

Radio Frequency Fingerprinting Exploiting Non-Linear Memory Effect

Yuepei Li, *Student Member, IEEE*, Yuan Ding, Junqing Zhang, George Goussetis, *Senior Member, IEEE*, and Symon K. Podilchak, *Senior Member, IEEE*

Abstract—Radio frequency fingerprint (RFF) identification distinguishes wireless transmitters by exploiting their hardware imperfection that is inherent in typical radio frequency (RF) front ends. This can reduce the risks for the identities of legitimate devices being copied, or forged, which can also occur in conventional software-based identification systems. This paper analyzes the feasibility of device identification exploiting the unique non-linear memory effect of the transmitter RF chains consisting of matched pulse shaping filters and non-linear power amplifiers (PAs). This unique feature can be extracted from the received distorted constellation diagrams (CDs) with the help of image recognition-based classification algorithms. In order to validate the performance of the proposed RFF approach, experiments are carried out in cabled and over the air (OTA) scenarios. In the cabled experiment, the average classification accuracy among systems of 8 PAs (4 PAs of the same model and the other 4 of different models) is around 92% at signal to noise ratio (SNR) of 10 dB. For the OTA line-of-sight (LOS) scenario, the average classification accuracy is 90% at SNR of 10 dB; for the non-line-of-sight (NLOS) scenario, the average classification accuracy is 79% at SNR of 12 dB.

Index Terms—Convolution neural network (CNN), non-linear memory effect, radio frequency fingerprint (RFF).

I. INTRODUCTION

IN the booming Internet of Things (IoT) era, wireless devices are used everywhere. According to a recent Cisco report [1], the number of IoT devices will be greater than three times the global population by 2023. Their prevalence brings great convenience to our everyday lives, such as smart sensing and home automation, public infrastructure, farming, etc. On the other hand, the explosive growth of the wireless connectivity is accompanied with significant cyber threats, like Sybil attacks [2], masquerade attacks, and resource depletion [3], etc. The conventional methods for device identification, a prerequisite for wireless network access control, commonly use IP and/or MAC addresses, electronic serial number (ESN), international mobile station equipment identity (IMEI) number, and mobile identification number

(MIN) as unique identifiers. However, malicious users can potentially modify these identifiers via software [4]. They can imitate legitimate transmitters or receivers to invade wireless systems, such as information interception and malicious manipulation. Therefore, a significant security concern of these conventional device identification methods is that multiple attackers can masquerade as legitimate users to steal sensitive and personal data once the device identities are compromised.

A robust and unmodifiable identification system is thereby imperative to evade such criminal activities. The radio frequency fingerprint (RFF) classification system has emerged and is a promising candidate for increased security, which identifies wireless devices through features intrinsically embedded in analogue radio frequency (RF) stages of wireless devices [5], [6]. The hardware imperfection is the result of uncontrollable factors in the manufacturing process for any wireless device. Thus, they are difficult to be modelled, predicted, modified, and copied.

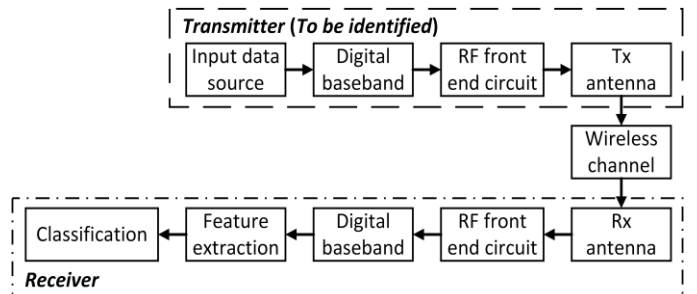


Fig. 1. Illustration of an RFF wireless device identification system block diagram.

The imperfection of the RF hardware in wireless transmitters leads to the distortions of the signal waveforms that are eventually captured by receivers. Here, the RFF technique exploits and extracts these distortions from the signal waveforms in various domains for transmitter classification/identification. An RFF system commonly operates in three stages, i.e., feature extraction, training, and device classification/identification. An RFF identification system block is illustrated in Fig. 1. Here a receiver, acting as an identifier, captures wireless signals radiated from the transmitter (to be identified) and extracts RFF features that originated from the unique RF imperfection in the transmitter. The RF imperfection exists in various RF components, e.g. phase noise of frequency carrier oscillators [7], non-linear I/O characteristics of power amplifiers (PAs) [8]–[11], antenna

Mr Yuepei Li's work was supported by Heriot-Watt University Doctor Training Programme (DTP) scholarship. Dr Yuan Ding's work was supported by EPSRC (UK) under the Grant EP/V002635/1.

Y. Li, Y. Ding, and G. Goussetis are with the Institute of Sensors, Signals and Systems (ISSS), Heriot-Watt University, UK (emails: y112@hw.ac.uk, yuan.ding@hw.ac.uk, g.goussetis@hw.ac.uk). Y. Li is also with the Institute of Digital Communications, School of Engineering, University of Edinburgh, UK.

J. Zhang is with the Department of Electrical Engineering and Electronic, University of Liverpool, UK (email: Junqing.Zhang@liverpool.ac.uk).

S. Podilchak is with the Institute of Digital Communications, School of Engineering, University of Edinburgh, UK (email: S.Podilchak@ed.ac.uk).

radiation patterns [12], [13], as well as their combinations [10], [14]–[20]. In a wireless communication system, the PAs are the main sources of the system non-linearity [8], [10], and their features in the entire link will be further researched for the RFF purpose in this paper.

The RFF features of legitimate wireless devices are commonly extracted under controlled environments [5], and they are learnt and classified using various machine learning algorithms. For example, the classification algorithm of the hypothesis test was employed to classify the RFF features originated from RF carrier oscillators [7] and RF PAs [8]. Other classification methods like deep neural network (DNN) [9] and convolution neural network (CNN) [15]–[20] were also found useful for some types of RF features. The effectiveness of the machine learning techniques relies significantly on the choice of the targeted RFF features. To date, achieving high classification accuracy within the low signal to noise ratio (SNR) region is still a challenge.

A. Related Works

The RFF feature originates from the components in RF front ends. In [7], the unique characteristics of the phase noise that is generated by RF carrier oscillators were extracted for RFF. Polak et al. in [8] exploited the unique I/O characteristics of PAs and digital to analogue converters (DACs). Here, the PA Volterra-series behavior model was employed. This reported RFF system had no classification errors when SNR is greater than 35 dB. In [9], the frequency domain features of the received symbols due to the non-linear RF front end amplitude modulation to amplitude modulation (AM/AM) characteristics were used for wireless device classification. It demonstrated the possibility of using the PA behavior model to emulate different RFFs. In [10], the RF front end non-linearity feature in the frequency domain was extracted. It was shown that the RFF performance is also affected by modulation schemes and digital shaping filters. Moreover, a concept of combining physically unclonable function (PUF)

and RFF was proposed to enhance the PHY layer security in [11]. It exploited a PUF-controlled PA spectral regrowth for RFF, which consists of a digital PUF and a PA whose bias was controlled by a DAC. In [12], the scattering modes of antennas as the RFF feature were exploited. The scattering signals for different types of antennas with varying load conditions were studied. However, this system can only be used for distinguishing different types of antennas. In [13], the authors studied the manufacturing tolerances in the antenna arrays used in wireless devices operating in millimeter wave (mmWave) bands, and the RFF feature was based on the unique beam patterns of different codebooks used by the devices. However, the antennas (or arrays) have to be static in the environment [12], [13], which becomes impractical in mobile wireless communication scenarios.

Image-based RFF has been widely studied. For instance, in [14] the received constellation diagrams (CDs) were converted into a type of pattern images in order to extract DC offset and I/Q imbalance features originated from the RF front end. A step further, works in [15] to [18] described a differential signal processing method which exposes the frequency feature of the received signals. While in [19], [20], the CDs, after being converted into contour images, were found useful to extract RF features of the entire RF chain.

B. Our Contributions

PAs' characteristics have been studied and considered for RFF applications in previous works. For example, the unique transient turn-on waveforms based on the PA or entire RF front end were studied in [21], [22], and PA behavior models were used for PA AM/AM characterizations and RFF analyses in [8]–[10].

It is well known that a PA experiences non-linear features, and the most prominent one is its non-linear AM/AM characteristic, which, commonly, is memoryless. Indeed, this feature can be used for RFF, which has already been studied in [8], [9]. However, it is not very distinctive as the AM/AM

TABLE I.
SOME EXISTING RFF WORKS AND OUR PROPOSED SYSTEM.

Work	RFF feature	RF device	Classification accuracy
[7]	PLL's phase noise	RF carrier oscillator	Cabled experiment: No classification error @ SNR = 35 dB.
[8]	I/O characteristics	PA + DAC	No classification error @ SNR = 35 dB.
[9]	Non-linear feature in frequency domain of the received symbols	Entire RF chain	Below 80% @ SNR < 15 dB.
[10]	Spectral regrowth	RF front end	Around 80% @ SNR = 15 dB.
[11]	Spectral regrowth	PUF + PA	144 devices can be classified with 95% accuracy in measurement.
[12]	Scattering modes of the antenna	Antenna	78% @ SNR = 10 dB for different types of antennas.
[13]	Antenna radiation beam pattern	Antenna	95% @ SNR = 15 dB.
[14]	DC offset and I/Q imbalance	Entire RF chain	80% to 90%, SNR was not reported.
[15]	I/Q phase mismatch, I/Q DC offset, I/Q gain imbalance and carrier frequency offset	Entire RF chain	93.8% @ SNR = 15 dB.
[18]	RF signal information in frequency domain	Entire RF chain	99.6% @ SNR = 5 dB
[19]	RF signal information in time domain	Entire RF chain	99% @ SNR = 0 dB
Our work	Non-linear memory effect	RRC + PA + RRC	Cabled experiment: 92% @ SNR = 10 dB; OTA experiment: 90% (LOS) @ SNR = 10 dB; 79% (NLOS) @ SNR = 12 dB.

differences among many PAs, especially when operating in more linear region, can be tiny. For example, in the work [8] no identification errors occur only when the SNR is 35 dB or higher for the cabled experiment; In [9], the classification accuracy of the non-linear RFF feature in frequency domain drops below 80% when the SNR is less than 15 dB. Our presented work here is not another attempt of using PAs' non-linear (either transient or steady states) property for RFF, but it, built upon our preliminary work in [23], provides a comprehensive study on the non-linear memory effect that is generated by the cascade of a transmit root-raised-cosine (RRC) filter, a non-linear PA, and a receive RRC filter. In our proposed system, PAs are deliberately driven in the non-linear region which leads to a more prominent non-linear memory effect on CDs and is explored as an RFF feature for device identification. A hybrid classification algorithm is also proposed to harvest the benefit of this memory effect.

