Deep Learning-Driven Black-Box Doherty Power Amplifier with Pixelated Output Combiner and Extended Efficiency Range

Han Zhou, *Member, IEEE*, Haojie Chang *Member, IEEE*, and David Widén

***Abstract*—This article presents a deep learning-driven inverse design methodology for Doherty power amplifiers (PA) with multi-port pixelated output combiner networks. A deep convolutional neural network (CNN) is developed and trained as an electromagnetic (EM) surrogate model to accurately and rapidly predict the S-parameters of pixelated passive networks. By leveraging the CNN-based surrogate model within a black-box Doherty framework and a genetic algorithm (GA)-based optimizer, we effectively synthesize complex Doherty combiners that enable an extended back-off efficiency range using fully symmetrical devices. As a proof of concept, we designed and fabricated two Doherty PA prototypes incorporating three-port pixelated combiners, implemented with GaN HEMT transistors. In measurements, both prototypes demonstrate a maximum drain efficiency exceeding 74% and deliver an output power surpassing 44.1 dBm at 2.75 GHz. Furthermore, a measured drain efficiency above 52% is maintained at the 9-dB back-off power level for both prototypes at the same frequency. To evaluate linearity and efficiency under realistic signal conditions, both prototypes are tested using a 20-MHz 5G new radio (NR)-like waveform exhibiting a peak-to-average power ratio (PAPR) of 9.0-dB. After applying digital predistortion (DPD), each design achieves an average power-added efficiency (PAE) above 51%, while maintaining an adjacent channel leakage ratio (ACLR) better than -60.8 dBc.**

***Index Terms*—Deep learning, Doherty power amplifier, genetic optimization, inverse design, load modulation, neural networks.**

I. Introduction

To meet the increasing demand for higher data rates, modern communication systems have embraced advanced modulation techniques. These methods, however, inherently produce signals with a high peak-to-average power ratio (PAPR). Meanwhile, energy efficiency has become a critical design consideration, as power consumption in mobile networks constitutes a substantial share of both operational costs and environmental impact. Excessive energy usage further complicates system design due to the need for additional cooling infrastructure to manage thermal dissipation. Among all radio front-end components, power amplifiers (PA) are typically the most power-intensive. Therefore, improving PA efficiency, particularly under backed-off output power conditions, can yield considerable energy savings and lead to more sustainable and cost-effective network deployments [1].

This research was supported in part by Swedish Innovation Agency (Vinnova) Grant 2024-02531, MULTIRACS, through the Eureka CELTIC Framework, and in part by VINNOVA, Sivers Semiconductors, and Chalmers University of Technology under Grant 2022-00863.

H. Zhou, H. Chang, and D. Widén are with Chalmers University of Technology, Sweden (e-mail: han.zhou@ieee.org).

Active load-modulated PA architectures have gained considerable attention for their ability to enhance efficiency at back-off power levels. Recent innovations in this area [2]–[8] have demonstrated notable improvements in back-off efficiency. However, their circuit complexity often hinders practical deployment, especially at higher frequencies. In contrast, the Doherty PA [9] exhibits a simpler circuit configuration, contributing to its widespread commercial adoption. Central to Doherty PA performance is the output combiner network, which govern the interaction between the main and auxiliary sub-amplifiers. This cooperative behavior directly impacts achievable back-off efficiency.

In parallel, inverse design methodologies have emerged as a powerful paradigm across diverse fields such as integrated photonics [10]–[12], meta-lens antennas [13]–[15], and scattering theory [16], [17]. The core principle is to define a target performance and then determine the optimal structure through a top-down approach. This often involves pixelating a designated layout area into an $N \times M$ binary matrix, where each pixel denotes the presence or absence of a specific material. Compared to conventional electromagnetic (EM) structures composed of lumped or distributed elements, inverse design enables exploration of a vastly larger design space. In recent years, deep learning has been increasingly applied to such problems for its ability to learn complex features and efficiently navigate high-dimensional design spaces [18].

In radio frequency (RF) circuit design [19], deep learning–based inverse design has been successfully applied to broadband PAs, passive circuits [20]–[24], and harmonic-tuned Class F PAs [25]. However, most prior work has focused on relatively standard structures such as power splitters, filters, and Class B PAs with conventional combiners. Doherty PAs, by contrast, require much more sophisticated three-port combiner networks that integrate load modulation, power combining, phase shifting, and impedance transformation within a compact footprint.

To address this gap, we extend the inverse design methodology to the synthesis of complex pixelated three-port Doherty combiner networks. Inverse design of a traditional Doherty combiner is particularly challenging because it is typically developed through an iterative process, resulting in multiple cascaded networks such as matching sections, parasitic compensation stages, offset lines, and post-matching circuits in addition to the load modulation network [26]–[29]. This complexity significantly limits the practicality of applying inverse design directly to conventional combiners. By contrast,

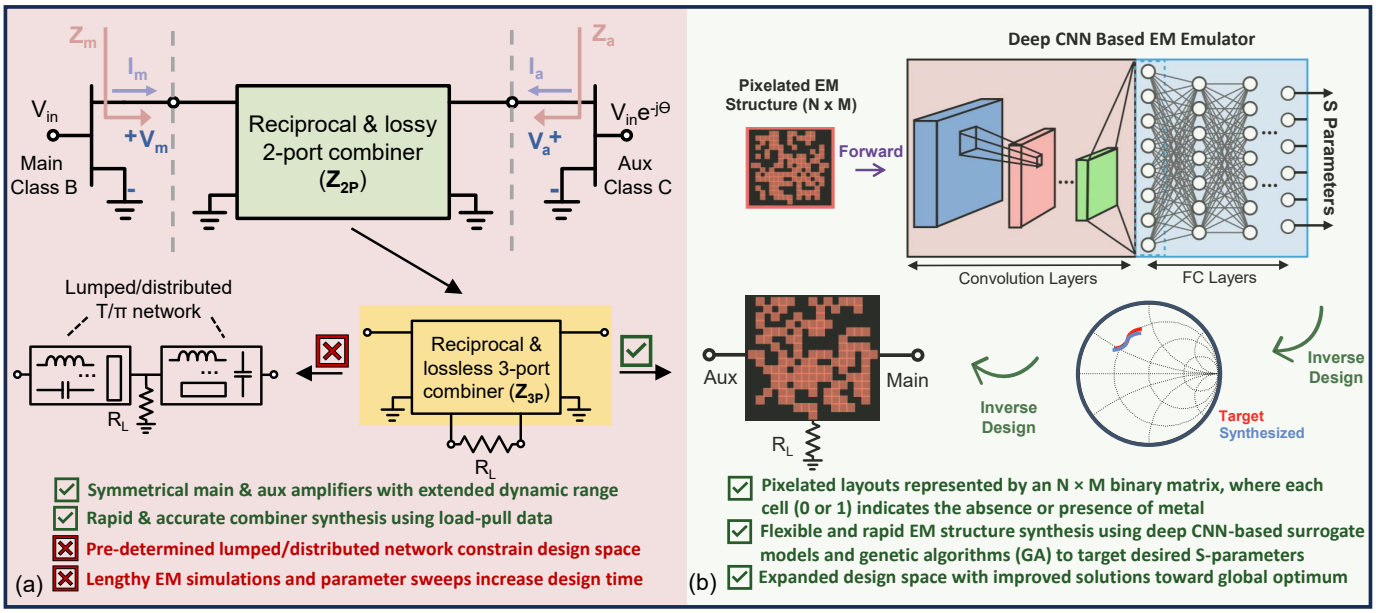


Fig. 1. Overview of the proposed deep learning–driven approach for the synthesis and design of black-box Doherty PAs with symmetrical size and extended back-off efficiency range. (a) Analytical black-box approach for deriving combiner parameters from load-pull data. (b) Deep learning–based top-down approach employing a pixelated Doherty combiner layout to explore the full design space.

adopting the black-box design approach of [30], [31] allows direct synthesis of the complete combiner from transistor load-pull data, eliminating the need to separately design and integrate each sub-network. This methodology greatly smooths the design process and enables extended efficiency enhancement within fully symmetrical configurations. Nevertheless, existing black-box approaches are typically limited to combiner structures composed of lumped or distributed circuit elements, such as transmission lines, capacitors, and inductors. These methods rely on bottom-up design strategies that involve parameter sweeps and repeated EM simulations. Although such techniques provide insight into circuit behavior, they inherently constrain the design space to predefined topologies with limited flexibility and a small number of tunable parameters. As a result, existing methodologies restrict the discovery of higher-performance solutions, require considerable simulation and optimization time, and ultimately impede the exploitation of the full performance potential of Doherty combiner networks.

Motivated by these limitations, this work introduces a data-driven inverse design framework that relaxes the topological constraints and mitigates the dependence on time-consuming EM simulations and tuning in existing Doherty combiner synthesis approaches. Building on this foundation, we present the first Doherty PA design that combines deep learning–driven inverse design with a pixelated output combiner network (Fig. 1). A deep convolutional neural network (CNN) surrogate model is trained to map pixelated EM structures to their scattering parameters (S-parameters) with high accuracy. Coupled with a genetic algorithm (GA), this surrogate accelerates the inverse design process by eliminating repeated full-wave EM simulations. Integrated with the black-box Doherty framework, our method achieves significant efficiency improvements at extended back-off levels using identical transistors directly synthesized from their load-pull data. The remainder of this paper is organized as follows. Section II illustrates the black-box combiner synthesis methodology and analyzes the Class B/C Doherty PA. Section III details the CNN architecture, data generation workflow, and GA strategy. Section IV demonstrates the implementation process, including two Doherty PA prototypes. Section V reports measurement results. Section VI concludes the paper.

