IEEE Open Journal of the
Solid-State Circuits Society

Received 15 April 2024; revised 12 July 2024; accepted 26 August 2024. Date of publication 30 August 2024; date of current version 16 September 2024.

Digital Object Identifier 10.1109/OJSSCS.2024.3451401

Enhancing RF Fingerprint Generation in Power Amplifiers: Unequally Spaced Multitone Design Approaches and Considerations

CHENGYU FAN (Graduate Student Member, IEEE), JUNTING DENG, ETHAN CHEN, AND VANESSA CHEN (Senior Member, IEEE)

(Invited Paper)

Electrical and Computer Engineering Department, Carnegie Mellon University, Pittsburgh, PA 15213, USA

CORRESPONDING AUTHOR: C. FAN (e-mail: chengyuf@andrew.cmu.edu)

This work was supported in part by the Army Research Office under Award W911NF-23-1-0073; in part by the Office of Naval Research under Award N000142412507; and in part by the National Science Foundation under Grant 1952907, Grant 1953801, and Grant 2028893.

ABSTRACT The rapid growth of Internet of Things (IoT) devices and communication standards has led to an increasing demand for data security, particularly with limited hardware resources. In addition to conventional software-level data encryption, physical-layer security techniques, such as device-specific radio frequency fingerprints (RFFs), are emerging as promising solutions. This article first summarizes prior arts on timestamped RFFs generation and reconfigurable power amplifier (PA) designs. Following that, an innovative 2-stage PA incorporating a reconfigurable class A stage with a Doherty amplifier, designed in 65-nm CMOS to generate 4096 timestamped RFFs without introducing in-band power variation, is presented. Multiple 3-bit resistive digital-to-analog converters (RDACs) are applied to control body biasing units within the two-stage PA, facilitating the generation of massive and distinguishable RFFs. Subsequently, time-varying unequally spaced multitone (USMT) techniques are proposed to further elevate the count of available timestamped RFFs from 4096 to 16 384. The validation results of RFFs utilizing 64-QAM WiFi-6E advertising packets, employing time-varying USMT transmitted within the 5.39–5.41-GHz band, confirm the successful generation of 16 384 distinct RFF patterns. Moreover, the measurement results demonstrate that more than 11 504 RFFs among the generated patterns can be classified with an accuracy exceeding 99%.

INDEX TERMS Combinatorial randomness (CR), Internet of Things (IoT), machine learning, neural networks, power amplifier (PA), radio frequency fingerprint (RFF), time-varying tone insertion, unequally spaced multitone (USMT).

I. INTRODUCTION

THE IMPORTANT security challenge associated with ubiquitous wireless connections has significantly increased due to new entry points from wireless devices for malicious acts [1], [2], [3]. While advanced security measures, such as physical-layer authentication using unique optical features [4] and directional beamforming [5], have been investigated, they may not be suitable for resource-constrained Internet of Things (IoT) nodes. Radio frequency fingerprint (RFF) technologies have been exploited to secure IoT devices [6], [7], [8], [9]. However, the variety of classifiable RFFs resulting from inherent manufacturing variations is very limited. In [10], spectral regrowth-based features from a power amplifier (PA) are tuned to generate 16 RFFs, and this number is further expanded to 220 in the previous work by using feature augmentation of a reconfigurable PA [11], [15]. However, such limited tunability is still vulnerable to interception by adversary receivers using prolonged observation time to mimic the transmission signals. Furthermore, as noted in [11] and [15], the performance of the combinatorial-randomness (CR)-based PA is found to be sensitive to process variation and

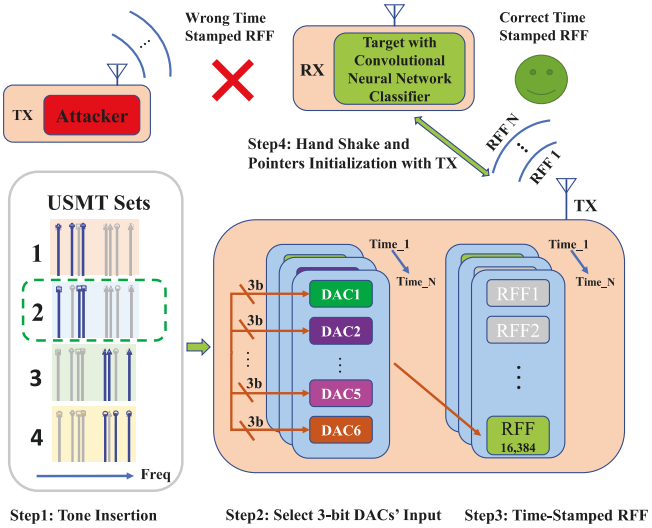

FIGURE 1. Time-varying RFF system utilizes various USMT frequency sets to enhance built-in RF features for generating timestamped signatures, effectively addressing security threats, such as malicious data injection or node cloning by modern intelligent transceivers.

layout-related parasitic elements due to the discrete structure of the PA. In prior research [16], to achieve the necessary level of PA tunability while maintaining PA performance, a significant number of digital-to-analog converters (DACs) are utilized to control the body voltage of PA devices. Moreover, adjustments to the power supply voltage and gate biasing voltage are frequently made to support the performance of the PA under complex operating conditions. Indeed, these techniques effectively enhanced the tunability of the PA and minimized performance variations during reconfiguration. Nevertheless, the requirement for a large number of redundant PA configurations to extract only a limited number of constant-performance PA configurations is undesirable. This work introduces a reconfigurable 2-stage Doherty PA aimed at addressing multiple challenges: minimizing performance variation caused by RFF generation, increasing the number of RFFs, and facilitating faster RFF authentication. Through these advancements, our approach enables the generation of over 16 000 timestamped RFFs. As depicted in Fig. 1, a reconfigurable transmitter integrated with timestamped RFF signatures is designed to continuously update RFFs. This is achieved by selecting various unequally spaced multitone (USMT) frequency sets and updating all six DAC-controlled body voltages. Time-varying USMT are transmitted between each data packet to facilitate rapid authentication at the receiver side, which is equipped with an intelligent classifier. During the handshaking process, the order of timestamped RFF signatures can be synchronized between transmitters and receivers to prevent the need for online exchange. This approach effectively mitigates security threats, such as malicious data injection or node cloning by modern intelligent receivers.

In addition to the generation of RFFs, the extraction of generated RFF features also plays a significant role in physical-layer security. Machine-learning-assisted RFF identification approaches have been investigated in the software field using different types of neural network models, including multilayer perceptron (MLP) [8], [17], [18], convolutional neural network (CNN) [8], [11], [17], [19], and long short-term memory (LSTM) [17], [20]. LSTM is a type of recurrent neural network well suited for time sequence features [17], whereas MLP and CNN are more versatile structures designed for RFF extraction with diverse features. According to experimental results comparing RFF identification using three types of models—MLP, CNN, and LSTM—CNN achieves the highest identification accuracy, while MLP incurs the most parameters due to its fully connected structure [17]. Additionally, the performance of LSTM heavily depends on the type of time sequence input feature [17]. The Bayesian neural network (BNN) [21], [22] represents another lightweight neural network structure suitable for deployment in time-sensitive systems with a relatively small pool of available RFFs. This suitability arises from the nonlinearity introduced by PA characteristics. Hence, given the extensive pool of over 16 000 distinctive RFFs generated from the proposed design in this study, CNN is selected to implement the intelligent classifier at the receiver side. This classifier comprises convolutional layers for feature extraction and fully connected layers for decision making, as elaborated in Section VI.

II. RFF GENERATION TECHNIQUES

A. COMBINATORIAL CLASS-E PAS [11], [15]

In Fig. 2, a Class-E PA featuring RFF capabilities is designed using the CR concept. The main power transistor is segmented into 12 selectable elements. As stated in Pelgrom's paper [27], the standard deviations of the threshold voltage (V_{th}) and transistor gain (β) parameters of MOSFETs are inversely proportional to the square root of the MOSFETs' area. Furthermore, research in [11] demonstrates that the magnitude and phase deviation of the second harmonic of the PA is more sensitive to variations in threshold voltage (V_{th}) compared to the fundamental tone. Hence, by minimizing the impact on the performance of the fundamental tone, the segmented PA structure can produce more distinguishable nonlinear behaviors compared to conventional concentrated PAs.

By selecting 9 out of the total 12 PA blocks during each authentication period, it is possible to generate 220 distinguishable PA combinations. Specific RFF features embedded in the transmitted signal are updated regularly, facilitating the generation of timestamped RFFs to enhance the system's security. In this study, the preambles of each data packet are utilized for training the CNN, ensuring that the authentication process remains independent of the transmitted data.

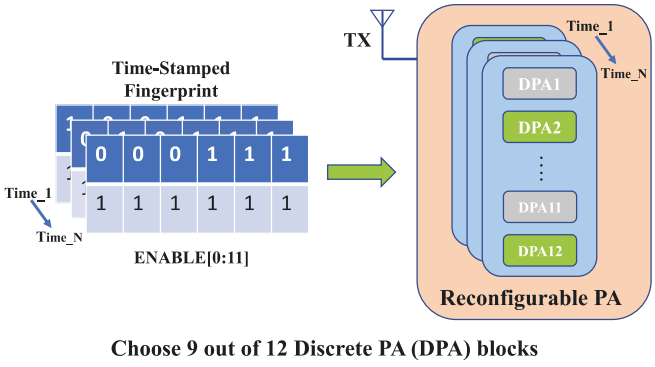

FIGURE 2. Advancement in the generation of RFFs using CR has increased the number of RFFs from 16 in previous designs to 220 in the current state of the art.

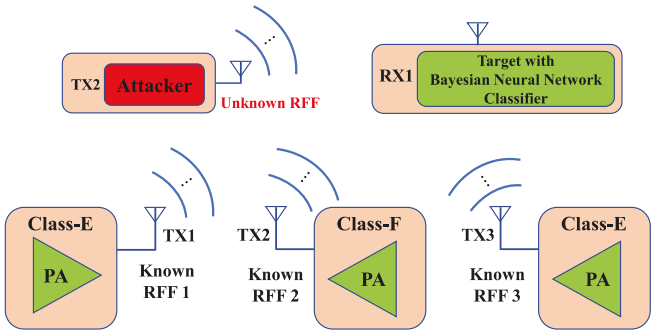

FIGURE 3. RFF generation utilizes process variations and device-specific harmonic components of different types of switching PAs.

