}^{(2)} = \cdots = \boldsymbol{H}^{(K)}\Delta\boldsymbol{H}^{(K)} \tag{20}$$

and then the block diagram in Fig. 7 can be simplified to the one shown in Fig. 8.

To derive the model for the required DPD system, two steps will be required. The first step is to extract the model for the reference DPD, $\boldsymbol{H}^{(\mathbf{ref})-\mathbf{1}}$, which can be obtained by using the conventional indirect learning approach for the selected RF chain, that is,

$$\boldsymbol{H}^{(\mathbf{ref})}\boldsymbol{H}^{(\mathbf{ref})-\mathbf{1}} = 1. \tag{21}$$

For simplicity, a memory polynomial (MP) model can be used for derivation

$$\widetilde{u}^{(\mathrm{ref})} = \sum_{p=0}^{P-1}\sum_{m=0}^{M} c_{p,m}^{(\mathrm{ref})} \cdot \widetilde{x}(n-m)\left|\widetilde{x}(n-m)\right|^{p}. \tag{22}$$

In the matrix form, (22) can be rewritten as

$$\boldsymbol{u}^{(\mathbf{ref})} = \mathbf{X}\boldsymbol{C}_{\mathbf{DPD}}^{(\mathbf{ref})} \tag{23}$$

where

$$\begin{aligned}\boldsymbol{u}^{(\mathbf{ref})} &= [\widetilde{u}^{\mathrm{ref}}(n), \widetilde{u}^{\mathrm{ref}}(n-1), \cdots, \widetilde{u}^{\mathrm{ref}}(n-N)]^{T}\\ \boldsymbol{C}_{\mathbf{DPD}}^{(\mathbf{ref})} &= \left[\widetilde{c}_{0,0}^{(\mathrm{ref})}, \widetilde{c}_{0,1}^{(\mathrm{ref})}, \cdots, \widetilde{c}_{P-1,M}^{(\mathrm{ref})}\right]^{T}\end{aligned}$$

and the matrix $\mathbf{X}$, shown at the bottom of this page. All the other blocks can be expressed in the similar way, for example,

$$\boldsymbol{z}^{(\mathbf{ref})} = \mathbf{U}^{(ref)}\boldsymbol{C}^{(\mathbf{ref})} \tag{24}$$

where $\boldsymbol{z}^{(\mathbf{ref})}$ and $\boldsymbol{C}^{(\mathbf{ref})}$ are the output and coefficient vectors, and $\mathbf{U}^{(\mathrm{ref})}$ is the matrix built from the nonlinear terms of the reference PA model.

$$\mathbf{X} = \begin{bmatrix} \widetilde{x}(n) & \widetilde{x}(n-1) & \cdots & \widetilde{x}(n-M)\left|\widetilde{x}(n-M)\right|^{P-1} \\ \widetilde{x}(n-1) & \widetilde{x}(n-2) & \cdots & \widetilde{x}(n-M-1)\left|\widetilde{x}(n-M-1)\right|^{P-1} \\ \vdots & \vdots & \ddots & \vdots \\ \widetilde{x}(n-N) & \widetilde{x}(n-N-1) & \cdots & \widetilde{x}(n-N-M)\left|\widetilde{x}^{(}n-N-M)\right|^{P-1} \end{bmatrix}$$

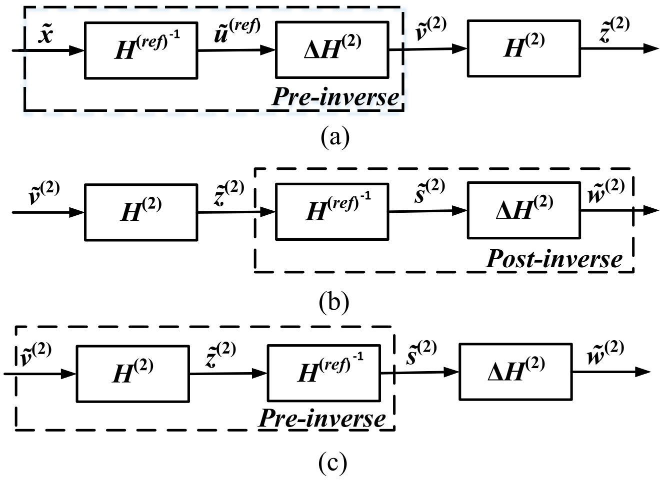


Fig. 9. System diagram for the derivation of tuning boxes.

Based on the indirect learning approach, if $\boldsymbol{H}^{(\mathbf{ref})-\mathbf{1}}$ and $\boldsymbol{H}^{(\mathbf{ref})}$ inverse each other, swapping input and output, (24) can be rewritten as

$$\boldsymbol{u}^{(\mathbf{ref})} = \mathbf{Z}^{(\mathrm{ref})}\boldsymbol{C}_{\mathbf{DPD}}^{(\mathbf{ref})}. \tag{25}$$

Because the model is linear-in-parameter, $\boldsymbol{C}_{\mathbf{DPD}}^{(\mathbf{ref})}$ can be obtained by using the least squares (LS), that is,

$$\boldsymbol{C}_{\mathbf{DPD}}^{(\mathbf{ref})} = \left(\mathbf{Z}^{(\mathrm{ref})^H}\mathbf{Z}^{(\mathrm{ref})}\right)^{-1}\mathbf{Z}^{(\mathrm{ref})^H}\boldsymbol{u}^{(\mathbf{ref})}. \tag{26}$$

The next step is to extract the model for the tuning box. For better illustration, we use the second-RF chain as an example and redraw it in Fig. 9.

Our goal is to linearize $\boldsymbol{H}^{(\mathbf{2})}$ to make $\tilde{z}(n) = \tilde{x}(n)$. To achieve this, the transfer function of the cascaded two boxes $\boldsymbol{H}^{(\mathbf{ref})-\mathbf{1}}$ and $\Delta\boldsymbol{H}^{(\mathbf{2})}$ together must be the preinverse of $\boldsymbol{H}^{(\mathbf{2})}$, shown in Fig. 9(a). According to the Pth-order inverse theory, the preinverse can be the same as the postinverse. In other words, the model coefficients in DPD module can be extracted from the postinverse, which is equivalent to preinverse. The cascaded two boxes can, thus, be moved after the box of $\boldsymbol{H}^{(\mathbf{2})}$, as shown in Fig. 9(b). By regrouping $\boldsymbol{H}^{(\mathbf{2})}$ and $\boldsymbol{H}^{(\mathbf{ref})-\mathbf{1}}$, we can find that the cascaded $\boldsymbol{H}^{(\mathbf{2})}$ and $\boldsymbol{H}^{(\mathbf{ref})-\mathbf{1}}$ can be treated as the preinverse of $\Delta\boldsymbol{H}^{(\mathbf{2})}$, shown in Fig. 9(c). To find $\Delta\boldsymbol{H}^{(\mathbf{2})}$, we can do the following.

First, we pass the signal $\tilde{v}^{(2)}(n)$ through $\boldsymbol{H}^{(\mathbf{2})}$ to obtain $\tilde{z}^{(2)}(n)$, that is,

$$z^{(2)} = \mathbf{V}^{(2)}\boldsymbol{C}^{(\mathbf{2})} \tag{27}$$

where $z^{(\mathbf{2})}$ and $\boldsymbol{C}^{(\mathbf{2})}$ are the output and coefficient vectors, and $\mathbf{V}^{(2)}$ is the matrix built from the nonlinear terms of the PA model using $\tilde{v}^{(2)}(n)$.