II. The Black-box Combiner Theory

In this section, the generalized combiner synthesis approach is presented. The parameters of a generic two-port combiner are first derived to analyze the Class B/C Doherty PA behavior with extended efficiency range under varying drive conditions. Then, the synthesis procedure of a two-port Doherty combiner using load-pull data is demonstrated, along with the condition required to transform it into a lossless three-port combiner with resistive output loading.

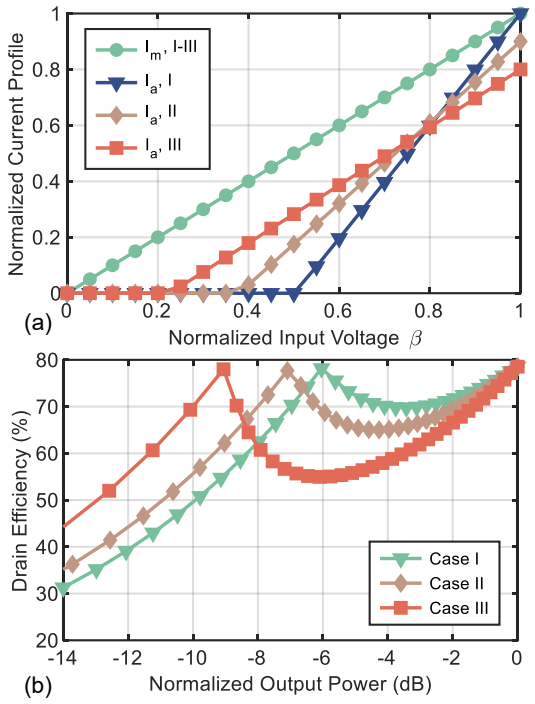
A. Black-Box Doherty PAs

Following the analysis in [31], transistors are modeled as ideal, piecewise, voltage-controlled current sources. Only the fundamental component is considered, with higher harmonics short-circuited, representing ideal Class B operation. A phase delay θ is assumed between the output currents of the main and auxiliary amplifiers. The output currents are given by [8]

$$I_{\mathrm{m}} = \beta i_{\mathrm{m,\,M}} \tag{1}$$

$$I_{\mathrm{a}} = \begin{cases} 0, & 0 \le \beta \le \beta_{\mathrm{B}} \\ \left(\dfrac{\beta - \beta_{\mathrm{B}}}{1 - \beta_{\mathrm{B}}}\right) i_{\mathrm{a,M}} \cdot e^{-j\theta}, & \beta_{\mathrm{B}} \le \beta \le 1 \end{cases} \tag{2}$$

where β and i denotes the normalized input voltage and fundamental current, respectively. In these expressions, the



(c)

	Case I	Case II	Case III
Back-Off V_{in} (β_B)	0.5	0.38	0.23
Current Ratio (**a**)	1.0	0.9	0.8
Back-Off P_{out} (**γ**)	6 dB	7 dB	9 dB

Fig. 2. Illustration of the black-box technique under different driving conditions. The output power is normalized to the corresponding saturated output power for each current-scaling ratio. (a) Normalized current drive profiles as a function of input voltage. (b) Theoretical performance comparison for different drive profiles (Cases I–III). (c) Corresponding current ratios and back-off levels for each case.

first subscript (m or a) refers to the main or auxiliary amplifier, respectively, and the second subscript (M or B) denotes operation at maximum or back-off power levels.

The combiner network is described by its impedance matrix $\mathbf{Z}_{2P}$, relating voltages and currents at the amplifier outputs

$$\begin{bmatrix} V_m \\ V_a \end{bmatrix} = \mathbf{Z}_{2P} \begin{bmatrix} I_m \\ I_a \end{bmatrix} = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12} & Z_{22} \end{bmatrix} \begin{bmatrix} I_m \\ I_a \end{bmatrix} \tag{3}$$

Assuming an optimal load of R_{opt} for the main amplifier, the auxiliary amplifier's optimal load reduces to R_{opt}/α, with α defined as the ratio of peak output currents between the main and auxiliary amplifiers. The combiner's impedance parameters are then derived as [32]

$$\begin{cases} Z_{11} = \dfrac{R_{opt}}{\beta_B} \\ Z_{12} = \left(1 - \dfrac{1}{\beta_B}\right)\left(\dfrac{R_{opt}}{\alpha}\right) e^{j\theta} \\ Z_{22} = \left(\dfrac{1}{\beta_B} + \alpha e^{-j2\theta} - 1\right)\left(\dfrac{R_{opt}}{\alpha^2}\right) e^{j2\theta} \end{cases} \tag{4}$$

To transform the lossy reciprocal two-port network into an equivalent lossless three-port combiner with resistive termination, the following condition must be satisfied [30]

$$\Re\{Z_{12}\}^2 = \Re\{Z_{11}\}\Re\{Z_{22}\} \tag{5}$$

The normalized output power back-off level, denoted by γ, and the phase delay θ are derived as [32]

$$\gamma = \frac{1+\alpha}{\beta_B} \tag{6}$$

$$\theta = k\pi \pm \arcsin\sqrt{\frac{\beta_B\left(\alpha - \beta_B + 1\right)}{1 - \beta_B^2}} \tag{7}$$

where k is an arbitrary integer.

Expressions (4–7) indicate that, for a given back-off level γ and current ratio α, a valid phase shift θ and corresponding combiner impedance matrix $\mathbf{Z}_{2P}$ can be determined. As illustrated in Fig. 2, the theoretical analysis suggests that this method enables the realization of desired back-off levels across various current ratios, facilitating full utilization of the main and auxiliary transistors.

B. Combiner Synthesized from Load-Pull Data

The theoretical analysis above demonstrates the fundamental operating principle of the black-box approach. However, at high frequencies, transistor parasitics and nonlinearities significantly influence PA performance. As a result, a combiner design methodology guided by nonlinear device behavior under large-signal excitation becomes essential for fully utilizing the device capabilities. The two-port combiner impedance matrix ($\mathbf{Z}_{2P}$) can be derived from load-pull data, as follows.

$$\begin{cases} Z_{11} = \dfrac{\left(Z_{a,M} + Z_{a,off}\right) Z_{m,B}\alpha^2 + \left(Z_{m,M} - Z_{m,B}\right) Z_{m,M}}{Z_{m,M} - Z_{m,B} + \left(Z_{a,M} + Z_{a,off}\right)\alpha^2} \\ Z_{12} = \dfrac{\left(Z_{m,M} - Z_{m,B}\right)\left(Z_{a,M} + Z_{a,off}\right)\alpha}{Z_{m,M} - Z_{m,B} + \left(Z_{a,M} + Z_{a,off}\right)\alpha^2} \\ Z_{22} = \dfrac{\left(Z_{a,M} + Z_{a,off}\right) Z_{a,M}\alpha^2 + \left(Z_{m,B} - Z_{m,M}\right) Z_{a,off}}{Z_{m,M} - Z_{m,B} + \left(Z_{a,M} + Z_{a,off}\right)\alpha^2} \end{cases} \tag{8}$$

The parameter α is given by

$$\alpha = \frac{i_{a,M}}{i_{m,M}} = \sqrt{\frac{\Re\{Z_{m,M}\}P_{a,M}}{\Re\{Z_{a,M}\}P_{m,M}}} e^{-j\theta} \tag{9}$$

where P represents the delivered output power and Z is the optimal impedance, both from load-pull data.

Furthermore, when selecting the design parameter for the back-off power level (γ_B), the following relationship must be satisfied to ensure power conservation

$$\gamma_B P_{m,B} = P_{m,M} + P_{a,M} \tag{10}$$

To implement a physically realizable, lossless three-port combiner ($\mathbf{Z}_{3P}$) terminated with a resistive load R_L, the synthesized lossy two-port network ($\mathbf{Z}_{2P}$) must satisfy the condition described in equation (5). Given equations (5), (8), and (9), the only remaining unknown is the phase parameter

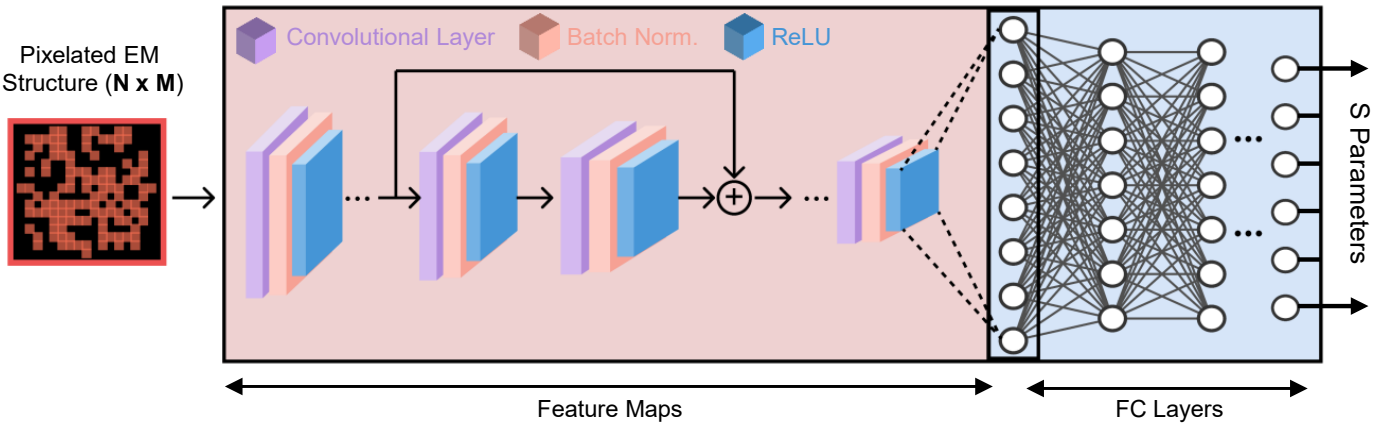


Fig. 3. Architecture of the trained deep CNN with residual connections. The input is a binary 15×15 matrix representing the pixelated EM layout structure, and the output is the predicted real and imaginary components of the corresponding S-parameters across the frequency range of interest.

