Research Repository UCD

| | |
|---|---|
| ***Title*** | Data-Clustering-Assisted Digital Predistortion for 5G Millimeter-Wave Beamforming Transmitters With Multiple Dynamic Configurations |
| ***Authors(s)*** | Yin, Hang, Yu, Zhiqiang, Yu, Chao, Zhu, Anding, et al. |
| ***Publication date*** | 2021-03 |
| ***Publication information*** | Yin, Hang, Zhiqiang Yu, Chao Yu, Anding Zhu, and et al. "Data-Clustering-Assisted Digital Predistortion for 5G Millimeter-Wave Beamforming Transmitters With Multiple Dynamic Configurations." IEEE, March 2021. https://doi.org/10.1109/tmtt.2020.3039747. |
| ***Publisher*** | IEEE |
| ***Item record/more information*** | http://hdl.handle.net/10197/12027 |
| ***Publisher's statement*** | © 2020 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. |
| ***Publisher's version (DOI)*** | 10.1109/tmtt.2020.3039747 |

Downloaded 2026-05-01 23:37:03

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd_oa)

© Some rights reserved. For more information

# Data-Clustering Assisted Digital Predistortion for 5G Millimeter-Wave Beamforming Transmitters with Multiple Dynamic Configurations

Hang Yin, *Student Member, IEEE*, Zhiqiang Yu, *Member, IEEE*, Chao Yu, *Member, IEEE*, Jianxin Jing, Xiao-Wei Zhu, *Member, IEEE*, Wei Hong, *Fellow, IEEE*, Anding Zhu, *Senior Member, IEEE*

***Abstract*—Motivated by data science, in this paper, a data-clustering assisted digital predistortion (DPD) is proposed to linearize millimeter-wave (mmWave) beamforming transmitters with multiple dynamic configurations. Based on the data analysis, similar transmitter states with different configurations can be clustered, resulting in significant reduction of linearization states. Model complexity within the cluster can be further reduced by utilizing penalty factor method. To validate the proposed concept, experiments were carried out on a 16-channel mmWave beamforming transmitter with configurable beam angle, operating frequency and input power. Total 216 transmitter states can be reduced to 8, 20, 40 states with unequal optimal model parameters for each state, without losing much performance. The proposed method can be extended to the scenario where a large scale of dynamic states can occur in complex transmitter structures in future wireless systems.**

***Index Terms*—Beamforming, clustering, digital predistortion (DPD), K-means, millimeter-wave (mmWave), multiple input multiple output (MIMO), power amplifier (PA)**

## I. Introduction

ONE of the key requirements for the next generation mobile network, the 5G, is to provide ubiquitous connections with multi-gigabit per second data transmissions to meet 1,000 times increased capacity. These high data rates drive the need for much wider spectral bandwidth and thus millimeter-wave (mmWave) bands will be deployed in 5G wireless systems. The propagation of mmWave signals however can be very poor and they are highly affected by atmospheric attenuation. Beamforming is a technique that uses multiple antennas to form a highly-focused beam pointing to a specific direction to increase communication quality and save energy. When combining with multiple input multiple output (MIMO) technique and using spatial coding, multiple data streams can be transmitted simultaneously into multiple channels to increase data capacity. The mmWave beamforming transmitters thus have been treated as one of the key technologies in 5G that can provide high link-level gains to overcome path losses, avoid undesirable interferences, and enable super high speed data rates [1]–[3].

Many beamforming architectures have been proposed, including fully analog, fully digital and hybrid beamforming [4]–[8]. To make the best use of communication resources, in 5G and future systems, the parameters of wireless transmitters may be dynamically configured according to the real time environment during deployment. For instance, the transmit power may be adjusted according to the data traffic to minimize the power assumption [9]. The operation frequency of the transmitter may change according to the availability of the spectrum [10]–[14] while the direction of the beam radiation will be dynamically steered to point to the user location [4]. As a result, the beamforming transmitter will exhibit complex dynamic behaviors from time to time due to changes of operating conditions. For example, dynamic power and frequency configurations of the transmitter will affect the operating states of the power amplifiers (PA), while the phase changes of the phase shift network would impact on the nonlinear characteristics observed at the receiver due to the mutual coupling effects [15].

Digital predistortion (DPD) has been widely adopted to keep PA operating with high efficiency without losing linearity [16]–[19]. It is expected that DPD will continue to be deployed in 5G transmitters. Some methods have been proposed to solve the related issues in DPD under MIMO scenarios in a fixed operating state. In [20], the feedback for over-the-air (OTA) acquisition is proposed to be placed inside the transmit array. In [21], a unified model is proposed to accurately model the PA nonlinearities under mismatch and mutual coupling caused by antenna array coupling. In [22], different kinds of non-idealities are proposed to be modeled by one integral neural network model. However, the dynamic behaviors of the transmitter induced by multiple dynamic configurations will bring huge challenges to the existing DPD techniques. As shown in Fig. 1, a DPD block is generally placed before the transmitter to generate the inverse of the transmitter characteristic so as to compensate the nonlinearity induced in the transmitter. Based on this linearization principle, once the transmitter characteristic changes, the DPD module must be adjusted in time accordingly, otherwise, the linearization

This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 62022025, Grant 61861136002 and Grant 61627801 and in part by Science Foundation Ireland (SFI) through the SFI-NSFC Partnership Program under Grant 17/NSFC/4850. (Corresponding author: Chao Yu.)

H. Yin, J. Jing and X.-W. Zhu are with the State Key Laboratory of Millimeter Waves, Southeast University, Nanjing, 210096, China (e-mail: yinhang_4@foxmail.com).

Z. Yu, C. Yu 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).

A. Zhu is with the School of Electrical and Electronic Engineering, University College Dublin, Dublin 4, D04 V1W8, Ireland (e-mail: anding.zhu@ucd.ie).

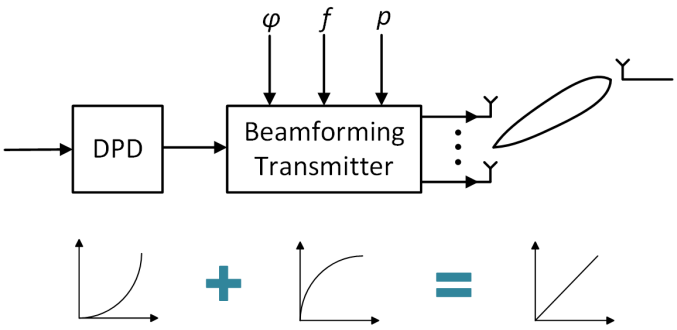



Fig. 1. DPD for Beamforming Transmitters.

will fail.

In [23], [24], the authors employed a feedback loop to couple the PA output signal for the purpose of the system monitoring and the DPD coefficients updating. This kind of closed-loop feedback is hard to implement directly in the MIMO beamforming system because a large number of PAs are involved and the closed-loop feedback is hard to be setup at the OTA interface. Another issue is the latency incurred by the coefficients extraction. Since the parameter extraction procedure takes some time, DPD coefficients updating may be slower than the change of operating state. Especially when sophisticated behavioral models are required to be employed, it results in much longer model extraction time. This will bring challenges to the scenarios that have high requirements on latency.

To avoid delay, multiple DPD coefficients can be pre-calculated and then a proper set of coefficients is selected according to current operation configuration. There are two types of configuration-based coefficients selection approaches including model-based and look-up-table (LUT) -based methods. In the model-based approach, the DPD coefficients are generated according to the model of the real-time configurations and transmitter behavior. Guo et al. [9] utilized the method to linearize a PA with dynamic configuration of input power. Generating coefficients under a large dimensions of dynamic configurations will be much more difficult because more complex behavioral models are required. In LUT-based approach, coefficients for all possible operating states are pre-extracted and stored in a LUT, then indexed by the real-time dynamic configurations. In this case, no coefficients generation is required, which leads to much time saving. This method has been utilized in [25] to linearize a PA excited by signals with dynamic configurations of power and carrier bandwidth and in [26] to index the sets of DPD coefficients for different ranges of angles. In [27], the author took both the feedback loop and DPD configuration into consideration by employing a LUT that can be dynamically pruned. Nevertheless, when the transmitter state varies due to more factors, the complexity of LUT will rise exponentially. For example, with three types of configurations, such as, power, frequency and phase, with each has 10 possible states, that will result in $10^3 = 1,000$ different transmitter operating states, which means that $1,000$ sets of coefficients are needed to be stored if "one set of coefficients for one state".