We then pass $\tilde{z}^{(2)}(n)$ through $\boldsymbol{H}^{(\mathbf{ref})^{-1}}$ to obtain $\tilde{s}^{(2)}(n)$, that is,

$$s^{(2)} = \mathbf{Z}^{(2)}\boldsymbol{C}_{\mathbf{DPD}}^{(\mathbf{ref})} \tag{28}$$

where $s^{(\mathbf{2})}$ and $\boldsymbol{C}_{\mathbf{DPD}}^{(\mathbf{ref})}$ are the output and coefficient vectors, and $\mathbf{Z}^{(2)}$ is the matrix built from the nonlinear terms of $\boldsymbol{H}^{(\mathbf{ref})-\mathbf{1}}$ using $\tilde{z}^{(2)}(n)$.

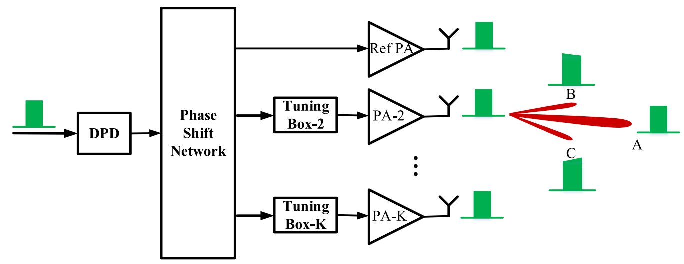


Fig. 10. Proposed full-angle DPD structure.

Swapping input and output, i.e., using $\tilde{s}^{(2)}(n)$ as input and $\tilde{v}^{(2)}(n)$ as output, we obtain

$$\boldsymbol{v}^{(2)} = \mathbf{S}^{(2)}\Delta\boldsymbol{C}^{(\mathbf{2})} \tag{29}$$

where $\Delta\boldsymbol{C}^{(\mathbf{2})}$ is the coefficients vector of $\Delta\boldsymbol{H}^{(\mathbf{2})}$ and it can be found by using LS, that is,

$$\Delta\boldsymbol{C}^{(\mathbf{2})} = \left(\mathbf{S}^{(2)^H}\mathbf{S}^{(2)}\right)^{-1}\mathbf{S}^{(2)^H}\boldsymbol{v}^{(2)}. \tag{30}$$

Following the same procedure, we can find the coefficients for all the other tuning boxes.

Based on this proposed architecture, the original input signal will be first predistorted by the reference DPD box, and then be phase-shifted for the purpose of beam-forming. The signals will be further processed by the tuning boxes, according to the nonlinearity of each channel and then be fed into the transmitter chains, and finally amplified and radiated into free space. After linearization, all the channels become linear and the signals combined in different angles, thus, will also be linear. Here, we denote this linearization approach as "full-angle DPD," as shown in Fig. 10.

It is worth mentioning that, after linearization, in the main beam direction, the output will be a scaled version of the original input, that is,

$$\tilde{y}_{\mathrm{receive}-m}(t) = K\tilde{x}(t) \tag{31}$$

while, in other directions, the output is a sum of delayed and rotated versions of the original signal, for example,

$$\tilde{y}_{\mathrm{receive}-s}(t) = \sum_{k=1}^{K}\tilde{x}(t+(k-1)a)e^{j\omega_c(k-1)a}. \tag{32}$$

This is because there are delays between different channels when combing in sidelobe direction. These delays can introduce linear distortion to the signal but it does not affect the system performance because sidelobe signals are usually not used as the received signal for users and the linear distortion does not create spectrum regrowth which does not interfere other channels.

As discussed earlier, the mutual coupling between channels can affect system performance. However, with the proposed approach, all PAs are simultaneously linearized. In other words, all the PA outputs become linear and the same. In this case, even if the coupling still exists, e.g., caused by antenna crosstalk, its effect is minimized since the crosstalk only introduces linear distortion while no nonlinear distortion will be generated. Therefore, this full-angle DPD also can

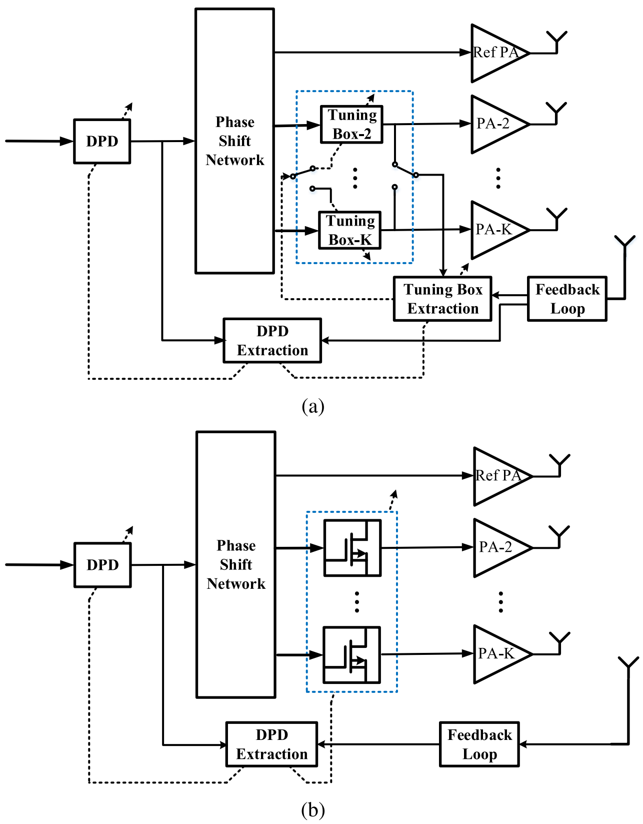


Fig. 11. Diagram for DPD implementation with (a) digital tuning box and (b) analog tuning box.

relieve mutual coupling issue. We will further validate this in Section V.

B. DPD Implementation

Depending on which system the proposed solution will be used in and how the tuning boxes are implemented, the building blocks of the proposed DPD are shown in Fig. 11.

In a fully digital beam-forming system, the DPD can be constructed as shown in Fig. 11(a), where both the common DPD block and the tuning boxes can be implemented in digital baseband. A feedback loop is required to capture the data from PA outputs. This can be conducted by using a shared receiver chain connected to the PA outputs or via over-the-air (OTA) test. First, the output of reference PA will be received by the feedback loop, which will be processed in the DPD extraction module to obtain the shared DPD coefficients. Then, the output of PA-2 will be captured and fed into the tuning box extraction module to obtain the coefficients for tuning Box-2. Following the same procedure, the DPD coefficients for other tuning boxes can be effectively obtained. After these two steps, all the coefficients can be updated. As the measurement results are shown in Section V, the differences between the channels usually are relatively small, and thus, the tuning box shown in Fig. 11(a) can be realized with a simple structure, such as a low-order MP model or a memoryless polynomial function.

In a hybrid beam-forming system, the DPD block is implemented in digital baseband while the tuning boxes can be implemented with analog circuits in RF, similar to that used in [26]. To include memory effects, an analog MP predistorter can be used [27]. In an analog implementation, the model extraction can be the same as that used in the digital one. As shown in Fig. 11(b), the tuning parameters can be extracted by using the DPD extraction block.