B. RFF GENERATION UTILIZING SWITCHING POWER AMPLIFIERS AND INTRINSIC PROCESS VARIATIONS [22]

To meet the energy efficiency requirements of modern wireless devices, Class-E and Class-F PAs are commonly deployed. In [22], Class-E and Class-F amplifiers are utilized for investigating RFF feature generation. In [12] and [13], it is noted that the drain impedance of the conventional Class-F PA can be manipulated up to the third harmonic by employing output matching networks featuring harmonic resonators. This capability facilitates the tunability of the PA output waveform's behavior. Research in [22] demonstrates that the second harmonic resonator in the output matching networks significantly affects the behavior of the third harmonic. Consequently, variations in the parasitic components within the matching networks lead to the generation of device-specific RFF features. In contrast to Class-F PAs, the harmonic content at the drain side of Class-E PAs is primarily induced by the soft-switching effect [14]. The soft-switching effect at the drain of the Class-E PA, which is highly dependent on process variation, results in distinct RFF features. As illustrated in Fig. 3, each PA transmits device-specific RF features, which are then extracted by the target receiver using the fast Fourier transform (FFT). Subsequently, the extracted RF features are forwarded to a BNN for identification and classification. In [22], six transmitters are tested and successfully classified with an accuracy of 89.5%.

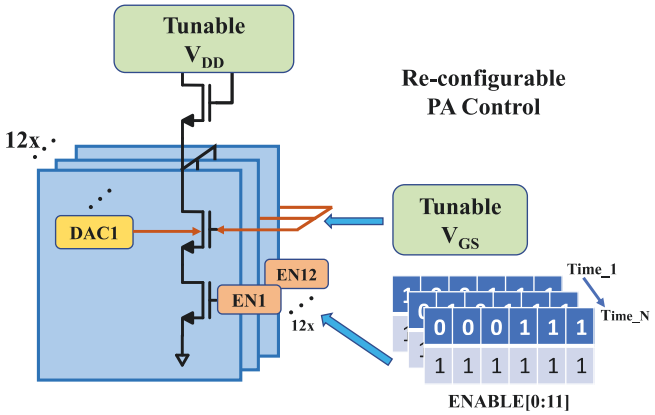

FIGURE 4. Highly tunable PA is employed to restore its performance in harsh conditions. Tunable power supply voltage and gate biasing voltage are utilized by integrating DACs to extend the programmability of the PA. A total of 2400 PA configurations are generated to cover operating temperature ranges from −197 °C to 80 °C.

C. DAC-CONTROLLED COMBINATORIAL PAS [16]

In [11] and [15], the integration of CR techniques with MOSFETs' intrinsic process variation enabled the creation of over 200 RFFs on a single chip. However, relying solely on process variation generated by the discrete PA structure and fabrication for enhancing the tunability of the PA is inherently limited. For instance, as the number of discrete PA blocks increases while maintaining a certain output power level, the size of each PA block decreases. Increasing device mismatch between each PA block can lead to overall PA performance degradation. Additionally, the discrete PA structure introduces extra parasitics due to complex and bulky interconnects required to support large currents, which can limit its application at higher frequency bands.

Hence, in [16], while preserving the same number of discrete PA blocks, the topology of each block is altered to integrate a significant number of DACs for controlling the body voltage of transistors. As depicted in Fig. 4, the body voltages of the input transistors are individually regulated by 3-bit DACs. Furthermore, by adjusting the supply voltage V_{dd} and the gate voltage of the input transistors, PA configurations that offer stable output power and high signal quality in specific operating environments are chosen. By employing the deep reinforcement learning techniques outlined in [16], nearly 2400 optimized PA configurations are selected to restore PA performance across operating temperatures ranging from −197 °C to 80 °C.

D. PUF-CONTROLLED PA SPECTRAL REGROWTH [10]

In [10], a differential PA structure consisting of Main and Auxiliary paths, as depicted in Fig. 5, is utilized. By adjusting the gate voltage of the Auxiliary PA, the out-of-band leakage power can be tuned to generate device-specific features. Since the Main path PA dominates the total power of the PA, tuning the Auxiliary path PA primarily influences the out-of-band leakage power. Consequently, the in-band power variation is effectively controlled and remains below

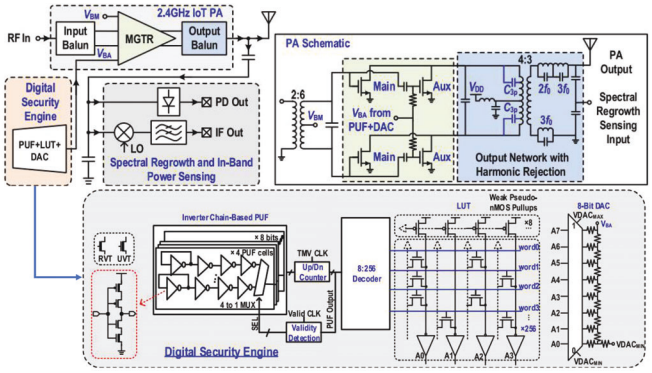

FIGURE 5. PUF-controlled PA is implemented to generate device-specific out-of-band leakage power features [10]. Sixteen distinctive PUF settings are applied to each PA chip, resulting in a total of 144 PUF settings measured from nine chips, achieving an identification accuracy of 95%.

1.5 dBm. A digital physical unclonable function (PUF) and an 8-bit DAC are employed to produce randomized Auxiliary PA biasing voltage combinations. However, the nonlinear relationship between out-of-band power leakage (OOBPL) and the Auxiliary path PA biasing voltage results in a nonuniform probability distribution of the OOBPL, which is undesirable for classification and security purposes. Hence, a hardware lookup table is implemented before the DAC to predistort the output of the PUF and restore the uniform probability distribution of the OOBPL. Each PA device is programmed with 16 PUF settings, enabling 16 types of PA configurations per chip. A total of nine PAs are utilized to create 144 PA configurations, resulting in an identification accuracy of 95% for these configurations.

III. DESIGN CONSIDERATIONS FOR MAXIMIZING THE NUMBER OF RFFS WHILE MAINTAINING SATISFACTORY PERFORMANCE

In the previous work [16], the tunability of the PA is significantly enhanced, and the PA performance is restored during extreme operating conditions. High PA tunability assists with RFF generation. However, this highly tunable PA structure depicted in Fig. 4 is challenging to sustain reliable PA performance and generate RFF features simultaneously. The DAC-controlled body voltage is directly connected to the body of the main power transistors. Varying the body voltage inevitably causes the PA's dc operating points to fluctuate during RFF reconfiguration. Additionally, the techniques used to boost tunability, such as modifying the power supply voltage and gate biasing voltage, also lead to variations in PA performance.

Although, as discussed in [11] and [15], the CR technique, without altering operation conditions, helps reduce the variation in PA output power, segmenting the PA into discrete blocks still introduces noticeable device mismatches. Based on measurement results in [11] during Bluetooth signal transmission, the average fluctuation in PA output power across RFF configurations is +/−2 dBm, which degrades the PA's performance.

The application of multigate transistors and PUF-controlled techniques in [10] effectively reduces the in-band power variation during PUF operation and generates identifiable out-of-band power features. However, an in-band power fluctuation of 1.2 dBm is still noticeable, and the number of identifiable PA patterns is limited by the PUF's tunability. Additionally, to extract the out-of-band features, high-sensitivity power detection and downconversion circuits are added to the receiver side, significantly increasing the hardware design complexity. The detection accuracy of the out-of-band power features can be significantly affected by the changing noise floors in the communication channel.

While paper [21] minimizes hardware resource consumption by employing only one lightweight machine-learning classifier, the presence of RF features at high-order harmonic frequencies unavoidably increases the complexity of receiver-side circuit design. For instance, the high-order harmonic components fall outside the desired frequency band. Additional downconversion circuits with a wideband analog-to-digital converter (ADC) are required to extract these high-frequency components.

Therefore, the following design objectives are established. First, to meet the requirements of complex modern communication standards, the entire PA system must generate prominent in-band RFFs without modifying standardized data packages. Second, minimizing the performance variation of the PA system caused by configuring RFFs is crucial, so generating RFFs by modifying the PA's operating points is not preferred. Lastly, to simplify data collection and machine-learning classifier training, the process of extracting features of RFFs must be streamlined, eliminating the need for collecting massive data to carry out high-dimensional RFF classification.

IV. PROPOSED 2-STAGE DOHERTY POWER AMPLIFIERS FOR RFF GENERATION

A 2-stage reconfigurable PA shown in Fig. 6(c) is proposed. The class A amplifiers at Auxiliary and Main paths compose the first stage, and the "Doherty amplifier" is applied in the second stage. Rather than directly generating RFF features by programming the body voltage, supply voltage, or gate biasing of a single-stage PA shown in Fig. 6(a) or a Doherty PA's main amplifier shown in Fig. 6(b), this design utilizes a reconfigurable amplifier added before main amplifier to synthesize a variety of RFF features. By stabilizing the dc operating point of the amplifier in the first stage and eliminating the need to skew the high-power Doherty amplifier for generating RFF features, the RFF configuration can be integrated without influencing the Doherty PA's steady-state operating point and output power level. Given the widespread application of Doherty topology in mobile devices for supporting high PAR communication standards, selecting the Doherty amplifier as the second stage illustrates the universality of the proposed circuit principles and signal processing techniques. However, it should be noted that

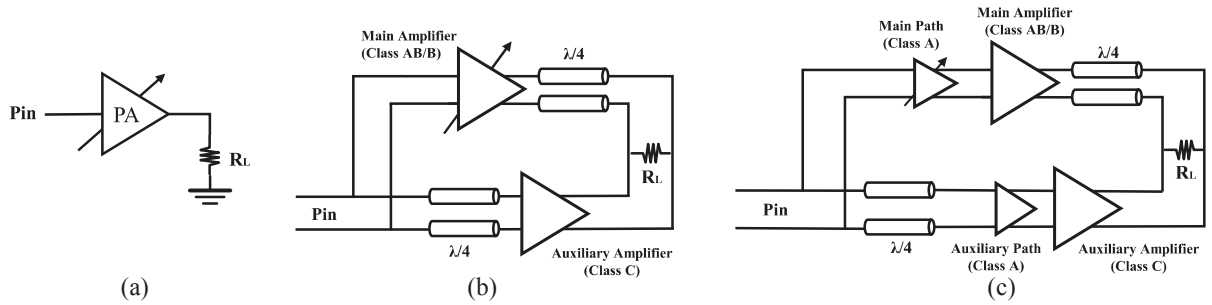

FIGURE 6. (a) Conventional single-stage PA with tunable RF features. (b) Integration of tunable RF features into a differential Doherty amplifier. (c) Proposed 2-stage differential Doherty amplifier integrating a tunable class A amplifier for RFF generation.