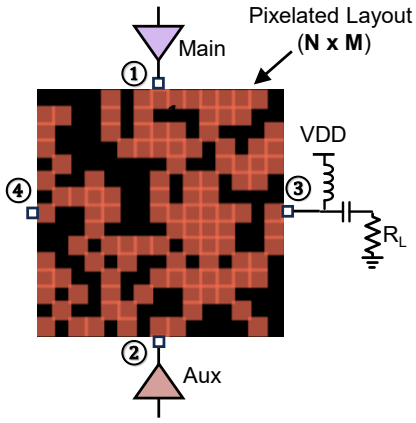


Fig. 4. The employed pixelated Doherty combiner networks, where the feed locations for the main (port 1), auxiliary (port 2), and output (port 3) are placed at the center of each corresponding edge of the 15-pixel array. The output port 3 is connected to the load (R_L), while port 4 is left open.

θ, which can be solved analytically or numerically. Once θ is determined, the complete impedance matrix $\mathbf{Z}_{2P}$ for the combiner can be fully specified.

III. Deep Learning-Driven Inverse Synthesis of Doherty Combiner Network

Building upon the theoretical analysis and equations in Section II, a two-port Doherty combiner matrix $\mathbf{Z}_{2P}$ can be directly derived from load-pull data, enabling designs with extended back-off efficiency while fully utilizing the employed transistors. However, physical realizations of such combiners have traditionally relied on predefined lumped or distributed topologies, often depending on expert intuition, iterative EM simulations, and manual parameter tuning [30], [31].

To overcome these limitations, this section introduces an inverse design methodology for synthesizing Doherty combiners using pixelated output networks. This approach enables exploration of a broader design space encompassing all fabricable structures. We begin by describing the pixelated layout design space and the corresponding training dataset. This is followed by the development of a deep CNN used as an EM surrogate model, as shown in Fig. 3. Lastly, the inverse synthesis process is detailed using a GA optimizer.

TABLE I
Detailed Configuration of the Deep CNN Architecture

Conv. Layer	Filter Size	Filter Number	Inputs
C_1	12×12	32	$\text{Matrix}_{\text{in}}$
C_2	12×12	32	C_1
C_3	10×10	32	C_2 + $\text{Matrix}_{\text{in}}$
C_4	10×10	32	C_3
C_5	8×8	32	C_4 + C_2
C_6	8×8	32	C_5
C_7	6×6	32	C_6 + C_4
C_8	6×6	32	C_7
C_9	4×4	32	C_8 + C_6
C_{10}	4×4	32	C_9
C_{11}	3×3	32	C_{10} + C_8
C_{12}	3×3	32	C_{11}
FC Layer	**Num. Neurons**	**Dropout**	**Act. Func.**
F_1	2048	0.25	LeReLU (0.01)
F_2	2048	0.25	LeReLU (0.01)
F_3	2048	0.25	LeReLU (0.01)
F_4	2048	0.25	LeReLU (0.01)
F_5	2048	0.25	LeReLU (0.01)
F_6	78	0	Tanh

A. Pixelated Doherty Combiner Network

To synthesize compact and low-loss Doherty combiner networks, a planar circuit layout is defined over a ground plane and discretized into a two-dimensional binary pixel grid. Each pixel is assigned a binary value: "1" represents the presence of metal, and "0" denotes a dielectric (non-metal) region. The layout is partitioned into a 15×15 grid, where each pixel measures 1.8×1.8 mm. This resolution is selected to balance EM coupling fidelity and the feasibility of model training [25]. A higher pixel resolution enables finer structural detail and access to a larger design space, potentially yielding improved performance at the expense of increased computational requirements. To ensure reliable

electrical connectivity, particularly across diagonal paths, a 20% metal overlap is applied to each metal pixel. This results in a total physical combiner size of $27.36\ \mathrm{mm} \times 27.36\ \mathrm{mm}$.

As illustrated in Fig. 4, the feed locations for the main (port 1) and auxiliary (port 2) amplifiers, as well as the output (port 3) connected to the load resistance (R_{L}), are placed at the center of each corresponding edge of the 15-pixel array. To enable DC biasing of the transistors, supply feeds (VDD) are introduced ahead of R_{L}, allowing the implementation of RF chokes and DC blocks.

The binary representation of this pixelated layout results in a total design space of 2^{225} possible EM structures. Given the vastness of this space, brute-force exploration is computationally infeasible in both time and memory. This challenge presents a compelling opportunity for deep learning. If a deep neural network can be trained as a surrogate EM model using a much smaller subset of simulation data (< 100000), it can then accurately and rapidly predict the scattering parameters of arbitrary pixelated layouts. This surrogate model effectively replaces time-intensive EM simulations governed by Maxwell's equations. When combined with a GA optimizer and guided by a target specification, the inverse design of pixelated structures can be effectively achieved.

B. Deep Convolutional Neural Network

The pixelated circuit layout can be naturally interpreted as a structured image, making CNNs, widely used in computer vision and image processing, well suited for this task. CNNs effectively model local spatial interactions and capture EM coupling between adjacent pixels by learning hierarchical features. In the following texts, we present the dataset generation process used for CNN training and describe the architecture of the proposed deep CNN model.

1) Data Generation and Augumentation: A dataset comprising diverse binary matrices representing pixelated EM structures and their corresponding S-parameters is essential for training the CNN. We fully automated the data generation process using Python scripts, which create pixelated layouts in Keysight Advanced Design System (ADS) and execute EM simulations using Momentum in ADS. The training circuits are generated by assigning each layout a random percentage of metal pixels, drawn from a normal distribution centered at 50% with a standard deviation of 15%. During data generation, we ensure that each circuit maintains connectivity between the ports by applying a depth-first search (DFS) algorithm to verify continuous electrical paths. We simulate a total of 77000 four-port pixelated EM structures, running four simulations in parallel. Each circuit simulation takes around 24 seconds. Using an Intel i5-14600KF CPU, the total time required to generate the dataset is approximately 129 hours.

Each simulated structure includes four ports, with one port terminated in an open condition and another port terminated with a $50\ \Omega$ load during augmentation. By applying rotations, translations, and mirror reflections [20], we generate additional three-port EM structures with a $50\ \Omega$ load termination, where the S-parameters are derived directly from the original four-port simulations. We augment the data such that for every simulated circuit, eight training data points are created, significantly reducing the effective dataset creation time. After augmentation, the total dataset size is approximately 1.5 GB.

Algorithm 1 Genetic Algorithm Optimization

Require: Desired S-parameter profile; trained CNN surrogate model; population size N; maximum iterations X etc.
Ensure: Optimal pixelated layouts with target S-parameters.

```
Select a target S-parameter profile.
Generate N binary matrices with direct connections, sampled from a normal metal pixel density distribution.
for each generation up to X iterations do
    Predict S-parameters of all candidate circuits using the trained deep CNN model.
    Calculate fitness based on deviation from the target S-parameter profile.
    if fitness of best candidate meets target then
        break loop; optimal solution found.
    end if
    Retain the top 10 circuits as elite individuals for the next generation.
    Add up to 30% new random circuits depending on current iteration count to maintain diversity.
    Perform tournament selection:
        • Split and recombine rows for crossover.
        • Flip 1−10% of pixels randomly for mutation.
        • Fill population up to N circuits.
end for
return The best pixelated circuit layout achieving or approaching the target S-parameter profile.
```

2) Deep CNN Architecture: As illustrated in Fig. 3, the deep CNN architecture consists of twelve convolutional layers and six fully connected (FC) layers. A residual network (ResNet) structure is employed to mitigate the vanishing gradient problem, which improves training stability and accelerates convergence [33]. Leaky rectified linear unit (LeReLU) functions are applied as activation functions after batch normalization [34], enabling the network to learn complex nonlinear relationships. In the convolutional layers, local spatial features are effectively extracted from the binary input matrix representing the pixelated EM layout. These feature maps are then flattened into a one-dimensional vector, which is passed to the FC layers for further processing. During training, dropout layers are inserted to reduce the risk of overfitting [35]. Each convolutional layer uses 32 filters, and each FC layer contains 2048 neurons with a dropout rate of 25%. The CNN takes binary matrices as input and produces real and imaginary components of the S-parameters as output. A detailed overview of the network architecture is provided in Table I.

We train the network using the mean absolute error (MAE) loss function, which quantifies the average absolute difference between predicted and actual values, and is defined by

$$L = \frac{1}{N} \sum_{i=1}^{N} |y_i - \hat{y}_i| \tag{11}$$

where y_i and $\hat{y}_i$ denote the actual and predicted values,

TABLE II
LOAD-PULL DATA OBTAINED AT 2.75 GHz

	Main Amplifier			Auxilary Amplifier		
	Z_{opt} (Ω)	P_{del} (dBm)	PAE (%)	Z_{opt} (Ω)	P_{del} (dBm)	PAE (%)
Peak Power Level	$14.3 + j1.6$	42.7	74	$14.3 + j1.6$	42.1	78
Back-Off Power Level	$7.2 + j15.3$	36.4	57	$0.25 + j21.1$ (Off)	-	-

respectively. The Adam optimizer [36] is used with an initial learning rate of 0.001, which we reduce by a factor of 0.93 every 10 epochs to aid convergence. The neural network is trained for 300 epochs using a batch size of 2790. Upon completion, the model achieves a mean absolute error (MAE) of 0.068, demonstrating high prediction accuracy.