In our previous work in [23], only AM/AM characteristics of PAs were studied. This paper further investigates the characteristics of PA's amplitude modulation effect on phase modulation (AM/PM) and demonstrates the feasibility of classification of the same PA models. Our contributions in this paper are summarized as follows:

- We study the distortion of constellation symbols (i.e. irregular shapes of each symbol cluster in I/Q plane) caused by the PA non-linear AM/AM and AM/PM characteristics with matched pulse shaping RRC filters. The resulting non-linear memory effect is explored as the RFF feature, which is validated through both simulations and experiments. In the simulations, the PA Bessel-Fourier behavior model accurately captures AM/AM and AM/PM characteristics that show excellent fitting performance with the measurement results. In this work, the non-linear memory effect exhibited on the received CD is extracted for device identification. In addition to the CD with all modulation symbols, a hybrid classification method that also exploits a sub-set constellation case is proposed to improve the classification performance. The received CDs are converted to a color-contoured 2-D image, named colored-constellation diagram (CCD), which facilitates the feature extraction.

- We explore the possibility of exploring the non-linear memory effect in the time domain which makes the DCTF approach ineffective because of the differential signal processing algorithm that exposes the frequency feature of the received signals in the I/Q plane. It is also worth noting that the DCTF in [15]–[18] and the contour stella image approach in [19], [20] are not mutually exclusive with our proposed CCD-based RFF approach for non-linear memory features extraction. They should be selected based on the dominant features that are unique in the transmitters.

- The targeted RFF feature is extracted and classified from the same and different PA models through two types of experiments: cabled and over the air (OTA) experiments. In the cabled experiment, the device under test (DUT) transmits modulated waveforms to the receiver via well-matched coaxial cables. In the OTA experiment, the performance of the proposed RFF system under the multipath environment,

including both line-of-sight (LOS) and non-line-of-sight (NLOS) scenarios, is evaluated. The average classification accuracy is around 92% at SNR of 10 dB among systems of 8 PAs (4 PAs of the same model and the other 4 of different models) in the cabled experiment. For the OTA experiment, the average classification accuracy of LOS scenario is found to be 90% at SNR of 10 dB, and 79% at SNR of 12 dB for NLOS scenario.

C. Organization

The rest of the paper is organized as follows. Section II elaborates on the origin of the RFF feature to be explored in this paper. Section III introduces the procedure of CCD generation from which the RFF features are extracted with the help of the CNN classification algorithm. In Section IV, an enhanced hybrid approach utilizing both full-set and sub-set CCDs is presented, and the experiment results are also obtained and discussed in this section. Finally, conclusions are drawn in Section V.

II. NON-LINEAR MEMORY EFFECT

A simplified modular illustration of a wireless communication link is shown in Fig. 2. A bit sequence \mathbf{b} to be transmitted is first mapped to complex symbols through the modulator with the resulting symbols denoted as s_m , where ' m ' is the symbol index. In order to reduce crosstalk in adjacent frequency bands, the digital pulse shaping RRC filters are typically applied at the transmitter and the receiver sides. In theory, two cascaded RRC filters behave as a digital raised-cosine (RC) filter.

After the modulated symbol s_m goes through the transmitter RRC filter, a baseband sample sequence \mathbf{u} is generated for frequency up-conversion, amplification, and radiation in the later stages at the transmitter.

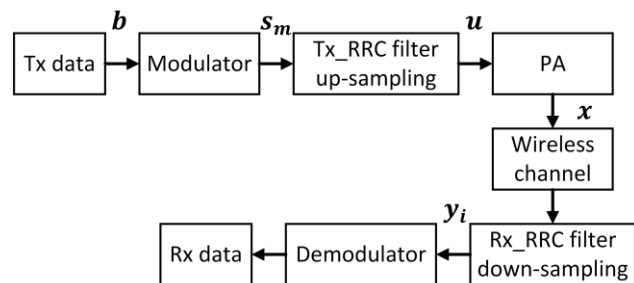


Fig. 2. Illustration of wireless communication system blocks.

A. RRC Filter

In practice, two cascaded RRC filters are not equivalent to a corresponding RC filter since the RRC filters are implemented as finite impulse response (FIR) filters instead of infinite impulse response (IIR) filters. Consequently, inter-symbol interference (ISI) exists with the two cascaded FIR RRC filters. The ISI can be approximated using a complex Gaussian distribution with zero mean [24], [25]. To facilitate our

discussion in this paper, as an example we study a 16-QAM transmission link with a pair of cascaded RRC filters which is shown in Fig. 3(a). The RRC filter time-domain response h_{RRC} is expressed as in (1),

$$h_{RRC}(t) = \frac{2 \cdot \beta}{\pi \cdot \sqrt{T}} \frac{\cos\left[\frac{(1+\beta)\pi \cdot t}{T}\right] + \frac{\sin(1-\beta)\pi \cdot t/T}{4 \cdot \beta \cdot t/T}}{1 - (4 \cdot \beta \cdot t/T)^2}, \quad (1)$$

where β ($0 \leq \beta \leq 1$) and T are the roll-off factor and the symbol period of the RRC filter, respectively. The number of RRC filter taps is given by $N_{tap} = D \cdot S$ [25], [26], where D represents the filter span, which determines the duration of filter response in number of symbols, and S refers to the up-sampling (or down-sampling) rate. T and β determine the filter bandwidth BW_{RRC} , i.e., $BW_{RRC} = (1+\beta)/T$.

In Fig. 3(b) and (c), the received noiseless constellations are shown. It can be seen that the generated ISI is greatly affected by the RRC filter parameters, especially β . In this example, when the $\beta = 0.1$, the received constellation symbols are blurred due to ISI, which is seen to be Gaussian distributed. In contrast, when the $\beta = 0.5$, the ISI is greatly suppressed. In a real wireless link, a PA is inserted in-between of two matched RRC filters, i.e. the first RRC and the PA are at the transmitter end and the second RRC is at the receiver end. When the PA operates in the non-linear region, the transmission link also exhibits memory effects. This is the result of the time spanning of the RRC filters and their non-zero ISI. This non-linear memory effect is studied and explored for RFF purposes. In this paper, unless otherwise stated, the parameters of the RRC filters are set as follows: $\beta = 0.5$, $D = 8$, $T = 100$ ms. It needs to be pointed out that though the choice of $\beta = 0.5$ has extremely low ISI with two RRC filters alone, see Fig. 3(c), the significant ISI occurs when a PA in-between a cascaded RRC filter system operates in the non-linear region, see further discussions in the following Section II.C.

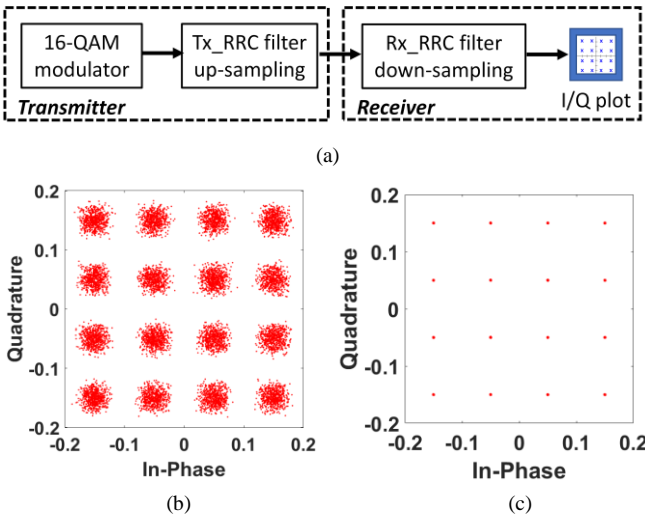


Fig. 3. (a) Simplified link with two cascaded RRC filters. (b) Received 16-QAM symbols with RRC parameters $\beta = 0.1$, $D = 8$, and $T = 100$ ms. (c) Received 16-QAM symbols with RRC parameters $\beta = 0.5$, $D = 8$, and $T = 100$ ms.

B. PA Behavior Model

The memoryless PA behavior models provide an analytical method to describe their AM/AM and AM/PM characteristics. There exist several different behavior models, such as Saleh's [27], Ghorbani's [28], Rapp's [29], and Bessel-Fourier [30], etc. As reported in [30], [31], the Rapp model fluctuates in the transition between linear and saturation regions; the Saleh model accurately fits only the linear region while the Ghorbani model is better suited for the non-linear region. In contrast, the Bessel-Fourier model, selected in our work, fits the entire PA operation region more accurately, despite the higher complexity in terms of more fitting coefficients. With the fitting order of the Bessel-Fourier method selected, the fitting coefficients can be calculated. In order to perform the fitting, the AM/AM and AM/PM conversions of the PA under test are first measured and the curves can then be fitted by determining coefficients using the behavioral model, which is expressed in (2).

$$B(\rho)e^{jF(\rho)} = \sum_{l=1}^L b_l J_l(\alpha \cdot l \cdot \rho) \quad (2)$$

Here in (2), ρ is the magnitude of the PA's input signal, $B(\cdot)$ and $F(\cdot)$ respectively represent the measured AM/AM and AM/PM conversions of the tested PA, and $J_l(\cdot)$ refers to the Bessel function of the first kind with L being the length of the Bessel series. b_l and α are the coefficients to be determined, and l is the index of the Bessel terms. In this work, L of 21 is used hereafter to obtain accurate fitting results.

In this paper, 8 PAs were selected for the study (see Table II), consisting of 6 different PA models. The AM/AM and AM/PM of each PA were measured for 2.4 GHz operating frequency and are fitted using the Bessel-Fourier model, see the fitting results in Fig. 4. The average absolute fitting errors are calculated, and they are extremely small, especially when $P_{in} = 20\log_{10}(\rho)$ is greater than -10 dBm. Generally, from Fig. 4 it can be observed that different PA models have individual AM/AM and AM/PM characteristics. In comparison, only the differences in AM/PM characteristics are noticeable for the PAs of the same model.