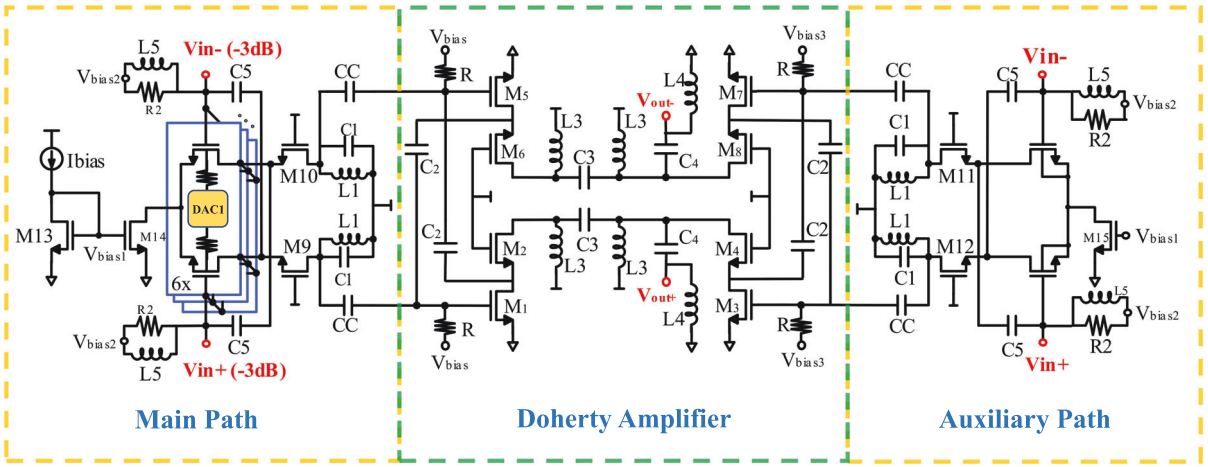

FIGURE 7. 2-stage PA is composed of the "auxiliary and main paths" as the first stage and the "Doherty amplifier" as the second stage. The amplifier at the main path is reconfigurable, and its body biasing is controlled by six units of 3-bit DACs.

the second stage could also employ other commonly used amplifier structures, such as PAs reviewed in Section II.

A. AUGMENTATION OF RFF FEATURES

Unique RFF features are typically generated from nonlinear behaviors of a reconfigurable PA. In previous works [11], [15], [16], complex deep reinforcement learning techniques and operating point adjustments are employed to generate and extract RFF features. The proposed 2-stage PA effectively reduces the difficulty of maintaining stable output power and signal quality while generating usable RFFs. As shown in Fig. 7, the proposed 2-stage amplifier's first stage is designed as an *LC*-based differential class A amplifier. The differential Class A amplifier's constant tail current source mitigates operating point variations from RFF selection, stabilizing the first stage's output and reducing its influence on the next stage. On the other hand, the choice of a Class A amplifier for the first stage results in reduced power of unique RF features compared to the Class B main amplifier. Therefore, augmentation of the RFF features is necessitated after passing through the second stage. In other words, the high power gain second stage amplifies both the conditioned input signal and the RFF features generated by the first stage of the PA. Therefore, the power of RFF features is boosted by the power gain of the second stage, which is around 17 dB in this design. The overall signal-to-"RFF feature power" ratio is not degraded but is further improved by the power gain of the first stage, which is 6 dB. Equations (1)–(3) show the mathematical explanation of the above theory

$$P\text{out}_{1\text{st}} = \text{Gain}_{1\text{st}} \cdot P\text{in} + P_{\text{RFF}} \tag{1}$$

$$P\text{out}_{2\text{nd}} = \text{Gain}_{1\text{st}} \cdot \text{Gain}_{2\text{nd}} \cdot P\text{in} + \text{Gain}_{2\text{nd}} \cdot P_{\text{RFF}} \tag{2}$$

$$\frac{P\text{out}_{SIG}}{P\text{out}_{\text{RFF}}} = \frac{\text{Gain}_{1\text{st}} \cdot \text{Gain}_{2\text{nd}} \cdot P\text{in}}{\text{Gain}_{2\text{nd}} \cdot P_{\text{RFF}}}. \tag{3}$$

B. CIRCUIT IMPLEMENTATION

The first stage utilizes an *LC*-based class A differential amplifier. Because the auxiliary amplifier in the second stage is class C biased, no RF features can be amplified when the input power level is low. Therefore, only the class A amplifier at the main path is designed to be highly tunable.

To provide a large number of RFFs, input transistors in the PA main path are divided into six groups during the layout design. Each group's body voltage is precisely programmed using six 3-bit RDACs, enabling the generation of eight programmable levels. However, to prevent the forward biasing from substrate to drain/source, four programmable levels covering a range from 0 to 0.45 V for each group are used. This configuration results in a total of 4096 selectable voltage combinations, providing fine control over the RFF generation process.

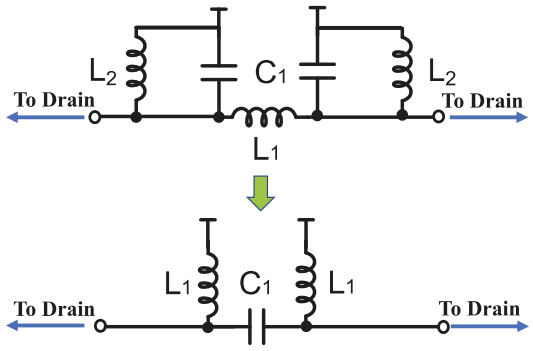

FIGURE 8. By modifying the series inductor type quarter wave phase shift network to series capacitor type network and combining the inductor load L2 at each side of the signal path with the quarter wave phase shift network, the number of used inductors is reduced from 3 to 2 and used capacitors is reduced from 2 to 1 for each side of the differential PA.

In contrast to relying on process variations enhanced by intentionally created mismatches for RFFs generation, as done in previous works [11], [15], [16], the separated transistor groups in this design are arranged as closely and symmetrically as possible. This arrangement aims to reduce transistor mismatch and maintain the overall consistency of the PA performance. Additionally, to further minimize the operating point variation during RFF selection, neither supply voltage nor gate biasing tuning techniques are employed. Instead, a tail current source M14 is added to provide a constant steady-state current. The M14 and M15 transistors are biased by the diode-connected device M13, which maintains a constant current value I_{bias}. Setting the I_{bias} value, transistor size ratio (w_{13}/w_{14}), and *L-C* load value determines the maximum input level for the second stage, which is equal to $I_{\text{bias}} \cdot [(w \cdot L1 \cdot Q1)//(gm \cdot \text{rds}^2)]$, where w is the operating frequency, $Q1$ is the quality factor of inductor $L1$, gm is the transconductance of input transistor, and rds is the drain–source on resistance. In this design, L1 is chosen to be 134 pH with a quality factor of 13.7, to resonate with the gate capacitance of the Doherty PA and form a purely resistive load at 5.4 GHz. To minimize the tail current source mismatch between main and auxiliary paths, transistors M13, M14, and M15 are interdigitatedly laid out. Consequently, the differential topology can effectively ensure the generation of extensive RFFs synthesized by the first stage while minimizing the risk of performance degradation on the transmitted signal.

The choice of the differential topology for the first stage requires the second stage also to be differential. Even though a differential design consumes more chip area, higher output power and advantageous 6-dB output back-off (OBO) point efficiency of the Doherty amplifier allow for the efficient augmentation of the 4096 RFFs generated at the first stage without compromising the overall performance of the PA system.

For conventional Doherty amplifiers shown in Fig. 6(b), a class AB/B main amplifier is applied with a class C amplifier. The class AB/B main amplifier is turned on before the class C auxiliary amplifier. As the input power

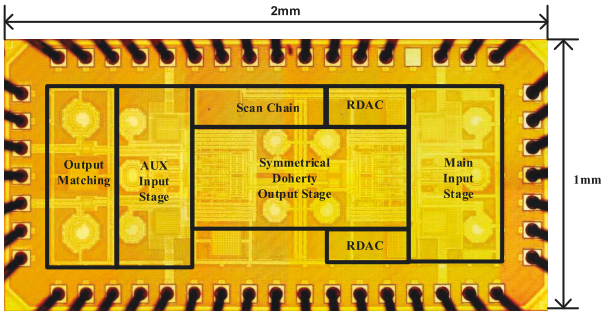

FIGURE 9. Chip micrograph of the proposed reconfigurable 2-stage PA.

increases, the main amplifier reaches its maximum output voltage and power efficiency. Then, the class C auxiliary amplifier is turned on to reduce the output impedance at the main amplifier. Even though the peak voltage at the main amplifier is constant as input power increases, the output power is linearly increased until the auxiliary amplifier is also saturated, which enhances the power efficiency at the back-off region. The load impedance at main and auxiliary PA can be derived using (4) and (5), where Z_0 is the characteristic impedance of the quarter-wave inverter, R_L is the load of Doherty amplifier, I_m and I_a are the current at the output of the amplifier

$$Z_{\text{main}} = Z_0 - \frac{Z_0}{R_L} \cdot \frac{I_m}{I_a} \tag{4}$$

$$Z_{\text{aux}} = \frac{I_m}{I_a} \cdot Z_0. \tag{5}$$

To achieve the desired efficiency at the 6-dB OBO point, the lower power gain of the Class C auxiliary PA compared to the main amplifier necessitates either a significantly larger auxiliary amplifier or higher input power to the auxiliary path. This adjustment reduces the load impedance of the main amplifier in the back-off region. In this design, the symmetrically operated Doherty amplifier is chosen to reduce circuit size, as both the main and auxiliary path transistors share the same dimensions. To recover the efficiency characteristics of the symmetrically operated Doherty PA at 6-dB OBO point, the unequal power division technique is implemented off-chip using a pair of microstrip Wilkinson power dividers to distribute power unevenly to the auxiliary and main amplifiers. In this configuration, the input power ratio of the main and auxiliary paths is set to 1:2 to relax the microstrip fabrication difficulty. The auxiliary path PA with higher input power mimics the ac behavior of a larger auxiliary path PA with lower input power. Therefore, the efficiency at the 6-dB back-off region resembles that of conventional Doherty PA designs. However, because the class A amplifier in the auxiliary path is driven by higher input power, it will saturate faster than the amplifier in the main path. As a result, this technique sacrifices the peak input level of the main amplifier, although it enhances the efficiency performance in the back-off region.