To resolve the above problem, a novel DPD concept motivated from the point view of data science is proposed. In this approach, we firstly collect all the behavioral data generated from different operating states of the transmitter and then use data clustering algorithms, such as K-means or range search-based algorithms, to find and cluster the states that exhibit similar nonlinear behaviors into a small groups where each group can share one DPD, which can significantly reduce the DPD model adaption complexity. After clustering, because the transmitter states are clustered into a small groups, instead of using a global model, we can optimize the model within the cluster to further reduce the system complexity and improve the performance. With a two-stage model optimization, the proposed data-clustering assisted DPD approach is capable of reducing the DPD system complexity significantly while achieving comparable performance as that using the conventional methods.

The rest of paper is organized as follows. The methodology for the data-clustering assisted DPD will be introduced in Section II, with its operation procedures described in Section III. Then experimental results will be presented in Section IV, followed by a brief conclusion in Section V.

## II. Data-Clustering Assisted DPD

The characteristic of the beamforming transmitter in Fig. 1 can be expressed as,

$$y = F(x; \varphi, f, p) \tag{1}$$

where $x$ and $y$ represents the input signal of the transmitter and the output signal received at the antenna, $\varphi$ is the phase shift, $f$ is the frequency, $p$ is the power configuration of the transmitter, respectively, while $F$ represents the nonlinear transfer function. From (1), we can see that the output of the transmitter $y$ not only depends on the input $x$ but also depends on the operation condition. Since DPD is only applied to the input signal, different DPDs must be used when the transmitter operation condition changes.

Fortunately, due to the multiple combinations of the operation conditions, there might be possibilities where the output may appear the same or similar when the transmitter is operated at different conditions when it is excited with the same input. For instance, $F(x; \varphi_1, f_5, p_7) \approx F(x; \varphi_4, f_2, p_1)$. In this case, the same DPD can be applied despite the transmitter is operated at different states. Now the task is how to find the similar states.

### A. Clustering Concept from Data Science

In data science, a variety of clustering algorithms have been developed for grouping similar data, such as hierarchical clustering, centroid-based clustering and density-based clustering [28]. Let's look at two typical clustering algorithms, range search and K-means, as examples.

The range search-based clustering is a hierarchical clustering algorithm based on the exhaustive search [28]. The core idea is to cluster the points that are within an allowable range. As illustrated in Fig. 2, firstly, a maximum search distance

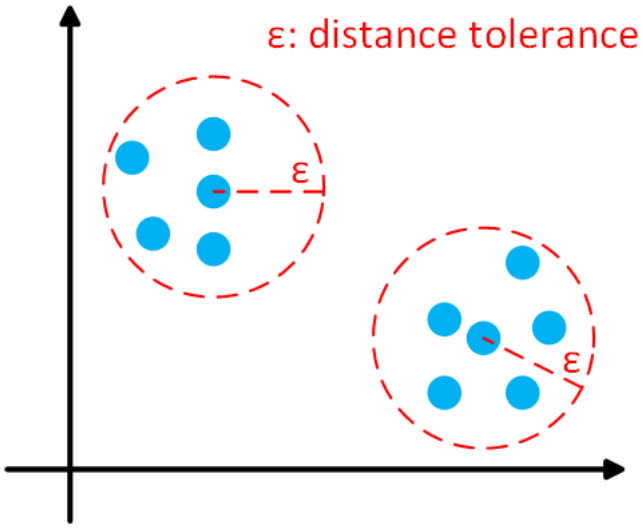



Fig. 2. Range Search Clustering.

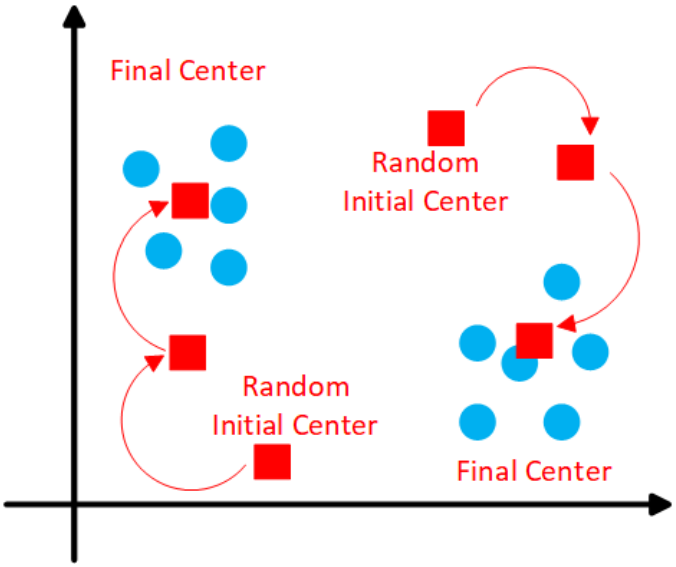



Fig. 3. K-means Clustering.

tolerance $\epsilon$ is set. Secondly, we randomly choose one point and set it in a cluster, then calculate its distance to all the other points and finally put all the points within distance tolerance to the same cluster. Do the same when traversing all the points until all of them are in clusters. The main feature of range search based clustering is that the maximum distance of two signals within a group can be fully ensured while the number of clusters varies depending on the data distribution.

The K-means clustering is a method of vector quantization that aims to partition multiple data points into $K$ clusters in which each sample belongs to the cluster with the nearest mean to the cluster centers. As shown in Fig. 3, firstly, a number of clusters ($K = 2$ for example) and cluster centers (solid squares) are chosen. Then for each point, it is put into the cluster which has a minimum signal-to-center distance. The cluster center location can be updated according to the means of all the points in these clusters. Convergence can be reached after several iterations, and the cluster centers and boundary will be decided. The main feature of K-means-based clustering algorithm is that the number of clusters can be fully ensured while the maximum signal-to-center distance within a cluster depends on the specific data sets.

To cluster data, the two algorithms have different focuses. The range search-based clustering algorithm has controllable distance within a cluster but not the total number of clusters, while K-means-based clustering algorithm has control on the number of clusters but not the distance. Both algorithms can be used to cluster transmitter states, which one is used depends on the system characteristics and practical requirements, e.g., system performance verses available resources. Nevertheless, if different transmitter states can be clustered into a small set of groups and each group uses a common DPD, the required number of DPDs will be significantly reduced and the system resource can be largely saved.

### B. Stage One Reduction: Clustering of States

To apply clustering for transmitter states, the key is how to define the blue dots in Fig. 2 and Fig. 3 and the distance between them. Since the transmitted signals are randomly generated and the system is with memory effect, how the nonlinear system state one, e.g., $F(\cdot;\varphi_1, f_1, p_1) = F_1$, is different from the nonlinear system state two, $F(\cdot;\varphi_2, f_2, p_2) = F_2$, cannot be well-defined by using a single data point but a group of time domain samples. In this work, we propose to calculate the difference between the outputs of $F_1(\cdot)$ and $F_2(\cdot)$ under the same stimuli sequence. Square error (SE), is proposed to measure the difference of two signals, expressed as:

$$\mathrm{SE}(y_1, y_2) = 10 * \log_{10} \sum_{n=1}^{N} |y_1(n) - y_2(n)|^2 (\mathrm{dB}) \quad (2)$$

where $n$ is the sample index and $N$ is the total number of samples used.

For better illustration, amplitude modulation to amplitude modulation (AM/AM), which is widely adopted for PA characteristics representation, is utilized to demonstrate the similar transmitter states in one cluster. As shown in Fig. 4, State 1 and State 2 may be clustered into the same group while State 3 stays in a different group. It is worth mentioning here that the above element-wise calculation is an augmented version of the methods [29], [30] that only utilize AM/AM as a two-dimensional figure for comparing different states.

### C. Stage Two Reduction: Model Optimization Within Cluster

After clustering, the large number of transmitter states can be clustered into a small number of clusters. Within one cluster, there may be $S$ states but they appear to behave similarly and thus one common DPD can be used to linearize them. To extract the common DPD for the cluster, there are mainly two approaches: the first one is to average the DPD coefficients after extracting the individual models and the second one is to average the signals and then extract the DPD coefficients. Both approaches can work well and which one to

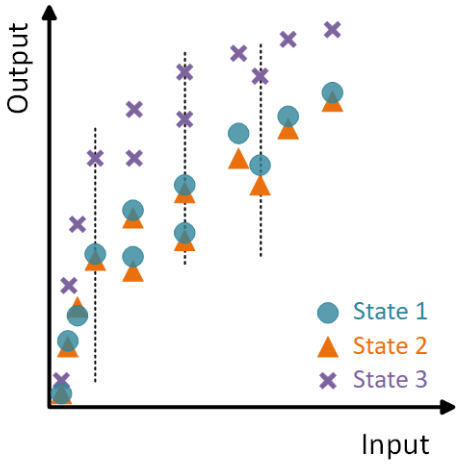



Fig. 4. Different AM/AMs under the same stimuli.

use depends on the application scenario. For instance, in real time operation, the transmit signals may not be repeatable, therefore averaging coefficients is preferred, while in system pre-calibration stage, averaging signals may be feasible.