TABLE II.
PAS USED IN THIS PAPER.

PA model	Manufacturer	Indexed in this work
ZJL4HG+	Mini-Circuits	PA-1
SKY67023-396LF	SKYWORKS	PA-2
ZJL-4G+	Mini-Circuits	PA-3
ZX60-V62+	Mini-Circuits	PA-4
TRF37A7FEVM (1)	Texas Instruments	PA-5
TRF37A7FEVM (2)	Texas Instruments	PA-6
ZX60-P103LN+ (1)	Mini-Circuits	PA-7
ZX60-P103LN+ (2)	Mini-Circuits	PA-8

C. Non-Linear Memory Effect

Revisiting the proposed wireless communication system

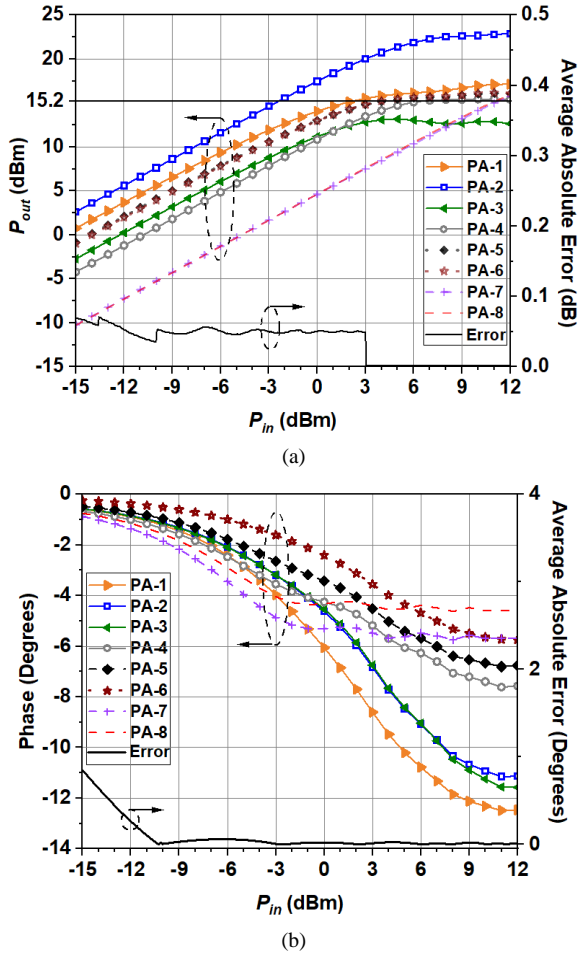


Fig. 4. Bessel-Fourier fitted (a) AM/AM and (b) AM/PM curves of 8 PAs listed in Table II. PAs were biased with typical/recommended drain voltages and currents stated in their datasheets.

shown in Fig. 2, and considering a non-linear PA described by (2), the transmitted signal x can be written as

$$\mathbf{x}(n) = B(\rho_u(n))e^{j(\theta_u(n)+F(\rho_u(n)))}, \quad (3)$$

where ρ_u and θ_u respectively represent the magnitude and the phase of \mathbf{u} , and n is the index of digital samples. At the receiver (identifier) side, the captured signals are first down-converted to the baseband, and then synchronized and demodulated.

Based on (3) and assumed wireless channels, the received signals with embedded non-linear memory effects can be obtained using simulation. Unless otherwise stated in this section, the AWGN channel is assumed in the links for performance simulation. The Bessel-Fourier fitted AM/AM and AM/PM curves are here used to obtain the transmitted signal \mathbf{x} . The average power of the PA output signals \mathbf{x} was kept identical to 15.2 dBm, seen in Fig. 4(a), for PA- $\{1, 2, 4, 5, 6, 7, 8\}$. The output signal power of the remaining PA-3 was set to 12 dBm since it has a saturation power of only 13 dBm. This \mathbf{x} is captured and processed at the receiver end, and the resulting CDs are plotted in Fig. 5(a), (c), and (e). Here only diagrams for PA- $\{1, 2, 5\}$ are shown as examples to

illustrate different PA models with similar or different characteristics, no additive white Gaussian noise (AWGN) and wireless channel are added in the simulation. It is clear that the system of PA-2 operates in a more linear region, making the resulting CD distinctively different from the other systems that operate in the saturated non-linear region. And importantly, comparing the CD in Fig. 3(c) for a link with no non-linear PA inserted, we can see the much-increased ISI arising from the cascade of the RRC, PA, and RRC, seen in Fig. 5(a) and (g). In our study, it is shown that the cascade of the RRC + PA + RRC generates the non-linear memory effect when PA operates in the non-linear region. This is because of the digital sample expansion in the time domain when RRC filters are involved. Due to the non-linear PA in-between, two matched RRC filters are unable to cancel out ISI. For PAs with different characteristics (e.g. the PA-1 and PA-2, see Fig. 4) the shapes of the entire 16-QAM CDs, as well as the shapes of each constellation symbol cluster, are different. While for the PAs with similar characteristics (e.g. the PA-1 and PA-5), the CDs become visually similar.

In our earlier discussion, the non-linear memory effect is generated due to the cascade of two RRC filters and a non-linear PA in-between. In the statistical form, this symbol-sequence dependent ISI results in irregular shapes of the constellation symbol clusters in the I/Q plane. This is evidenced in Fig. 5(c) and (d), wherein the CD of the non-linear memoryless link only experiences AM/AM compression, i.e., the magnitudes of symbols of larger power are suppressed more; while the CD of the non-linear memory link, in addition to the AM/AM compression, also suffers from the non-linear expansion to each constellation symbol, i.e., irregular shapes of each constellation cluster. The memory effect can magnify the differences between different RF PAs, thus improving the classification accuracy in the low SNR region.

The memory effect indicates that the shape of each constellation symbol cluster is dependent on the symbol sequence transmitted. In the statistical sense, it is dependent on the number and indices of symbols in the data stream. Therefore, when choosing a sub-set constellation in the transmission, the shapes of symbol clusters become different, see the evidence in Fig. 5(a) and (b), as well as Fig. 5(g) and (h). Here we show that the shape of each symbol cluster in the sub-set constellation is not identical to those corresponding symbol clusters in the full-set constellation. Those different constellation shapes can thus be used to further improve identification accuracy. In principle, there are many sub-set constellation options, such as the one shown in Fig. 5(i). After some performance comparison, it is found that most sub-set constellation choices give a similar performance boost. Thus, in our work we select only one sub-set option to demonstrate the concept and the performance improvement. This is illustrated using an example in Fig. 6. Here a second symbol stream is generated from the module *Tx_data_2*, followed by a conditioner (≥ 8 used as an example in this paper), which is employed to remove the unmatched symbols at the transmitter side. In this case, if the symbol index is smaller than 8, they

are removed from the sequence. Subsequently, a sub-set of the constellation can be obtained. Similarly, different sub-set constellation symbols can be generated by changing the criteria of the ‘conditioner’.

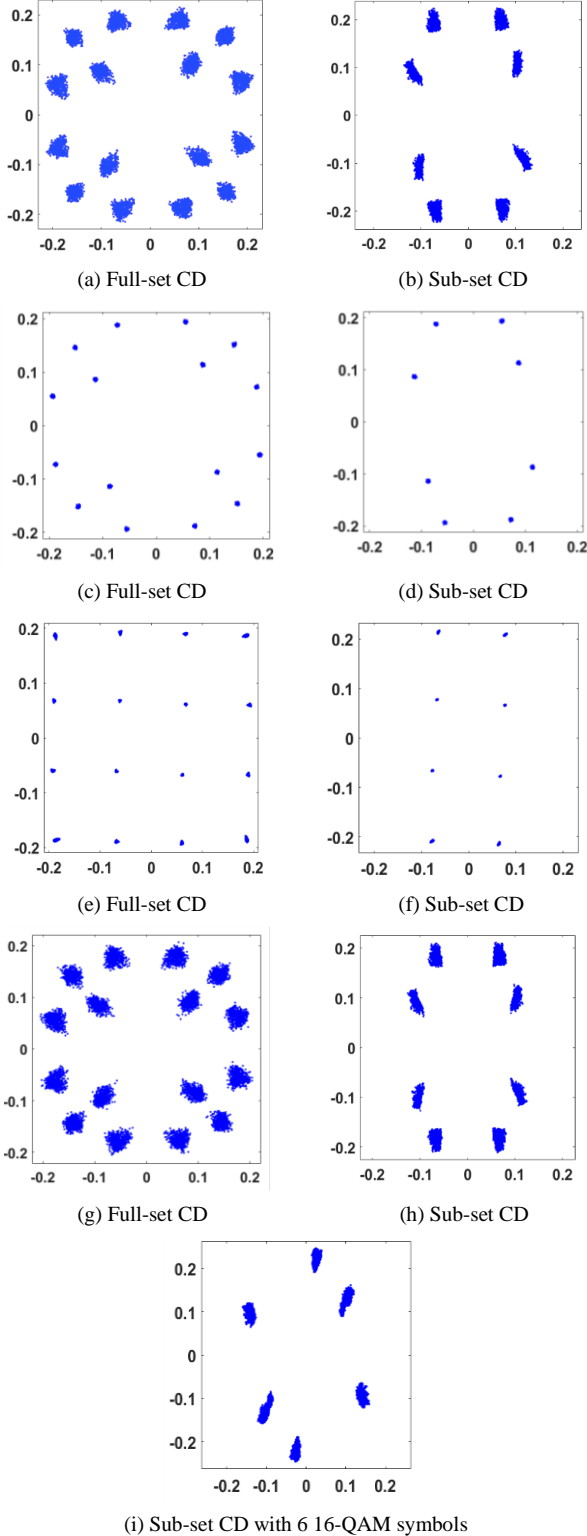


Fig. 5. Simulated CDs. (a) & (b) corresponding to the PA-1 memory effect; (c) & (d) corresponding to PA-1 memoryless effect; (e) & (f) corresponding to the PA-2 memory effect; (g) & (h) corresponding to PA-5 memory effect; (i) corresponding to the PA-1 memory effect. No AWGN is included in the simulation.

The corresponding sub-set constellations are presented in Fig. 5(b), (f), and (h). It needs to be highlighted that the shapes of constellation clusters in the sub-set case are different from those corresponding clusters in the full-set case, confirming that the ISI depends on the transmitted symbols. Though the full-sets and sub-sets of constellations for the PA-1 and PA-5 look similar, they are different after some processing that is discussed in the following section.