Additionally, to save space occupied by bulky lumped elements, the *C-L-C* quarter-wave phase shift network

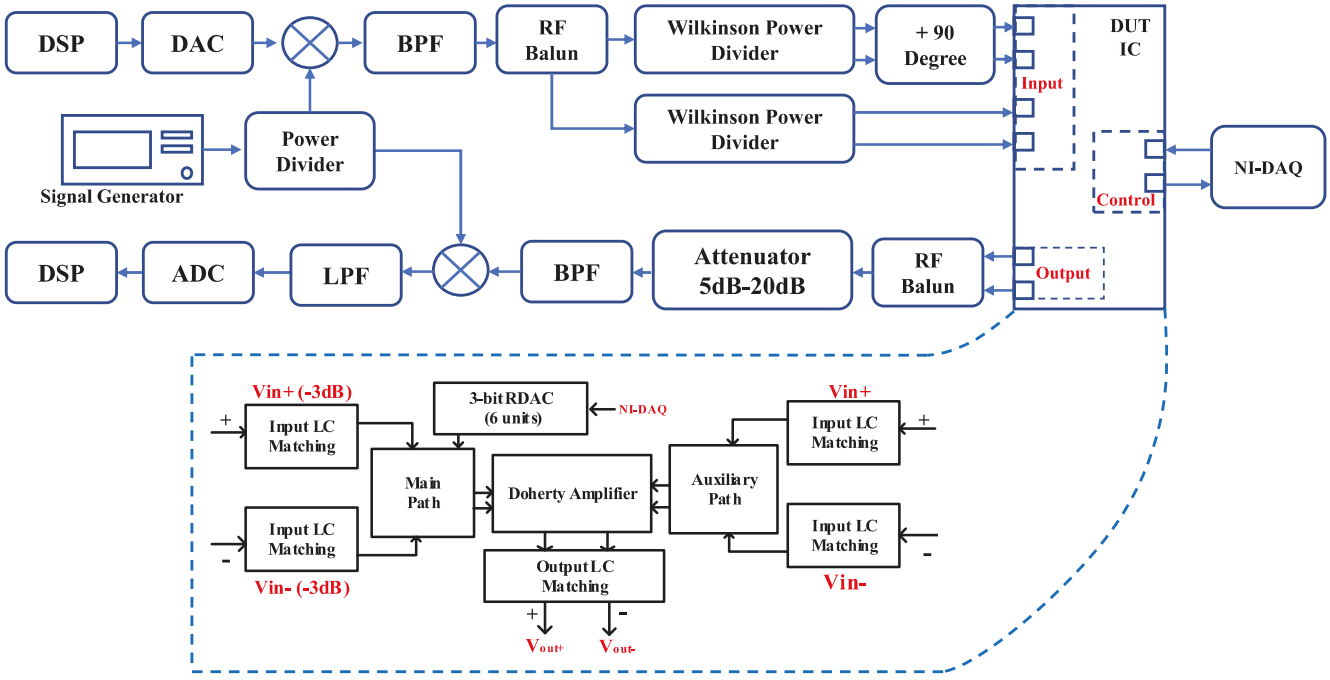

FIGURE 10. Test setup for both Wi-Fi 6E packet classification and USMT-assisted RFF classification is presented. A high-level block diagram describing the signal path and blocks' connection of the proposed 2-stage PA is added to assist understanding.

between the main and auxiliary amplifiers is integrated with the inductor–capacitor (*L-C*) loads at each drain side of the signal path. Conventionally, *L-C* loads are designed separately at the main and auxiliary amplifier and a series inductor-type quarter wave phase shift network is added to share the dc path between two amplifiers. As explained in Fig. 8, the quarter-wave phase shift network is designed to be a series capacitor type, and the inductor L2 in each of the *L-C* loads is combined with the inductor L1 in the quarter-wave phase shift network. The value of the capacitor C1 and inductor L1 in the equivalent *L-C-L* network can be derived using (6) and (7), where w is the operating frequency, Z_0 is the characteristic impedance of the quarter wave inverter, and C_{para} is the capacitance at each side of the Doherty amplifier

$$C1 = \frac{1}{w \cdot Z_0} \tag{6}$$

$$L1 = \frac{\frac{1}{w^2 \cdot C_{para}} \cdot \frac{Z_0}{w}}{\frac{1}{w^2 \cdot C_{para}} + \frac{Z_0}{w}} = \frac{Z_0}{w + Z_0 \cdot w^2 \cdot C_{para}}. \tag{7}$$

The equivalent *L-C-L* network reduces the number of consumed inductors from 3 to 2 and capacitors from 2 to 1 at each side of the differential Doherty amplifier. Additionally, using two inductors to provide dc power individually reduces drain-side power degeneration and ensures compliance with the electromigration rule. Harmonics at each side of the Doherty amplifier are reduced by the equivalent band pass *L-C-L* tank. The output *LC* matching networks and off-chip filters further reduce the harmonics leaked to the load.

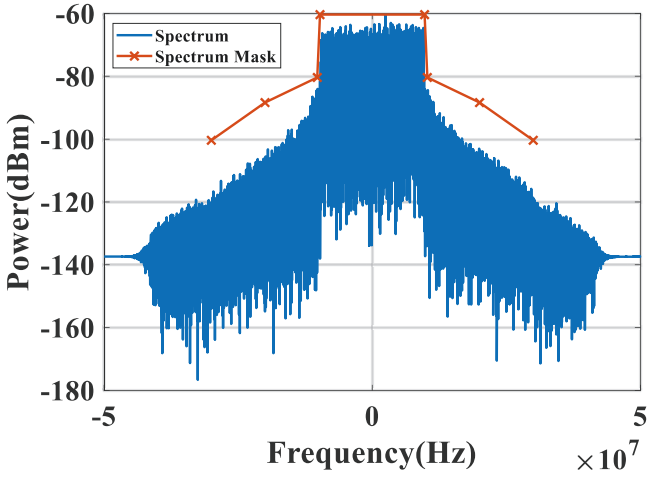

FIGURE 11. Measured power spectrum of the transmitted 64-QAM WiFi-6E signal and the corresponding spectrum mask. The peak transmitting power is 13 dBm.

C. TEST SETUP AND VERIFICATION OF THE GENERATION OF 4096 STABLE INHERENT RFF PATTERNS

Before applying the time-varying USMT techniques to boost the number of RFF patterns, verification of the inherently generated 4096 RFF patterns and their quality is conducted. To ensure comparability with previous research results in [11] and [15], the same RFF extraction method and CNN algorithm are used. The proposed 2-stage reconfigurable PA is fabricated using 65-nm CMOS technology and measured using the setup shown in Fig. 9 to verify the successful generation and extraction of 4096 inherent RFF patterns. Fig. 10 shows the micrograph of the fabricated PA chip.

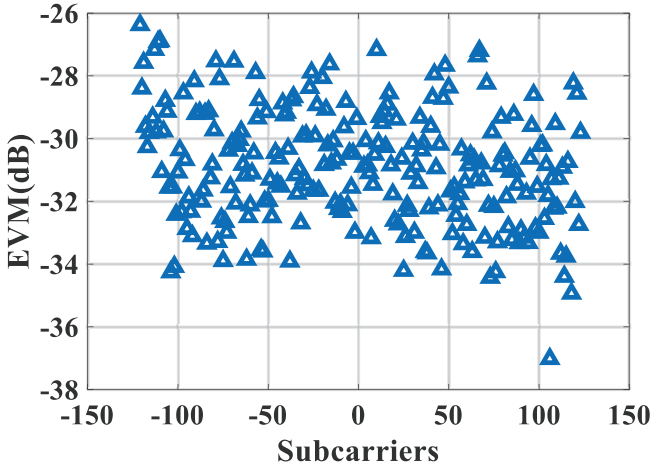

FIGURE 12. Measured EVM of a single data package and the EVM of all OFDM subcarriers.

The 20-MHz 5/6 coding 64-QAM Wi-Fi 6E packets are generated in the digital domain before being sent into an RF DAC. In this configuration, the IF frequency of the transmitted Wi-Fi packets is chosen to be 1 GHz. A signal generator with an equal power divider is added to provide the 4.4-GHz LO frequency for both up and down conversion. To reduce the intermodulation products leaked from mixers, a BPF is added to the transmitting path, while an LPF is added to the receiving path. An SMD RF balun is added to the transmitting path to accomplish single-ended to differential signal conversion. A pair of microstrip unequal Wilkinson power dividers are designed to achieve the 1:2 ratio power division for the main and auxiliary path PAs. A pair of quarter-wavelength microstrip lines is connected in series with the first stage for a 90° phase shift. On the receiving side, the differential signal is converted to a single-ended signal using the same SMD RF balun. A 5–20-dB power attenuator is applied to prevent voltage saturation of the RF ADC and to fully utilize the ADC's resolution. An additional BPF is added to reduce harmonics leaked from the PA.

Fig. 11 shows the measured spectrum of the amplified 64-QAM WiFi-6E signal alongside the corresponding spectrum mask. The amplified signal's spectrum fully complies with the required spectrum mask. In Fig. 12, the error vector magnitude (EVM) of the OFDM subcarriers in the amplified signal is depicted, with values concentrated between −28 and −34 dB. Fig. 13 displays the 5.4-GHz single-tone test results measured at the output node of the RF balun using the same setup. The Doherty PA demonstrates an output power of 9 dBm with a drain efficiency of 12% at the 6-dB OBO point and an output power of 15 dBm with a drain efficiency of 18% at the 1-dB compression point. The peak power is around 17 dBm with a peak efficiency of 20%. Both the EVM and spectrum figures indicate that the proposed PA exhibits high linearity at the target peak output power.