C. Genetic Algorithm-Based Inverse Synthesis

The inverse synthesis of pixelated Doherty combiners is performed using a GA to efficiently explore the vast design space. As illustrated in Algorithm 1, the process begins by defining a desired S-parameter profile, which serves as the target specification. We then generate an initial population of $N = 4000$ binary matrices representing connected circuit layouts, sampling metal pixel densities from a normal distribution to ensure a diverse starting point.

The trained CNN surrogate model is used to rapidly predict the S-parameters of all candidate circuits in each generation, and fitness scores are calculated based on the deviation from the target S-parameter profile. If the fitness of the best candidate circuit meets the target performance or after $X = 240$ iterations, the algorithm is terminated. Otherwise, the top 10 circuits with the highest fitness scores are retained as elite individuals for the next generation.

To maintain diversity, up to 30% of the population is replaced with new random circuits depending on the current iteration count. This encourages exploration during the early stages of the optimization while focusing on refinement in later generations. Tournament selection is performed to identify parent pairs for crossover, where rows of parent matrices are split at random points and recombined to produce offspring. We also apply a mutation step by flipping between 1% and 10% of the pixels in the offspring circuits, introducing additional variability and helping to avoid convergence to local minima. This iterative process continues until the termination criterion is satisfied, gradually refining the population toward circuits that match the desired S-parameter response.

IV. DEEP LEARNING-DRIVEN DOHERTY PA DESIGN

In this section, we present the design of two prototype Doherty PAs using the proposed black-box synthesis methodology augmented by deep learning. The objective is to demonstrate the complete design procedure and validate the effectiveness of the proposed approach.

To ensure a fair and focused evaluation, the input networks of both Doherty PAs are kept completely identical. The only difference lies in the output Doherty combiners, which are individually synthesized using the proposed deep learning-aided inverse design strategy. This setup highlights the diversity of synthesis solutions achievable with the method, as well as its robustness and flexibility.

For both the main and auxiliary branches, we employ the same commercially available 10-W GaN HEMT transistor form (CG2H40010F, MACOM) as the active device. This choice ensures symmetry in the device characteristics.

A. Black-Box Doherty PA Design Based on Load-Pull Data

We begin by designing both amplifier branches using the MACOM transistors (CG2H40010F). The input network, including matching, stabilization, and biasing circuits, is identically designed and fixed for both Doherty prototypes. Low-impedance open stubs connected to the transistor gates, together with series RC networks at the inputs, are employed to simultaneously ensure unconditional stability and match the optimal source impedances of the transistors. In addition, resistors are incorporated in the gate bias lines to further enhance stability. A Wilkinson power divider is then used to equally distribute the input signal to the main and auxiliary amplifier branches. Next, we select the desired back-off power level ($\gamma_{\text{B}} = 9$ dB) and configure the gate bias of the main amplifier. Note that main amplifier's gate bias involves a trade-off between gain and efficiency: biasing the main amplifier closer to Class AB operation improves gain performance but can reduce efficiency. In our design, we set the main amplifier's gate bias to achieve a quiescent current of 40 mA. For the auxiliary amplifier, we set the gate bias in Class C operation. Selecting the auxiliary gate bias may require iterative adjustments to ensure that the auxiliary amplifier activates at the desired 9 dB back-off level.

Load-pull characterization of the main and auxiliary amplifiers is performed at both the peak output and the specified back-off power level. These simulations are conducted using the bias conditions established in the preceding design phase. It should be noted that when selecting the optimal load impedance, multiple degrees of freedom exist to trade off efficiency, output power, and linearity. In this work, we prioritize maximizing efficiency at the back-off power level and maximizing delivered output power at the peak power level. The extracted optimal impedances and corresponding power performance for both Doherty prototypes at peak and back-off conditions are summarized in Table II. Using equations (8)–(10), together with equation (5), we derive the phase parameter solution of $\theta = 133.4^\circ$ and the corresponding two-port combiner impedance matrix ($\mathbf{Z}_{\text{2P}}$) as

$$\mathbf{Z_{2P}} = \begin{bmatrix} 1.35 + j6.94 & -5.37 + j14.02 \\ -5.37 + j14.02 & 21.27 + j16.10 \end{bmatrix} \tag{12}$$

The obtained $\mathbf{Z}_{\text{2P}}$ matrix fully specifies the desired Doherty combiner behavior in a reciprocal two-port black-box form,

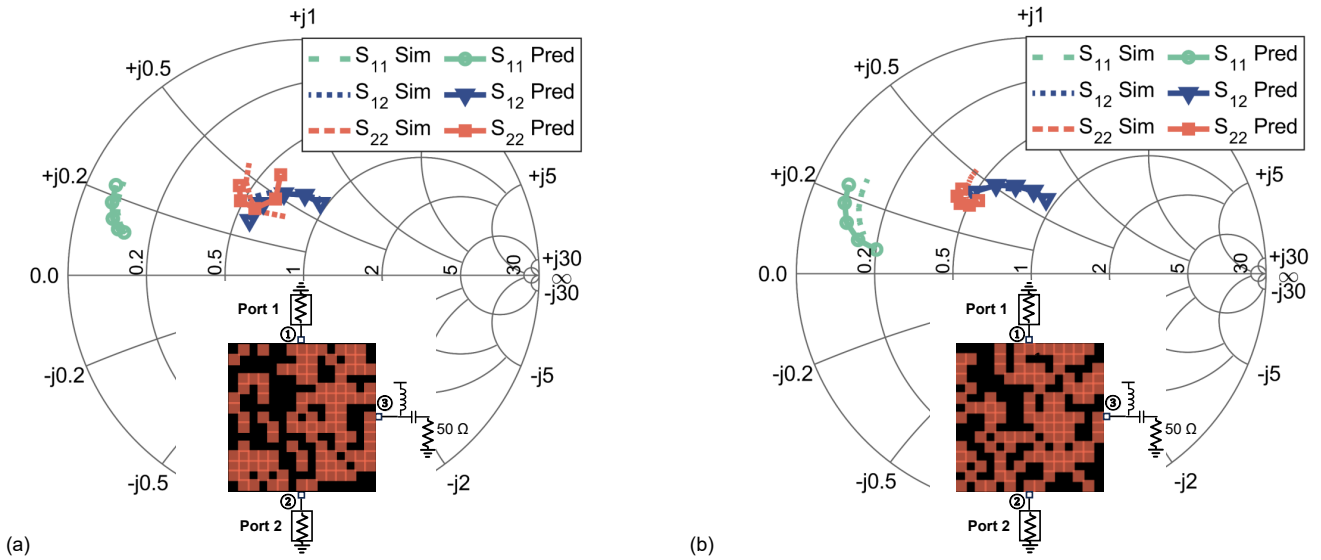
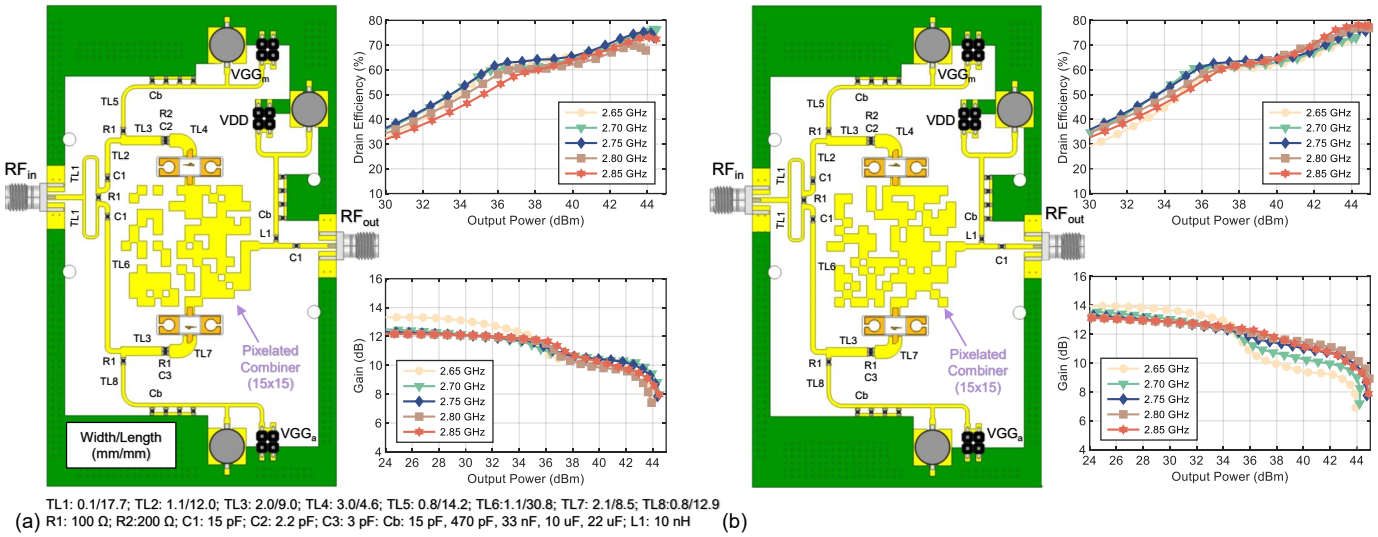


Fig. 5. The two synthesized pixelated Doherty combiner networks, (a) and (b), along with their EM-simulated S-parameter results compared to the corresponding responses predicted by the deep learning approach, over the frequency range of 2.55–2.95 GHz.