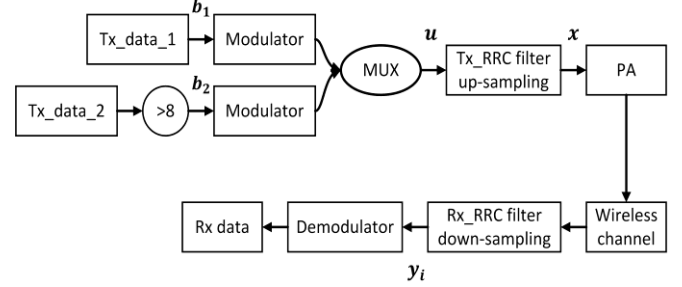


Fig. 6. Illustration of an example wireless communication system for the generation of full-set and sub-set constellations.

III. CLASSIFICATION

A. Colored Constellation Diagram (CCD)

When the data sequence is fixed or known, it is possible to extract RFF features directly from raw I/Q samples in a time sequence. This, comparing to the features in CCDs, will also include some features in the time domain, which will inevitably enhance the RFF classification. While for our design in this work, we aim to keep the method more generic, i.e., we do not assume the data sequence is fixed or known. This provides the possibility of using data payloads that consist of random data sequences for RFF. This random sequence makes the extraction of time-domain feature infeasible. Thus, only the shape and density of constellation clusters in I/Q plane, namely CCDs, can be exploited for RFF classification. Algorithm 1 elaborates the procedure of the CCD generation. After scaling the q -th ($q = 1, \dots, Q$, here $Q = 16$ for the full-set case and $Q = 8$ for the sub-set case) constellation symbol cluster in its coordinate system with the origin setting to the corresponding $p_r^{(q)}$, the maximum symbol offset in the q -th constellation symbol cluster can be obtained, expressed as $r_{\max}^{(q)}$. Selecting $r^{(q)} = r_{\max}^{(q)}/2$ as the reference magnitude, the ratio $R^{(q)}$ between the reference $r^{(q)}$ and the magnitude of any received symbols in the q -th constellation symbol cluster can be computed, which is expressed in (4).

$$A(n_q)^{(q)} = g\left[R(n_q)^{(q)}\right] = g\left[\frac{r^{(q)}}{|(p_r^{(q)} - y(n_q)^{(q)})|}\right] \quad (4)$$

Here $g[\cdot]$ refers to a scale function corresponding to the heatmap, $y(n_q)$ is the data point in q -th constellation, and A is the scale of the color density from 0 to 1.

After interpolation, a color contoured plot for the q -th constellation cluster is generated. All interpolated constellation clusters are then merged to construct a CCD. In our study, the resulting CCDs are saved as a JPEG image file with the pixel size of (height, width) = (400, 500), and the dots per inch (dpi) of each CCD is set to 200. This colored representation helps reveal the subtle differences between constellation clusters in systems with very similar PA behaviors, see those for PA-1 and PA-5 in Fig. 7.

Algorithm 1. CCD Generation

- 1: Received data stream detection (full-set or sub-set).
 - 2: Locate Q constellation symbol cluster.
 - $(Q$: No. of constellation clusters, 16 for full-set and 8 for sub-set)
 - for** $q = 1: Q$ **do**
 - 3: Averaging all symbols in q -th constellation cluster to calculate the reference constellation point $p_r^{(q)}$.
 - 4: Calculate the maximum offset $r_{\max}^{(q)}$ of the received symbols in the q -th constellation cluster, and define $r^{(q)} = r_{\max}^{(q)}/2$.
 - 5: Calculate ratio $R^{(q)}$ between $r^{(q)}$ and the magnitude of each symbol in the q -th constellation cluster.
 - 6: Scale all $R^{(q)}$ into the range from 0 to 1, and plot them into their corresponding mesh grids. (Note: A square mesh of size 100×100 is pre-defined).
 - 7: Interpolate to generate a smoothly contoured plot for the q -th cluster
 - end for**
 - 8: Combine all clusters together.
 - 9: Save the entire constellation plot as JPEG image (CCD outputs).
-

B. CNN Classification

A CNN image classification algorithm is utilized to extract and classify the transmitter devices based on the RFF feature exhibited in the CCDs. Generally, an image classification system takes an input image, performs feature weight calculation, and outputs a class or probability of the input image belonging to a pre-trained class. The developed CNN classification system consists of 3 convolutional layers, 2 max-pooling layers and 1 fully connected layer, see details in Table III. The input image size is set to (400, 500, 3), corresponding to (height, width, RGB). The image size in the convolutional layer is set to (3×3) in order to capture the details of RFF in the CCDs. The image filter corresponds to the number of neurons in the three convolution layers, and they are set to 16, 32, and 64, respectively. The convolutional layers contain Rectified Linear Unit (ReLU) activation function, which returns the input elements if they are positive, otherwise, 0 is returned. The max-pooling layers help reduce the size of the input image by combining the outputs of neuron clusters into a single neuron in the prior layer. The fully connected layer uses the softmax activation function to perform the classification among the target devices. According to Table III, hyperparameters, such as minimum batch sizes, the maximal epochs and the initial learning rates, etc, are trained by executing LeNet and learn to improve from the fitness function, which represents accuracy for each solution.

The supervised CNN algorithm commonly includes training and classification stages, see details in Fig. 8. In the CNN training stage, the receiver collects the signals transmitted

from the target devices under high SNR conditions, which can be performed in both cabled and OTA measurements (see discussions in Section IV). After the signals are synchronized and RRC filtered, the training samples are generated. The SNR values of training samples can be varied by adding a variable attenuator in the transmission link or by adding artificial AWGN in the received signal processing stage. The received signals from target DUTs are processed following Algorithm 1 to generate CCDs. In a similar process, the test samples/CCDs are obtained for the same DUTs, and they are ready for classification.

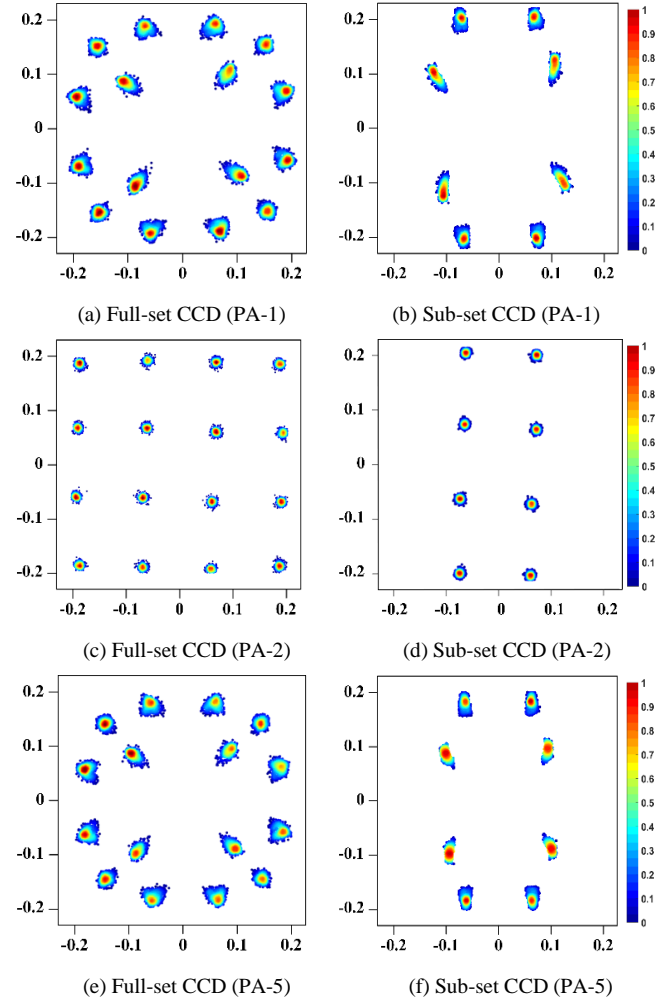


Fig. 7. Full-set and sub-set CCDs for systems with PA-1, PA-2, and PA-5. Different from the constellations in Fig. 6, here the SNR of 30 dB is used in the simulation.

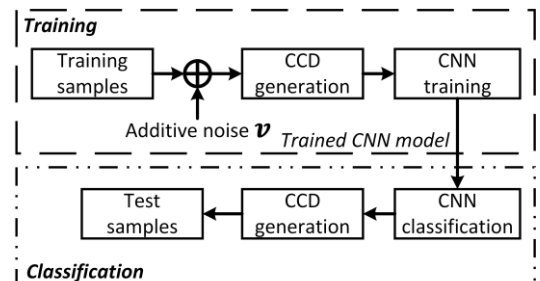


Fig. 8. CNN training and classification process.

TABLE III.
PROPOSED CNN ARCHITECTURE.

Layer	Matrix dimension	Trainable parameters
Input	Image size: (400 × 500 × 3)	-
Convolution 2-D + ReLU function	Filter size: (10 × 10), No. filter: 16	1616
Max-pooling	(2 × 2)	-
Convolution 2-D + ReLU function	Filter size: (10 × 10), No. filter: 32	51232
Max-pooling	(2 × 2)	-
Convolution 2-D + ReLU function	Filter size: (10 × 10), No. filter: 64	204864
Fully connected	Final image size: (100 × 125 × 3)	800001
Hyperparameter	Range	Optimal value
Optimizer	(Rmsprop, Adam, Sgd)	Sgd
Initial learning	(0.001, 0.01, 0.1)	0.001
Mini-batch size	(32, 64, 128)	64
Maximal epochs	(20, 40, 60)	60
Filter size	(5, 10, 15, 20)	10
No. filter 1 st layer	(4, 8, 16)	16
No. filter 2 nd layer	(16, 32, 64)	32
No. filter 3 rd layer	(32, 64, 128)	64

IV. EXPERIMENTAL EVALUATION

In this section, the performance of the proposed RFF classification system is evaluated in the cabled and the OTA measurements.