To extract the RFFs, all transmitted signals containing the corresponding 4096 types of RFF features are recorded. The

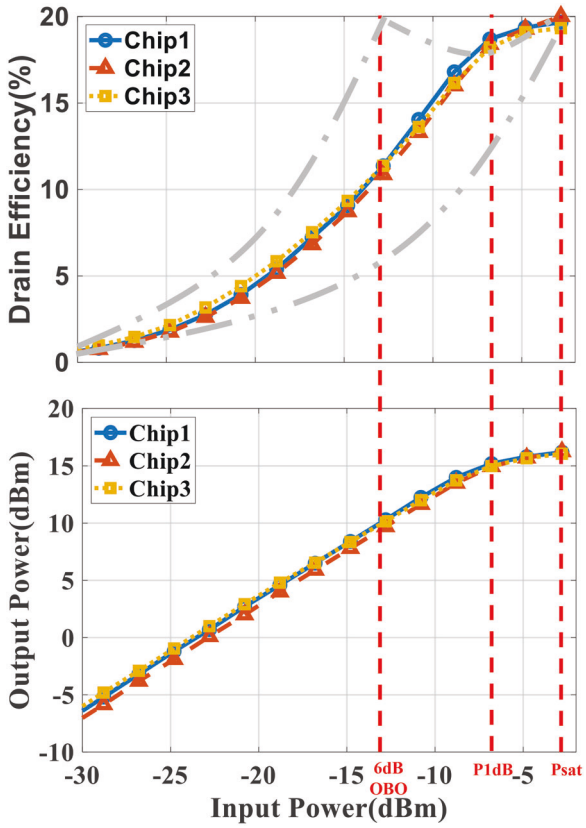

FIGURE 13. Measured PA output power versus input power at 5.4 GHz across three chips (bottom) shows less than 1-dBm power variation at the peak output power (OP1dB = 15 dBm). The top figure illustrates the measured drain efficiency of the Doherty amplifier at 5.4 GHz across the same three chips. For comparison, the efficiency behavior of the conventional Doherty amplifier is indicated by a single dot gray line, while the efficiency behavior of the conventional Class A/AB amplifier is marked by a double dot gray line.

standardized legacy preambles and high-efficiency long training field in each WiFi data package are then sent to a CNN with an input size of (256, 2, 480). This input size represents the batch size, 2-channel IQ signals, and N samples in the time domain, facilitating RFF extraction. Details regarding the application of CNN and data processing are elaborated in Section VI. As shown in Fig. 14, utilizing initially collected signals with a signal-to-noise ratio (SNR) of 42 dB, all 4096 patterns of RFFs are accurately identified with an accuracy exceeding 99%. To further challenge the classifier, additional random Gaussian noise is intentionally introduced at various noise power levels. Despite a degradation of SNR to 10 dB, a classification accuracy of over 90% is maintained for more than 4000 RFFs. With an increase in SNR degradation to 15 dB, over 2048 RFFs still exhibit a classification accuracy exceeding 90%. Hence, the RFF features produced by the main path of the 2-stage PA are readily distinguishable using a straightforward classifier. These high-quality RFF features demonstrate robustness and can withstand a broad spectrum of SNR degradation.

V. INCORPORATING TIME-VARYING UNEQUALLY SPACED MULTITONE TECHNIQUES FOR BOOSTING RFFS BY FOURFOLD

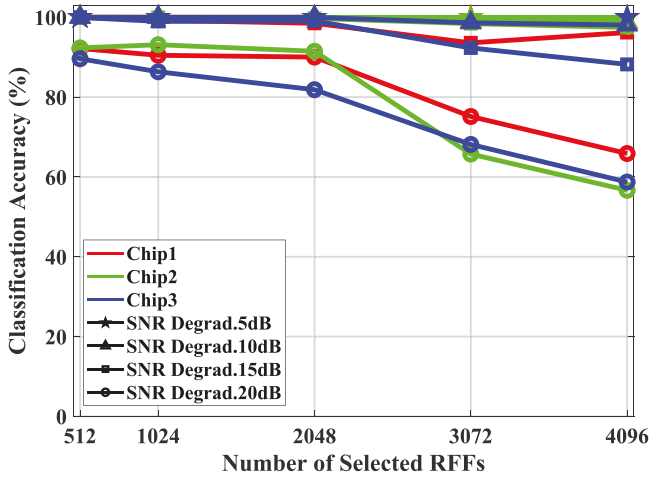

FIGURE 14. Classification accuracy versus the number of selected RFFs. The original signal SNR is 42 dB. SNR degradation levels from 5 to 20 dB are achieved by adding random Gaussian noise to the original signal to challenge the classifier. The classification of RFFs is tested across three different chips to verify the generality of the generated RFF features.

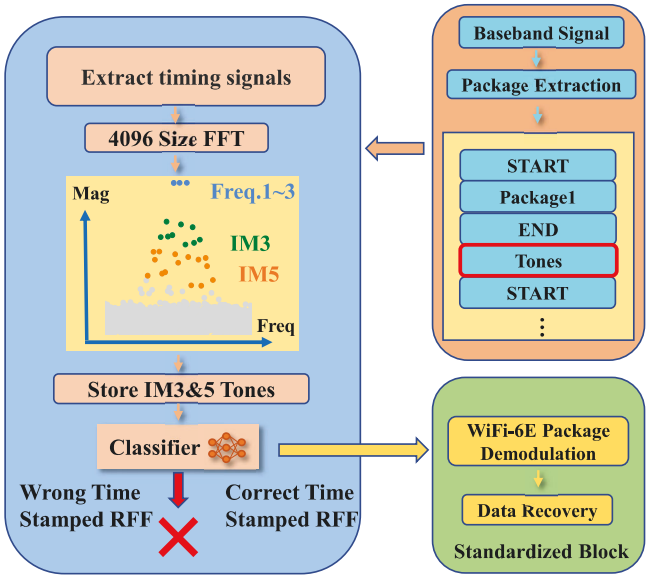

FIGURE 15. Proposed time-varying USMT technique. Tones are transmitted between each data package to prevent the addition of extra preamble blocks. A 4096-size FFT is used to extract all dominant intermodulation products. By eliminating the time-consuming partial demodulation of received data during RFF authentication, faster data processing time is achieved.

A. PROPOSED TIME-VARYING UNEQUALLY SPACED MULTITONE TECHNIQUE

In the previous RFF extraction process, each received WiFi-6E data package has to be partially demodulated before the legacy preambles and the high-efficiency long training field in the standard WiFi-6E package can be recovered for RFF authentication. However, the demodulation process can indeed introduce significant delays, particularly when handling numerous received packages. For example, if an attacker sends 100 packages, and each package necessitates partial demodulation for RFF feature verification, the overall process can be time consuming. Additionally, since no power supply or gate-biasing tuning is utilized, a noninvasive method is needed to effectively maximize the feature space of the reconfigurable PA for boosting RFFs.

To address the aforementioned challenges, the proposed solution utilizes the time-varying USMT method, as illustrated in Fig. 15. By employing three unequally spaced frequency tones generated within the targeted transmitting channel [29], a total of 27 nonoverlapping third-order (IM3) and fifth-order (IM5) intermodulation products can be generated at specific frequency locations. To prevent intermodulation products from overlapping the initial frequency tones, initial tones' generation must follow the rules defined by (8). In this equation, $f_1 = l \cdot f_\eta$ is the lowest frequency tone, f_η is the FFT frequency grid, k is the rank of the tone, and $p_k = [0, 1, 3, 9, \ldots, 3^{k-2}]$ is the unequally added frequency shift. The value of m controls the equally added frequency shift, while the value of l determines the initial tone's frequency. Both m and l must be properly chosen to ensure that all initial tones and IM3 and IM5 products are constrained within the same frequency band as the Wi-Fi signal for hardware reuse

$$f_k = f_1 + (k-1) \cdot m \cdot f_\eta + p_k \cdot f_\eta. \tag{8}$$

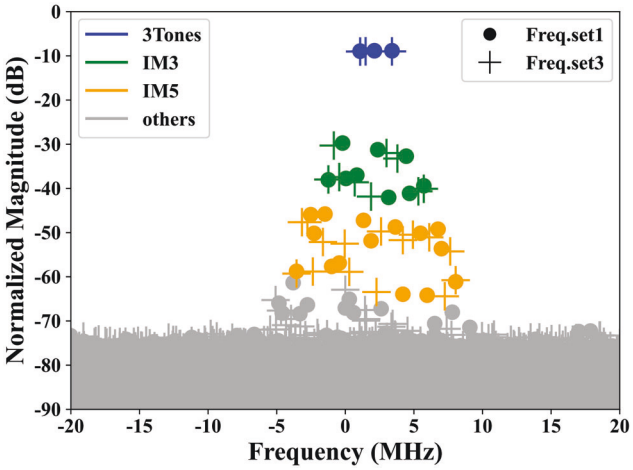

FIGURE 16. Frequency sets 1 and 3 of the USMT are generated in the right half band of the 20-MHz channel. By tuning the intermediate tone, distinctive intermodulation products are generated at different frequencies.

The nonlinear behavior of the designed PA can be characterized by the in-band carrier to third- and fifth-order intermodulation ratio (C/IM), as depicted in (9). Furthermore, the relationship between EVM and carrier to third and fifth intermodulation ratio is shown in (10) [30], [31]. Consequently, the dominant nonlinear PA behaviors are captured by extracting the power of intermodulation products using the FFT

$$C/IM(dB) = 10\log\left(\frac{P_{\text{Tones}}}{P_{IMs}}\right) \tag{9}$$

$$20\log(EVM_{\%}) = 40 - C/IM(dB). \tag{10}$$

The unequally spaced tones are transmitted between each WiFi-6E package. As illustrated in Fig. 15, without the

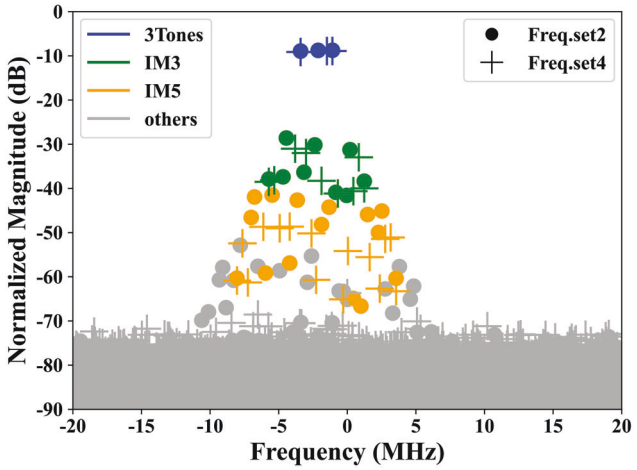

FIGURE 17. USMT frequency set2 and set4 are generated in the left half band of the 20-MHz channel. By mirroring USMT sets 1 and 3 according to the center frequency, distinctive intermodulation products are generated at different frequencies.