Because the transmitter states are clustered into a small groups, instead of using a global model, we can optimize the model within the cluster to further reduce the system complexity. There exists some methods like exhaustive search in [31], [32], compressed sensing in [33] and genetic algorithm in [34]. In this work, we propose a penalty factor method.

A cost function $J$:

$$J = \mathrm{NMSE}(y, \hat{y}) + \alpha_1 P + \alpha_2 Q \tag{3}$$

can be defined to optimize the model size, where NMSE is normalized mean square error, $y$ is the actual signal to be modeled, and $\hat{y}$ is the modeled signal by a DPD model, such as a memory polynomial (MP) model (with "even-order" terms) [19] with nonlinear order $P$, memory depth $Q$ and coefficients of $c_{qp}$, as shown in the following equation:

$$y(n) = \sum_{q=0}^{Q} \sum_{p=1}^{P} c_{qp} x(n-q)|x(n-q)|^{p-1} \tag{4}$$

Meanwhile, $\alpha_1$ and $\alpha_2$ are the penalty factors for nonlinear order and memory depth, respectively. $\alpha_1$ and $\alpha_2$ are always positive and can be adjusted according to actual needs. The optimizing procedure can be described as follows: set up the maximum value for $P$ and $Q$ as $P_m$ and $Q_m$, separately, then search for $P$ and $Q$ within the range that can minimize $J$ to an optimal value.

### D. System Architecture

The complete system architecture is proposed as shown in Fig. 5, where the system consists of four blocks: beamforming transmitter, data acquisition, training block and DPD block. In the training block, two stages of model reductions are conducted. Firstly, a large number of nonlinear transmitter states can be grouped into a small number of clusters, i.e., stage one reduction, by analyzing signals collected with the data acquisition block under various configurations. Secondly, the DPD model in each cluster can be further optimized, i.e., stage two reduction. After that, the clustered and optimized DPD coefficients will be copied to the DPD block, meanwhile, the trained cluster mapping will be generated.

In real time system operation, the original input signal will be predistorted according to the set of DPD coefficients corresponding to the cluster attribution of current transmitter configurations. The main difference between the proposed system and the conventional LUT-based DPD updating method in [25] is on the mapping of different DPD coefficients. Specifically, the mapping of DPD coefficients in the conventional method is one-to-one mapping between DPD coefficients and the transmitter state, in other words, the number of coefficients sets is the same as transmitter states, which results in a serious challenge to the hardware memory resources, especially for the transmitter with more dynamic configurations. In contrast, the mapping of the proposed cluster-based method is a many-to-one mapping. By clustering states that can utilize the same set of DPD coefficients, the number of coefficient sets can be much smaller than the number of transmitter states and the memory cost for storing coefficients can be dramatically reduced.

## III. System Operation

The objective of the proposed system is to resolve the linearization issues of the beamforming transmitter with multiple dynamic configurations. It is a general approach that can be applied to many scenarios and under many different operation conditions. For simplicity, in this section, we apply this approach in linearizing the main beam only, similar to the "beam-oriented" DPD concept proposed by Liu. et al [15], [35] in which the relationship between the transmitted signal and the distorted signal observed at the main beam is seen as a single input single output system.

The operation workflow of the system in Fig. 5 consists of two parts: system training and system running. In system training, the input and output data of the transmitter are collected. The transmitter states are clustered into a small group and within each group, a DPD model is extracted and optimized. In system running, the input signal will be predistorted by the selected DPD block that is selected according to which cluster the current configuration belongs to. The detailed procedures are described in the following paragraphs.

### A. System Training

*1) Data Collection:* To simplify the expression, the configuration of transmitter operating state with phase shift $\varphi$ of phase shift network, carrier frequency $f$ of the mixer and power gain $p$ of the pre-amplifier is named as a configuration $(\varphi, f, p)$. If the transmitter is excited multiple times with the same baseband modulated signal $\boldsymbol{x}$, with configuration $(\varphi_i, f_j, p_k)$, where $i = 1, 2, \cdots, i_m$, $j = 1, 2, \cdots, j_m$ and $k = 1, 2, \cdots, k_m$, respectively, then totally $M = i_m * j_m * k_m$ different output signals $\{y_m\}$, $m = 1, 2, \cdots, M$ can be collected.

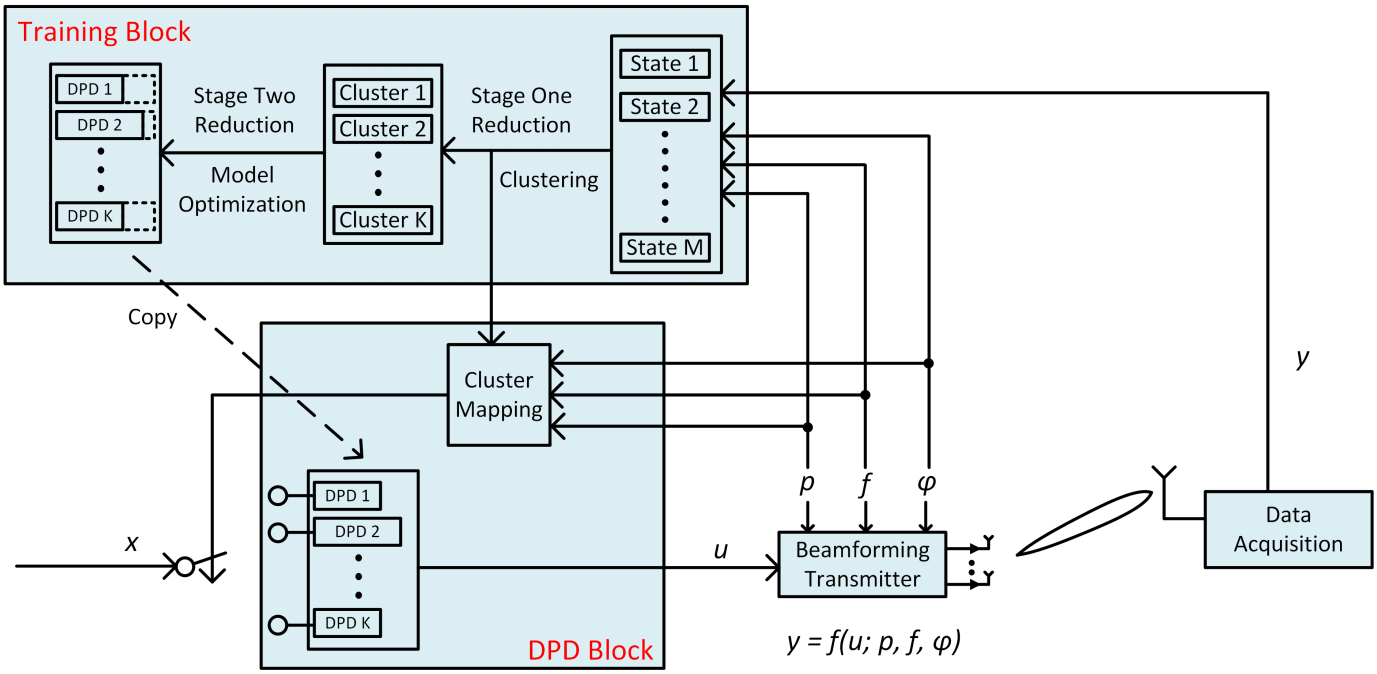



Fig. 5. Proposed System Architecture.

*2) Data Clustering:* After collecting $M$ transmitter output signals $\{y_m\}$, cluster them by range search or K-means algorithm by setting a signal distance tolerance $\epsilon$ or by setting $K$ number of clusters as shown in the detailed procedure Algorithm 1 and Algorithm 2, respectively.

To avoid confusion, it is also worth mentioning here that in [36] K-means is utilized to cluster different segments of PA input signal for better model accuracy, while in proposed method, K-means is utilized to cluster similar transmitter states on different transmitter output signals under the same stimuli.