A. Experimental Set-up

Recap in Table II, 8 PAs, operating at 2.4 GHz, were used to conduct the experiments. The average transmit signal power was kept identical to 15.2 dBm, except for the PA-3 at 12 dBm. The sampling rate was set to 2 MHz. The parameters of RRC filter were set to $\beta = 0.5$, $D = 8$, $T = 100$ ms. With the deliberately chosen long symbol duration T , the ISI is dominantly contributed by the non-linear memory effect of the RF chains, reducing the potential ISI effect from the multi-path channels.

- SNR measurement: The SNR levels of the received signal were obtained by first measuring the background noise on the signal analyzer and then comparing the received signal power with this background noise power.

- Training dataset: The pre-processed data (full-set and the sub-set) were collected for each DUT with sufficiently high SNR, i.e. 40 dB, under both cabled and OTA experiments. In this way, the AWGN can be subsequently applied to the high SNR received data (in MATLAB), artificially producing numerous training samples with varied SNRs. In our study, the SNRs of the training samples were set from 0 dB to 30 dB with a 1 dB step. 10^4 random symbols, forming a packet, were captured to generate each CCD. Two training datasets were obtained for the full-set and the sub-set cases. In each dataset, a total of 1240 packets (or equivalently CCDs) per target DUT were generated and used for the CNN training. The samples in

the dataset were randomly divided into 80% and 20% for training and validation.

- Test dataset: Similarly, in the testing stage, the received symbols were collected for each DUT with varied SNRs and they were converted to CCDs. For each target DUT the test dataset comprises 200 packets per SNR value for both full-set and sub-set cases.

B. Hybrid Classification Method

The hybrid classification is implemented using two trained CNN classifiers associated with full-set and sub-set, respectively. This approach can also help learning algorithms be flexibly adapted to changes in different conditions. In this work, we exploit the non-linear memory effect that is dependent on symbol sequences. Hence, we choose data streams consisting of different symbols to generate and exploit the different features exhibited in full-set and sub-set CCDs. In fact, there are many different criteria for choosing sub-set constellations, such as the one shown in Fig. 5(i). This separately trained hybrid classification approach permits flexible extension to any other hybrid constellation combinations. Here the softmax function is applied in the last layer of the CNN which outputs probabilities to the trained classes. The mathematical expression of the softmax function is written as

$$\eta_i = \sigma_i(\vec{C}) = \frac{e^{C_i}}{\sum_{k=1}^K e^{C_k}} \quad \text{for } i = 1, 2, \dots, K, \quad (5)$$

where σ refers to the softmax function, whose output η_i is the probability of the i -th class, and K is the number of the trained classes. $\vec{C} = (C_1, C_2, \dots, C_K)$ represents the input vector to the softmax function. In this paper, two vectors \vec{C}_f and \vec{C}_s (test samples of full-set and sub-set) are separated inputs to the softmax function. The probabilities of full-set (η_f) and sub-set (η_s) cases are then obtained for the i -th class. Ultimately, the final probability can be computed as $\eta h_i = (\eta_f + \eta_s)/2$, and the class with the highest probability is selected.

C. Cabled Measurement

The cabled measurement set-up is shown in Fig. 9. The transmitter, consisting of a signal generator, a PA under test, and a variable attenuator, is directly connected to the receiver via a well-matched coaxial RF cable.

It is worth noting that further experiment has been carried out to validate that other hardware modules that existed in our experiment links, such as low-noise amplifier (LNA), frequency down-conversion mixer, and band pass filter, do not play a noticeable role in the RFF classification.

In order to explicitly illustrate the contribution of the memory effect on RFF classification performance, a corresponding memoryless link, i.e., removing RRC filters at transmitter and receiver ends, was also studied. To facilitate the discussion, the non-linear memory link (namely

RRC+PA+RRC) is labelled as Model#1 and the non-linear memoryless link (namely PA alone) is referred to as Model#2. The measured full-set and sub-set 16-QAM constellations are shown in Fig.10. It is clear that each constellation symbol cluster in the sub-set CD is identical to the corresponding symbol cluster in the full-set constellation, experiencing no memory effect.

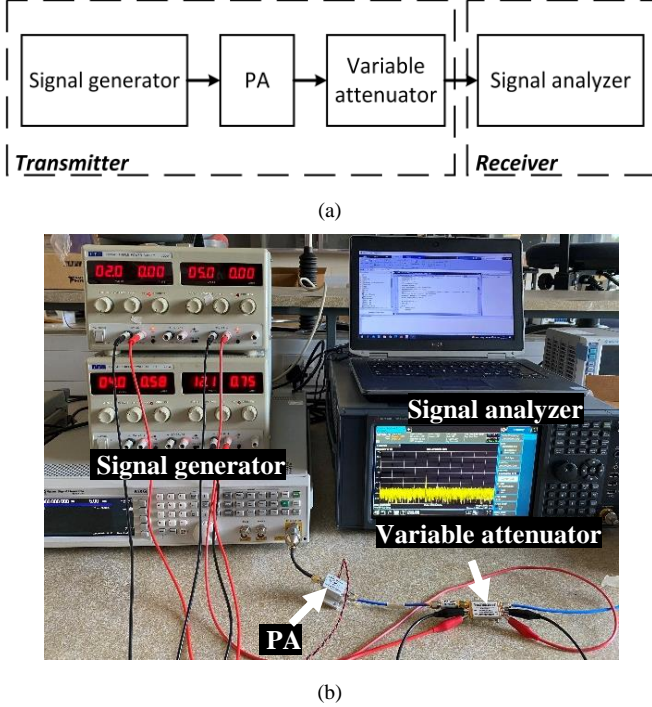


Fig. 9. (a) Block diagram and (b) photo of the cabled experiment.

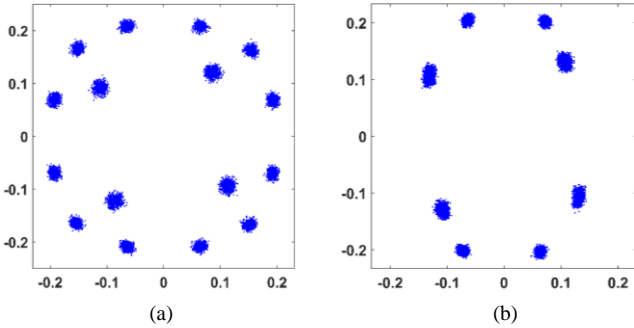


Fig. 10. Normalized (a) full-set and (b) sub-set constellation diagrams in cabled experiment for Model#1. PA-1 was used in the experiment at SNR of 23 dB.

The signal generator's output power was set accordingly to maintain the average transmit power of the PAs under test. The variable attenuator was used here to vary the SNRs of the received sample for the classification test. The links with SNRs of {0, 3, 5, 8, 10, 13, 15, 17, 19, 21, 23} dB were measured. The test CCD samples were subsequently generated from measured symbols of all PA models. Example CCDs for the PA-{1, 2, 5, 7, 8} are presented in Fig. 11, where the SNR is 23 dB. As expected, different PA models lead to significant different CCDs, and there are still visible differences among those of the same PA model, e.g. PA-7 and PA-8.

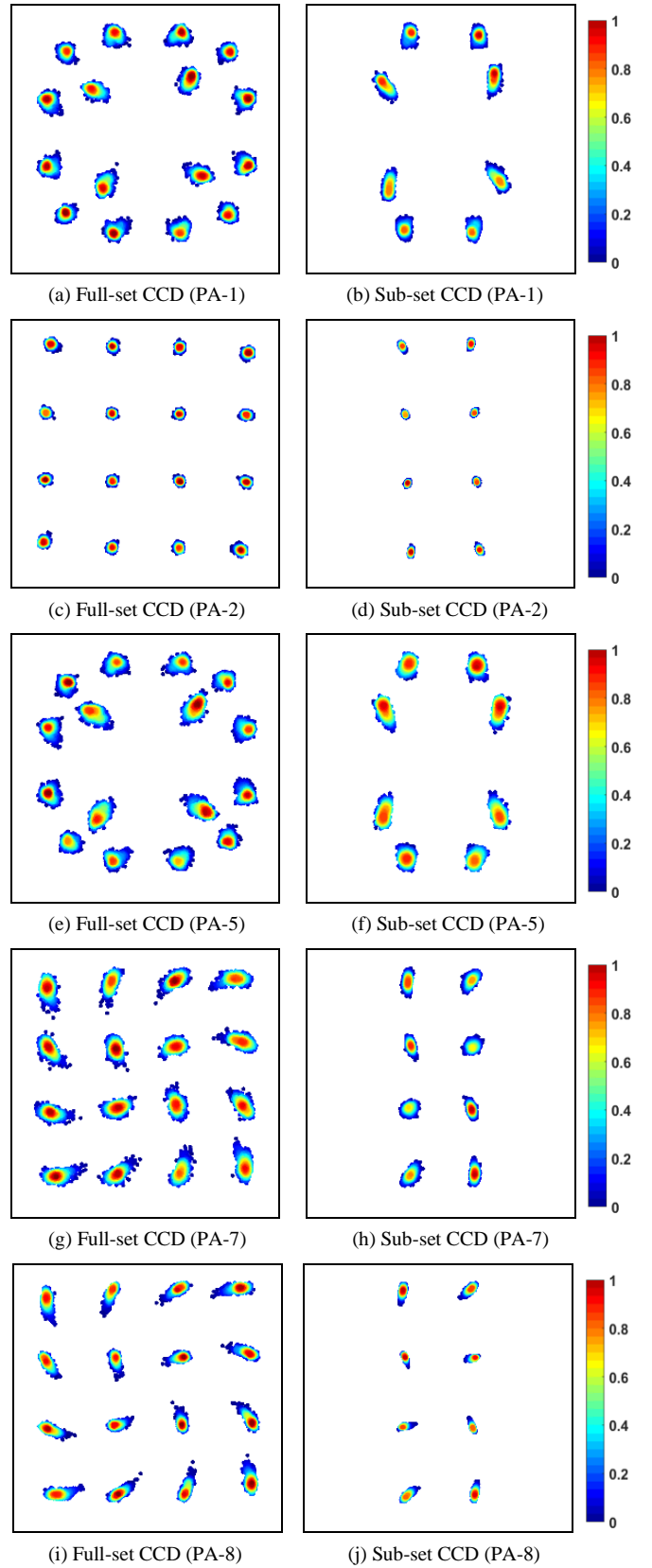
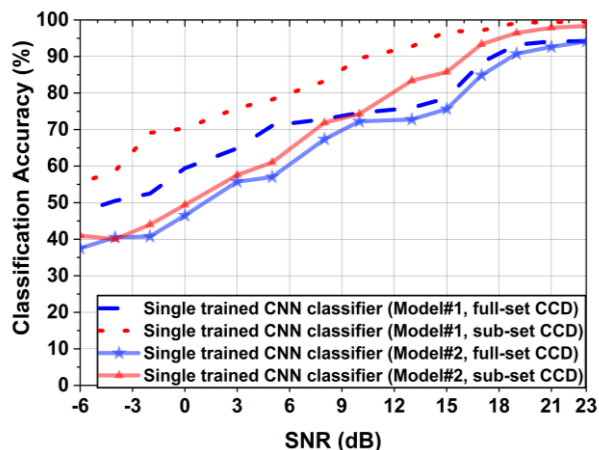