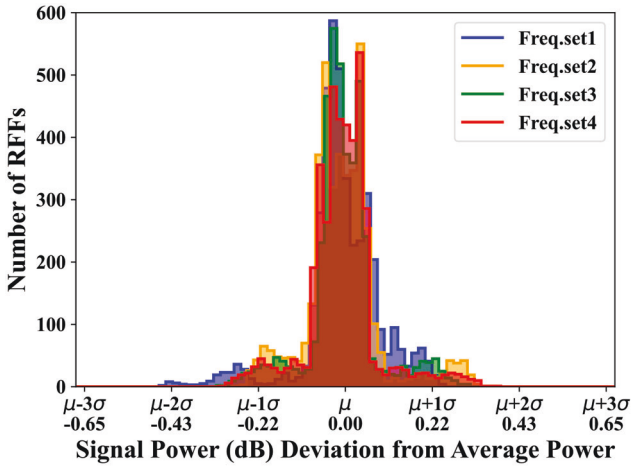

FIGURE 18. Measured power deviation of the transmitted signal from the average signal power is evaluated across all inherent 4096 RFF patterns and all 4 USMT sets. The PA's power performance remains consistent across all 16 384 RFF patterns, with a standard deviation of 0.22 dB.

need for additional correlation blocks, the existing preamble blocks at the beginning and end of each WiFi-6E package can be utilized to locate the recorded unequally spaced tones. Subsequently, a single 4096-size FFT is employed to extract the IM tones at known frequency locations. Since no legacy preambles or high-efficiency data preambles within the WiFi-6E package are utilized, only a correlation process and a 4096-size FFT are necessary to extract all the information for RFF classification. Consequently, the time delay for RFF data recovery is significantly reduced. Without the need for data demodulation prior to RFF classification, the time required to complete one RFF information extraction is reduced from 0.260856 to 0.090922 s, which is over three times faster.

Since the USMT method does not alter the operational conditions of the PA, the PA performance remains uncompromised, while the number of RFFs is significantly increased. As depicted in Figs. 16 and 17, four sets of unequally spaced tones are designed within the 20-MHz channel. By adjusting the lowest and highest frequency tones, the upper and lower frequencies of the intermodulation products are determined. Furthermore, by tuning the intermediate tone while maintaining the peak power of the in-band signal during RFF selection, noticeable differences are observed between the intermodulation products. This indicates that the nonlinear behavior of the PA is frequency-dependent, particularly for higher-order intermodulation tones. Leveraging this discovery, 4 sets of unequally spaced tones are designed to increase the number of RFF patterns from 4096 to 16 384.

B. MINIMIZING PA PERFORMANCE VARIATION

One of the most important objectives of this proposed 2-stage amplifier is to reduce the PA power variation while generating a large variety of RFFs. Since no dc operating point variation is directly introduced for RFF feature generation, the proposed PA's power performance is not affected. As shown in Fig. 18, the measured PA in-band power variation versus RFF configurations is depicted, with the transmitted signal's power at each of the 4096 RFF patterns plotted. Additionally, to account for the power variation caused by Time-varying USMT frequency, all four frequency sets of USMT are tested with each of the 4096 inherent RFFs. The measured standard deviation of 0.22-dB in-band power variation indicates that the power variation is well controlled for all 16 384 RFF patterns.

Multiple chips' performance data are also measured and compared using the same testing setup to ensure the die-to-die performance variation is well controlled. This allows for the categorization of RFFs from multiple chips using a single machine-learning classifier. Fig. 13 shows the output power versus the input power measured at 5.4 GHz across three chips, revealing less than 1-dBm power variation at the peak output power (OP1dB = 15 dBm).

VI. RF FINGERPRINT CLASSIFICATION

A. CONVOLUTIONAL NEURAL NETWORK

At the receiver equipped with an intelligent classifier, the arrived USMT signal is converted to the frequency domain by FFT before classification. This approach is preferred over directly executing classification on the time-domain signal to achieve good accuracy while keeping the low complexity of the classifier. In the frequency domain, the 27 intermodulation products construct the identifiable RFFs, as depicted in Fig. 19 and represented by $\tilde{\mathbf{m}}27 = m0 \sim 2, m_{3\sim11}, m_{12\sim26}$, where $m_{0\sim2}$ denotes the three predetermined tones, $m_{3\sim11}$ and $m_{12\sim26}$ refer to the magnitudes of 9 IM3 and 15 IM5 components, respectively. The classifier analyzes the 27 intermodulation products to identify the RFF, resulting in a predicted RFF index. Trusted communication is established only when the predicted RFF matches the target timestamped RFF.

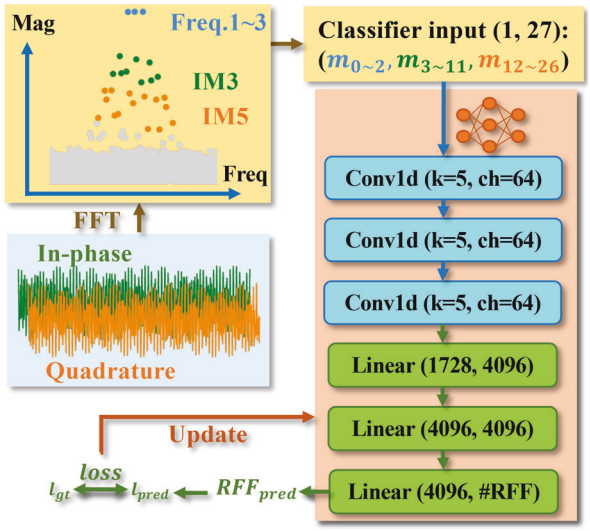

FIGURE 19. Classification diagram with the processing of FFT. A CNN is utilized as the classifier to identify RFFs from the frequency-domain features.

A classifier is a function that projects the input features to the labels represented by one-hot coding [25]. A CNN with a softmax output layer (11) is implemented as the N_{RFF}-class classifier, depicted in Fig. 19. The input features with a dimension of $(1, 27)$ are mapped to predicted labels l_{pred}.

Each of the elements in l_{pred} holds a likelihood value $q_c \in [0, 1], c = \{0, 1, \ldots, N_{\text{RFF}} - 1\}$ after softmax processing, representing the likelihood of the received signal being identified as each of the potential RFFs. The received signal is specifically identified as carrying the RF fingerprint RFF_{pred} that holds the maximum likelihood in l_{pred}. One-hot encoded l_{gt} denotes the ground-truth label of class c_{gt} with only element $l_{gt}^{c_{gt}} = p_{c_{gt}} = 1, c_{gt} = \{0, 1, \ldots, N_{\text{RFF}} - 1\}$ and others remaining zeros

$$q_c = \frac{e^{l_{\text{pred}}^c}}{\sum_{c=0}^{N_{\text{RFF}}-1} e^{l_{\text{pred}}^c}}, c = \{0, 1, \ldots, N_{\text{RFF}} - 1\}. \quad (11)$$

The cross-entropy loss function is the most commonly used criterion for classification tasks [25]. Cross-entropy measures the difference between the predicted distribution from the ground-truth label l_{gt} after each inference. Due to the one-hot encoded labels, the loss value is determined by the predicted likelihood of the ground-truth class c_{gt} as (12) shows. The loss value decreases as the predicted $q_{c_{gt}}$ increases. Therefore, by minimizing the loss function with the gradient-descent-based optimizer, the expected outputs fit the true results gradually until the average classification accuracy over all available RFFs ceases to improve for a certain number of iterations, which is set to 100 in this work

$$\begin{aligned} \text{loss} &= -\frac{1}{N_{\text{RFF}}} \sum_{c=0}^{N_{\text{RFF}}-1} p_c \log(q_c) \\ &= -\log(q_c), c = c_{gt}. \end{aligned} \quad (12)$$

The CNN-based classifier comprises three convolutional layers and three fully connected layers, as illustrated in Fig. 19. Each of these layers is followed by a batch normalization (BN) layer and a ReLU activation layer, except for the last fully connected layer that outputs the predicted RFF index. Specifically, the first three convolutional layers share the same kernel size $k = 5$ and number of output channels $ch = 64$. The output of all convolutional layers maintains the feature dimension of 27 so that the size of the output at the last convolutional layer is $(64, 27)$. The fully connected layers share the same output dimension of 4096 while the last layer generates a vector l_{pred} of N_{RFF} dimensions, where N_{RFF} refers to the number of all potential RFFs. In this work, $N_{\text{RFF}} = 4096$. Therefore, the classification model contains 40 673 600 32-bit parameters in convolutional and fully connected layers in total, which requires 163 MB of memory to store. Due to the model size and the large amount of RF signals, a powerful product, NVIDIA H100-80 GB with multi-instance GPU (MIG), is utilized in this work. MIG can partition H100 into seven instances. All classification experiments are executed on an MIG device consisting of 4/7 GPU instances and 40/80 GB memory.

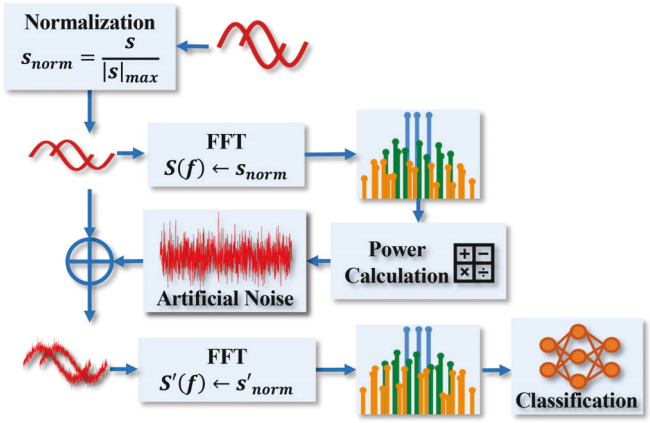

FIGURE 20. Enhance RFF distinctiveness evaluation by introducing additional Gaussian white noise to normalized signals before classification, enabling assessment under varied SNR conditions.