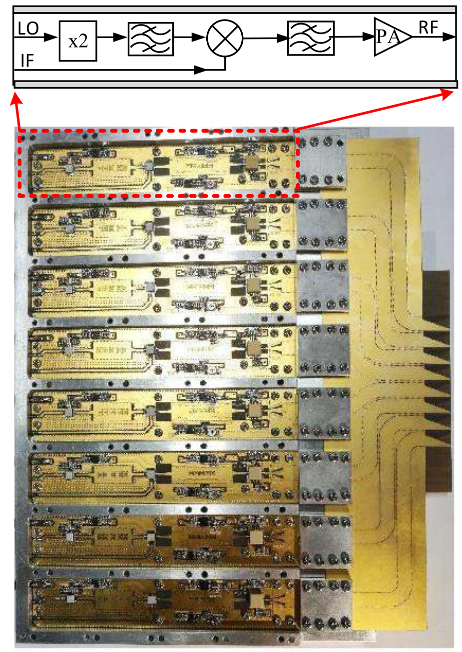


Fig. 12. Designed mmWave mMIMO array.

Compared with the analog implementation, the digital version is expected to have higher precision and with more flexible adaptions, but it is only suitable for the systems where each RF chain has its own digital baseband unit. The analog implementation may have lower accuracy but it can be directly inserted in front of PA in the RF chain, which can be used in any beam-forming systems. There is a tradeoff between analog and digital implementations, mainly depending on the application.

IV. System Design and Test Bench Setup

To validate the proposed idea, an mmWave mMIMO transmitter was developed and a test bench was setup.

A. mmWave mMIMO System Design

To cover 5G frequency band (24.75–28.5 GHz), an mmWave mMIMO system was designed, as shown in Fig. 12, where each transmitter chain contains a frequency multiplier, a passive mixer, two substrate integrated waveguide (SIW) bandpass filters, and a PA. The mixer is used to upconvert the signal from the intermediate frequency (IF) to RF. Since the passive mixer requires both a high-frequency and high-power local oscillator (LO) driver, a frequency multiplier is employed to provide LO signal with good noise performance. Also, an LO SIW bandpass filter is used at the output of the frequency multiplier to achieve high LO-leakage rejection. Similarly, an RF SIW filter is employed to suppress the harmonics and spurs of the mixer's RF output. Finally, the RF signal is fed into the PA. Eight transmitter chains in individual housings are mounted in parallel inside one metal shielded enclosure. SIW transmission lines serve as low-loss

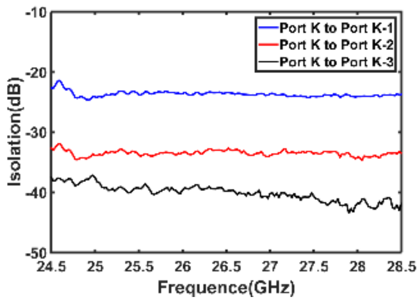


Fig. 13. Isolation measurement of the mMIMO array.

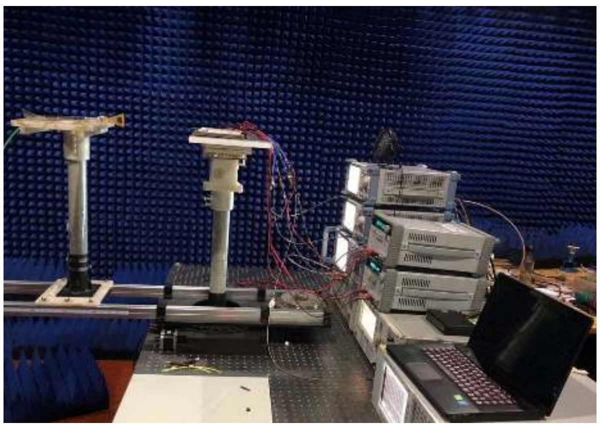


Fig. 14. Test bench setup.

interconnections between the mmWave transmit frontend and the antenna array. The transmission lines were bent to satisfy the element space requirement of the antenna array. Here, the array distance was set as 6 mm for the designed frequency band. For wideband operation, a tapered slot antenna array [28] was chosen due to its inherent property of high isolation between antenna elements. Furthermore, compared to the conventional design, we have optimized the shape of the elements and the distance between the elements to further reduce the mutual coupling. The RF substrate is Taconic TLY-5 with a permittivity of 2.2 and the thickness of 0.254 mm.

The isolation characteristics were measured using a Keysight PNA-X N5247A network analyzer and the results are shown in Fig. 13, where we can see that the isolation between the nearest neighbor ports (e.g., port 1 and port 2) is more than 23 dB at 27 GHz. The isolation between other distant ports is even higher. Therefore, the mutual couplings between different ports are relatively low.

B. Test Bench Setup

Based on the designed system, a DPD test bench can be set up as shown in Fig. 14. Due to limited signal sources available, only four of the eight transmitter chains shown in Fig. 12 were used. Four baseband four-carrier 20 and 40-MHz long term evolution (LTE) input signals with a peak-to-average power ratio of 7.27 dB were generated with the software MATLAB in PC. These signals passed through the common DPD module before being fed into different tuning boxes. These signal were then downloaded to four signal channels provided by

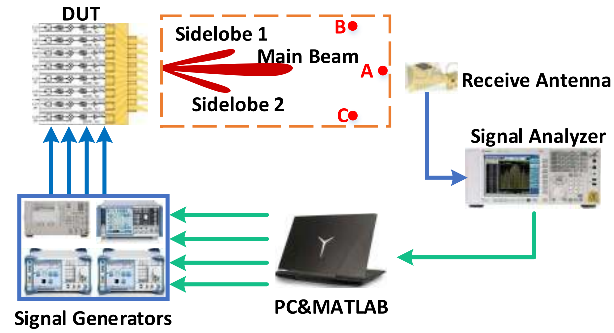


Fig. 15. Test setup for single-DPD method validation.

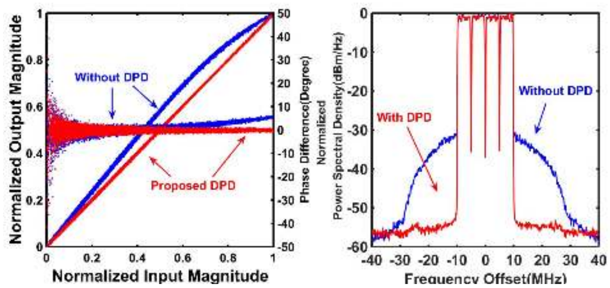
one dual-channel signal generator (R&S SMW200A) and two single-channel signal generators (R&S SMBV100A). Next, the four signals were upconverted to IF at 5.5 GHz and fed into the designed RF frontend. In this module, all signals were again upconverted to 27 GHz by four 10.75-GHz LO signals generated by a signal generator (Keysight E8267D) with a 1-to-4 power divider and fed into four Class-AB PAs with the average output power of around 14-dBm and 1-dB compression point of 21 dBm. Finally, the outputs of the transmitter chain were fed into the antenna elements to form the desired radiation pattern. In the receiver side, a horn antenna was employed for OTA test and a spectrum analyzer (Keysight N9030A) was utilized to capture the output through OTA test. Both the outputs and the input were sent back to the PC for DPD model extraction.

V. Measurement Results

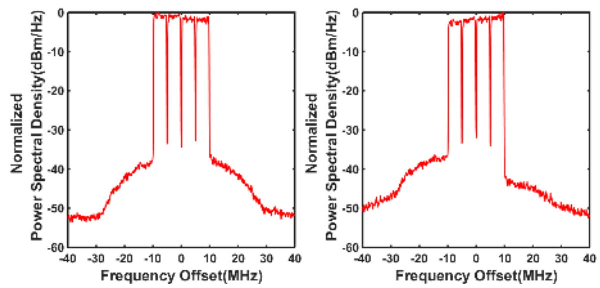
To verify the proposed idea, several experiments have been carried out and the results are presented in this section.

