A Multi-Head, Segmented, and Real-Imaginary Separated Algorithm for Automated ANN Model Extraction and Packaging in MMIC Design

***Abstract*—The integration of machine learning into the MMIC circuit design process, notably in the training and application of Radio Frequency (RF) scattering parameter ANN models, holds substantial promise. We propose an automated extraction algorithm for artificial neural network models, designed for high-precision, low-overhead models with multiple physical significances, and paired with a model compression method to ensure compatibility with standard computer simulation software. Results reveal that our algorithm produces highly accurate models, corroborated by their alignment with HFSS simulation outcomes. Upon model compression, ANN models can be successfully transformed into macro models within ADS, facilitating simulation alongside peripheral circuits. This method provides a valuable approach for researchers aiming to harness neural network models within the MMIC design process.**

***Index Terms*- MMIC, machine learning, neural network, model compression, EDA.**

I. Introduction

Monolithic Microwave Integrated Circuits (MMIC) [1] integrate all microwave components onto a single semiconductor substrate [2]. They are extensively employed across several applications, including wireless local area networks (WLANs) [3] and automotive radars [4], due to their compact size, light weight, and superior performance. However, along with technological advancements in integrated circuit manufacturing, MMIC design faces stark challenges from quantum effects [5], thermal effects [6], and electromagnetic interference [7] due to device miniaturization, and burgeoning market demand for high-performance microwave circuits [2].

Researchers attempt to address these challenges by combining machine learning techniques with circuit design and testing [8] [9] [10], profoundly optimizing the design efficiency. This study introduces an automated, multi-port, segment-wise, and separate real-imaginary parts Artificial Neural Network (ANN) model extraction algorithm, enhancing model accuracy by preprocessing, clustering simulation data, and applying a separate real-imaginary and segmented machine learning model training algorithm.

To facilitate integration with electronic design automation (EDA) software, we introduced model compression techniques reducing the complexity of the neural network structure. This resultantly bridged the gap between the neural network models and the EDA software, simplifying the invocation of neural network models by circuit engineers. Currently, this technology is utilized in the design of an 8-port power amplifier combiner structure. After carrying out model compression, we encapsulated the ANN model in Advanced Design System (ADS) software, creating a macro model applicable for joint simulation with surrounding circuits.

II. Methods and Principles

A. Technical process

We introduce a method for deriving segmented, multi-port, and real-imaginary-separated ANN models, as illustrated in Fig. 1. This procedure creates an automatic simulation library, utilizes logic provided by High Frequency Structure Simulator(HFSS) for simulation invocation, conducts batch electromagnetic simulations and organizes the results into a database. Following data normalization and cleaning, the algorithm facilitates the training of segmented ANN models. These high-precision models are then compressed simplifying the neural network before their incorporation into ADS and MATLAB for utilization.

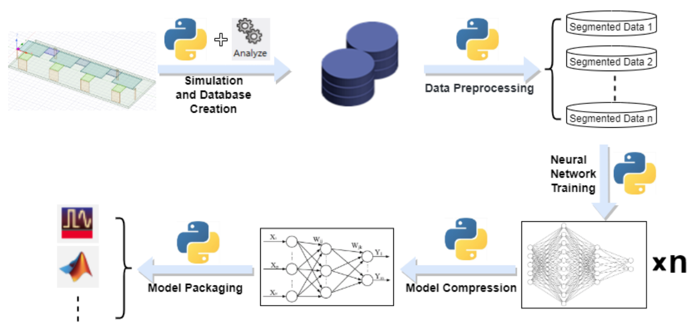

Fig. 1. The outlined process provides an automated method for extracting ANN models. It constitutes automatic simulation database creation, portioning for output, model optimizations via compression algorithms, and finally, encapsulation for simulation applications such as ADS and MATLAB.

B. Autonomous Algorithm for Segmental Training of Multi-port Wideband Data and Distinct Real and Imaginary Components

Circuit simulations frequently use dB scattering parameters, as they provide intuitive comparison of signal strength and attenuation [11]. However, this approach neglects phase-specific attributes essential for complex networks [12], [13]. Challenges in MMIC design and component interdependency further obstruct efficiency and accuracy [14]. Additionally, factors such as sparse connections within wideband data and certain theoretically isolated MMIC structures affect model accuracy when used as indiscriminate training input [15] [16].