B. DATASET OVERVIEW

As described in the proposed time-varying USMT strategy, four sets of USMT frequency are employed in this work to increase the number of classifiable RFFs by fourfold, from 4096 to 16 384. The USMT signal is sampled at 62.5 MHz with 40 960 samples for each of the RFFs to construct the measurement dataset. Consequently, the measurement dataset has a shape of $(N_{Freq.set} \times N_{1set}, T_{sample}) = (16384, 40960)$, where $N_{Freq.set}$ refers to the four frequency sets, N_{1set} denotes 4096 available RFFs given one frequency set, and T_{sample} represents 40 960 samples in the time-series signal.

To validate the generalizability of the proposed RFFs' distinctiveness across different manufacturing implementations, two datasets are collected from two independent chips with corresponding RFF indexes across two datasets, each having a size of 16 384 as described above. Therefore, the total measurement dataset consists of 16 384 × 2 = 32 768

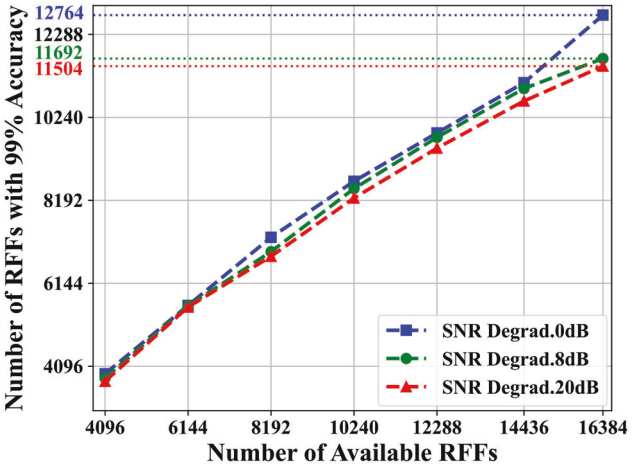

FIGURE 21. Plot illustrating the relationship between the number of available RFFs and the number of RFFs classified with 99% accuracy. The original raw signal has an SNR of 41 dB, with additional 8 and 20 dB random Gaussian noises added to challenge the classifier and evaluate the quality of RFF features.

time-domain signals with the RFF index $\text{RFF}_{idx} \in \{0, 1, \ldots, 16383\}$.

Starting from arbitrary positions within the 40 960 samples, an artificial dataloader generates the intermodulation products $\tilde{\mathbf{m}}_{27}$ with 4096 continuous samples. By performing the 4096-FFT every $\Delta sample$, there are $\lfloor (40\ 960 - 4096)/\Delta sample \rfloor = \lfloor 36\ 864/\Delta sample \rfloor$ possible $\tilde{\mathbf{m}}_{27}$ contributing to the input dataset for the classifier. Here, $\Delta sample$ is empirically set to 100 and 300 in this work. A split ratio of (0.6, 0.4) is used to divide the training and validation data, with the first 60% of the 40 960 samples allocated for training and the remaining 40% for validation.

The original measurement dataset unavoidably contains noise due to environmental factors, resulting in a peak SNR of 41 dB. To assess the distinctiveness of the proposed RFFs under different SNR levels, additional Gaussian white noise is artificially introduced to the measurement dataset before applying the classification FFT, as depicted in Fig. 20. This added noise induces an SNR degradation of P_d dB, where s and s_{norm} represent the time-series signals and their corresponding normalized signals, respectively. $S(f)$ denotes the spectrum in the frequency domain after applying the FFT on s_{norm}. The SNR degradation is computed by comparing the overall original noise power to the artificial noise power. The artificial noise causing a certain level of degradation is added to the normalized signal s_{norm}, resulting in a new spectrum $S'(f)$. Following data preprocessing, the intermodulation products $\tilde{\mathbf{m}}_{27}$ generated from $S'(f)$ are utilized as inputs for the classifier.

C. RESULTS OF SNR DEGRADATION

The proposed time-varying USMT technique boosts the number of potential RFF patterns from 4096 to 16 384, as shown in Fig. 21. However, because the increase in RFFs is attributed to the frequency dependency of intermodulation products, a small portion of the potential RFFs might not be usable. To evaluate the quality of RFFs generated using the time-varying USMT technique, only RFFs with 99% classification accuracy are chosen for deployment in this work. The original SNR of the raw data is 41 dB, and two additional levels of SNR are manually created by adding random Gaussian noise to the original raw data in Fig. 20, where the SNR degradation is $P_d = \{8, 20\}$ dB. Compared to previous works outlined in the comparison table (Table 1), the two-stage PA incorporating the time-varying USMT technique successfully generates more than 50 times the number of high-quality RFFs, totaling 11 504 RFFs.

D. FACTORS INFLUENCING RFF CLASSIFICATION

The C/IM ratio, SNR, and received signal power are three dominant factors influencing the accuracy of RFF classification. The C/IM ratio is highly correlated with the EVM value. According to the IEEE 802.11ax standard, the maximum EVM value for reliable 64-QAM Wi-Fi 6E package demodulation is −25 dB. A lower C/IM ratio means that the power of received intermodulation products is close to the power of the generated tones. For a certain level of transmitted power, a lower C/IM ratio will give us higher classification accuracy but increase the EVM, potentially causing the data package demodulation to fail. Compared to USMT-assisted RFF classification, the RFF classification using Wi-Fi data packets is more sensitive to C/IM ratio variation because the packets have to be demodulated before RFFs can be extracted.

Higher SNR increases the detectability of IM products. As shown in Fig. 14, we tested the classification accuracy from 42 to 22 dB, covering the typical working environment for 5/6 coding 64-QAM Wi-Fi 6E communication. For classification using Wi-Fi data packets, the accuracy of classification decreases as SNR decreases. To recover classification accuracy, the number of usable RFFs has to be decreased. For USMT-assisted classification, the classification accuracy is less susceptible to decreasing SNR but still shows a proportional relationship to SNR variations.

Received signal power is correlated with SNR. Higher received signal power increases SNR and classification accuracy. Because both excitation tones and intermodulation tones are used for classification, IM tones are also considered as signals for USMT RFF classification. Therefore, all the above three factors are intercorrelated, and the accuracy of RFF classification results from balancing these factors.

VII. CONCLUSION

In this article, various design strategies based on prior research on RF fingerprint generation and reconfigurable PA control are discussed. A novel RF fingerprinting system featuring a highly tunable 2-stage PA designed in 65-nm CMOS is presented. This system is capable of generating 4096 built-in high-quality RFFs while ensuring consistent PA performance across all RFF configurations. Additionally, the time-varying USMT technique is introduced to further extend the number of high-quality RFFs (classified with 99% accuracy) from 4096 to at least 11 504. All RFFs are

TABLE 1. Performance summary and comparison.

	OJCAS'23 [16]	TCAS-1'22 [11]	ISSCC'21 [10]	This Work
Structure	Combinatorial PA	Class E PA	PA	2-Stage PA
Technology	65nm	65nm	45nm SOI	65nm
Supply Voltage (V)	1.2	1.2	0.5/1	1.2
Number of RFFs per PA	2400 (PA configurations)	220	16	11,504
Classification Accuracy	–	95%	95%	99%
In Band Power Variation	– (Minimized by RL)	>1dB	1.5dB	STD = 0.22dB (11,504 RFFs)
RFF Source	PA +DAC	PA	PA & PUF+DAC	PA Input Stage+ DAC & Frequency Hopping
Frequency	10.5GHz	2.4GHz	2.5GHz	5.4GHz
Peak Power(dBm)	2.1	18.4*	4.7	16
Peak Efficiency(%)	11.2	–	32-36	20/28*
EVM (%)	–	–	2.8-4.7	2-3.98 (Including Off-chip Buffer)
Modulation	Single Tone	Bluetooth_LE	Bluetooth EDR	64QAM WIFI-6E

* Post Layout Simulated

successfully detected and classified, even with a wide range of additional SNR degradation. The in-band output power variation is well controlled, with a standard deviation of only 0.22 dB. The measurement results demonstrate that the proposed architecture, which separates the RFF generator from conventional PA while utilizing USMT-based feature augmentation methods, enables a large variety of RFFs while achieving minimized in-band power variations and high-speed RFF classification.

ACKNOWLEDGMENT

The authors extend their gratitude to Yuyi Shen and John Kan of Carnegie Mellon University for their valuable discussions.

REFERENCES

[1] W. Iqbal, H. Abbas, M. Daneshmand, B. Rauf, and Y. A. Bangash, "An in-depth analysis of IoT security requirements, challenges, and their countermeasures via software-defined security," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 10250–10276, Oct. 2020, doi: 10.1109/JIOT.2020.2997651.

[2] A. S. K. Pathan, H.-W. Lee, and C. S. Hong, "Security in wireless sensor networks: Issues and challenges," in *Proc. 8th Int. Conf. Adv. Commun. Technol.*, 2006, pp. 1043–1048, doi: 10.1109/ICACT.2006.206151.

[3] N. Neshenko, E. Bou-Harb, J. Crichigno, G. Kaddoum, and N. Ghani, "Demystifying IoT security: An exhaustive survey on IoT vulnerabilities and a first empirical look on Internet-scale IoT exploitations," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2702–2733, 3rd Quart., 2019, doi: 10.1109/COMST.2019.2910750.

[4] X. Lu, L. Hong, and K. Sengupta, "CMOS optical PUFs using noise-immune process-sensitive photonic crystals incorporating passive variations for robustness," *IEEE J. Solid-State Circuits*, vol. 53, no. 9, pp. 2709–2721, Sep. 2018, doi: 10.1109/JSSC.2018.2850941.

[5] X. Lu, S. Venkatesh, B. Tang, and K. Sengupta, "4.6 Space-Time Modulated 71-to-76GHz mm-Wave transmitter array for physically secure directional wireless links," in *Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC)*, San Francisco, CA, USA, 2020, pp. 86–88, doi: 10.1109/ISSCC19947.2020.9062929.

[6] K. Sa, D. Lang, C. Wang, and Y. Bai, "Specific emitter identification techniques for the Internet of Things," *IEEE Access*, vol. 8, pp. 1644–1652, 2020, doi: 10.1109/ACCESS.2019.2962626.

[7] M. Köse, S. Taşcioğlu, and Z. Telatar, "RF fingerprinting of IoT devices based on transient energy spectrum," *IEEE Access*, vol. 7, pp. 18715–18726, 2019, doi: 10.1109/ACCESS.2019.2896696.

[8] S. S. Hanna and D. Cabric, "Deep learning based transmitter identification using power amplifier nonlinearity," in *Proc. Int. Conf. Comput., Netw. Commun. (ICNC)*, Honolulu, HI, USA, 2019, pp. 674–680, doi: 10.1109/ICCNC.2019.8685569.