A. Single-DPD Method Validation

As mentioned in Section II, the single-DPD method treats the PAs in different chains as one virtual PA and linearizes it using one DPD, which results in that the main beam and sidelobe located at different angles are not simultaneously linearized. To verify it, a single BO-DPD test was implemented as shown in Fig. 15, where the device under test (DUT) generates one main beam with proper phase assignments in signal generators. The main beam is pointed at position A, and two sidelobes are captured at position B and C. The model is set with the nonlinear order of seven and the memory length of two. Fig. 16 shows the measured linearization performance, including amplitude modulation (AM)-to-AM (AM/AM) curve, AM-to-phase modulation (AM/PM), and normalized power spectral density. It can be seen that the nonlinear distortion at the main beam can be effectively removed as shown in Fig. 16(a), but the ones at the sidelobes still exist as shown in Fig. 16(b). The detailed measurement results are listed in Table I, where it can be seen that the adjacent channel power ratio (ACPR) value of the main beam can be improved from −34.7/−35.0 to −54.0/−53.7 dBc, and normalized mean square error (NMSE) reaches −38.3 dB. However, the ACPRs for the sidelobes, it only reaches −40.7/−40.0 and −37.6/−43.5 dBc, respectively. As derived in Section II, linear distortions are expected to appear in sidelobe even with linear



(a)



(b)

Fig. 16. Results of the single BO-DPD method for OTA main beam and sidelobe tests. (a) Main beam at Position A. (b) Sidelobe at Position B (left) and Position C (right).

TABLE I
MEASURED PERFORMANCE WITH MAIN BEAM AND SIDELOBES

	ACPR (dBc)	NMSE (dB)
Position A wo. BO-DPD	-34.7/-35.0	-16.6
Position A w. BO-DPD	-54.0/-53.7	-38.3
Position B w. BO-DPD	-40.7/-40.0	N/A
Position C w. BO-DPD	-37.6/-43.5	N/A

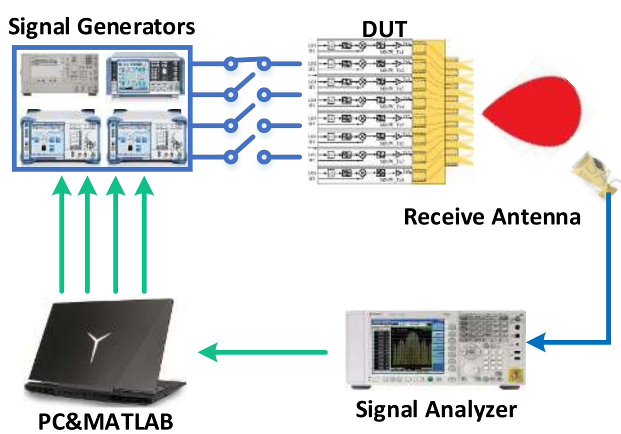


Fig. 17. Test setup for single RF chain validation.

RF chains, NMSE values are, therefore, not meaningful and, thus, are omitted in Table I.

B. Proposed Method With Single RF Chain Validation

In this section, the proposed method has been validated for each RF chain by using the test setup shown in Fig. 17. The baseband signal was fed into each RF chain one by one and the output signal was received by an OTA receiving antenna. Both the input and output signals were processed in the PC. In other words, the DPD coefficients were calculated one by one, but these modules run together to form the beam.

First, one of the RF chains must be selected as the reference channel. In our tests, we found that the complexity of the tuning boxes is more or less the same no matter which chain is used as the reference. PA can be linearized using a smaller number of parameters if the nonlinearity is relatively weak. In order to minimize the total DPD complexity, the TX chain with the weakest nonlinearity was chosen as the reference chain, that is, RF chain 2 in this test. However, this may not always be the case. In practical operation, it might need further verification or optimization to find the best reference. Second, MP models were selected to linearize the reference channel with the nonlinear order of P and memory depth of M, denoted as “reference DPD.” Third, the nonlinear order and memory depth for the tuning box can be set according to the system requirement. It is worth mentioning that all the DPD coefficients are estimated by utilizing the indirect learning approach and operated in a closed loop with three to four iterations.

For comparison, we have done five tests with both 20- and 40-MHz signals: 1) all PAs were tested without DPD; 2) all PAs were linearized with only reference DPD; 3) all PAs were linearized with an independent conventional DPD; 4) all PAs were linearized with proposed DPD, including one shared DPD and several tuning boxes with memory; and 5) all PAs were linearized with proposed DPD, including one shared DPD and several tuning boxes without memory.

The measurement results of the 20-MHz signal are illustrated in Fig. 18, where we can see that the PAs can only be partly linearized if only the reference DPD is employed, depending on the similarity with the reference PA. The proposed DPD can achieve almost the same linearization performance as that of using a DPD in each chain. In addition, the DPD with memory tuning box can obtain slightly better performance than the one with a memoryless tuning box. The detailed performance for each PA is listed in Table II. The ACPR value for the proposed DPD with memory tuning box and memoryless tuning box can both reach below -54 and -50 dBc in 20 and 40 MHz, respectively, which is almost equal to the one achieved by using the conventional DPD in each chain. The NMSE values of the one with memoryless tuning box are around 2 dB less than the one with memory tuning box in 20-MHz measurement. With the bandwidth increase, the performance of the proposed method without memory degrades. The complexity comparison, including the coefficient number and the number of floating point operations (FLOPs), has been made between the proposed and conventional method by using the metric in [29], as shown in Table III. From the table, it can be seen that, compared to the conventional DPD, the proposed DPD can significantly reduce the number of the model coefficients in both 20- and 40-MHz scenarios, no matter whether the tuning box is memoryless or memory. Although the method with memoryless tuning box is only employed very few coefficients, it can still obtain very good performance, which has been validated in Table II. It is worth mentioning that the complexity reduction for the proposed method is mainly achieved by using