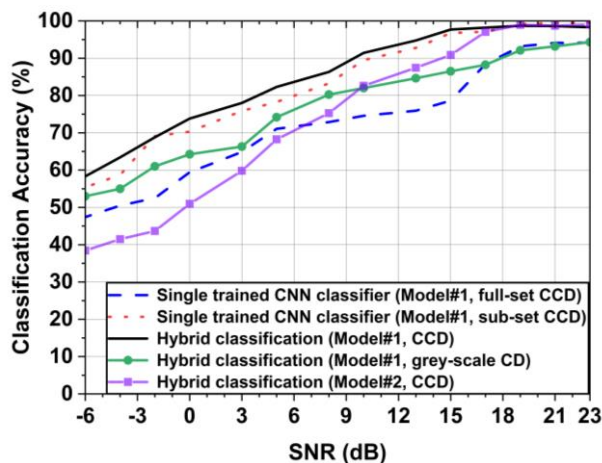
Fig. 11. CCDs associated with the PA-{1, 2, 5, 7, 8}, generated from the cabled measurements at SNR of 23 dB.

The classification accuracy comparisons between these two models for full-set and sub-set cases are plotted in Fig. 12(a). As expected, the memory link (Model#1) achieves better

classification accuracy, especially in low SNR region. The performance improvement for the Model#1 is about 10% for the full-set case and about 18% for the sub-set case. When combining full-set and sub-set for a hybrid training as proposed in this paper, the overall RFF classification accuracy improvement in low SNR region is about 20%, seen in Fig. 12(b). From these results we can conclude that the PA alone with memoryless effect can be used for RFF classification as discussed in the reported work [8], while the memory effect generated by the extra cascaded RRC filters can efficiently boost the classification accuracy at low SNR region wherein the RFF classification normally struggles.



(a)



(b)

Fig. 12. Classification accuracy comparison between the link Model#1 and the link Model#2 for (a) single trained CNN classifier (full-set & sub-set CCD) and (b) hybrid trained CNN classifier in cabled measurement.

In Fig.12(b), the average classification accuracies of the single trained CNN classifier (full-set and sub-set) and the dual-trained CNN classifier are compared for the proposed hybrid classification method. The hybrid approach, exploiting the RFF feature in both full-set and sub-set of CCDs, offers an SNR advantage of more than 9 dB, compared to the result of a single trained CNN classifier based on the full-set case.

In addition, to highlight the role of CCD method for this feature extraction, a comparison of classification accuracies

among the CNN classifier based on CCD training dataset with that of raw I/Q constellation training dataset has been conducted. The raw I/Q CDs have the same image pixel sizes as the CCD. The average classification accuracy based on our proposed CCD method has around 10% advantage. The colored representation helps reveal the subtle differences between constellation clusters in systems with very similar PA behaviors.

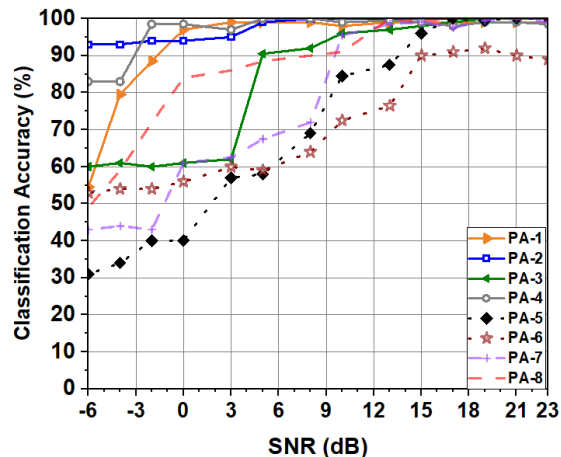


Fig. 13. Classification accuracy based on the hybrid classification method in cabled measurement (Model#1).

It can be observed that in Fig. 12 and Fig. 13, the average classification accuracy for Model#1 is around 92% when the SNR is greater than 10 dB. Among them, high accuracy rate for different PA models (PA-1 to PA-4) is achieved, while the accuracy rate for the same PA models ({PA-5, PA-6} and {PA-7, PA-8}) reduces to 80% and 70%. In the low SNR region, the quality of the RFF feature is more affected by the Gaussian noise. The average classification accuracy drops, especially for those of the same PA models that sit at around 60% at the SNR of 3 dB. The classification accuracy of the PA-3 dramatically decreases when the SNR is below 5 dB, which may be caused by the different output power settings resulting in similar non-linear characteristics with those of the PA-4. The average classification accuracies for hybrid and single-trained CCDs drop to around 60% and 50% when SNR is -6 dB. This performance drop is as expected. Reducing SNR even further becomes challenging since the constellation clusters blur around and CCD generation is unreliable. In low SNR region we can find the random choice occurs frequently in PAs of the same model ({PA-5, PA-6} and {PA-7, PA-8}). The systems of different PA brands behave reasonably well.

The confusion matrices are also calculated, and three cases for high (23 dB), medium (15 dB) and low (5 dB) SNRs are shown in Fig. 14. In the figures, the numbers displayed in the grids refer to the numbers of test samples (out of 200) being correctly (in the trace) or incorrectly (off diagonal) classified. From the confusion matrices, a similar conclusion can be drawn. The PAs of different models (PA-1 to PA-4) maintained high accuracy when the SNR is lower than 10 dB, and conversely, misclassification occurred in the PAs of the

same model ($\{PA-5, PA-6\}$ and $\{PA-7, PA-8\}$). In contrast, different models of PAs achieved sufficiently high classification accuracy even when the SNR is low, as they have more distinct non-linear memory effects.

PA-1	200							
PA-2		200						
PA-3			200					
PA-4			3	197				
PA-5				200				
PA-6				22	178			
PA-7		2				198		
PA-8								200
	PA-1	PA-2	PA-3	PA-4	PA-5	PA-6	PA-7	PA-8

(a)

PA-1	200							
PA-2		200						
PA-3			196	4				
PA-4				200				
PA-5				192	8			
PA-6	2			19	179			
PA-7						200		
PA-8		1						199
	PA-1	PA-2	PA-3	PA-4	PA-5	PA-6	PA-7	PA-8

(b)

PA-1	199				1			
PA-2		199						1
PA-3			181	18		1		
PA-4				200				
PA-5	9			116	75			
PA-6	2			80	118			
PA-7		5				135	60	
PA-8		3		1			18	178
	PA-1	PA-2	PA-3	PA-4	PA-5	PA-6	PA-7	PA-8

(c)

Fig.14. Confusion matrices with SNR of (a) 23 dB, (b) 15 dB, and (c) 5 dB in the cabled measurement.

D. OTA Measurement

The OTA measurements were performed in a laboratory, see the layout and photo in Fig. 15. Both LOS and NLOS scenarios were investigated. Two vertically polarized microstrip patch antennas were equipped at transmitter and

receiver ends, with a realized gain of approximately 2 dBi at the operation frequency of 2.4 GHz.

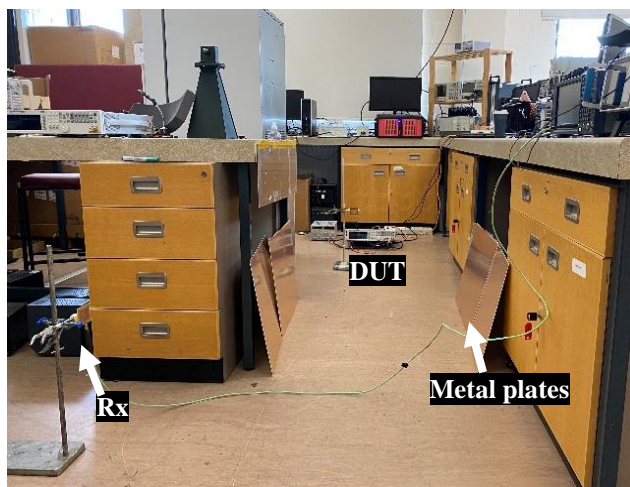
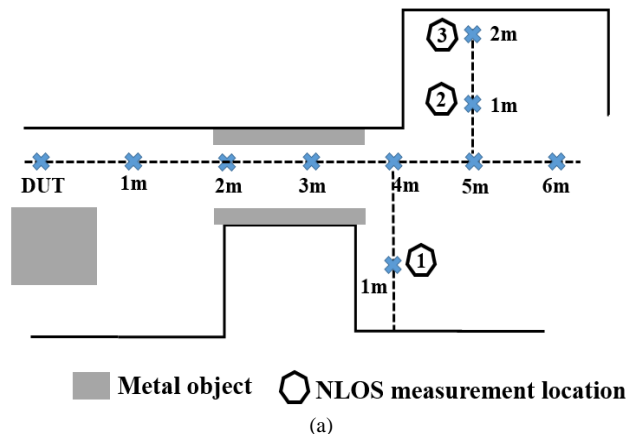


Fig. 15. (a) Layout diagram and (b) photo of OTA experiment.

1) LOS Scenarios

The target DUTs and the receiver were positioned in the same room with a direct LOS link of distance set from 1 m to 6 m. A multipath environment was purposely created by placing some metal objects/plates nearby. The SNRs associated with these 6 LOS locations (from 1 m to 6 m) were measured to be $\{19, 15, 10, 8, 5, 2\}$ dB.

2) NLOS Scenarios

A rich multipath environment was also explored by placing the receiver at 3 labelled locations with furniture and metal objects/plates blocking direct LOS links, as seen in Fig. 15(a). The measured SNRs of the three NLOS scenarios were around 12, 8, and 5 dB for locations 1, 2, and 3, respectively.