TL1: 0.1/17.7; TL2: 1.1/12.0; TL3: 2.0/9.0; TL4: 3.0/4.6; TL5: 0.8/14.2; TL6:1.1/30.8; TL7: 2.1/8.5; TL8:0.8/12.9
R1: 100 Ω; R2:200 Ω; C1: 15 pF; C2: 2.2 pF; C3: 3 pF: Cb: 15 pF, 470 pF, 33 nF, 10 uF, 22 uF; L1: 10 nH

Fig. 6. Full circuit schematic of the proposed deep learning-driven pixelated Doherty PAs: a) prototype 1 and (b) prototype 2. The corresponding full EM simulation results show the large-signal performance in terms of drain efficiency and gain versus output power over the frequency range of $2.65 - 2.85$ GHz.

specifying the impedance relationships required for achieving proper load modulation and matching. By implementing this impedance matrix as a lossless three-port combiner with all lossy resistive components consolidated into a single port, the Doherty PA can achieve the correct load trajectories across the entire back-off range, enabling an efficiency enhancement at $\gamma_{\mathrm{B}} = 9$ dB. Successfully implementing this combiner network represents a critical step, as the associated design space is vast and requires systematic exploration.

B. Pixelated Doherty Combiner Synthesis

Given the desired two-port impedance matrix derived earlier, we first convert this matrix into its corresponding scattering matrix, which specifies the target S-parameter behavior of the combiner network. To realize a physical implementation matching this target response, we use the GA in combination with the trained deep CNN surrogate model to search the vast pixelated design space efficiently. Unlike traditional approaches that require time-intensive EM simulations for every candidate structure, our deep learning-based method allows the CNN surrogate model to rapidly predict the S-parameters of thousands of pixelated layouts within a second per optimization step. This dramatically accelerates the design process compared to conventional brute-force EM sweeps.

During the optimization, we evaluate each candidate's predicted S-parameters by calculating a fitness value based on how closely they match the target scattering parameters. The fitness function is defined as

$$F = \frac{1}{e + \epsilon} \tag{13}$$

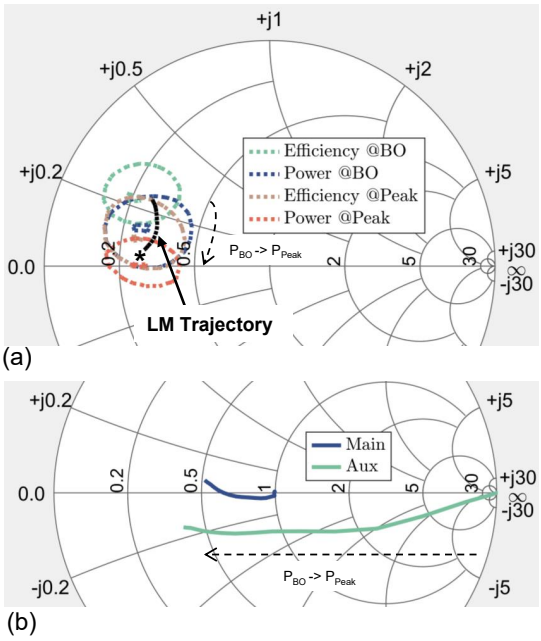


Fig. 7. (a) Power-delivered contours (maximum and -0.3 dB) and efficiency contours (maximum and -3% of the main amplifier for prototype 1 at peak and back-off output power, together with its load-modulation (LM) trajectory. (b) Load-modulation trajectories of the main and auxiliary amplifiers at the current-generator plane at the design center frequency.

where the error term e is defined as

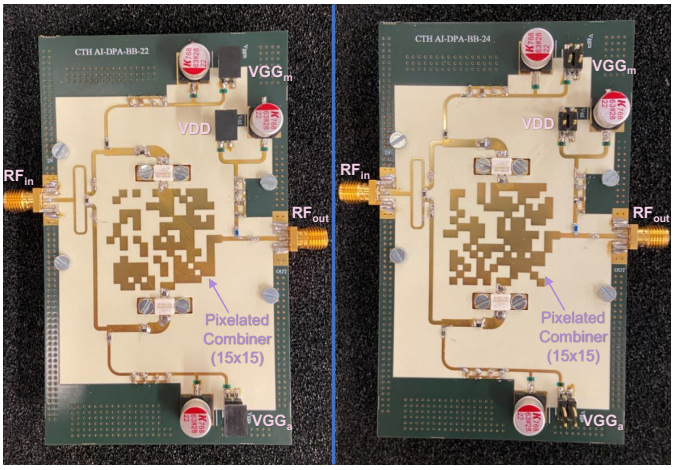
$$e = \max\left(\sum\left|\Re\{\hat{S}_i\} - \Re\{\bar{S}_i\}\right|, \sum\left|\Im\{\hat{S}_i\} - \Im\{\bar{S}_i\}\right|\right) \tag{14}$$

with $\left|\hat{S}_i - \bar{S}_i\right|$ representing the absolute errors between the predicted and target values for both the real and imaginary parts of S_{11}, S_{12}, and S_{22}. A small constant $\epsilon = 10^{-5}$ is included in the denominator to ensure numerical stability during fitness calculation.

Based on this optimization strategy, we synthesize two pixelated Doherty combiner networks, as shown in Fig. 5. The two solutions illustrate the stochastic nature of the GA, revealing that multiple distinct pixelated combiner networks can satisfy the same S-parameter targets for Doherty operation. Note that all three ports are terminated with $50\ \Omega$ during the simulations. The agreement observed between the EM-simulated S-parameters and the predictions generated by the CNN surrogate model demonstrates reasonable accuracy of the proposed deep learning–assisted design methodology. It is acknowledged that the prediction accuracy is subject to limitations, particularly when modeling S-parameters with small magnitudes. This agreement could be further improved by expanding the training dataset, enhancing the neural network architecture, and increasing the number of training epochs, given sufficient computational resources. Overall, these results confirm that the proposed approach effectively captures the complex EM interactions within the pixelated structures and enables efficient exploration of a large design space.

C. Simulation Results

Fig. 6 demonstrated the complete circuit schematics of the two fabricated Doherty PA prototypes. EM simulations of all transmission lines are performed using Momentum within Keysight ADS. We use accurate models from Modelithics for all lumped components, and transistor models provided by Macom are employed to ensure reliable simulation results.



(a) (b)

Fig. 8. Photograph of the fabricated deep learning-driven Doherty PAs with pixelated output combiners: (a) prototype 1 and (b) prototype 2.

The simulated efficiency and gain of the two prototypes versus output power at different operating frequencies are presented in Fig. 6. Excellent efficiency is observed across both saturation and back-off operating points within the targeted frequency band. At the center frequency of 2.75 GHz, both prototypes achieve a peak drain efficiency exceeding 78%, delivering more than 44.2 dBm saturated output power. Moreover, drain efficiency remains above 60% at the designated 9-dB back-off point.

To further illustrate the large-signal behavior, Fig. 7(a) presents the delivered-power contours, drain-efficiency contours, and load-modulation trajectory of the main amplifier for prototype 1 as a representative example at both peak and back-off operating points. The delivered-power contours indicate the maximum output power and the output-power contour within 0.3-dB of this maximum, while the efficiency contours denote the maximum achieved drain efficiency and the corresponding 3% degradation boundary. In addition, the intrinsic load-modulation trajectory is shown in Fig. 7(b), clearly revealing the characteristic load-modulation behavior of a Doherty PA. Overall, both prototypes demonstrate consistent large-signal performance across the entire design bandwidth, thereby validating the robustness and practical viability of the proposed synthesis approach.

V. Measurement Results

A photograph of the two fabricated Doherty PA prototypes is shown in Fig. 8. The prototype circuits are fabricated on a 20-mil-thick Rogers 4350B substrate, each with a compact size of 60×110 mm^2. The manufactured printed circuit boards (PCBs) and the employed CG2H40010F GaN HEMT transistors are mounted on brass fixtures, which also serve as heat sinks for thermal management. We perform small-signal, continuous-wave (CW), and modulated-signal measurements

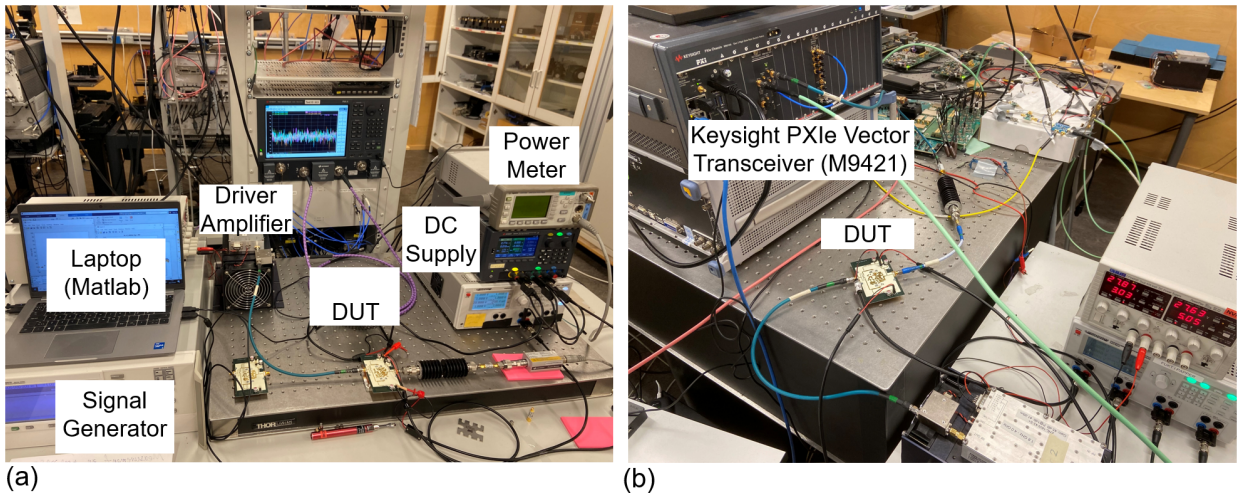


Fig. 9. Measurement setup: (a) small-signal and CW measurements, and (b) modulated-signal measurements.