Our method addresses these issues by decomposing complex numbers into real and imaginary components for neural network training [17], maintaining scattering properties. Particularly for smaller anticipated dB parameters, converting complex data into dB parameters (Eq. (1)) is crucial. Adding ANN training on dB data can enhance accuracy when magnitude discrepancy is substantial.

$$dB = 20 * log(\sqrt{complex.real^2 + complex.imag^2}) \quad (1)$$

ANN model training, utilizing datasets and model structure data, partitions complex scattering parameters into two separate real variables, facilitating an ANN model that discretely processes real and imaginary details as outlined

The multi-port wideband data partitioning training algorithm proposed in this study facilitates a more refined data processing strategy. It deploys an intelligent segmentation approach towards data, prior to model training. On one hand, it reduces the complexity of model attributes contained within single-segment data by segmenting simulation data based on frequency range, thereby enhancing the Artificial Neural Network (ANN)'s capacity to capture the diverse circuit traits across different frequency bands. On the other hand, it prescribes the division of independent models based on the number of circuit ports. Each model is trained against data from a singular port, escalating the accuracy of training outcomes.

The advanced algorithm detailed in this research paper, notably boosts the training precision of the final model, even within the constraints of a sparse data set. Concomitantly, it diminishes training time and computational expenditure. This partitioned training algorithm sufficiently considers the physical characteristics of the circuit, hence optimizing the overall prediction logic process, rendering it more efficient.

C. Model Compression and Packaging Methodsm

Our approach utilizes ANN models for MMIC design optimization, albeit processing complexities [18]. To mitigate these, efficient model compression strategies, inclusive of pruning and fine-tuning [19], are implemented to enhance computational efficacy.

We have introduced a proficient "Three-Step" algorithm for pruning and fine-tuning neural networks, which not only streamlines the process of model pruning, but also accomplishes automation in the realm of structured pruning. The steps are as follows:

1) Train the original model.
2) Pruning the model structure.
3) Fine-tune the model to compensate accuracy.
4) Output the revised model.

Ultimately, we achieve cross-software joint simulation of circuit models by establishing macro models of the ANN model structure within ADS. This method amplifies the reusability, manageability, and convenience of ANN models, offering innovative avenues for the optimization of electronic design.

III. Results and Discussion

A. Test Subject

The efficacy of the proposed algorithm is ascertained through testing on an eight-port power amplifier combiner, a task typically demanding impracticable computation time due to a high quantity of variable parameters. By employing an automated library building algorithm and leveraging an HFSS-provided simulation interface, We harnessed 625 full-frequency simulation instances as the prime dataset for our training purposes.

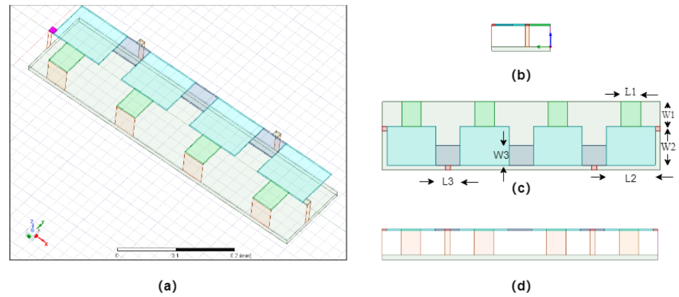

Fig. 2. The eight-port combiner structure, where (a) is a model example, and (b) to (d) illustrate the three-view diagram of the model.

B. Training of the Model Involving Separation of Real and Imaginary Components

The proposed algorithm automates intricate data splitting, normalization, and model training processes while preserving relevant information for accurate predictions. Despite an increase in training time, it enhances the predictive model's physical interpretability by separately dealing with real and imaginary components, as substantiated in TableI.

TABLE I: Comparative Analysis of Outcomes from Direct Training and Separate Training Approaches.