**Algorithm 1** Cluster Similar Transmitter States Utilizing Range Search

**Input:** $M$ transmitter output signals $\{y_m\}$, $m = 1, 2, \cdots, M$, tolerance of signal distance $\epsilon$

**Output:** cluster result vector $v = [v(1), v(2), \cdots, v(M)]$, each element in the vector represents the cluster attribution

Given transmitter output signals $\{y_m\}$ and tolerance $\epsilon$, initialize $v$ to a zero vector, initialize $K$ to zero
**for all** $y_p \in \{y_m\}$ **do**
  **if** $v(p) = 0$ **then**
    $K \leftarrow K + 1$
    $v(p) \leftarrow K$
    **for all** $y_q \in \{y_m\}$, $q \neq p$ **do**
      **if** $\mathrm{SE}(y_q, y_p) < \epsilon$ **then**
        $v(q) \leftarrow K$
      **end if**
    **end for**
  **end if**
**end for**
**return** $v$

**Algorithm 2** Cluster Similar Transmitter States Utilizing K-means

**Input:** $M$ transmitter output signals $\{y_m\}$, $m = 1, 2, \cdots, M$, number of desired clusters $K$

**Output:** cluster result vector $v = [v(1), v(2), \cdots, v(M)]$, each element in the vector represents the cluster attribution

Given transmitter output signals $\{y_m\}$ and number of desired clusters $K$, initialize $c$ to a zero vector, randomly generate $K$ different cluster center signals $\{V_k\}$
**repeat**
  %% Clustering
  **for all** $y_p \in \{y_m\}$ **do**
    $v(p) \leftarrow \min_K \mathrm{SE}(y_p, V_k)$
  **end for**
  %% Cluster Center Updating
  **for all** $V_i \in \{V_k\}$ **do**
    count $\leftarrow 0$
    centersum $\leftarrow 0$
    **for all** $y_p \in \{y_m\}$ **do**
      **if** $v(p) = i$ **then**
        count $\leftarrow$ count + 1
        centersum $\leftarrow$ centersum + $y_p$
      **end if**
    **end for**
    $V_i \leftarrow$ centersum/count
  **end for**
**until** *convergence*
**return** $v$

*3) Model Optimization:* After the clustering, $M$ transmitter states can be clustered to $K$ categories, then the next step is to extract one set of common DPD coefficients for the states within the cluster. We choose the method of averaging coefficients as mentioned in Section II-D. In this case, the averaged coefficients and the stimuli sequence $x$ will be firstly used to generate a virtual signal $y_v$, then the optimizing will runs on $x$ and $y_v$ utilizing the proposed penalty factor method.

### B. System Running

After training, the system can run online and select pre-extracted coefficients according to the cluster attribution of dynamic power, frequency and beam configurations of the transmitter. Signals will be collected again at the OTA on the mentioned total $M$ states, named as calibrated states, utilizing $K$ sets of optimized DPD coefficients to validate performance of clustering-based DPD method.

*1) Uncalibrated States:* Since the power, frequency and beam are preset on discrete calibrated states, it is also important to pay attention to the states within the configuration range but not exactly at the exact calibrated discrete values, named as uncalibrated states such as power configuration $\frac{p_1+p_2}{2}$, because the uncalibrated states are likely to appear in real communication systems in which the configurations is expected to be a continuous number. Then nearest neighbour rule is employed for linearizing the uncalibrated states. Specifically, when the configuration is unknown $(\varphi, f, p)$, a nearest state $(\varphi_i, f_j, p_k)$ will be found according to the nearest states distance firstly:

$$\underset{i,j,k}{\operatorname{argmin}}(|\varphi - \varphi_i| + |f - f_j| + |p - p_k|) \tag{5}$$

then the cluster mapping will select the set of DPD coefficients which state $(\varphi_i, f_j, p_k)$ belongs to.

*2) Online Calibration:* During real time deployment, the behavior of the transmitter may change due to unpredicted variations of operating environment, e.g., temperature changes or aging, and thus the DPD system may need be re-calibrated. It is therefore worth discussing the calibration procedure during system running here.

Firstly, if the whole system need be re-calibrated, for the data acquisition, an OTA based feedback loop can be utilised. For instance, in [37], a receiver antenna at a fixed location was proposed to capture multiple sets of data in real time and these data can be used to reconstruct the output signal at the each channel or at the main beam direction. By employing this approach, the input and output data can gathered and then be used for model clustering and model extraction.

Secondly, in some cases, it is not necessary to re-calibrate the whole system. For instance, if the cluster attributes will not change from their original assignments after the operating condition changes. In other words, the transmitter behavior changes in the same way in all the cases within the cluster. Under this circumstance, the data captured at one condition may be used to extract the new coefficients and then applied to all the cases within the cluster. In other cases, if the cluster attributes change, e.g., the states in one cluster may become different when environmental condition changes, a thorough data collection and new clustering and model extraction will be necessary.

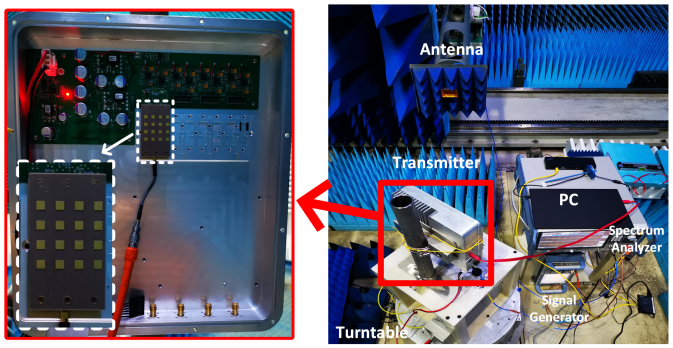



Fig. 6. Test Bench Setup with Transmitter Details

TABLE I
KEY PARAMETERS OF THE TRANSMITTER

| | |
|---|---|
| Beamforming IC | AWMF-0158 |
| Frequency Range | From 26.5 GHz to 29.5 GHz |
| Output P1dB (Per Channel) | ≈ 15dBm in Average |
| Gain of the Element Antenna | ≈ 5.8 dBi |
| Array EIRP at P1dB | 44.8 dBm |

## IV. EXPERIMENTAL RESULTS

### A. Test Bench Setup

To validate the proposed idea, a test bench was set up as shown in Fig. 6, where a R&S SMW 200A vector signal generator was used to generate a baseband long term evolution (LTE) signal and up-convert it to around 28 GHz with reconfigurable frequency and power. The signal was then sent to an in-house made beamforming transmitter which includes a 16-channel (4x4) beamforming array with key parameters listed in Table I.

The phase and gain of each transmitting channel can be controlled by Matlab and the subarray was calibrated before the test. The transmitter was set on the antenna turntable and the receive horn antenna was set at a fixed location. The signal captured by the receive antenna was sent to spectrum analyzer R&S FSW for validation of the proposed system.

A 20-MHz LTE signal with 6.2 dB PAPR and a sampling rate of 160 MHz was downloaded into the signal generator. To emulate the dynamic configurations of the transmitter, the beam angle, carrier frequency and power were reconfigured into 6 cases in each category, respectively. The input power of the signal varied from -4.0 dBm to -2.5 dBm with the step of 0.3 dB and the center frequency changed from 27.925 GHz to 28.050 GHz with the step of 25 MHz while the phase shift was also configured with their corresponding main beam pointing at $0^\circ$ to $25^\circ$ with the step of $5^\circ$, as listed in Table II. In total, $6 * 6 * 6 = 216$ different configurations were generated, and 216 different output signals $\{y_m\}$, $m = 1, 2, \cdots, 216$ were collected at the horn antenna and sent to the spectrum analyzer. For the purpose of cross validation, half of $x$ and half of $\{y_m\}$ were used for clustering and DPD extraction (index from 1 to 8192) and the other half for DPD validation (index from 8193 to 16384) under the cross validation. In this experiment, we focused on the controllable reduction of the

TABLE II
TRANSMITTER CONFIGURATIONS

| Attributes | Configurations |
|---|---|
| Power (dBm) | -4.0, -3.7, -3.4, -3.1, -2.8, -2.5 |
| Freq. (GHz) | 27.925, 27.950, 27.975, 28.000, 28.025, 28.050 |
| Beam (°) | 0, 5, 10, 15, 20, 25 |
| $6 * 6 * 6 = 216$ transmitter configurations in total | |

number of states, so the K-means clustering algorithm was adopted for experimental validation.

### B. DPD Validation on Calibrated States

*1) DPD Construction:* To validate and highlight features of the proposed DPD system, five DPD approaches for the 216 states are experimentally conducted. They are:

1) Full DPD with 216 sets of coefficients
2) Uniform DPD with 8 sets of coefficients
3) K-means DPD 1 with 8 sets of coefficients
4) K-means DPD 2 with 20 sets of coefficients
5) K-means DPD 3 with 40 sets of coefficients

In the "Full DPD", a separate DPD for each transmitter state is used, and thus 216 sets of coefficients are extracted for linearization of 216 states. Since each state has a dedicated DPD, it is expected that the best DPD performance can be delivered in this case. In comparison, the "Uniform DPD" uses a much smaller number of DPDs, where different states of the transmitter are distributed uniformly based on the configuration coordinate. The states are clustered into a small number groups based on their nearest neighbours and each group shares one DPD. In other words, the DPDs are selected uniformly across the states over the configuration coordinate space. In this work, three parameters of beam angle, frequency and power, are used and there are 6*6*6 states in total. For the 3D configuration, 8 clusters are selected within the 6*6*6 space, as shown in Fig. 7, in which every 27 (3*3*3) states employs the same set of DPD coefficients.