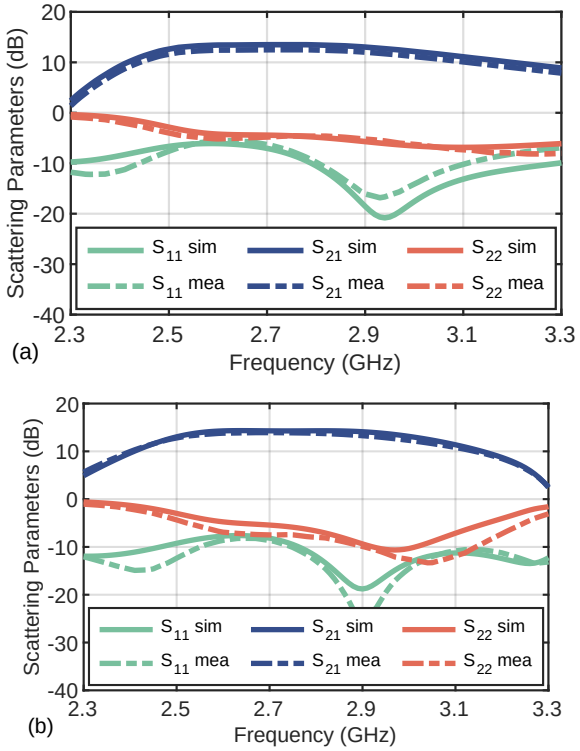


Fig. 10. Measured and simulated S-parameters of the fabricated Doherty PAs: (a) prototype 1 and (b) prototype 2.

to thoroughly characterize the fabricated Doherty prototypes. To ensure consistency, the drain supply voltage is set to 28 V for both prototypes throughout all measurements. Specifically, a quiescent current of 40 mA is set to the main amplifiers, while the gate bias of the auxiliary amplifiers is fixed at -7 V.

The measurement setup is shown in Fig. 9. Small-signal measurements are performed using a Keysight PNA-X network analyzer. For CW measurements, a signal generator provides the CW excitation, and the output power is measured using a power meter. To improve measurement accuracy, a low-pass filter is inserted at the output to suppress harmonic components and prevent them from entering the power meter. Moreover, a broadband linear driver amplifier supplies sufficient input power to drive the prototype circuits. For modulated-signal measurements, a Keysight PXIe vector transceiver (VXT M9421) generates and captures the modulated signals.

A. Small Signal

The performance of the two Doherty prototypes was first assessed through small-signal measurements to determine their frequency responses. As shown in Fig. 10, the strong correlation between simulated and measured data confirms proper operation. Within the 2.5–3.1 GHz range, both designs achieve a measured small-signal gain (S_{21}) exceeding 10 dB, while maintaining an input return loss (S_{11}) better than -8 dB across the entire band.

B. Continuous Wave

Fig. 11 presents the measured drain efficiency and gain as functions of output power for frequencies spanning 2.65 to 2.85GHz in 50-MHz steps. At the center design frequency of 2.75 GHz, Doherty prototype 1 achieves a measured peak drain efficiency of 74% with a maximum output power of 44.1 dBm, and a measured drain efficiency of 55% at 9-dB back-off power level. Meanwhile, Doherty prototype 2 reaches a peak drain efficiency of 76% with a peak output power of 44.8 dBm, and achieves a measured drain efficiency of 51% at the 9-dB back-off level. The performance of both prototypes versus frequency is also shown in Fig. 11, where the drain efficiency and power-added efficiency (PAE) at both the peak output power level and the 9-dB back-off level are presented, along with the peak output power. Across the frequency range from 2.65 to 2.85 GHz, the measured peak and 9-dB back-off efficiency for both prototypes exceed 68% and 46%, respectively.

These results clearly demonstrate that excellent high-efficiency enhancement is achieved around the 9-dB back-off power level, accompanied by a power gain compression

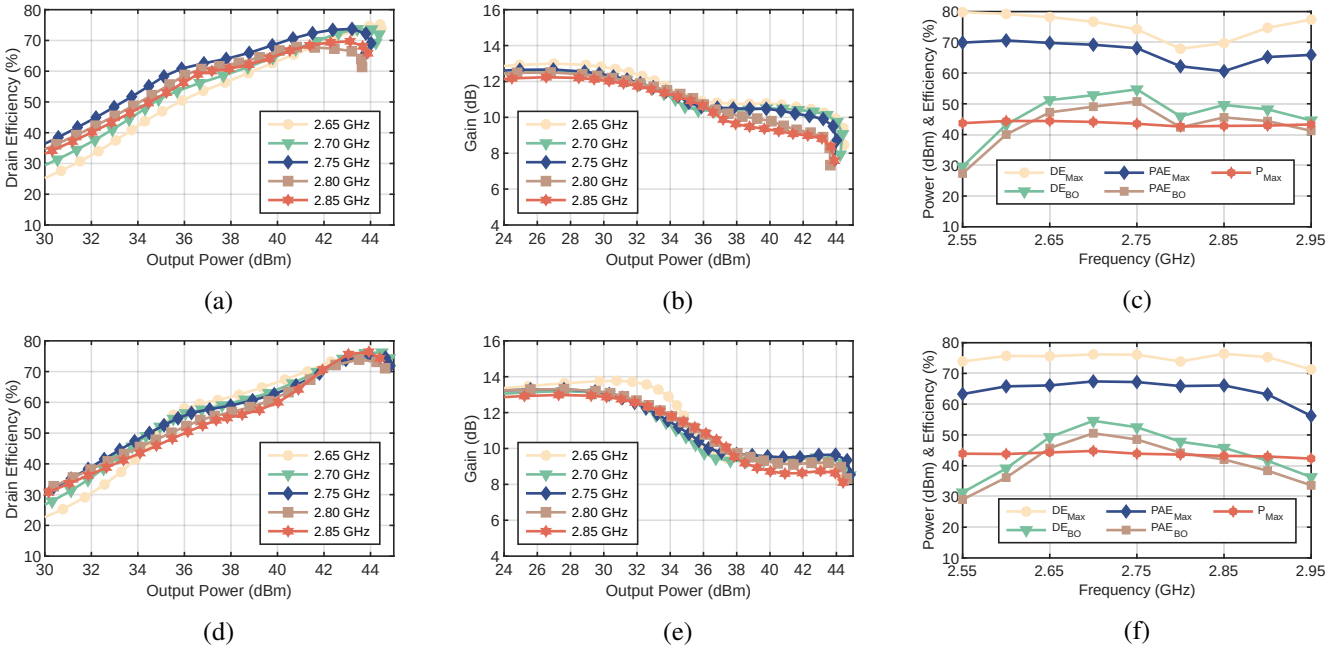


Fig. 11. Measured drain efficiency and gain of the Doherty prototype circuits versus output power, along with measured peak and back-off drain efficiency, PAE, and output power versus frequency: (a–c) prototype 1 and (d–f) prototype 2.

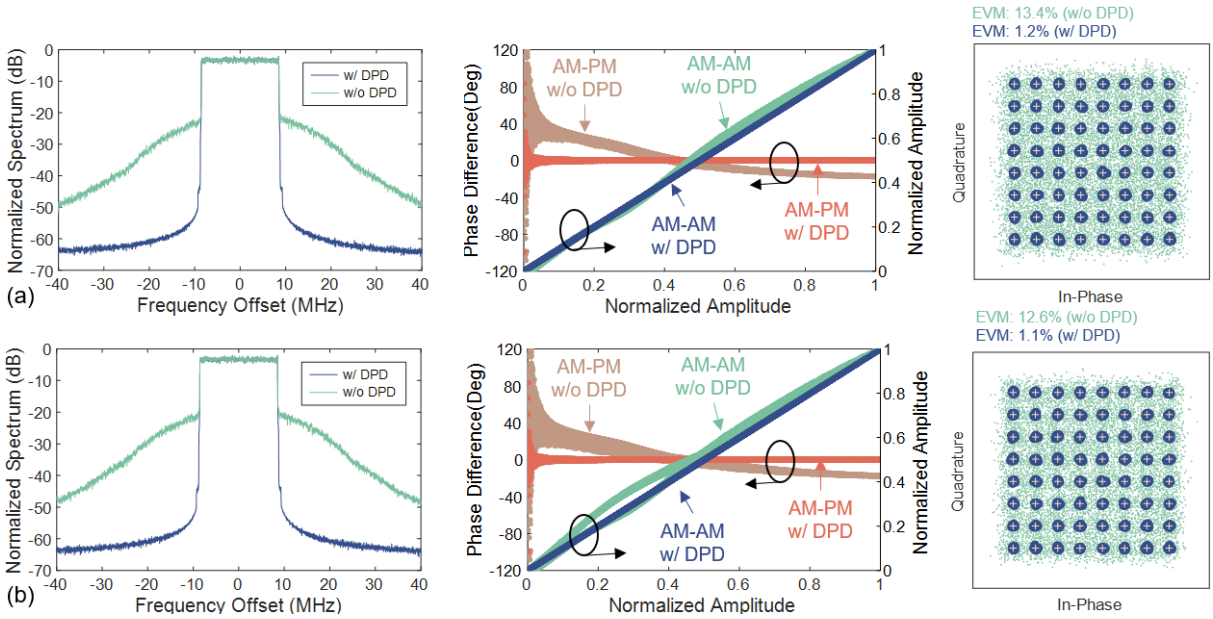


Fig. 12. Measured output spectrum, AM–AM/AM–PM characteristics, and constellation diagram under a 20-MHz, 9-dB PAPR 5G NR-like communication signal at 2.75 GHz: (a) Prototype 1 and (b) Prototype 2.

of approximately 3 dB. Furthermore, both prototype circuits exhibit consistent CW performance across the entire measured frequency band, confirming the effectiveness and robustness of the proposed design methodology.