3) OTA Classification Results

We first study the classification performance when the CNN is trained using CCDs obtained in the cabled measurements to elucidate the impacts of the antenna and the wireless channel on the overall RFF system. When comparing the accuracies for the LOS and the NLOS cases with those in the corresponding cabled experiment under the same SNRs

(seen in Fig. 16), there is a significant degradation in performance. Because this RFF feature is originated from the transmitter itself, the proposed RFF system should be location, or channel, independent. However, in the OTA experiments the influence of antenna impedance mismatching, the multipath channel, as well as interference from ambient wireless systems operating at 2.4 GHz, as expected, inevitably blur the RFF feature and deteriorate the RFF performance.

In order to include antenna and wireless channel effects, the OTA training dataset was subsequently used. For the OTA training dataset, the receiver collected the transmitted signals from the target DUTs in the LOS far-field measurements with a sufficiently high SNR of greater than 40 dB.

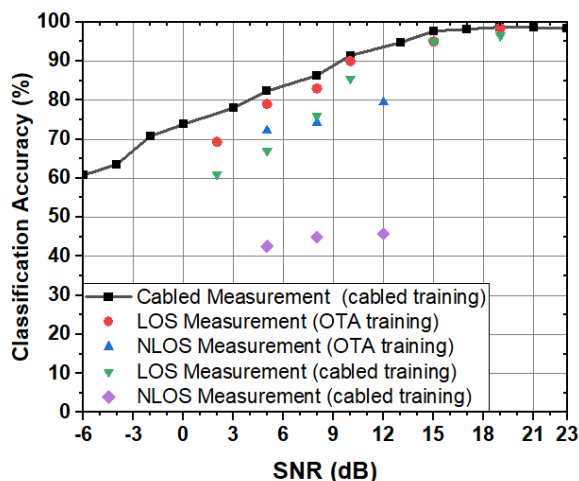


Fig. 16. Classification accuracies in cabled and OTA measurements.

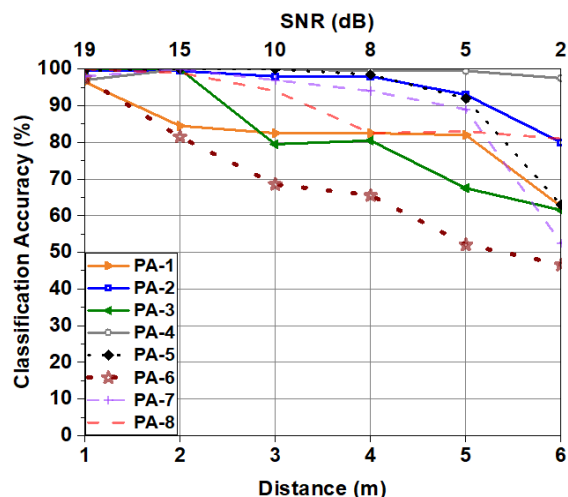


Fig. 17. Classification results in the LOS experiment (based on the OTA training dataset).

The average classification accuracies, shown also in Fig. 16, were found to be significantly improved, compared with the OTA classification result based on the cabled training dataset. The detailed classification performance among 8 PAs (4 PAs of the same model and 4 PAs of different models) is illustrated in Figs. 17 and 18. As compared with the results shown in Fig. 12, here in the OTA case no excessive performance

degradation was observed when the SNR is greater than 10 dB, except those for the PA-1 and PA-2. When the SNR is below 10 dB, the slight gap between the results obtained in the cabled and the LOS scenario may be due to the multipath interference (as the training samples were extracted with much shorter transmitter and receiver distance). With this negative impact of antenna loading and multipath channels, the real-world performance is less robust compared with more ideal simulation and cabled experiment. This can be alleviated by using OTA training dataset which includes some of these effects. The investigation of this difference and some insights can be considered as one contribution of our work. Overall, the average classification accuracy for the LOS scenario is 90% at SNR of 10 dB, and 79% at SNR of 12 dB for the NLOS scenario.

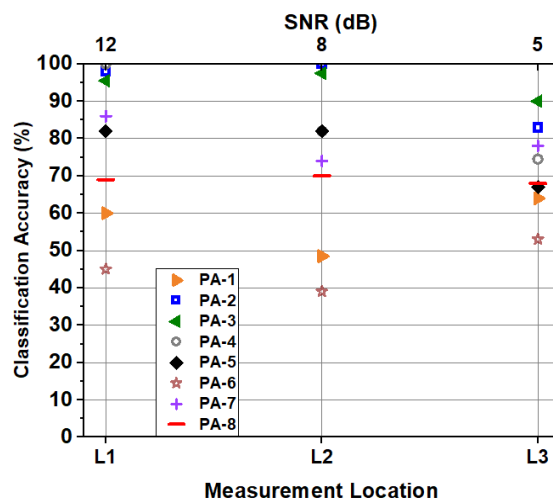


Fig. 18. Classification results in the NLOS experiment (based on the OTA training dataset).

V. CONCLUSION

This paper presented the non-linear memory effect caused by the cascade of transmitter RRC, transmitter non-linear PA and receiver RRC. This non-linear memory effect was extracted from the received CDs, more specifically the newly constructed CCDs, for the RFF device classification. We adopted a CNN-based algorithm for classification and proposed a hybrid classification method to improve the performance. Cabled and OTA measurements were conducted to evaluate the performance of the RFF system. In the cabled experiment, the average classification accuracy among the systems of 8 PAs (4 PAs of the same model and the other 4 of different models) was beyond 92% when SNR was higher than 10 dB. For the OTA experiment, with the SNR of 10 dB the RFF classification system achieved an average accuracy of 90% for the LOS scenario, and 79% at an SNR of 12 dB for the NLOS scenario, when the OTA training data were used. It is possible to use much deeper CNN for RFF classification. To validate this, a deeper CNN configuration VGG-16 was implemented in the study. It has been found that in some cases a small percent improvement of classification accuracy can be obtained compared with the 4-layer CNN used in this paper.

We believe that no observable step-change is because of the relatively simple RFF features in I/Q domain, though benefits of using deeper CNN might be increased with a greater number of device samples.

It is worth noting that in our study apart from the PAs, the remaining modules, such as signal generators and antennas, were kept identical. These modules, however, are different in practical systems. They may contribute to the RFF features in positive or negative fashions, which is an interesting topic worth future research.

REFERENCES

- [1] Cisco Annual Internet Report (2018–2023) White Paper, Cisco System Inc., USA, 2021.
- [2] J. Douceur, “The sybil attack,” in *Proc. 1st IPTPS*, Cambridge, MA, Mar. 2002, pp. 251–260.
- [3] D. B. Faria and D. R. Cheriton, “Detecting identity based attacks in wireless networks using signalprints,” in *Proc. 5th ACM Workshop Wireless Security*, Los Angeles, California, USA, 2006, pp. 43–52.
- [4] J. Yang, Y. Chen, and W. Trappe, “Detecting sybil attacks in wireless and sensor networks using cluster analysis,” in *Proc. 5th IEEE Int. Conf. Mobile Ad Hoc Sensor Syst.*, Atlanta, GA, USA, Sep. 2008, pp. 834–839.
- [5] Q. Xu, R. Zheng, W. Saad, and Z. Han, “Device fingerprinting in wireless networks: challenges and opportunities,” *IEEE Commun. Surveys & Tutorials*, vol. 18, no. 1, pp. 94–104, Firstquarter 2016.
- [6] P. Scanlon, I. Kennedy, and Y. Liu, “Feature extraction approaches to RF fingerprinting for device identification in femtocells,” *Bell Labs Tech. J.*, vol. 15, no. 3, pp. 141–151, Dec. 2010.
- [7] A. Polak and D. Goeckel, “Wireless device identification based on RF oscillator imperfections,” *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 12, pp. 2492–2501, Dec. 2015.
- [8] A. Polak, S. Dolatshahi, and D. Goeckel, “Identifying wireless users via transmitter imperfections,” *IEEE J. Sel. Areas Commun.*, vol. 29, no. 7, pp. 1469–1479, Aug. 2011.
- [9] S. S. Hanna and D. Cabric, “Deep learning based transmitter identification using power amplifier nonlinearity,” in *Proc. Int. Conf. Comput., Netw. Commun.*, Honolulu, HI, USA, 2019, pp. 674–680.
- [10] W. Wang, Z. Sun, S. Piao, B. Zhu, and K. Ren, “Wireless physical-layer identification: modeling and validation,” *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 9, pp. 2091–2106, Sep. 2016.
- [11] Q. Zhou, Y. He, K. Yang, and T. Chi, “Exploring PUF-controlled PA spectral regrowth for physical-layer identification of IoT nodes,” in *Proc. IEEE Int. Conf. Solid-State Circuit*, San Francisco, CA, USA, 2021, pp. 204–206.
- [12] Y. Ma and Y. Hao, “Antenna classification using Gaussian mixture models (GMM) and machine learning,” *IEEE Open J. Antennas Propag.*, vol. 1, pp. 320–328, 2020.
- [13] S. Balakrishnan, S. Gupta, A. Bhuyan, P. Wang, D. Koutsonikolas, and Z. Sun, “Physical layer identification based on spatial-temporal beam features for millimeter-wave wireless networks,” *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 1831–1845, 2020.
- [14] K. Sankhe, “No radio left behind: radio fingerprinting through deep learning of physical-layer hardware impairments,” *IEEE Trans. Cognitive Commun. & Netw.*, vol. 6, no. 1, pp. 165–178, Mar. 2020.
- [15] L. Peng, J. Zhang, M. Liu, and A. Hu, “Deep learning based RF fingerprint identification using differential constellation trace figure,” *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 1091–1095, Jan. 2020.
- [16] Y. Xing, A. Hu, J. Zhang, L. Peng, and G. Li, “On radio frequency fingerprint identification for DSSS systems in low SNR scenarios,” *IEEE Commun. Lett.*, vol. 22, no. 11, pp. 2326–2329, Nov. 2018.
- [17] L. Peng, A. Hu, J. Zhang, Y. Jiang, J. Yu, and Y. Yan, “Design of a hybrid RF fingerprint extraction and device classification scheme,” *IEEE Internet Things J.*, vol. 6, no. 1, pp. 349–360, Feb. 2019.
- [18] Y. Peng, P. Liu, Y. Wang, G. Gui, B. Adebisi and H. Gacanin, “Radio frequency fingerprint identification based on slice integration cooperation and heat constellation trace figure,” *IEEE Wireless Commun. Lett.*, vol. 11, no. 3, pp. 543–547, Mar 2022.
- [19] Y. Lin, Y. Tu, Z. Dou, L. Chen and S. Mao, “Contour stella image and deep learning for signal recognition in the physical layer,” *IEEE Trans. Cognitive Commun. & Netw.*, vol. 7, no. 1, pp. 34–46, Mar. 2021.
- [20] B. Tang, Y. Tu, Z. Zhang and Y. Lin, “Digital signal modulation classification with data augmentation using generative adversarial nets in cognitive radio networks,” *IEEE Access*, vol. 6, pp. 15713–15722, 2018.
- [21] J. Toonstra and W. Kinsner, “Transient analysis and genetic algorithms for classification,” in *Proc. IEEE WESCANEX, Commun, Power, Comput. Conf.*, vol. 2, May 1995, pp. 432–437.
- [22] M. Köse, S. Taşcioglu, and Z. Telatar, “RF fingerprinting of IoT devices based on transient energy spectrum,” *IEEE Access*, vol. 7, pp. 18715–18726, 2019.
- [23] Y. Li, Y. Ding, G. Goussetis, and J. Zhang, “Power amplifier enabled RF fingerprint identification,” in *Proc. IEEE Texas Symp. WMCS*, 2021, pp. 1–6.
- [24] K. Mukumoto and T. Wada, “Realization of root raised cosine rolloff filters using a recursive FIR filter structure,” *IEEE Tran. Commun.*, vol. 62, no. 7, pp.2456–2464, Jul. 2014.
- [25] J. G. Proakis, *Digital Communications, 4th ed.*, New York, USA: McGraw-Hill, ch. 9, 2001.
- [26] M. Cañavate Sánchez, A. Segneri, S. Kosmopoulos, Q. Zhu, T. Tsiptsis, A. Georgiadis, and G. Goussetis, “Novel data pre-distorter for APSK signals in solid-state power amplifiers,” *IEEE Trans. Circuits & Syst.*, vol. 66, no. 10, pp. 4044–4054, Oct. 2019.
- [27] A. Saleh, “Frequency-independent and frequency dependent non-linear models of TWT amplifiers,” *IEEE Trans. Commun.*, vol. 29, no. 11, pp. 1715–1720, 1981.
- [28] A. Ghorbani and M. Sheikhan, “The effect of solid state power amplifiers (SSPAs) non-linearities on MPSK and M-QAM signal transmission,” in *Proc. 6th Int. Conf. Digital Proc. Signals Commun.*, 1991, pp. 193–197.
- [29] C. Rapp, “Effects of HPA-nonlinearity on a 4- DPSK/OFDM-signal for a digital sound broadcasting system,” in *Proc. 2nd Eur. Conf. Satellite Commun.*, 1991, pp. 179–184.
- [30] J. Fuenzalida, O. Shimbo, and W. Cook, “Time domain analysis of intermodulation effects caused by non-linear amplifiers,” *Comsat. Technical Rev.*, vol. 3, no. 1, pp. 88–138, 1973.
- [31] M. Cañavate Sánchez, A. Segneri, A. Georgiadis, S. Kosmopoulos, G. Goussetis, and Y. Ding, “System performance evaluation of power amplifier behavior models,” in *IET Active Passive RF Dev. Seminar*, London, UK, Apr. 30, 2018.