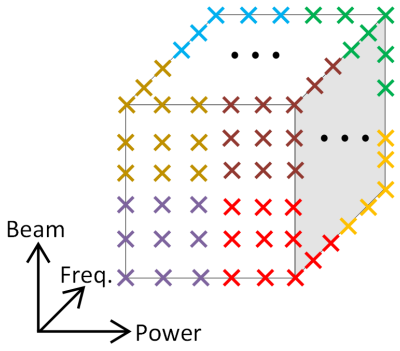



Fig. 7. Uniform DPD with 8-Octant Separation

The rest three DPDs are using the proposed method with different settings on the number of clusters, namely, 8 clusters (K-means DPD 1), 20 clusters (K-means DPD 2) and 40 clusters (K-means DPD 3), respectively. "K-means DPD 1" has 8 clusters to provide a fair comparison to the "Uniform DPD" that also utilizes 8 sets of coefficients.

*2) Clustering Result:* The clustering result on each of the 216 states by different DPD methods is illustrated by matrix as shown in Fig. 8. As the annotations show, the horizontal axis represents power and beam angle, and the vertical axis represents frequency and different DPD methods. Beam 1 to beam 6 correspond to main beams pointing from 0° to 25°, ranging 5°, frequency 1 to frequency 6 correspond to frequencies from 27.925 GHz to 28.050 GHz, ranging 25 MHz and power 1 to power 6 correspond to power from -4.0 dBm to -2.5 dBm with the step of 0.3 dB, respectively. The "Uniform DPD" in Fig. 8 can be seen as a slicing of the cubic in Fig. 7. For each DPD method, the DPD performance on 216 states is organized into a 6*36 matrix, and five different DPD methods (Uniform DPD, K-means DPD 1, K-means DPD 2, K-means DPD 3 and Full DPD) are organized by column, so coefficients selections in Fig. 8 are constructed into a 30*36 matrix. The number in each element of the matrix represents the cluster attribution. For example, for "Full DPD", total 216 sets are utilized, and for "K-means DPD 2", the numbers show how the 20 clusters are distributed for the 216 states. Also, it is worth pointing out that the number for different methods are numbered within the specific method, in other words, the set of DPD coefficients with some number in one method is independent to that in the other method.

Interestingly, it can be found from the star marker in the "K-means DPD 3" row of Fig. 8 that the configurations clustered into category 4 are (beam, frequency, power) of (2, 1, 6), (3, 6, 2), (3, 5, 3), (3, 3, 6), (4, 5, 4), (4, 5, 5), (4, 6, 3), (4, 6, 4) and (6, 5, 6), which means that the transmitter with different combinations of power, frequency and beam direction can be grouped into one cluster, which provides an effective verification of the cluster feasibility. Besides, it indicates that states with large configuration coordinate distance can also be clustered, giving proof to the invalidity of reference methods such as "Uniform DPD" in which the clustering just happens among configuration coordinates.

*3) Model Optimization Result:* A MP model [19] with "even-order" terms is used as a demonstration to fast validate the proposed idea. Since the bandwidth of stimuli $x$ is 20 MHz and the sample rate is 160 MHz, the uniform hyperparameter for optimization start at each state is set as nonlinear order of 9 and memory depth of 4. This means that the modeling optimization procedure according to the trade-off criterion (between NMSE and model hyperparameters) in (3) would start from $P_m = 9$ and $Q_m = 4$. The factors $\alpha_1$ and $\alpha_2$ were all set as "1" in our experiment. Due to space limitation, only optimization results of "K-means DPD 1" with 8 clusters are given, as shown in Table III. It can be seen from Table III that NMSE value increases and J value decreases after the optimization, which gives proof the validity of the optimization target of trade-off between model performance and complexity. Meanwhile, model parameters can be significantly reduced, and still great performance can be achieved after the trade-off. It is worth mentioning that the model can be further simplified if we reduce the linearity requirement. This can be conducted according to the requirement of the real system.

*4) Performance Comparison:* The results of the performance for different DPD methods on linearization NMSE and

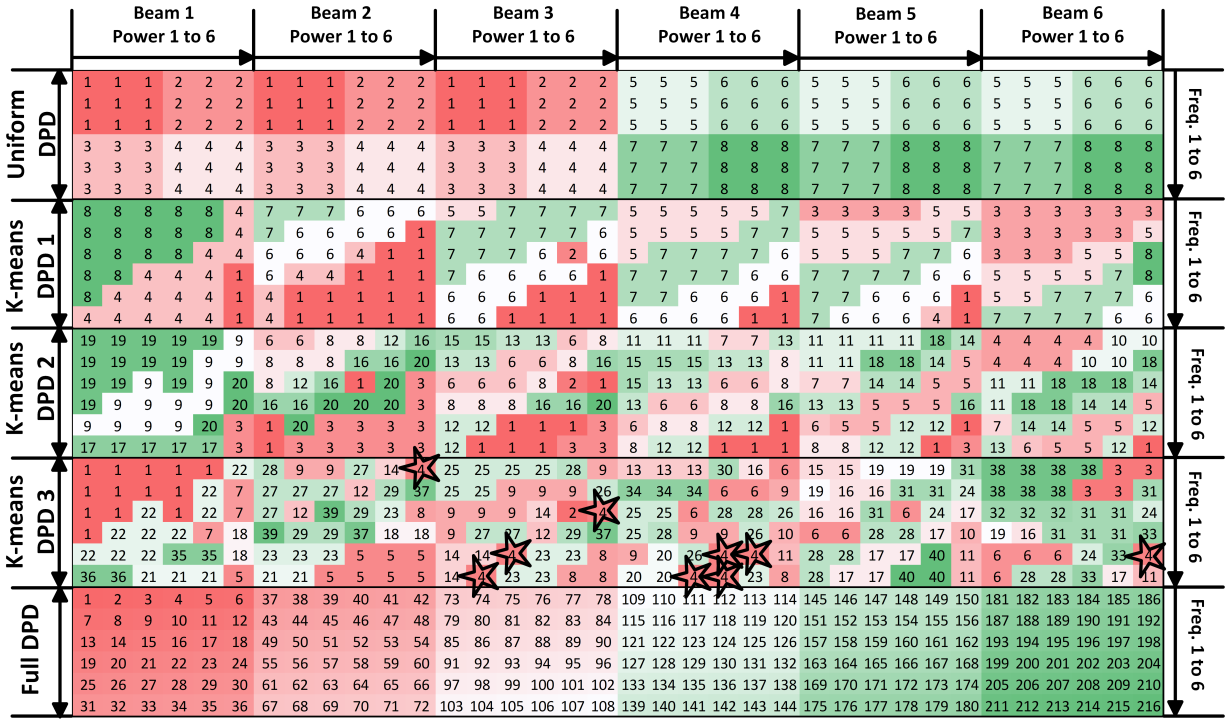



Fig. 8. Clustering Result of All 216 States Using Different DPD Methods.

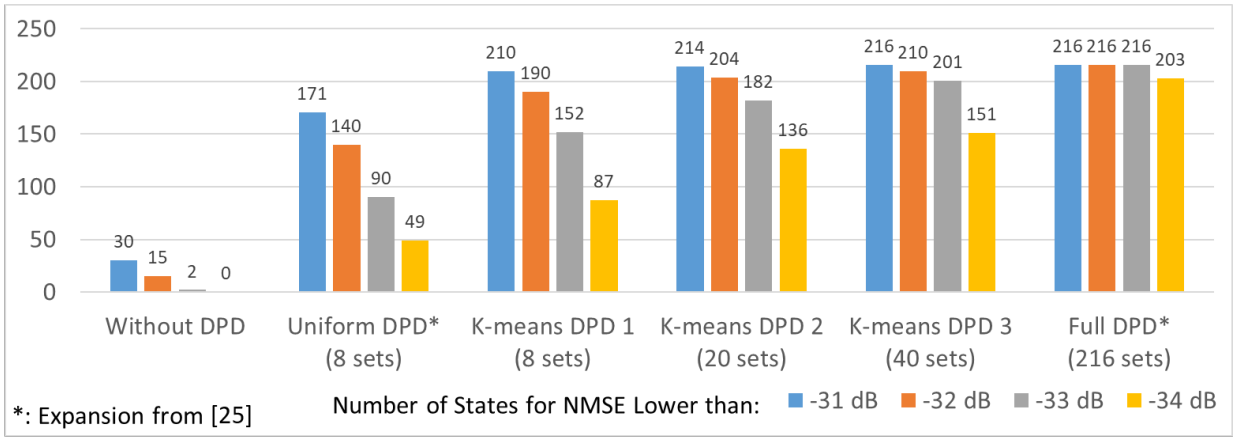



Fig. 9. NMSE Distribution of Different Methods.