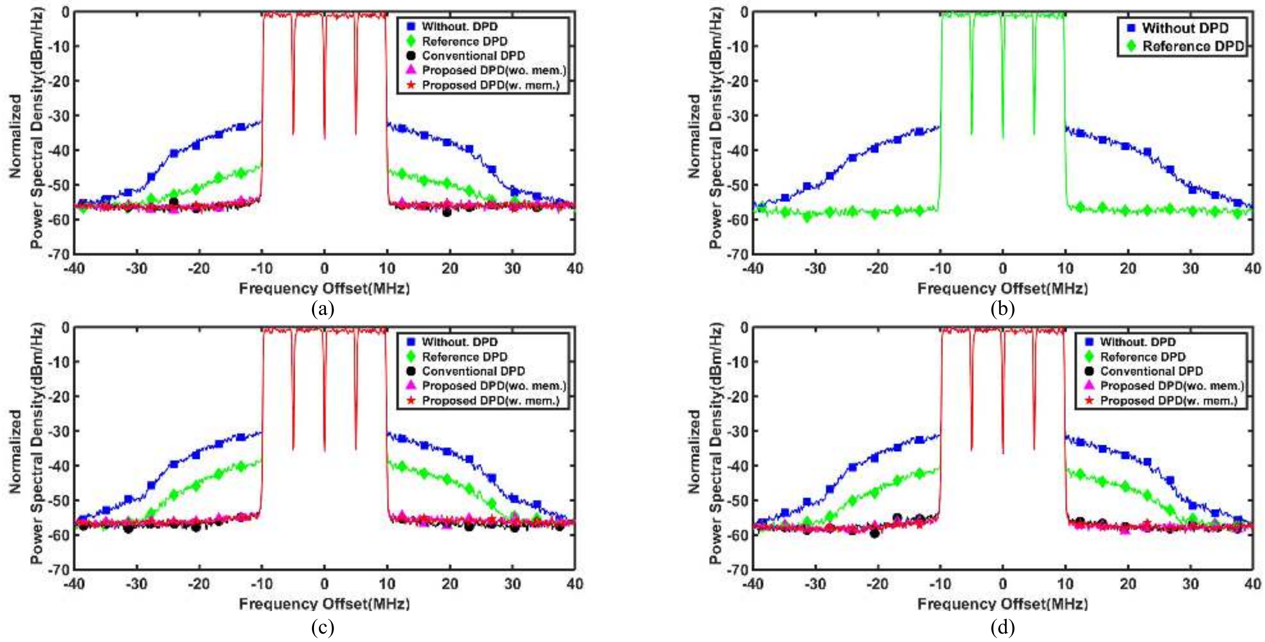


Fig. 18. Measured power spectral density of TX chains. (a) RF Chain 1. (b) RF Chain 2 (reference chain). (c) RF Chain 3. (d) RF Chain 4.

TABLE II
MEASURED PERFORMANCE FOR SINGLE RF CHAIN VALIDATION

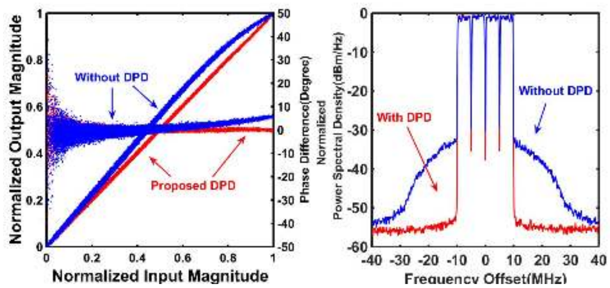
	Without DPD		Ref DPD		Conventional DPD		Proposed DPD Without Memory		Proposed DPD With Memory	
	ACPR	NMSE	ACPR	NMSE	ACPR	NMSE	ACPR	NMSE	ACPR	NMSE
Validation with 20 MHz signal										
RF Chain 1	-35.4/-36.2	-17.4	-48.3/-48.7	-34.7	-55.2/-55.3	-41.0	-54.8/-54.9	-39.3	-55.0/-54.9	-41.4
RF Chain 2 (Ref)	-36.8/-37.4	-20.3	-56.7/-56.2	-41.0	-	-	-	-	-	-
RF Chain 3	-33.9/-34.6	-16.7	-42.5/-42.8	-25.9	-55.2/-55.6	-41.7	-54.7/-54.6	-39.8	-54.8/-54.8	-41.5
RF Chain 4	-34.7/-35.6	-18.7	-44.5/-44.9	-29.0	-55.8/-56.2	-41.4	-56.4/-56.5	-38.0	-56.4/-56.5	-41.4
Validation with 40 MHz signal										
RF Chain 1	-35.8/-36.7	-20.3	-47.3/-47.7	-32.3	-52.0/-52.1	-39.3	-50.8/-50.4	-35.3	-51.5/-51.5	-38.3
RF Chain 2 (Ref)	-36.3/-36.3	-19.0	-53.0/-52.8	-40.1	-	-	-	-	-	-
RF Chain 3	-33.6/-34.1	-16.5	-44.1/-45.3	-28.0	-50.6/-50.7	-39.7	-50.2/-50.1	-32.9	-50.6/-50.5	-39.2
RF Chain 4	-33.8/-34.4	-16.9	-44.0/-43.4	-28.0	-51.2/-51.4	-38.9	-50.6/-50.6	-31.3	-51.4/-50.8	-37.2

TABLE III
DPD COMPLEXITY COMPARISON

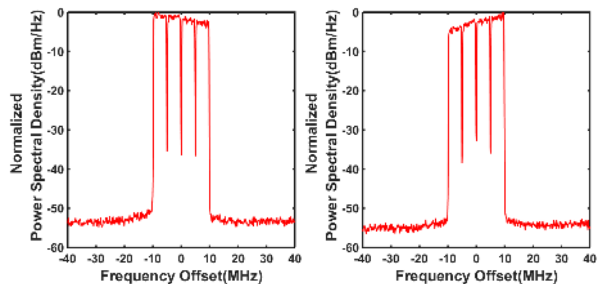
	Conventional DPD		Proposed DPD (with memery)			Proposed DPD (without memery)		
Channel Number	Model Parameter	Coeff. No. (FLOPs)	Model Parameter	Coeff. No. Ref. DPD (FLOPs)	Coeff. No. Tuning Box (FLOPs)	Model Parameter	Coeff. No. Ref. DPD (FLOPs)	Coeff. No. Tuning Box (FLOPs)
Validation with 20 MHz signal								
1	P=7,M=2	21(180)	P=3,M=1	N/A	6(24)	P=3,M=0	N/A	3(12)
2	P=7,M=2	21(180)	P=7,M=2	21(180)	N/A	P=7,M=2	21(180)	N/A
3	P=7,M=2	21(180)	P=5,M=1	N/A	10(64)	P=5,M=0	N/A	5(32)
4	P=7,M=2	21(180)	P=5,M=1	N/A	10(64)	P=5,M=0	N/A	5(32)
Validation with 40 MHz signal								
1	P=7,M=2	21(180)	P=3,M=1	N/A	6(24)	P=3,M=0	N/A	3(12)
2	P=7,M=2	21(180)	P=7,M=2	21(180)	N/A	P=7,M=2	21(180)	N/A
3	P=7,M=3	28(240)	P=5,M=1	N/A	10(64)	P=5,M=0	N/A	5(32)
4	P=7,M=2	21(180)	P=5,M=1	N/A	10(64)	P=5,M=0	N/A	5(32)

the cascade of two boxes, in which one box is responsible for the shared DPD and the other is responsible for tuning PA differences, and the complexity of the tuning block depends on the PA difference.

In this paper, we employed a digital tuning box to compensate for the difference between different RF chains. As mentioned in Section III-B, it is possible to replace the digital tuning boxes with an analog one, since the characteristics for



(a)



(b)

Fig. 19. Proposed DPD for OTA main-beam and sidelobe test. (a) Main beam at Position A. (b) Sidelobe at Position B (left) and Position C (right).

TABLE IV
MEASURED PERFORMANCE WITH OTA SIDELOBE LINEARIZATION

	ACPR (dBc)	NMSE (dB)
Position A wo. DPD	-36.0/-36.5	-18.4
Position A w. DPD	-53.8/-53.8	-38.4
Position B w. DPD	-51.0/-51.6	N/A
Position C w. DPD	-52.0/-51.5	N/A

memoryless or memory nonlinearity can also be effectively realized by using analog circuits.

C. Proposed Method With OTA Sidelobe Linearization

In this part, the proposed method with sidelobe linearization was validated, which used the same test setup as shown in Fig. 15. The measurement results are demonstrated in Fig. 19. Similar to the single-DPD method as shown in Fig. 16, the output in the direction of the main beam can be effectively linearized at position A as shown in Fig. 19(a). However, compared to the sidelobe performance shown in Fig. 16, it can be seen that the nonlinearity of the sidelobe at positions B and C can be effectively removed, which can realize the full-angle linearization. As it has been derived in Section III, some linear memory effects remain in the sidelobe after linearization, appearing as a frequency-dependent response in the spectra plots, as shown in Fig. 19(b). The detailed performance for the measurement has been listed in Table IV.