C. Modulated Signal

Modulated-signal measurements are carried out to validate the proposed Doherty design methodology under realistic wireless communication conditions. Digital predistortion (DPD) is implemented using an iterative learning control (ILC) algorithm [46], which iteratively refines the input waveform to produce an optimally linearized output from the prototypes. The test signals consisted of 20-MHz 5G new radio (NR)-like waveforms with a PAPR of 9.0 dB. Fig. 12 shows the measured adjacent channel leakage ratio (ACLR) for the two prototypes, which are -24.8 dBc and -24.5 dBc at 2.75 GHz without digital predistortion (DPD). When DPD is applied, the ACLR improves significantly to -60.8 dBc and -61.1 dBc at the same frequency. After DPD, both prototypes deliver an average output power of about 36 dBm while maintaining an average drain efficiency above 51% and an error vector magnitude (EVM) below 1.2%. The measured AM-AM and AM-PM

TABLE III
BENCHMARKING OF STATE-OF-THE-ART HIGH-EFFICIENCY PAS.

Reference	Architecture	Freq (GHz)	$P_{\max}$ (dBm)	OPBO (dB)	DE / PAE @ $P_{\max}$ (%)	DE / PAE @ OPBO (%)	Signal BW (MHz)	PAPR (dB)	DE @ P_{ave} (%)	ACLR (dBc)	Size
[30] 2016	Sym. 2-way DPA	1.95	44.0	9	68/60	52/44	20	8.5	55	-49.0	0.062
[37] 2023	Sym. 2-way DPA	3.20	43.2*	6	54*/N.A.	46*/N.A.	100	6.0	42.2	-49.1	0.037
[38] 2019	Dual-In Sym. 2-way DPA	2.00	42.9	8	70*/68*	65/62	50	9.5	53.3	-47.1	N.A.
[39] 2022	Har. Inj. Sym. 2-way DPA	1.80	43.2*	9	68*/N.A.	64*/N.A.	5	9.6	N.A.	-23.5**	N.A.
[40] 2023	Class-EF Sym. 2-way DPA	2.60	45.2	6	76/75*	74/74*	N.A.	N.A.	N.A.	N.A.	N.A.
[41] 2023	Dual-In Sym. 2-way DPA	2.40	43.4	6	70/N.A.	58/N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
[31] 2022	Sym. 3-way DPA	2.14	45.3	10	69/57	55/45	20	8.5	56.6	-49.8	0.091
[42] 2022	RF-Input LMBA	2.40	44.1	6	54/N.A.	47/N.A.	10	8.6	44.0	-40.5**	N.A.
[43] 2020	RF-Input SLMBA	3.30	43.2*	10	71*/59*	48*/42*	200	10.0	43.6	-43.9	0.086
[44] 2024	Double-Balanced SLMBA	2.10	40.0	15	65/N.A.	54/N.A.	20	13.0	50.2	-23.5**	0.098
[45] 2024	RF-Input CLMA	3.40	42.3	6	57/N.A.	53/N.A.	20	7.0	51.0	-51.6	N.A.
This Work	**DL DPA Prototype 1**	**2.75**	**44.1**	**9**	**74/68**	**55/51**	**20**	**9.0**	**53.5**	**-61.1**	**0.06**
	DL DPA Prototype 2	**2.75**	**44.8**	**9**	**76/69**	**52/48**	**20**	**9.0**	**51.2**	**-60.8**	**0.06**

N.A. stands for data not available; *Estimated from graph; **No DPD performed; Size is defined as the prototype circuit area normalized to the wavelength at the center frequency.

characteristics, both with and without DPD applied, are The measured AM–AM and AM–PM responses, with and without DPD, are also presented, demonstrating that both prototypes achieve excellent linearity once DPD is applied.

D. Performance Comparison

The performance of the proposed deep learning-driven Doherty PA is summarized in Table III, alongside recently reported two-way and three-way Doherty PAs, as well as LMBAs and CLMAs. As can be observed, the designed prototype circuits achieve some of the highest efficiencies at deep power back-off levels, even when compared to other symmetrical two-way Doherty PAs employing dual-input or harmonic injection/tuning techniques. The efficiency is also competitive with that of symmetrical three-way Doherty architectures. In addition, the performance of the proposed design surpasses several RF-input LMBAs and CLMAs that target Doherty-like behavior. Although sequential LMBAs (SLMBAs) demonstrate improved efficiency under extreme power back-off conditions (beyond 10 dB), these architectures often experience excessive compression in the main amplifier and demand a larger circuit footprint.

Overall, the proposed deep learning-driven Doherty PAs offer a compelling combination of efficiency, compactness, and design flexibility. The measured linearity results verify that the prototypes satisfy the rigorous linearity and spectral requirements of contemporary wireless communication systems.

VI. CONCLUSION

A deep learning-driven inverse design methodology for Doherty PAs with pixelated output combiner networks and extended back-off efficiency is proposed. In this approach, a deep CNN is trained as an EM surrogate model to accurately and efficiently predict the S-parameters of arbitrary pixelated structures. By integrating the surrogate model with a GA optimizer within a black-box Doherty framework, the method enables efficient synthesis of complex three-port combiner networks using fully identical transistors.

To validate the proposed methodology, two Doherty PA prototypes are developed using GaN HEMT transistors. Both prototypes achieve peak drain efficiencies above 74% with output power exceeding 44.1 dBm at the center frequency of 2.75 GHz, and maintain efficiencies greater than 52% at the 9-dB back-off level. Furthermore, under a 20-MHz 5G NR-like signal with an 9.0-dB PAPR, the prototypes deliver average efficiencies exceeding 51% with ACLR better than -60.8 dBc after DPD deployment. These results demonstrate that the proposed deep learning-based methodology effectively enables rapid and accurate design of efficient, linear Doherty PAs suitable for modern wireless communication systems.

ACKNOWLEDGMENT

The authors thank Modelithics, Inc., Tampa, FL, USA, for the use of Modelithics models utilized under the University License Program. The authors also acknowledge MACOM for supplying the GaN HEMT transistors used in this work. Furthermore, We would like to thank Prof. María José Madero Ayora of the University of Seville for fruitful discussions and for her support with 5G NR-like communication signal generation and modulated-signal testing.

REFERENCES

[1] H. Wang, P. M. Asbeck, and C. Fager, "Millimeter-wave power amplifier integrated circuits for high dynamic range signals," *IEEE Microw. Mag.*, vol. 1, no. 1, pp. 299–316, 2021.

[2] D. J. Shepphard, J. Powell, and S. C. Cripps, "An efficient broadband reconfigurable power amplifier using active load modulation," *IEEE Microw. Wireless Compon. Lett.*, vol. 26, no. 6, pp. 443–445, 2016.

[3] J. Guo, P. Gong, and K. Chen, "Theory and design of pseudo-Doherty load-modulated double balanced amplifier with intrinsic insensitivity to antenna VSWR," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 72, no. 5, pp. 2048–2060, 2025.

[4] Y. Cao, H. Lyu, and K. Chen, "Continuous-mode hybrid asymmetrical load- modulated balanced amplifier with three-way modulation and multi-band reconfigurability," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 69, no. 3, pp. 1077–1090, 2022.

[5] P. Saad, R. Hou, R. Hellberg, and B. Berglund, "A 1.8–3.8-GHz power amplifier with 40% efficiency at 8-dB power back-off," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 11, pp. 4870–4882, 2018.

[6] G. Lv, W. Chen, X. Chen, F. M. Ghannouchi, and Z. Feng, "A 1.8–5.4-GHz GaN MMIC distributed efficient power amplifier with reactance compensation and adaptive biasing," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 71, no. 4, pp. 1531–1543, 2024.

[7] E. Liu, H. Zhou, C. Fager, and H. Wang, "A Ka-band Doherty-like non-load modulated power amplifier," *IEEE J. Solid-State Circuits*, vol. 60, no. 5, pp. 1584–1593, 2025.

[8] H. Zhou, J.-R. Perez-Cisneros, B. Langborn, T. Eriksson, and C. Fager, "A wideband and highly efficient circulator load modulated power amplifier architecture," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 70, no. 8, pp. 3117–3129, 2023.

[9] W. H. Doherty, "A new high efficiency power amplifier for modulated waves," *Proc. Inst. Radio Eng.*, vol. 24, no. 9, pp. 1163–1182, Sep. 1936.

[10] M. H. Tahersima, K. Kojima, T. Koike-Akino, D. K. Jha, B. Wang, C. Lin, and K. Parsons, "Deep neural network inverse design of integrated photonic power splitters," *Sci. Rep.*, vol. 9, 2019.

[11] W. Ma, Z. Liu, Z. A. Kudyshev, A. Boltasseva, W. Cai, and Y. Liu, "Deep learning for the design of photonic structures," *Nat. Photonics*, vol. 15, pp. 77 – 90, 2020.

[12] S. Molesky, Z. Lin, A. Piggott, W. Jin, J. Vučković, and A. W. Rodriguez, "Inverse design in nanophotonics," *Nat. Photonics*, vol. 12, pp. 659–670, 2018.

[13] P. Liu, L. Chen, and Z. N. Chen, "Prior-knowledge-guided deep-learning-enabled synthesis for broadband and large phase shift range metacells in metalens antenna," *IEEE Trans. Antennas Propag.*, vol. 70, no. 7, pp. 5024–5034, 2022.