satellite communication.

Yuepei Li (Student member, IEEE) received the M.Eng degree in electrical & electronics engineering from Heriot-Watt University, UK, in 2019. He is currently pursuing the Ph.D. degree with the School of Engineering and Physical Sciences, Heriot-Watt University, and The University of Edinburgh, UK. His primary research interests include the areas of wireless security solution PHY layer radio frequency fingerprinting (RFF) and the miniaturized antennas for small



Yuan Ding (Member, IEEE) received the bachelor's degree in electronic engineering from Beihang University, Beijing, China, in 2004, the master's degree in electronic engineering from Tsinghua University, Beijing, in 2007, and the Ph.D. degree in electronic engineering from the Queen's University of Belfast, Belfast, U.K., in 2014. He was a Radio Frequency (RF) Engineer with the Motorola Research and Development Centre, Beijing, from 2007 to 2009, before joining Freescale Semiconductor, Inc., Beijing, as an RF Field

Application Engineer, responsible for high-power base-station amplifier design from 2009 to 2011. He is currently an Assistant Professor with the Institute of Sensors, Signals and Systems, Heriot-Watt University, Edinburgh, U.K. His research interests are in antenna arrays, physical layer security, and 5G-related areas. Dr. Ding was a recipient of the IET Best Student Paper Award at the Loughborough Antennas and Propagation Conference (LAPC) 2013 and the Young Scientists Awards in the General Assembly and Scientific Symposium, 2014 XXXIst International Union of Radio Science (URSI).



Junqing Zhang received the B.Eng and M.Eng degrees in Electrical Engineering from Tianjin University, China in 2009 and 2012, respectively, and the Ph.D degree in Electronics and Electrical Engineering from Queen's University Belfast, UK in 2016. From Feb. 2016 to Jan. 2018, he was a Postdoctoral Research Fellow with Queen's University Belfast. From Feb. 2018 to May 2020, he was a Tenure Track Fellow (Assistant Professor) with the University of Liverpool, UK. Since June 2020, he is a Lecturer (Assistant

Professor) with the University of Liverpool. His research interests include the Internet of Things, wireless security, physical layer security, key generation, and radio frequency fingerprint identification.



George Goussetis (Senior Member, IEEE) received the Diploma degree in electrical and computer engineering from the National Technical University of Athens, Athens, Greece, in 1998, the B.Sc. degree in physics (First Class) from University College London, London, U.K., in 2002, and the Ph.D. degree from the University of Westminster, London, in 2002. In 1998, he joined Space Engineering, Rome, Italy, as an RF Engineer, and in 1999, the Wireless Communications Research Group, University of Westminster, as a Research

Assistant. From 2002 to 2006, he was a Senior Research Fellow with Loughborough University, Loughborough, U.K. He was a Lecturer (Assistant Professor) with Heriot-Watt University, Edinburgh, U.K., from 2006 to 2009, and a Reader (Associate Professor) with Queen's University Belfast, Belfast, U.K., from 2009 to 2013. In 2013, he joined Heriot-Watt University as a Reader and was promoted to a Professor in 2014, where he currently directs the Institute of Sensors Signals and Systems. He has authored or co-authored over 500 peer-reviewed articles five book chapters one book and four patents. His research interests are in the areas of microwave and antenna components and subsystems. Dr. Goussetis has held a research fellowship from the Onassis Foundation in 2001, a research fellowship from the U.K. Royal Academy of Engineering from 2006 to 2011, and European Marie-Curie experienced researcher fellowships in 2011 and 2012 and again in 2014 and 2017. He was a co-recipient of the 2011 European Space Agency Young Engineer of the Year Prize, the 2011 EuCAP Best Student Paper Prize, the 2012 EuCAP Best Antenna Theory Paper Prize, and the 2016 Bell Labs Prize. He has served as an Associate Editor for IEEE Antennas and Wireless Propagation Letters.



Symon K. Podilchak (Member, IEEE) received the B.A.Sc. degree in engineering science from the University of Toronto, Toronto, ON, Canada, in 2005, and the M.A.Sc. and Ph.D. degrees in electrical engineering from Queen's University, Kingston, ON, Canada, in 2008 and 2013, respectively. From 2013 to 2015, he was an Assistant Professor with Queen's University. In 2015, he joined Heriot-Watt University, Edinburgh, U.K., as an Assistant Professor, and became an Associate Professor in 2017. He is currently a Senior Lecturer with the

School of Engineering, The University of Edinburgh, Edinburgh, U.K. His research interests include surface waves, leaky-wave antennas, metasurfaces, UWB antennas, phased arrays, and RF integrated circuits. He has had industrial experience as a computer programmer, and has designed 24 and 77 GHz automotive radar systems with Samsung and Magna Electronics. His recent industry experiences also include the design of high frequency surface-wave radar systems, professional software design and implementation for measurements in anechoic chambers for the Canadian Department of National Defence and the SLOWPOKE Nuclear Reactor Facility. He has also designed compact antennas for wideband military communications, highly compact circularly polarized antennas for CubeSats with COM DEV International and The European Space Agency, and new wireless power transmission systems for Samsung. Dr. Podilchak and his students were the recipient of many best paper awards and scholarships, most notably research fellowships from the IEEE Antennas and Propagation Society (AP-S), the IEEE Microwave Theory and Technology Society (MTT-S), the European Microwave Association, and six young scientist awards from the International Union of Radio Science. He was the recipient of the Postgraduate Fellowship from the Natural Sciences and Engineering Research Council of Canada (NSERC). In 2011, 2013, 2020,

and 2021, he and his students received student paper awards at the IEEE International Symposium on Antennas and Propagation; in 2012, the Best Paper Prize for Antenna Design at the European Conference on Antennas and Propagation for his work on CubeSat antennas; and in 2016, the European Microwave Prize for his research on surface waves and leaky-wave antennas. In 2017 and 2019, he was bestowed a Visiting Professorship Award at Sapienza University, Rome, Italy, and from 2016 to 2019, his research was supported by a H2020 Marie Skłodowska-Curie European Research Fellowship. He was recognized as an outstanding reviewer of IEEE Transactions on Antennas and Propagation, in 2014 and 2020. In 2021, he was the recipient of the COVID-19 Above and Beyond Medal for leading research on remote microwave sterilization of the coronavirus. He was also the Founder and First Chairman of the IEEE AP-S and IEEE MTT-S Joint Chapters in Canada and Scotland, in 2014 and 2019, respectively. In recognition of these services, he was presented with an Outstanding Volunteer Award from IEEE in 2015; and in 2020 and 2021, MTT-S and AP-S, respectively, recognized this Scotland Chapter for its activities and it was awarded the winner of the Outstanding Chapter Awards hosted by these two IEEE societies. He was also the recipient of the Outstanding Dissertation Award for his Ph.D. He was a Lecturer with the European School of Antennas and an Associate Editor for IET Electronic Letters and IEEE Antennas and Wireless Propagation Letters. He is currently a Guest Associate Editor of IEEE Open Journal of Antennas.