D. Proposed Method With OTA Beam Steering

In this part, the DUT with the beam-steering operation was implemented to verify if all the signals at beam directions can be linearized, as shown in Fig. 20. Three beam directions were employed for the test with specific phase configurations in the baseband. When the phase difference was tuned, the beam will be steered, and in the whole procedure, the DPD configuration remains the same. The measured AM/AM, AM/PM, and spectral density have been illustrated in Fig. 21. From the figure, it can be seen that the nonlinear distortion at different beams can be effectively removed without updating the DPD coefficients, which is very promising for the application of fast beam steering. The detailed performance for the measurement has been listed in Table V.

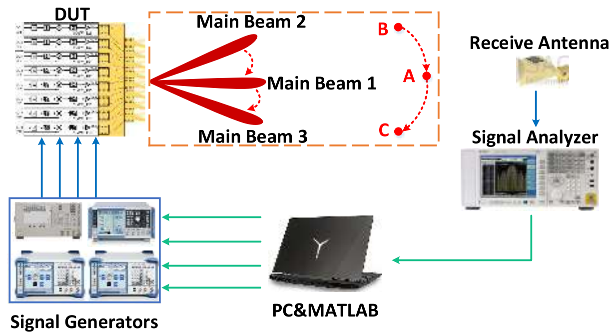
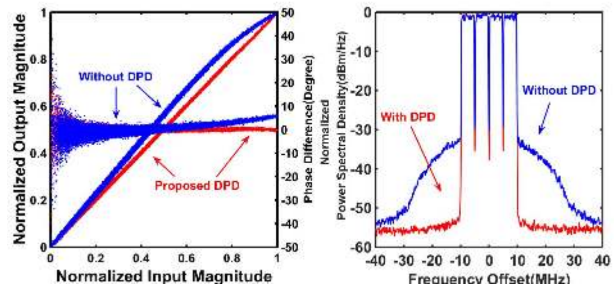
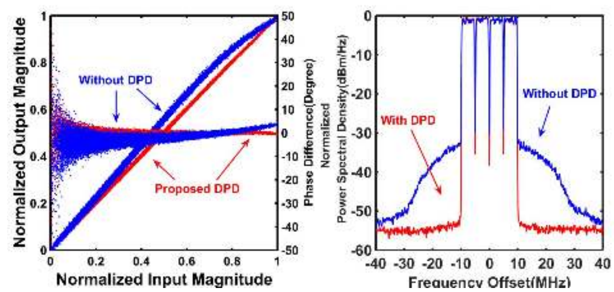


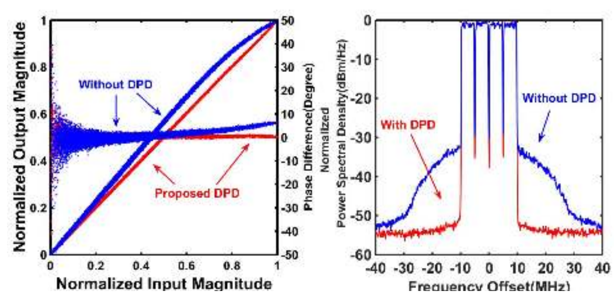
Fig. 20. Test setup for proposed DPD with OTA beam steering.



(a)



(b)



(c)

Fig. 21. Proposed DPD for OTA test with Beam steering. (a) Main beam at Position A. (b) Main beam at Position B. (c) Main beam at Position C.

TABLE V
MEASURED PERFORMANCE WITH OTA BEAM STEERING

	ACPR (dBc)	NMSE (dB)
Position A wo. DPD	-36.0/-36.5	-18.4
Position A w. DPD	-53.8/-53.8	-38.4
Position B wo. DPD	-36.1/-35.7	-18.9
Position B w. DPD	-53.3/-53.4	-37.6
Position C wo. DPD	-35.8/-36.5	-18.1
Position C w. DPD	-52.6/-52.7	-38.4

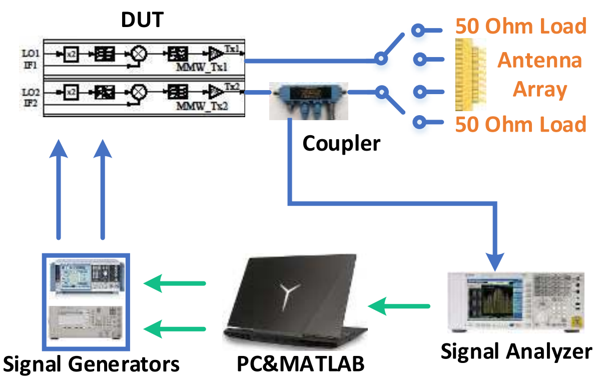


Fig. 22. Test setup for mutual coupling validation.

E. Mutual Coupling Validation

Because multiple channels are involved, the mutual coupling effect between antenna elements and RF chains can change the load of the PAs and affect the PA characteristics, and finally degrade the linearization performance. Several papers have addressed this issue, such as [22]. In this paper, since all the channels become linear after linearization, if only linear mutual coupling exists, e.g., that caused by antenna crosstalk, the linearization performance can still be maintained.

In this paper, we have done two tests to validate the mutual coupling effect as shown in Fig. 22: 1) both the output of RF chain 1 and the coupler output of RF chain 2 were connected to 50-Ω load and 2) both the output of RF chain 1 and the coupler output of RF chain 2 were connected to the antenna array, which generates the mutual coupling. The MP model with nonlinear order of seven and memory length of two was used. First, two sets of DPD coefficients will be extracted separately to linearize both RF chains. In the first test, only RF chain 1 was ON, and thus, the linearized output can be obtained with the corresponding DPD. Then, in the second test, both RF chains were ON. The signal at RF chain 2 generated the mutual coupling signal to RF chain 1 through antenna array. The measured performance with AM/AM curve, AM/PM curve, and power spectral density of the output is shown in Fig. 23. The detailed performance for the measurement has been listed in Table VI. From the table, it can be found that the ACPRs are almost the same. The NMSE value slightly drops from −41.3 to −40.7 dB for the one with an antenna array. Also, the coupling between two neighbor ports has been measured with the value of −23 dB. In the two cases, it proves that, although there is some mutual coupling effect, the linearization performance will not be affected significantly when the DPD coefficients are extracted by the proposed method.

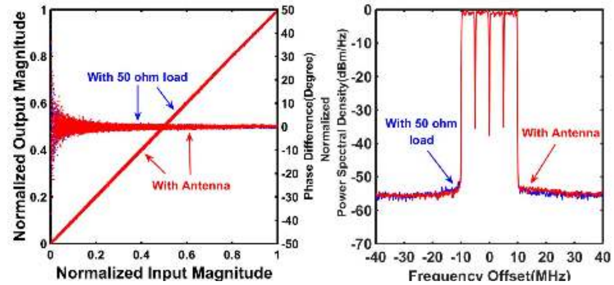


Fig. 23. Measured performance for mutual coupling validation.

TABLE VI
MEASURED PERFORMANCE FOR COUPLING VALIDATION

	Coupling (dB)	ACPR (dBc)	NMSE (dB)
With 50 ohm load	N/A	-53.8/-53.8	-41.3
With Antenna Array	-23	-54.3/-53.2	-40.7

VI. CONCLUSION

In this paper, a novel DPD architecture has been proposed to resolve the full-angle linearization for mmWave mMIMO beam-forming transmitters with low complexity. Based on the experimental validation, the proposed method can efficiently realize the linearization for both main beam and sidelobes, which appears to be a promising solution for 5G mMIMO applications.