[14] P. Liu and Z. N. Chen, "Full-range amplitude–phase metacells for sidelobe suppression of metalens antenna using prior-knowledge-guided deep- learning -enabled synthesis," *IEEE Transactions on Antennas and Propagation*, vol. 71, no. 6, pp. 5036–5045, 2023.

[15] R. Chen, C. W. L. Lee, P. K. Tan, R. Solanki, and T. H. Gan, "High-efficiency metalens antenna design through a controlnet diffusion generation model," *IEEE Antennas Wirel Propag Lett.*, vol. 24, no. 4, pp. 938–942, 2025.

[16] K. Xu, L. Wu, X. Ye, and X. Chen, "Deep learning-based inversion methods for solving inverse scattering problems with phaseless data," *IEEE Trans. Antennas Propag.*, vol. 68, no. 11, pp. 7457–7470, 2020.

[17] Y. Zhou, Y. Zhong, Z. Wei, T. Yin, and X. Chen, "An improved deep learning scheme for solving 2-D and 3-D inverse scattering problems," *IEEE Trans. Antennas Propag.*, vol. 69, no. 5, pp. 2853–2863, 2021.

[18] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.

[19] F. Feng, W. Na, J. Jin, J. Zhang, W. Zhang, and Q.-J. Zhang, "Artificial neural networks for microwave computer-aided design: The state of the art," *IEEE Trans. Microw. Theory Tech.*, vol. 70, no. 11, pp. 4597–4619, 2022.

[20] E. A. Karahan, Z. Liu, and K. Sengupta, "Deep-learning-based inverse-designed millimeter-wave passives and power amplifiers," *IEEE J. Solid-State Circuits*, vol. 58, no. 11, pp. 3074–3088, 2023.

[21] E. A. Karahan, Z. Liu, A. Gupta, Z. Shao, J. Zhou, U. Khankhoje, and K. Sengupta, "Deep-learning enabled generalized inverse design of multi-port radio-frequency and sub-Terahertz passives and integrated circuits," *Nat. Commun.*, vol. 15, no. 1, p. 10734, Dec. 2024.

[22] J. Zhou, E. A. Karahan, S. Ghozzy, Z. Liu, H. Jalili, and K. Sengupta, "AI-enabled design space discovery and end-to-end synthesis for rfics with reinforcement learning and inverse methods demonstrating mm-wave/sub-THz PAs between 30 and 120 GHz," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, vol. 68, 2025, pp. 1–3.

[23] S. Hang Chai, H. Chae, H. Yu, D. Z. Pan, and S. Li, "A D-band InP power amplifier featuring fully AI-generated passive networks," *IEEE Microw. Wireless Technol. Lett.*, vol. 35, no. 6, pp. 824–827, 2025.

[24] C. Chu, J. Xu, Y. Liu, J. Zeng, A. Wang, T. Torii, S. Shinjo, K. Yamanaka, and H. Wang, "AI-assisted template-seeded pixelated design for multi-metal-layer high-coupling EM structures: A Ku-band 6G FR3 PA in 22nm FDX+," in *2025 IEEE/MTT-S International Microwave Symposium - IMS 2025*, 2025, pp. 922–925.

[25] H. Zhou, H. Chang, D. Widén, L. Fornstedt, G. Melin, and C. Fager, "AI-assisted deep-learning-based design of high-efficiency class F power amplifiers," *IEEE Microw. Wireless Technol. Lett.*, pp. 1–4, 2025.

[26] X. Zhou, W. S. Chan, T. Sharma, J. Xia, S. Chen, and W. Feng, "A Doherty power amplifier with extended high-efficiency range using three-port harmonic injection network," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 69, no. 7, pp. 2756–2766, 2022.

[27] J. Pang, C. Chu, J. Wu, Z. Dai, M. Li, S. He, and A. Zhu, "Broadband GaN MMIC Doherty power amplifier using continuous-mode combining for 5G Sub-6 GHz applications," *IEEE J. Solid-State Circuits*, vol. 57, no. 7, pp. 2143–2154, 2022.

[28] M. Li, J. Pang, Y. Li, and A. Zhu, "Bandwidth enhancement of Doherty power amplifier using modified load modulation network," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 67, no. 6, pp. 1824–1834, 2020.

[29] W. Chen, S. Zhang, Y. Liu, Y. Liu, and F. M. Ghannouchi, "A concurrent dual-band uneven Doherty power amplifier with frequency-dependent input power division," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 61, no. 2, pp. 552–561, 2014.

[30] M. Özen, K. Andersson, and C. Fager, "Symmetrical Doherty power amplifier with extended efficiency range," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 4, pp. 1273–1284, Apr. 2016.

[31] H. Zhou, J.-R. Perez-Cisneros, S. Hesami, K. Buisman, and C. Fager, "A generic theory for design of efficient three-stage Doherty power amplifiers," *IEEE Trans. Microw. Theory Techn.*, vol. 70, no. 2, pp. 1242–1253, 2022.

[32] H. Zhou, "Efficient and wideband load modulated power amplifiers for wireless communication," Ph.D. dissertation, Chalmers University of Technology, Gothenburg, Sweden, 2023. [Online]. Available: https://research.chalmers.se/en/publication/537868

[33] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778.

[34] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proc. of the 27th Int. Conf. on Machine Learning (ICML)*, 2010, pp. 807–814.

[35] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, Jan. 2014.

[36] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv:1412.6980*, 2014.

[37] Z. Xiao, F. You, C. Shen, Q. He, B. Pang, J. Guo, P. Hao, Y. Fan, M. Song, C. Li, and S. He, "A Doherty power amplifier based on AM-AM/PM cancellation combining network synthesized by back-off complex load impedance," *IEEE Microw. Wireless Technol. Lett.*, vol. 33, no. 9, pp. 1333–1336, 2023.

[38] C. Liang, P. Roblin, and Y. Hahn, "Accelerated design methodology for dual-input Doherty power amplifiers," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 10, pp. 3983–3995, 2019.

[39] X. Zhou, W. S. Chan, T. Sharma, J. Xia, S. Chen, and W. Feng, "A Doherty power amplifier with extended high-efficiency range using three-port harmonic injection network," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 69, no. 7, pp. 2756–2766, 2022.

[40] C. Liu, X. Li, and F. M. Ghannouchi, "A highly-efficient Doherty power amplifier with generalized parallel-circuit class-EF mode," *IEEE Trans. Circuits Syst. II, Exp. Briefs.*, vol. 70, no. 8, pp. 2819–2823, 2023.

[41] Y. Yao, W. Shi, J. Pang, Z. Dai, and M. Li, "Design of a dual-input Doherty power amplifier with selectable output port," *IEEE Trans. Circuits Syst. II, Exp. Briefs.*, vol. 70, no. 4, pp. 1405–1409, 2023.

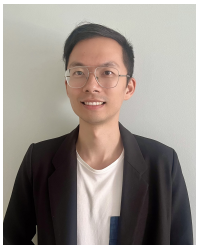
[42] K. Vivien, P. E. de Falco, O. Venard, G. Baudoin, P. Pierre-Charles-Félix, and T. Barton, "Load modulated balanced amplifier design method based on complex impedance trajectories," *IEEE J. Microw.*, vol. 2, no. 1, pp. 199–213, 2022.

[43] J. Pang, Y. Li, M. Li, Y. Zhang, X. Y. Zhou, Z. Dai, and A. Zhu, "Analysis and design of highly efficient wideband RF-input sequential load modulated balanced power amplifier," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 5, pp. 1741–1753, 2020.

[44] S. F. B. Faruquee, J. Guo, P. Gong, and K. Chen, "Hybrid load-modulated double-balanced amplifier (H-LMDBA) with four-way load modulation and $>$ 15-dB power back-off range," in *Proc. IEEE Wireless Microw. Technol. Conf. (WAMICON)*, 2024, pp. 1–4.

[45] H. Zhou, H. Chang, and C. Fager, "Analysis and design of RF-input Doherty-like circulator load modulated amplifier," in *Proc. 54th Eur. Microw. Conf. (EuMC)*, 2024, pp. 27–30.

[46] J. Chani-Cahuana, P. N. Landin, C. Fager, and T. Eriksson, "Iterative learning control for RF power amplifier linearization," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 9, pp. 2778–2789, 2016.



Han Zhou (Member, IEEE) received his B.Sc. degree in space science and astrophysics from the Harbin Institute of Technology, China, in 2016. He received his M.Sc. and Ph.D. degrees in electronic engineering from the Chalmers University of Technology, Sweden in 2018 and 2023, respectively. In 2022, he was a visiting researcher with the IDEAS Group at ETH Zurich, Switzerland.

He is currently a postdoctoral researcher at Chalmers University of Technology. His research interests include highly efficient, wideband, and linear power amplifier architectures for future wireless transmitters, the design of RF/mm-wave integrated circuits for wireless communication and sensing, and AI-assisted, deep learning-driven design for circuits and systems.

Dr. Zhou received the IEEE Microwave Theory and Techniques Society (MTT-S) Graduate Fellowship Award in 2023 and the EuMC Young Engineer Prize at the 52nd European Microwave Conference in 2022. He was also awarded a grant from the Ericsson Research Foundation in 2021.

Haojie Chang received the Ph.D. degree in Nanjing University of Science and Technology, China, in 2023. In 2022, he became a visiting researcher in Microwave Electronics Laboratory, Chalmers University of Technology, Sweden and from 2023 he became a post-doc in Chalmers. His current research interests include sub-THz circuits, antenna and packaging design, as well as advanced high-efficiency power amplifier theory and architectures.

David Widén (Graduate Student Member, IEEE) received the B.Sc. degree in engineering physics from the Chalmers University of Technology, Gothenburg, Sweden, in 2024, where he is currently pursuing the M.Sc. degree in electrical engineering. His research interests include analog circuits and wireless communication.