[9] S. Rajendran, Z. Sun, F. Lin, and K. Ren, "Injecting reliable radio frequency fingerprints using metasurface for the Internet of Things," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 1896–1911, 2021, doi: 10.1109/TIFS.2020.3045318.

[10] Q. Zhou, Y. He, K. Yang, and T. Chi, "12.3 exploring PUF-controlled PA spectral regrowth for physical-layer identification of IoT nodes," in *Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC)*, San Francisco, CA, USA, 2021, pp. 204–206, doi: 10.1109/ISSCC42613.2021.9365941.

[11] Y. Shen, J. Xu, J. Yi, E. Chen, and V. Chen, "Class-E power amplifiers incorporating fingerprint augmentation with combinatorial security primitives for machine-learning-based authentication in 65 nm CMOS," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 69, no. 5, pp. 1896–1909, May 2022, doi: 10.1109/TCSI.2022.3141336.

[12] F. H. Raab, "Class-F power amplifiers with maximally flat waveforms," *IEEE Trans. Microw. Theory Techn.*, vol. 45, no. 11, pp. 2007–2012, Nov. 1997, doi: 10.1109/22.644215.

[13] R. A. Beltran, "Class-F and inverse class-F power amplifier loading networks design based upon transmission zeros," in *Proc. IEEE MTT-S Int. Microw. Symp. (IMS2014)*, Tampa, FL, USA, 2014, pp. 1–4, doi: 10.1109/MWSYM.2014.6848269.

[14] K. C. Tsai, "CMOS power amplifiers for wireless communications," Dept. Elect. Eng. Comput. Sci., Univ. California, Berkeley, CA, USA, Rep. UCB/EECS-2007-161, 2007.

[15] V. Chen, J. Xu, Y. Shen, and E. Chen, "RF fingerprint classification with combinatorial-randomness-based power amplifiers and convolutional neural networks: Secure analog/RF electronics and electromagnetics," *IEEE Solid-State Circuits Mag.*, vol. 14, no. 4, pp. 28–36, Nov. 2022, doi: 10.1109/MSSC.2022.3200302.

[16] J. Xu, Y. Shen, J. Yi, E. Chen, and V. Chen, "Deep reinforcement learning on FPGA for self-healing cryogenic power amplifier control," *IEEE Open J. Circuits Syst.*, vol. 4, pp. 176–187, Jun. 2023, doi: 10.1109/OJCAS.2023.3282929.

[17] G. Shen, J. Zhang, A. Marshall, L. Peng, and X. Wang, "Radio frequency fingerprint identification for LoRa using deep learning," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 8, pp. 2604–2616, Aug. 2021, doi: 10.1109/JSAC.2021.3087250.

[18] B. Chatterjee, D. Das, S. Maity, and S. Sen, "RF-PUF: Enhancing IoT security through authentication of wireless nodes using in-situ machine learning," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 388–398, Feb. 2019, doi: 10.1109/JIOT.2018.2849324.

[19] K. Sankhe et al., "No radio left behind: Radio fingerprinting through deep learning of physical-layer hardware impairments," *IEEE Trans. Cogn. Commun. Netw.*, vol. 6, no. 1, pp. 165–178, Mar. 2020, doi: 10.1109/TCCN.2019.2949308.
[20] D. Roy, T. Mukherjee, M. Chatterjee, E. Blasch, and E. Pasiliao, "RFAL: Adversarial learning for RF transmitter identification and classification," *IEEE Trans. Cogn. Commun. Netw.*, vol. 6, no. 2, pp. 783–801, Jun. 2020, doi: 10.1109/TCCN.2019.2948919.
[21] E. Chen, J. Xu, J.-G. Zhu, and V. Chen, "Wireless Bayesian neural networks with self-assembly DNA memory and spin-torque oscillators," in *Proc. IEEE 63rd Int. Midwest Symp. Circuits Syst. (MWSCAS)*, Springfield, MA, USA, 2020, pp. 269–272, doi: 10.1109/MWSCAS48704.2020.9184674.
[22] J. Xu, Y. Shen, E. Chen, and V. Chen, "Bayesian neural networks for identification and classification of radio frequency transmitters using power amplifiers' nonlinearity signatures," *IEEE Open J. Circuits Syst.*, vol. 2, pp. 457–471, Jul. 2021, doi: 10.1109/OJCAS.2021.3089499.
[23] L. Van der Maaten and G. Hinton, "Visualizing high-dimensional data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, pp. 2579–2605, Nov. 2008.
[24] A. C. Belkina, C. O. Ciccolella, R. Anno, R. Halpert, J. Spidlen, and J. E. Snyder-Cappione, "Automated optimized parameters for T-distributed stochastic neighbor embedding improve visualization and analysis of large datasets," *Nat. Commun.*, vol. 10, p. 5415, Nov. 2019. [Online]. Available: https://doi.org/10.1038/s41467-019-13055-y
[25] Z. Zhang and M. Sabuncu, "Generalized cross entropy loss for training deep neural networks with noisy labels," in *Proc. 32nd Conf. Neural Inf. Process. Syst.*, 2018, p. 31.
[26] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
[27] M. J. M. Pelgrom, A. C. J. Duinmaijer, and A. P. G. Welbers, "Matching properties of MOS transistors," *IEEE J. Solid-State Circuits*, vol. 24, no. 5, pp. 1433–1439, Oct. 1989, doi: 10.1109/JSSC.1989.572629.
[28] S. C. Cripps, *Advanced Techniques in RF Power Amplifier Design*. Norwood, MA, USA: Artech House, 2002.
[29] S. Laurent, J. P. Teyssier, R. Quere, J. Sombrin, and M. Prigent, "Linearity characterization of RF circuits through an unequally spaced multi-tone signal," in *Proc. 88th ARFTG Microw. Meas. Conf. (ARFTG)*, Austin, TX, USA, 2016, pp. 1–4, doi: 10.1109/ARFTG.2016.7839729.
[30] J. P. Teyssier, J. Sombrin, R. Quéré, S. Laurent, and F. Gizard, "A test set-up for the analysis of multi-tone intermodulation in microwave devices," in *Proc. 84th ARFTG Microw. Meas. Conf.*, Boulder, CO, USA, 2014, pp. 1–3, doi: 10.1109/ARFTG.2014.7013408.
[31] J. Sombrin, "On the formal identity of EVM and NPR measurement methods: Conditions for identity of error vector magnitude and noise power ratio," in *Proc. 41st Eur. Microw. Conf. (EuMC)*, 2011, pp. 337–340.
[32] E. Kaymaksut and P. Reynaert, "Transformer-based uneven Doherty power amplifier in 90 nm CMOS for WLAN applications," *IEEE J. Solid-State Circuits*, vol. 47, no. 7, pp. 1659–1671, Jul. 2012, doi: 10.1109/JSSC.2012.2191334.
[33] N. Ryu, S. Jang, K. C. Lee, and Y. Jeong, "CMOS Doherty amplifier with variable balun transformer and adaptive bias control for wireless LAN application," *IEEE J. Solid-State Circuits*, vol. 49, no. 6, pp. 1356–1365, Jun. 2014, doi: 10.1109/JSSC.2014.2313561.
[34] X. Li et al., "A 110-to-130 GHz SiGe BiCMOS Doherty power amplifier with a slotline-based power combiner," *IEEE J. Solid-State Circuits*, vol. 57, no. 12, pp. 3567–3581, Dec. 2022, doi: 10.1109/JSSC.2022.3211828.

CHENGYU FAN (Graduate Student Member, IEEE) received the B.S. degree in electrical and computer engineering from the University of Illinois at Urbana–Champaign, Champaign, IL, USA, in 2021, and the M.S. degree in electrical engineering from Carnegie Mellon University, Pittsburgh, PA, USA, in 2022, where he is currently pursuing the Ph.D. degree with the EECS Lab.

He is interested in RFIC design with a focus on the application of high-frequency communication.

Mr. Fan is a recipient of the ISSCC Analog Devices Outstanding Student Designer Award in 2024.

JUNTING DENG received the B.S. degree in communication engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2020, and the M.S. degree in interdisciplinary information studies from The University of Tokyo, Tokyo, Japan, in 2022. She is currently pursuing the Ph.D. degree with Carnegie Mellon University, Pittsburgh, PA, USA, with interests in machine learning algorithms, self-healing systems design, and signal processing for wireless applications.

ETHAN CHEN is a Research Scientist with the Energy-Efficient Circuits and Systems Lab, Carnegie Mellon University, Pittsburgh, PA, USA. His research interests include neuromorphic computing, hardware security, and biomedical interfaces.

VANESSA CHEN (Senior Member, IEEE) received the Ph.D. degree in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA, USA, in 2013.

Before joining Carnegie Mellon University as an Assistant Professor, she was affiliated with The Ohio State University, Columbus, OH, USA. During her doctoral studies at Carnegie Mellon University from 2010 to 2013, she conducted research on algorithm-assisted approaches for enhancing energy efficiency and ultrahigh-speed ADCs with on-chip real-time calibration, and interned with IBM T. J. Watson Research Center, Yorktown Heights, NY, USA, in 2012. Prior to academia, she held positions as a Circuit Designer with Qualcomm, San Diego, CA, USA, and Realtek, Hsinchu, Taiwan, focusing on self-healing RF/mixed-signal circuits. Her research focuses on AI-enhanced circuits and systems, which include intelligent sensory interfaces, RF/mixed-signal hardware security, and ubiquitous sensing and computing systems.

Dr. Chen has received the NSF CAREER Award, the IBM Ph.D. Fellowship, and the Analog Devices Outstanding Student Designer Award. She has been involved in various technical program committees, including the IEEE International Solid-State Circuits Conference, the IEEE Symposium on VLSI Circuits, the IEEE Custom Integrated Circuits Conference, the IEEE Asian Solid-State Circuits Conference, and the IEEE/ACM Design Automation Conference. She has also served as an Associate Editor for several IEEE journals, including IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—PART I: REGULAR PAPERS, IEEE TRANSACTIONS ON BIOMEDICAL CIRCUITS AND SYSTEMS, and IEEE OPEN JOURNAL OF CIRCUITS AND SYSTEMS. Additionally, she has contributed as a Guest Editor for the *ACM Journal on Emerging Technologies in Computing Systems*.