Training Method	Training Time (min)	Accuracy	Amplitude	Phase	Model Cost (ms)
Direct Training	144	99.00%	√	×	874
Separate Training	205	99.89%	√	√	879

The algorithm in this study endows the trained neural network model with enhanced physical implications. The separately trained models post-decomposition can accurately forecast both the magnitude and phase information of scattering parameters.

C. Training of Model with Frequency Segmentation

Our algorithm shrewdly partitions the training dataset based on input frequencies, thereby augmenting the precision of model training. As illustrated in the figure, the data segmentation strategy enhances model accuracy. Fig. 3 evidences this fact: the segmentation strategy surpasses under equivalent training conditions.

The proposed algorithm's multi-port training conducts segmentation based on the data object's port parameters count.

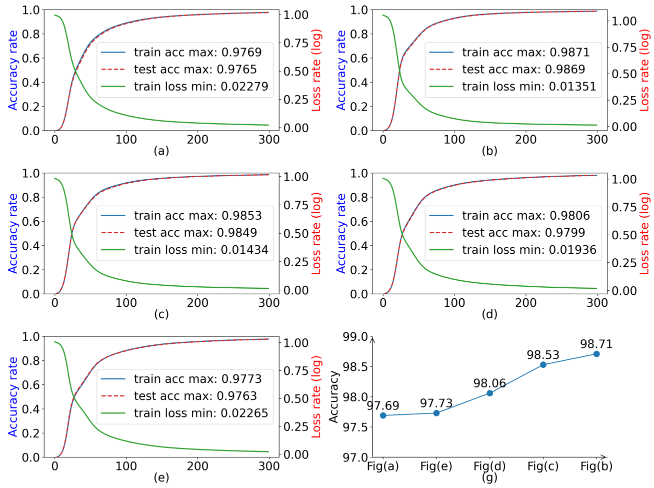

Fig. 3. Figure (a) illustrates a non-segmented data training process, while figures (c) to (e) demonstrate the segmented data training process. Figure (f) provides a comparative view of model accuracy.

Segmenting training data by output ports significantly enhances model training accuracy;based on the final result, a solo port-trained model can achieve accuracy up to 99%.

D. Model Compression

The "three-step" compression algorithm, engineered into this study's neural network design, notably preserves the accuracy intrinsic to the original model. Subsequent to compression, a consequential reduction in computational parameters occurs, mitigating the complexity encompassed within the macro-model integrated into ADS. These achievements have been demonstrated to facilitate the studies involving neural network models within the ADS, as illustrated in Fig. 4.

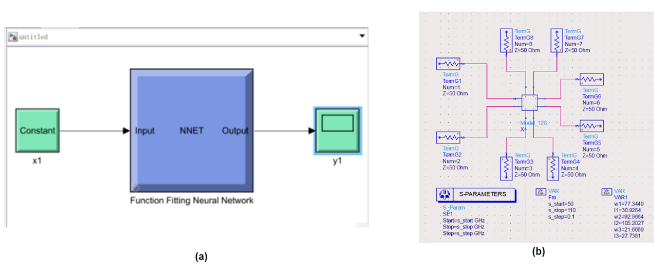

Fig. 4. Figure (a) represents the application of the encapsulated neural network in MATLAB, and Figure (b) depicts its integration into ADS.

IV. Conclusion

This study successfully employs a broad-band multi-port segmented data ANN training algorithm to optimize training logic, intensify accuracy, and merge machine learning with MMIC automated design. However, the necessity of manual intervention and input stands as a significant limitation, indicating future work towards algorithm refinement, improved automation, and exploring effective ADS encapsulation methods, reducing manual input.

References

[1] David N McQuiddy, John W Wassel, J Bradner Lagrange, and Wiliam R WiSSeman. Monolithic microwave integrated circuits: An historical perspective. *IEEE Transactions on Microwave Theory and Techniques*, 32(9):997–1008, 1984.

[2] Robert A Pucel. Looking back at monolithic microwave integrated circuits. *IEEE Microwave Magazine*, 13(4):62–76, 2012.

[3] Rory L Van Tuyl. Unlicensed millimeter wave communications. a new opportunity for mmic technology at 60 ghz. In *GaAs IC Symposium IEEE Gallium Arsenide Integrated Circuit Symposium. 18th Annual Technical Digest 1996*, pages 3–5. IEEE, 1996.