It is worth mentioning that mmWave mMIMO systems intend to employ a large number of antennas to transmit signals with very wide bandwidths, e.g., hundreds of megahertz modulated signals. In this paper, we designed an mmWave front end operated at 24.75–28.5 GHz with a modulated bandwidth of 500 MHz. However, due to the limited availability of test instruments, we only can conduct DPD tests with four channels using 20- and 40-MHz modulated signals. Although narrowband signals and a smaller number of channels were used, the test results still effectively validated the proposed idea. This is because the main contribution of this work is proposing a two-box DPD architecture that first compensates the PA differences in multiple RF chains with low-complexity tuning boxes and then linearizes all PAs by using a single-shared DPD block. This architecture can be scaled to any number of channels and with any signal bandwidths. With a larger antenna array and a further increase of signal bandwidth, the PAs may exhibit stronger nonlinearity and the overall system may become much more complex, but the DPD system structure can remain the same. In other words, the proposed solution, namely, the two-box DPD structure, still works in these systems. Compared to the conventional approaches, the complexity reduction can even be further achieved in a larger array since the tuning box is simpler than the conventional DPD boxes, more tuning boxes are used, more reductions can be obtained.

REFERENCES

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

[2] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101–107, Jun. 2011.

[3] 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.

[4] W. Hong *et al.*, "Multibeam antenna technologies for 5G wireless communications," *IEEE Trans. Antennas Propag.*, vol. 65, no. 12, pp. 6231–6249, Dec. 2017.

[5] G. W. Kant, P. D. Patel, S. J. Wijnholds, M. Ruiter, and E. van der Wal, "EMBRACE: A multi-beam 20,000-element radio astronomical phased array antenna demonstrator," *IEEE Trans. Antennas Propag.*, vol. 59, no. 6, pp. 1990–2003, Jun. 2011.

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

[7] B. Yang, Z. Yu, J. Lan, R. Zhang, J. Zhou, and W. Hong, "Digital beamforming-based massive MIMO transceiver for 5G millimeter-wave communications," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 7, pp. 3403–3418, Jul. 2018.

[8] Y. Hu *et al.*, "A digital multibeam array with wide scanning angle and enhanced beam gain for millimeter-wave massive MIMO applications," *IEEE Trans. Antennas Propag.*, vol. 66, no. 11, pp. 5827–5837, Nov. 2018.

[9] F. M. Ghannouchi and O. Hammi, "Behavioral modeling and predistortion," *IEEE Microw. Mag.*, vol. 10, no. 7, pp. 52–64, Dec. 2009.

[10] N. Kelly, W. Cao, and A. Zhu, "Preparing linearity and efficiency for 5G: Digital predistortion for dual-band Doherty power amplifiers with mixed-mode carrier aggregation," *IEEE Microw. Mag.*, vol. 18, no. 1, pp. 76–84, Feb. 2017.

[11] J. Kim and K. Konstantinou, "Digital predistortion of wideband signals based on power amplifier model with memory," *Electron. Lett.*, vol. 37, no. 23, pp. 1417–1418, Nov. 2001.

[12] D. R. Morgan, Z. Ma, J. Kim, M. G. Zierdt, and J. Pastalan, "A generalized memory polynomial model for digital predistortion of RF power amplifiers," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3852–3860, Oct. 2006.

[13] A. Zhu, J. C. Pedro, and T. J. Brazil, "Dynamic deviation reduction-based Volterra behavioral modeling of RF power amplifiers," *IEEE Trans. Microw. Theory Techn.*, vol. 54, no. 12, pp. 4323–4332, Dec. 2006.

[14] A. Zhu, "Decomposed vector rotation-based behavioral modeling for digital predistortion of RF power amplifiers," *IEEE Trans. Microw. Theory Techn.*, vol. 63, no. 2, pp. 737–744, Feb. 2015.

[15] S. A. Bassam, M. Helaoui, and F. M. Ghannouchi, "Crossover digital predistorter for the compensation of crosstalk and nonlinearity in MIMO transmitters," *IEEE Trans. Microw. Theory Techn.*, vol. 57, no. 5, pp. 1119–1128, May 2009.

[16] P. M. Suryasarman and A. Springer, "A comparative analysis of adaptive digital predistortion algorithms for multiple antenna transmitters," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 62, no. 5, pp. 1412–1420, May 2015.

[17] S. Amin, P. Landin, P. Händel, and D. Rönnow, "Behavioral modeling and linearization of crosstalk and memory effects in RF MIMO transmitters," *IEEE Trans. Microw. Theory Techn.*, vol. 62, no. 4, pp. 810–823, Apr. 2014.

[18] Z. A. Khan, E. Zenteno, P. Händel, and M. Isaksson, "Digital predistortion for joint mitigation of I/Q imbalance and MIMO power amplifier distortion," *IEEE Trans. Microw. Theory Techn.*, vol. 65, no. 1, pp. 322–333, Jan. 2017.

[19] A. Abdelhafiz, L. Behjat, F. M. Ghannouchi, M. Helaoui, and O. Hammi, "A high-performance complexity reduced behavioral model and digital predistorter for MIMO systems with crosstalk," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 1996–2004, May 2016.

[20] S. Choi and E.-R. Jeong, "Digital predistortion based on combined feedback in MIMO transmitters," *IEEE Commun. Lett.*, vol. 16, no. 10, pp. 1572–1575, Oct. 2012.

[21] M. Abdelaziz, L. Anttila, A. Brihuega, F. Tufvesson, and M. Valkama, "Digital predistortion for hybrid MIMO transmitters," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 3, pp. 445–454, Jun. 2018.

[22] K. Hausmair, P. N. Landin, U. Gustavsson, C. Fager, and T. Eriksson, "Digital predistortion for multi-antenna transmitters affected by antenna crosstalk," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 3, pp. 1524–1535, Mar. 2018.

[23] Q. Luo, C. Yu, and X. Zhu, "A dual-input canonical piecewise-linear function-based model for digital predistortion of multi-antenna transmitters," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2018, pp. 559–562.

[24] X. Liu *et al.*, "Beam-oriented digital predistortion for 5G massive MIMO hybrid beamforming transmitters," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 7, pp. 3419–3432, Jul. 2018.

[25] Q. Luo, C. Yu, and X. Zhu, "Digital predistortion of phased array transmitters with multi-channel time delay," in *Proc. IEEE Topics Conf. RF/Microw. Power Amplif. Radio Wireless Appl.*, Jan. 2018, pp. 54–57.

[26] N. Rostomyan, J. A. Jayamon, and P. M. Asbeck, "15 GHz Doherty power amplifier with RF predistortion linearizer in CMOS SOI," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 3, pp. 1339–1348, Mar. 2018.

[27] F. Roger, "A 200 mW 100 MHz-to-4 GHz 11^{th}-order complex analog memory polynomial predistorter for wireless infrastructure RF amplifiers," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2013, pp. 94–95.

[28] K. Ebnabbasi, D. Busuioc, R. Birken, and M. Wang, "Taper design of Vivaldi and co-planar tapered slot antenna (TSA) by Chebyshev transformer," *IEEE Trans. Antennas Propag.*, vol. 60, no. 5, pp. 2252–2259, May 2012.

[29] A. S. Tehrani, H. Cao, S. Afsardoost, T. Eriksson, M. Isaksson, and C. Fager, "A comparative analysis of the complexity/accuracy tradeoff in power amplifier behavioral models," *IEEE Trans. Microw. Theory Techn.*, vol. 58, no. 6, pp. 1510–1520, Jun. 2010.