TABLE III
MODEL OPTIMIZATION RESULTS STARTING FROM $P_m = 9$ AND $Q_m = 4$

| K-means DPD 1 | | | | | | |
|---|---|---|---|---|---|---|
| Cluster | Optimization Start | | | Optimization Result | | |
| | (P, Q) | NMSE (dB)* | J | (P, Q) | NMSE (dB)* | J |
| 1 | (9, 4) | -63.79 | -51.79 | (8, 4) | -62.80 | -51.80 |
| 2 | (9, 4) | -48.27 | -36.27 | (4, 2) | -46.98 | -41.98 |
| 3 | (9, 4) | -56.48 | -44.48 | (4, 2) | -52.19 | -47.19 |
| 4 | (9, 4) | -58.41 | -46.41 | (7, 3) | -57.02 | -48.02 |
| 5 | (9, 4) | -60.83 | -48.83 | (7, 4) | -59.87 | -49.87 |
| 6 | (9, 4) | -54.71 | -42.71 | (4, 2) | -50.87 | -45.87 |
| 7 | (9, 4) | -56.73 | -44.73 | (4, 2) | -52.63 | -47.63 |
| 8 | (9, 4) | -60.78 | -48.78 | (7, 4) | -60.00 | -50.00 |
| *NMSE represents for the modeling accuracy | | | | | | |

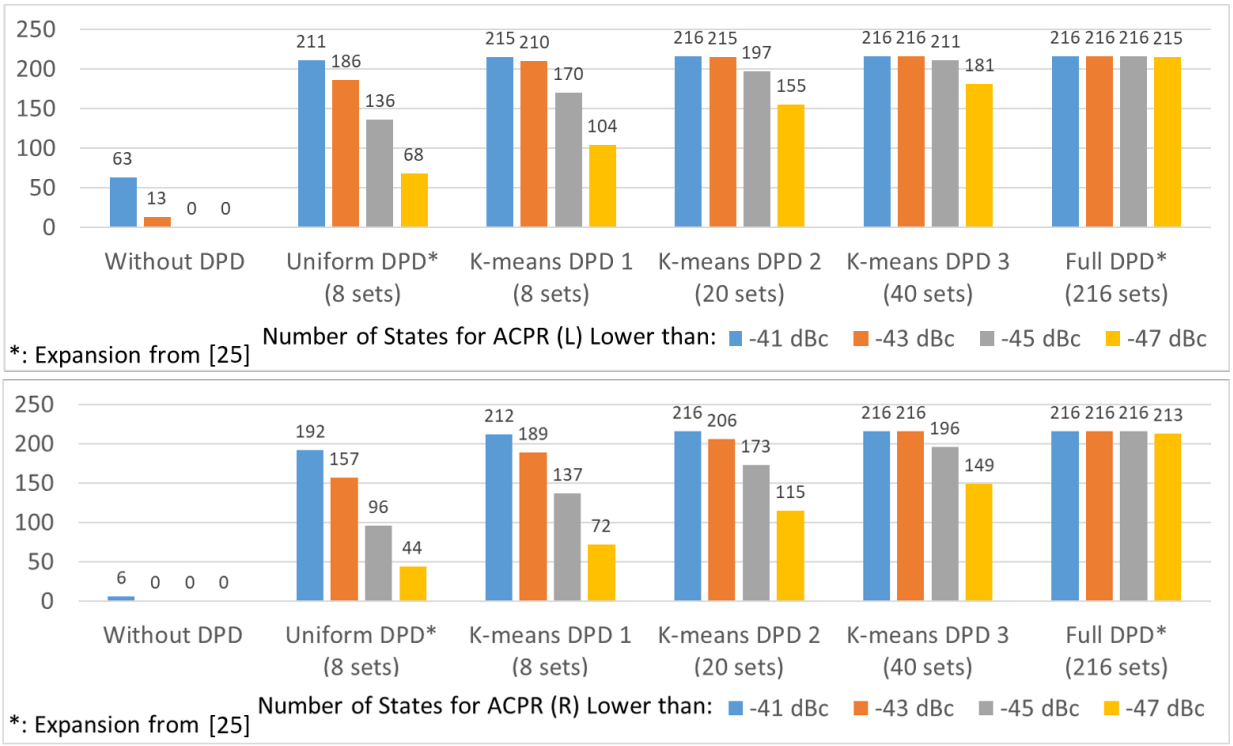



Fig. 10. ACPR Distribution of Different Methods.

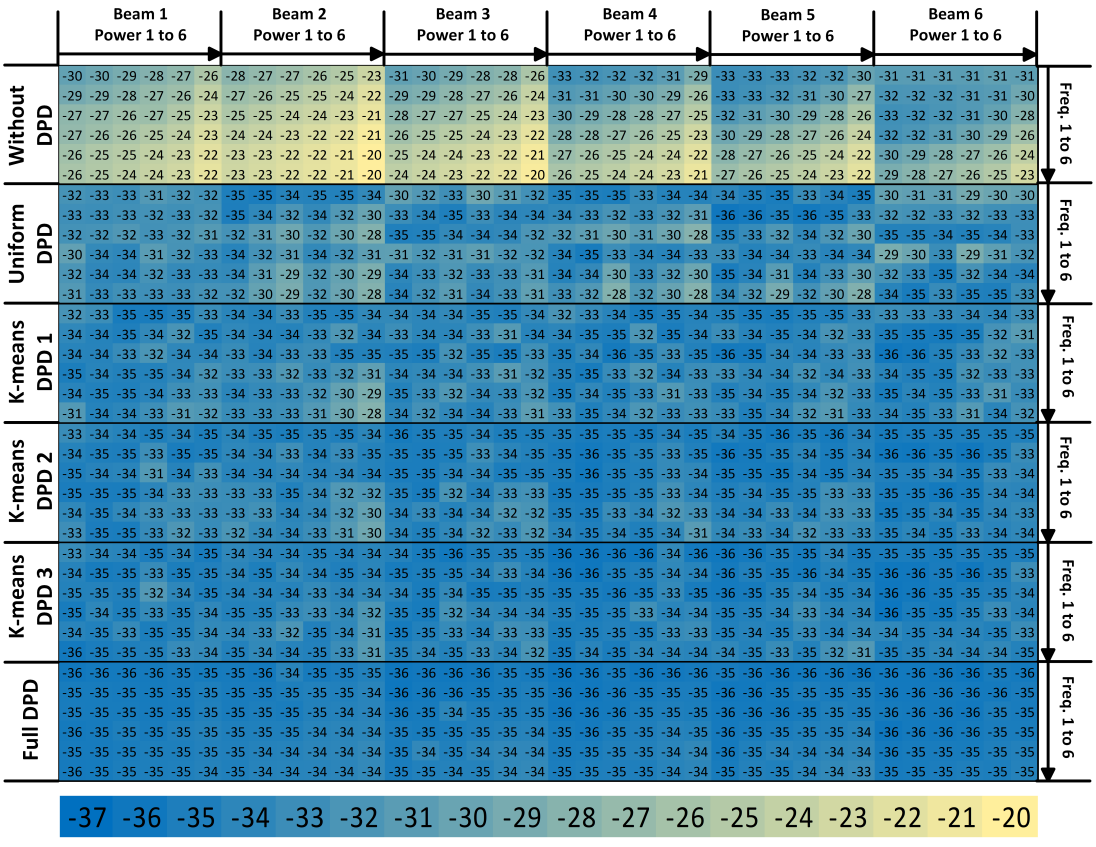



Fig. 11. NMSE on Each of the 216 States.