[4] Nicholas P Morenc. Mmics for automotive radar applications. In *1996 IEEE MTT-S International Microwave Symposium Digest*, volume 1, pages 39–41. IEEE, 1996.

[5] Lieven Vandersypen and Antoni van Leeuwenhoek. 1.4 quantum computing-the next challenge in circuit and system design. In *2017 IEEE International Solid-State Circuits Conference (ISSCC)*, pages 24–29. IEEE, 2017.

[6] Chiara Ramella, Marco Pirola, Andrea Reale, Mariarcangela Ramundo, Paolo Colantonio, Matthias Auf Der Maur, Vittorio Camarchia, Anna Piacibello, and Rocco Giofrè. Thermal-aware gan/si mmic design for space applications. In *2019 IEEE International Conference on Microwaves, Antennas, Communications and Electronic Systems (COMCAS)*, pages 1–6. IEEE, 2019.

[7] Richard Leach. Failures and anomalies attributed to electromagnetic interference. In *Space Programs and Technologies Conference*, page 3654, 1995.

[8] Yunong Ye, Yifan Wu, Jiayu Chen, Guodong Su, Junchao Wang, and Jun Liu. Intelligent design of hairpin filters based on artificial neural network and proximal policy optimization. *Applied Sciences*, 13(16):9379, 2023.

[9] Yiding Wei, Jun Liu, Dengbao Sun, Guodong Su, and Junchao Wang. From netlist to manufacturable layout: An auto-layout algorithm optimized for radio frequency integrated circuits. *Symmetry*, 15(6):1272, 2023.

[10] Naiyin Zhang, Zhenya Liu, and Junchao Wang. Machine-learning-enabled design and manipulation of a microfluidic concentration gradient generator. *Micromachines*, 13(11):1810, 2022.

[11] S Arslanagić, Troels Vejle Hansen, N Asger Mortensen, Anders Heidemann Gregersen, Ole Sigmund, Richard W Ziolkowski, and Olav Breinbjerg. A review of the scattering-parameter extraction method with clarification of ambiguity issues in relation to metamaterial homogenization. *IEEE Antennas and Propagation Magazine*, 55(2):91–106, 2013.

[12] Fayu Wan, Ningdong Li, Blaise Ravelo, Qizheng Ji, Binhong Li, and Junxiang Ge. The design method of the active negative group delay circuits based on a microwave amplifier and an rl-series network. *IEEE Access*, 6:33849–33858, 2018.

[13] Hong-Yeh Chang, Tian-Wei Huang, Huei Wang, Yu-Chi Wang, Pane-Chane Chao, and Chung-Hsu Chen. Broad-band hbt bpsk and iq modulator mmics and millimeter-wave vector signal characterization. *IEEE transactions on microwave theory and techniques*, 52(3):908–919, 2004.

[14] William Batty, Carlos E Christoffersen, Alexander B Yakovlev, John F Whitaker, Amir Mortazawi, Ayman Al-Zayed, Mete Ozkar, Sean C Ortiz, Ronald M Reano, Kyoung Yang, et al. Global coupled em-electrical-thermal simulation and experimental validation for a spatial power combining mmic array. *IEEE transactions on microwave theory and techniques*, 50(12):2820–2833, 2002.

[15] Robert E Collin. *Foundations for microwave engineering*. John Wiley & Sons, 2007.

[16] Djilali Hammou, Emilia Moldovan, and Serioja Ovidiu Tatu. Novel mhmic millimeter wave power divider/combiner. In *2011 24th Canadian Conference on Electrical and Computer Engineering (CCECE)*, pages 000280–000283. IEEE, 2011.

[17] Yu-chen Wu and Jun-wen Feng. Development and application of artificial neural network. *Wireless Personal Communications*, 102:1645–1656, 2018.

[18] Gavin J Bowden, Holger R Maier, and Graeme C Dandy. Real-time deployment of artificial neural network forecasting models: Understanding the range of applicability. *Water Resources Research*, 48(10), 2012.

[19] Song Han, Jeff Pool, John Tran, and William Dally. Learning both weights and connections for efficient neural network. *Advances in neural information processing systems*, 28, 2015.