Chao Yu (S'09–M'15) received the B.E. degree in information engineering and the M.E. degree in electromagnetic fields and microwave technology from Southeast University (SEU), Nanjing, China, in 2007 and 2010, respectively, and the Ph.D. degree in electronic engineering from University College Dublin (UCD), Dublin, Ireland, in 2014.

He is currently an Associate Professor with the State Key Laboratory of Millimeter Waves, School of Information Science and Engineering, SEU, and also with Purple Mountain Laboratories, Nanjing. His current research interests include microwave and millimeter wave power amplifier modeling and linearization and 5G massive multiple-input-multiple-output (MIMO) RF system design.

Jianxin Jing received the B.E. degree in applied physics from the University of Electronic Science and Technology of China, Chengdu, China, in 2017, where he is currently pursuing the M.E. degree at the State Key Laboratory of Millimeter Waves.

His current research interests include modeling and linearization of 5G massive multiple-input–multiple-output (MIMO) systems.

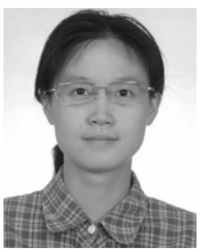
Han Shao was born in Xuzhou, China, in 1994. He received the B.S. degree in communication engineering from the Nanjing University of Science and Technology, Nanjing, China, in 2016. He is currently pursuing the M.S. degree in microwave technology at the State Key Laboratory of Millimeter Waves, Southeast University, Nanjing.

His current research interests include microwave and millimeter wave circuits and systems.

Zhi Hao Jiang (S'07–M'13) was born in Nanjing, China, in 1986. He received the B.S. degree in radio engineering from Southeast University, Nanjing, in 2008, and the Ph.D. degree in electrical engineering from Pennsylvania State University, State College, PA, USA, in 2013.

From 2013 to 2016, he was a Post-Doctoral Fellow with the Computational Electromagnetics and Antennas Research Laboratory, Department of Electrical Engineering, Pennsylvania State University. He is currently a Professor with the State Key Laboratory of Millimeter Waves, School of Information Science and Engineering, Southeast University, and also with Purple Mountain Laboratories, Nanjing. He has authored or coauthored more than 50 papers in peer-reviewed journals, more than 60 papers in conference proceedings, and 8 book chapters. He has also coedited *Electromagnetics of Body-Area Networks: Antennas, Propagation, and RF Systems* (Wiley/IEEE Press, 2016). He holds seven granted U.S. patents and one granted Chinese patent. His current research interests include microwave/millimeter-wave antennas and circuits, millimeter-wave systems, impedance surfaces, and metamaterials.

Dr. Jiang has served as the TPC Co-Chair or a TPC member for multiple international conferences. He was a recipient of the High-Level Innovative and Entrepreneurial Talent by Jiangsu, China, in 2017, the Thousands of Young Talents by China Government in 2016, the Honorable Mention of the 2013 IEEE AP-S International Symposium on Antennas and Propagation Student Paper Contest, and the 2012 A. J. Ferraro Outstanding Doctoral Research Award in Electromagnetics. He serves as an Associate Editor for *IET Communications* and as a Reviewer for more than 40 journals.



Pinpin Yan (M'11) received the B.S. degree in radio engineering and the M.S. and Ph.D. degrees in electromagnetic field and microwave technique from Southeast University, Nanjing, China, in 2000, 2004, and 2009, respectively.

Since 2000, she has been with the State Key Laboratory of Millimeter Waves, Southeast University, where she is currently an Associate Professor with the School of Information Science and Engineering. Her current research interests include microwave and millimeter-wave circuit design and monolithic microwave-integrated circuit design.

Xiao-Wei Zhu (S'88–M'95) received the M.E. and Ph.D. degrees in radio engineering from Southeast University, Nanjing, China, in 1996 and 2000, respectively.

Since 1984, he has been with Southeast University, where he is currently a Professor with the School of Information Science and Engineering. He has authored or coauthored more than 90 technical publications. He holds more than 15 patents. His current research interests include RF and antenna technologies for wireless communications, microwave and millimeter-wave theory and technology, and power amplifier (PA) nonlinear character and its linearization research, especially on wideband and high-efficiency GaN PAs.

Dr. Zhu was a recipient of the 2003 Second-Class Science and Technology Progress Prize of Jiangsu Province, China. He is the President of the Microwave Integrated Circuits and Mobile Communication Sub-Society and the Microwave Society of CIE. He is also the Secretary of the IEEE MTT-S/AP-S/EMC-S Joint Nanjing Chapter.

Wei Hong (M'92–SM'07–F'12) received the B.S. degree in radio engineering from the University of Information Engineering, Zhengzhou, China, in 1982, and the M.S. and Ph.D. degrees in radio engineering from Southeast University, Nanjing, China, in 1985 and 1988, respectively.

Since 1988, he has been with the State Key Laboratory of Millimeter Waves, Southeast University, where he has been the Director of the State Key Laboratory of Millimeter Waves since 2003. In 1993 and from 1995 to 1998, he was a Short-Term Visiting Scholar with the University of California at Berkeley, Berkeley, CA, USA, and the University of California at Santa Cruz, Santa Cruz, CA, USA, respectively. He is currently a Professor with the School of Information Science and Engineering, Southeast University, and also with Purple Mountain Laboratories, Nanjing. He has been involved in numerical methods for electromagnetic problems, millimeter wave theory and technology, antennas, RF technology for wireless communications, and so on. He has authored or coauthored more than 300 technical publications and 2 books.

Dr. Hong was an elected IEEE MTT-S AdCom member from 2014 to 2016. He is currently a Fellow of the CIE. He was a recipient of several awards including National Natural Prizes (twice), the First-Class Science and Technology Progress Prizes (thrice) issued by the Ministry of Education of China and Jiangsu Province Government, and the Foundations for China Distinguished Young Investigators and for "Innovation Group" by the NSF of China. He is the Chair of the IEEE MTT-S/AP-S/EMC-S Joint Nanjing Chapter. He is the Vice President of the CIE Microwave Society and Antenna Society. He served as an Associate Editor for the IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES from 2007 to 2010.

Anding Zhu (S'00–M'04–SM'12) received the B.E. degree in telecommunication engineering from North China Electric Power University, Baoding, China, in 1997, the M.E. degree in computer applications from the Beijing University of Posts and Telecommunications, Beijing, China, in 2000, and the Ph.D. degree in electronic engineering from University College Dublin (UCD), Dublin, Ireland, in 2004.

He is currently a Professor with the School of Electrical and Electronic Engineering, UCD. He has authored or coauthored more than 100 peer-reviewed journals and conference papers. His current research interests include high-frequency nonlinear system modeling and device characterization techniques, high-efficiency power amplifier design, wireless transmitter architectures, and digital signal processing, and nonlinear system identification algorithms, especially on behavioral modeling, and linearization of RF power amplifiers for wireless communications.

Dr. Zhu is currently an elected member of the IEEE MTT-S AdCom and an Editorial Board member of the IEEE Future Networks Tech Focus. He is the Vice Chair of Publications and Electronic Information Committees. He was the General Chair of the 2018 IEEE MTT-S International Microwave Workshop Series on 5G Hardware and System Technologies (IMWS-5G). He was the Secretary of the IEEE MTT-S AdCom in 2018. He is an Associate Editor of *IEEE Microwave Magazine*.