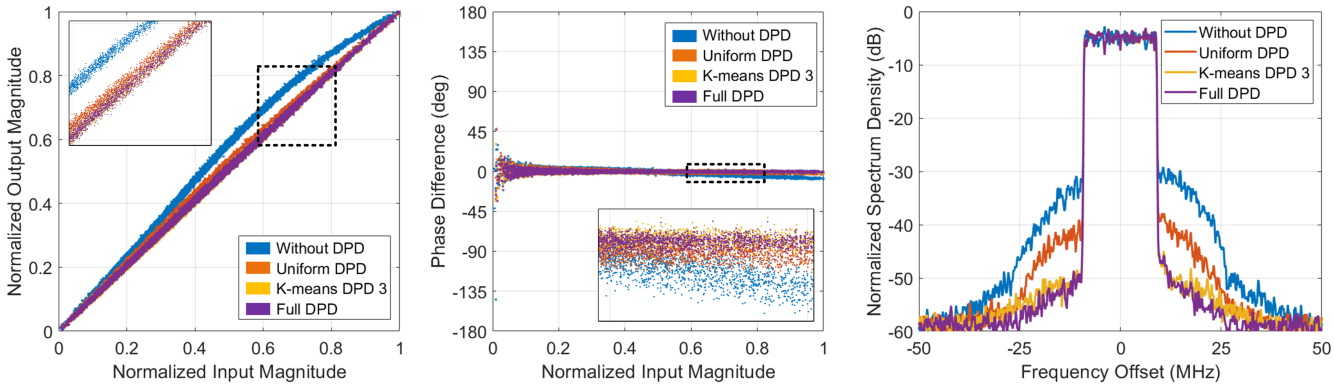



Fig. 12. An example of AM/AM, AM/PM and spectrum at state (beam 4, freq. 6, power 6).

TABLE IV
DPD PERFORMANCE COMPARISON ON UNCALIBRATED STATES (NMSE (DB)/ACPR-L (DBC)/ACPR-R (DBC))

| Test Case | 1 | 2 | 3 |
|---|---|---|---|
| Configuration (beam, freq., power) | 23°, 27.980 GHz, -3.9 dBm | 4°, 27.930 GHz, -2.6 dBm | 4°, 27.940 GHz, -3.8 dBm |
| Nearest State | 25°, 27.975 GHz, -4.0 dBm | 5°, 27.925 GHz, -2.5 dBm | 5°, 27.950 GHz, -3.7 dBm |
| Without DPD | -32.99 / -43.49 / -40.80 | -23.65 / -36.16 / -33.18 | -26.93 / -40.14 / -36.22 |
| Proposed DPD | -35.50 / -48.65 / -48.29 | -34.21 / -49.38 / -46.89 | -33.66 / -45.61 / -44.68 |
| Full DPD | -35.53 / -49.40 / -49.52 | -34.77 / -50.11 / -49.34 | -34.50 / -48.00 / -46.74 |

adjacent channel power ratio (ACPR) distribution are shown in Fig. 9 and Fig. 10, respectively. From Fig. 9 and Fig. 10, it can be seen that all five DPD methods have achieved significant linearization results, compared with the case without DPD. The "Full DPD" achieved the best DPD performance among the 216 states on both NMSE and ACPR. For "Uniform DPD" and "K-means DPD" that both utilize 8 sets of coefficients, "K-means DPD 1" outperformed "Uniform DPD" on NMSE and ACPR, which indicates the proposed metric for nonlinear behavior similarity and clustering is better than the reference method that utilizes configuration coordinate for clustering. Comparing "K-means DPD 1, 2 and 3", we can that the system with a larger set of coefficients obtained better performance and "K-means DPD 3" achieved almost the same performance as that of the "Full DPD". In a real system, the number of clusters can be selected according to the available resources and system linearity requirements. Due to space limitation, only NMSE at each of the 216 states by different methods is given in Fig. 11.

*5) Linearization Data Plots:* An example of AM/AM, AM/PM (amplitude modulation to phase modulation) and power spectrum are given in Fig. 12 to demonstrate the validity of clustering based DPD methods. A chosen state is the case with beam 4, freq. 6, power 6. As illustrated in Fig. 12, we can see that the proposed "K-means DPD 3" can achieve comparable performance to "Full DPD", while outperform "Uniform DPD".

### C. DPD Validation on Uncalibrated States

The above tests have validated the DPD performance on 216 calibrated states with beam direction from 0° to 25°, frequency from 27.925 GHz to 28.050 GHz and input power from -4.0 dBm to -2.5 dBm, ranging 5°, 25 MHz and 0.3 dBm, respectively. In a real system, new states may occur. In this section, we show how the proposed DPD can perform with non-precalibrated states. A few test cases are conducted, such as the configuration (23°, 27.980 GHz, -3.9 dBm), shown in Table IV. It is worth mentioning in Fig. 9 that for the "Uniform DPD" method, only $\frac{90}{216} \approx 41.67\%$ percentage of states achieved NMSE lower than -33 dB, indicating that DPD performance on many states is not valid, compared with other methods. Since DPD coefficients on calibrated states were selected according to the nearest state as shown in (5), the "Uniform DPD" will not be valid on many uncalibrated states as well. Thus, the "Uniform DPD" method was not tested on uncalibrated states. It can be found in Table IV that the "Full DPD" and the proposed method ("K-means DPD 3") can both achieve valid DPD results on both NMSE and ACPR.

## V. CONCLUSION

In this paper, a novel concept of data-clustering assisted DPD is proposed to linearize beamforming transmitters with multiple dynamic configurations. By clustering similar transmitter states and sharing one set of DPD coefficients within a cluster, the number of to-be-linearized states can be largely reduced but with comparable linearization performance. The system complexity can be further reduced by utilizing the penalty factor method for DPD extraction within the cluster.

In future wireless systems, a large number of channels, e.g., massive MIMO, will be deployed in beamforming transmitters and much more complex behaviors are likely to be exhibited over more transmitter configurations dimensions than power, frequency and beam direction, such as bandwidth, bias voltage and temperature. It is expected that the proposed concept

can be further extended to these more complex scenarios, which makes it a promising solution for linearizing future communication systems.

## REFERENCES


[1] P. Popovski, K. F. Trillingsgaard, O. Simeone, and G. Durisi, "5G wireless network slicing for eMBB, URLLC, and mMTC: A communication-theoretic view," *IEEE Access*, vol. 6, pp. 55 765–55 779, Sep. 2018.

[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] C. Wang *et al.*, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 122–130, Feb. 2014.

[4] W. Roh *et al.*, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 106–113, Feb. 2014.

[5] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. on Inform. Theory*, vol. 48, no. 6, pp. 1277–1294, Jun. 2002.

[6] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Proces.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.

[7] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: A comprehensive survey," *IEEE Commun. Surveys Tutor.*, vol. 18, no. 3, pp. 1617–1655, Feb. 2016.

[8] T. S. Rappaport *et al.*, "Broadband millimeter-wave propagation measurements and models using adaptive-beam antennas for outdoor urban cellular communications," *IEEE Trans. on Antennas and Prop.*, vol. 61, no. 4, pp. 1850–1859, Apr. 2013.

[9] Y. Guo, C. Yu, and A. Zhu, "Power adaptive digital predistortion for wideband RF power amplifiers with dynamic power transmission," *IEEE Trans. on Microw. Theory and Techn.*, vol. 63, no. 11, pp. 3595–3607, Nov. 2015.

[10] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Commun. Surv. Tutor.*, vol. 11, no. 1, pp. 116–130, Mar. 2009.

[11] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[12] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 79–89, May 2007.

[13] J. Mitola and G. Q. Maguire, "Cognitive radio: Making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[14] Y. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. on Wireless Commun.*, vol. 7, no. 4, pp. 1326–1337, Apr. 2008.

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

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

[17] 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.

[18] 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. on Signal Process.*, vol. 54, no. 10, pp. 3852–3860, Oct. 2006.

[19] L. Ding *et al.*, "A robust digital baseband predistorter constructed using memory polynomials," *IEEE Trans. on Commun.*, vol. 52, no. 1, pp. 159–165, Jan. 2004.

[20] X. Liu, W. Chen, L. Chen, F. M. Ghannouchi, and Z. Feng, "Linearization for hybrid beamforming array utilizing embedded over-the-air diversity feedbacks," *IEEE Trans. on Microw. Theory and Techn.*, vol. 67, no. 12, pp. 5235–5248, 2019.

[21] S. K. Dhar, A. Abdelhafiz, M. Aziz, M. Helaoui, and F. M. Ghannouchi, "A reflection-aware unified modeling and linearization approach for power amplifier under mismatch and mutual coupling," *IEEE Trans. on Microw. Theory and Techn.*, vol. 66, no. 9, pp. 4147–4157, 2018.

[22] P. Jaraut, M. Rawat, and F. M. Ghannouchi, "Composite neural network digital predistortion model for joint mitigation of crosstalk, $I/Q$ imbalance, nonlinearity in MIMO transmitters," *IEEE Trans. on Microw. Theory and Techn.*, vol. 66, no. 11, pp. 5011–5020, 2018.

[23] C. D. Presti, D. F. Kimball, and P. M. Asbeck, "Closed-Loop digital predistortion system With fast real-time adaptation applied to a handset WCDMA PA module," *IEEE Trans. on Microw. Theory and Techn.*, vol. 60, no. 3, pp. 604–618, Mar. 2012.

[24] R. N. Braithwaite, "Closed-Loop digital predistortion (DPD) using an observation path with limited bandwidth," *IEEE Trans. on Microw. Theory and Techn.*, vol. 63, no. 2, pp. 726–736, Feb. 2015.

[25] O. Hammi, A. Kwan, and F. M. Ghannouchi, "Bandwidth and power scalable digital predistorter for compensating dynamic distortions in RF power amplifiers," *IEEE Trans. on Broadc.*, vol. 59, no. 3, pp. 520–527, Sep. 2013.

[26] E. Ng, Y. Beltagy, G. Scarlato, A. Ben Ayed, P. Mitran, and S. Boumaiza, "Digital predistortion of millimeter-wave RF beamforming arrays using low number of steering angle-dependent coefficient sets," *IEEE Trans. on Microw. Theory and Techn.*, vol. 67, no. 11, pp. 4479–4492, 2019.

[27] R. N. Braithwaite, "A self-generating coefficient list for machine learning in RF power amplifiers using adaptive predistortion," in *2006 Europ. Microw. Conf.*, 2006, pp. 1229–1232.

[28] M. Kantardzic, "Cluster analysis," in *Data Mining: Concepts, Models, Methods, and Algorithms*. 2011, pp. 249–279.

[29] H. Yin, Z. Chen, and C. Yu, "Pattern recognition of RF power amplifier behaviors with multilayer perceptron," in *2018 Intern. Conf. on Microw. and Millim. Wave Techn. (ICMMT)*, May 2018, pp. 1–3.

[30] H. Yin *et al.*, "Pattern sensing based digital predistortion of RF power amplifiers under dynamical signal transmission," in *2019 IEEE MTT-S Intern. Microw. Conf. on Hardw. and Syst. for 5G and Beyond (IMC-5G)*, 2019, pp. 1–3.

[31] F. Mkadem and S. Boumaiza, "Physically inspired neural network model for RF power amplifier behavioral modeling and digital predistortion," *IEEE Trans. on Microw. Theory and Techn.*, vol. 59, no. 4, pp. 913–923, 2011.

[32] D. Wang, M. Aziz, M. Helaoui, and F. M. Ghannouchi, "Augmented real-valued time-delay neural network for compensation of distortions and impairments in wireless transmitters," *IEEE Trans. on Neural Netw. and Learning System*, vol. 30, no. 1, pp. 242–254, 2019.

[33] J. Reina-Tosina, M. Allegue-Martínez, C. Crespo-Cadenas, C. Yu, and S. Cruces, "Behavioral modeling and predistortion of power amplifiers under sparsity hypothesis," *IEEE Trans. on Microw. Theory and Techn.*, vol. 63, no. 2, pp. 745–753, 2015.

[34] S. Wang, M. Roger, J. Sarrazin, and C. Lelandais-Perrault, "Hyperparameter optimization of two-hidden-layer neural networks for power amplifiers behavioral modeling using genetic algorithms," *IEEE Microw. and Wirel. Compon. Lett.*, vol. 29, no. 12, pp. 802–805, 2019.

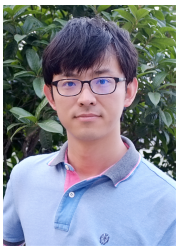
[35] X. Liu, W. Chen, L. Chen, and Z. Feng, "Beam-Oriented digital predistortion for hybrid beamforming array utilizing over-the-air diversity feedbacks," in *2019 IEEE MTT-S Intern. Microw. Sympos. (IMS)*, Jun. 2019, pp. 987–990.

[36] S. Afsardoost, T. Eriksson, and C. Fager, "Digital predistortion using a vector-switched model," *IEEE Trans. on Microw. Theory and Techn.*, vol. 60, no. 4, pp. 1166–1174, 2012.

[37] X. Wang, Y. Li, C. Yu, W. Hong, and A. Zhu, "Digital predistortion of 5G massive MIMO wireless transmitters based on indirect identification of power amplifier behavior with OTA tests," *IEEE Trans. on Microw. Theory and Techn.*, vol. 68, no. 1, pp. 316–328, 2020.


**Hang Yin** was born in Shenyang, China, in 1999. He received the B.E. degree in information science and engineering from Southeast University, Nanjing, China, in 2018, where he is currently pursuing the M.E. degree at the State Key Laboratory of Millimeter Waves.

His current research interests include the modeling and linearization (especially AI based) of radio frequency power amplifiers and 5G massive multiple-input–multiple-output (MIMO) systems.



**Zhiqiang Yu** (M'13) received the B.S. degree from the Nanjing University of Science and Technology, Nanjing, China, in 2002, and the Ph.D. degree from Southeast University, Nanjing, in 2013.

From 2002 to 2007, he was a Research Staff in airborne radar transmitter with the Nanjing Institute of Electronics, China Electronics Technology Group Corporation, Nanjing. He is currently a Lecturer with the School of Information Science and Engineering, Southeast University. His current research interests include microwave and millimeter-wave transceiver systems, beamforming networks, and phased arrays for mobile communication.

**Chao Yu** (S'09–M'15) received the B.E. degree in information engineering and 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, Southeast University. His research interests include microwave and millimeter wave power amplifier modeling and linearization, and 5G massive 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, and the M.E. degree in electromagnetic fields and microwave technology from Southeast University, Nanjing, China, in 2020.

He is currently an Algorithm Engineer with Huawei Technologies Co., Ltd., Shanghai, China. His current research interests include modeling and linearization for wireless base stations.

**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.

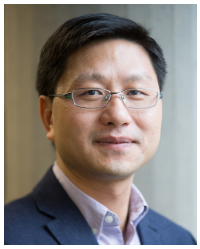
Since 1984, he has been with Southeast University, where he is currently a Professor with the School of Information Science and Engineering. He has authored or coauthored over 100 technical publications. He holds over 20 patents. His research interests include RF and antenna technologies for wireless communications, as well as microwave and millimeter-wave theory and technology, and also power amplifier (PA) nonlinear character and its linearization research with a particular emphasis on wideband and high-efficiency GaN PAs.

Dr. Zhu is the Microwave Society of CIE, and the secretary of the IEEE MTT-S/AP-S/EMC-S Joint Nanjing Chapter. He was the recipient of the 2003 Second-Class Science and Technology Progress Prize of Jiangsu Province, China.

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

Since 1988, he has been with the State Key Laboratory of Millimeter Waves and serves for the director of the lab since 2003, and is currently a professor of the School of Information Science and Engineering, Southeast University. In 1993, 1995, 1996, 1997 and 1998, he was a short-term Visiting Scholar with the University of California at Berkeley and at Santa Cruz, respectively. He has been engaged in numerical methods for electromagnetic problems, millimeter wave theory and technology, antennas, RF technology for wireless communications etc. He has authored and co-authored over 300 technical publications and two books. He twice awarded the National Natural Prizes, thrice awarded the first-class Science and Technology Progress Prizes issued by the Ministry of Education of China and Jiangsu Province Government etc. Besides, he also received the Foundations for China Distinguished Young Investigators and for "Innovation Group" issued by NSF of China.

Dr. Hong is a Fellow of IEEE, Fellow of CIE, the vice presidents of the CIE Microwave Society and Antenna Society, the Chair of the IEEE MTT-S/AP-S/EMC-S Joint Nanjing Chapter, and was an elected IEEE MTT-S AdCom Member during 2014-2016. He served as the Associate Editor of the IEEE Trans. on MTT from 2007 to 2010.



**Anding Zhu** received 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. His research interests include high-frequency nonlinear system modeling and device characterization techniques, high-efficiency power amplifier design, wireless transmitter architectures, digital signal processing, and nonlinear system identification algorithms. He has published more than 140 peer-reviewed journal and conference articles.

Prof. Zhu is an elected member of MTT-S AdCom, the Chair of the Electronic Information Committee and the Vice Chair of the Publications Committee. He is also the Chair of the MTT-S Microwave High-Power Techniques Committee. He served as the Secretary of MTT-S AdCom in 2018. He was the General Chair of the 2018 IEEE MTT-S International Microwave Workshop Series on 5G Hardware and System Technologies (IMWS-5G) and a Guest Editor of the IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES on 5G Hardware and System Technologies. He is currently an Associate Editor of the IEEE Microwave Magazine and a Track Editor of the IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES.