IMS Deep Learning Enabled Generalized Synthesis of Multi-Port Electromagnetic Structures and Circuits for mmWave Power Amplifiers

Emir Ali Karahan[#], Zheng Liu[#*], Kaushik Sengupta[#]
[#]Princeton University Electrical and Computer Engineering, USA
[*]Now with Texas Instruments Kilby Lab, USA
karahan@princeton.edu

2024 IEEE/MTT-S International Microwave Symposium - IMS 2024 | 979-8-3503-7504-6/24/$31.00 ©2024 IEEE | DOI: 10.1109/IMS40175.2024.10600352

***Abstract*— This work reports an artificial intelligence (AI) assisted method for implementation of multiport electromagnetic (EM) structures. Utilization of AI in mm-wave circuit design is rapidly gaining attraction. While many previous works relied on surrogate modelling of a template, non-intuitive design spaces that provide exponentially larger degrees of freedom can uncover solutions that are not limited to the given template. Although it can be challenging to train accurate predictors for such design spaces, this issue can be tackled with a correct modelling approach. Within this context, we demonstrate the extension of previously laid out design framework to multiport EM structure synthesis [1]. This aims to facilitate the application of AI assisted methods for more comprehensive scenarios such as multiport matching networks, power dividers, diplexers and so on. To demonstrate the flexibility of this method, we show examples for different design goals. Lastly, we report a compact mm-wave amplifier with 3 dB gain bandwidth of 23.6-37.3 GHz, implemented with three port asymmetric power dividers and combiners that emerged through the AI-enabled method.**

***Keywords*— mm-wave, 5G, inverse design, machine learning, SiGe, power combining.**

I. INTRODUCTION

In high frequency systems, the design of EM structures is critical for optimum circuit performance. This is especially prevalent in mm-wave systems where gain and power that can be harvested from an active device is quite limited. Traditionally, the design of EM structures to implement a variety of functions relied on physics knowledge, intuition, or experience. Moreover, the manual design method requires considerable domain knowledge from the perspective of a designer. While this is advantageous in terms of gaining insight into circuit operation, it is not certain that resulting EM structures will yield optimal results. Furthermore, each EM design requires a time intensive design cycle. Given these considerations, one can suggest that we should pursue new design paradigms that can ease the designers' job, accelerate design cycles, and improve circuit performance. In parallel to this, device and EM modelling, inverse design algorithms, and heuristic optimizations have seen a surge of interest [1]–[8]. Furthermore, moving towards non-intuitive design spaces could open up new possibilities for circuit and system performance. Particularly for EM structures, instead of relying on templates, it is worthwhile to investigate topology optimization/inverse design methods. These approaches have been successfully deployed in the field of nanophotonics [9], [10]. In this regard, discretizing a surface onto small pixels lends itself as a convenient and manufacturable method for constructing large design spaces. Yet the bottleneck is that possible number of designs exponentially increase as we go to finer pixelations. For example, a 25 by 25 grid yields $\approx 10^{188}$ possible designs. It is impractical to sweep over such a design space by doing electromagnetic simulations.

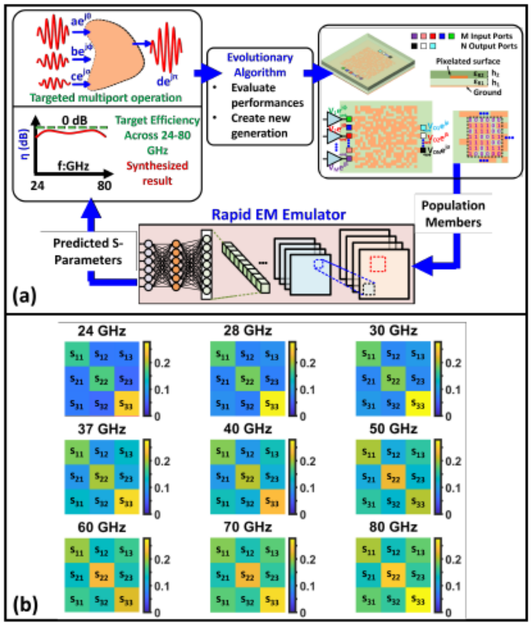

Fig. 1. (a) Concept and method for inverse design of multiport EM structures. While here an example based on power combining is given, it is possible to synthesize different functions on demand. (b) The choice of CNN architecture is based on overall error performance. Average prediction error of S-Matrix terms across frequency for the CNN architecture of choice. The comparison is conducted between roughly ≈10K HFSS simulations and CNN predictions. The average error for this test does not exceed ≈0.28.

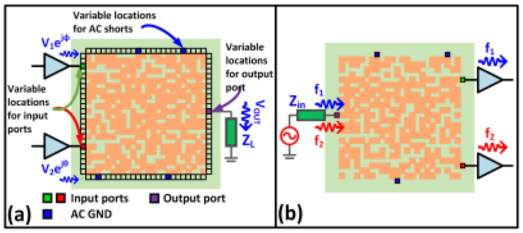

Fig. 2. (a) Design space consists of pixelated structures where ports and shorting pins can be placed along the edges. In this paper, we consider design of power combiners/dividers which are useful for multistage mm-wave designs such as power amplifiers. (b) It is also possible to use the same predictor model for multiple design scenarios such as diplexers.

Authorized licensed use limited to: Princeton University. Downloaded on April 25,2025 at 18:29:43 UTC from IEEE Xplore. Restrictions apply.

Here, we suggest a deep neural network-based EM emulator to replace simulations by rapidly predicting S-Parameters over a broad range of frequencies. With this modelling approach, the design process can be accelerated significantly. *Motivated by this, this work proposes a generalized deep learning enabled inverse design method to synthesize multiport EM structures on demand.* Section II discusses the deep learning-based modelling, inverse synthesis flow, and design examples. Section III demonstrates the application of this method to mm-wave systems as well as measurement results.

II. DEEP LEARNING BASED INVERSE DESIGN OF MM-WAVE PASSIVE NETWORKS

A. *Deep Learning Based Inverse Design Method*

According to the Universal Approximation Theorem, a neural network (NN) with at least one arbitrarily large hidden layer can learn any function [11]. While this is an empowering theoretical result, in practice, the ability of NNs would be limited with training data and computation. In this aspect, instead of relying on a vanilla NN, it is wiser to consider different architectures and machine learning (ML) tools. For the modelling problem we are interested in, a convolutional neural network (CNN) would be preferable for a number of reasons. Firstly, image-like representation of pixelated structures lends itself conveniently. It is worth noting that CNNs can outperform other ML tools for many image processing problems [11]. Secondly, we can draw some analogy between CNN operation and EM simulations, since in both cases localized interactions are prominent.

As illustrated in Fig. 1-a, CNN based EM emulator/predictor is the enabling component of algorithmic design method we are utilizing. The accuracy of this CNN is therefore critical. Main factors we identified for the prediction performance is large training dataset and optimal architecture.

In this work we consider 3-port EM structures that are discretized over a 25×25 grid within a 400×400 μm box, as shown in Fig. 2-a. A dense discretization consisting of square pixels would be able to approximate almost any structure that could be implemented in the given area and chip layer. Moreover, compared to circular pixels, rectangular shapes can pass design rule checks more easily. Based on these, we use square pixels. The bounding box size is the result of the tradeoff between the lowest operation frequency and chip area. While 400 μm is only ≈$\lambda/16$ at 24 GHz, using large chip area is generally undesirable. Therefore, the box size could be made a design parameter depending on the target frequency band and design goals. Ports and AC shorts can be arbitrarily placed on the edges, allowing great flexibility.

Once a predictor is trained it can be re-used for different synthesis runs, as illustrated in Fig. 2-b. Training dataset for this design space consists of simulation of randomly generated pixelated structures. Thus, it is desirable to reduce the computational burden required for these simulations. In this aspect, we adopt 2 step transfer learning [12]. The initial set of simulations were conducted with coarse mesh settings, hence large number of simulations are more easily achievable. Roughly 236K simulation results were divided into train-validation-test sets of 200-18-18K. A randomly initialized CNN is trained with this simulation data. At the second step, more accurate simulations with fine mesh were used to tune this initial model. Moreover, data augmentation based on geometric transformations is applied to training sets to further alleviate the problem [13]. Roughly 150K simulation results were augmented to ≈250K and divided into train-validation-test sets of 210-20-20K for this step. We observed that without conducting the initial training, the second set of simulations did not suffice for a good predictor.

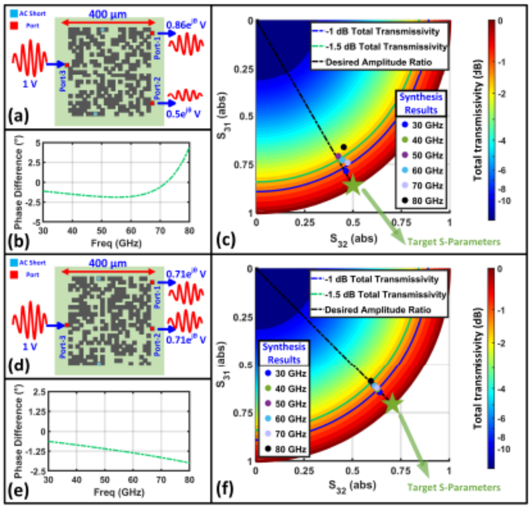

Fig. 3. Design examples for different set of S-Parameter targets. Once the synthesis is completed, EM structures were simulated in HFSS. (a) Broadband un-equal amplitude splitter designed with deep learning assisted inverse synthesis method. (b) Phase difference between 2 output ports is close to the 0° target. (c) Ideal response for the given unequal splitting ratio (-6 dB – -1.3 dB) lies on the 0 dB total transmissivity circle. Considering the mismatch and insertion losses, realistic response should lay over the black dotted line, which indicates perfect amplitude balance. Moreover, total transmissivity should be maximized. This example achieves greater than -1.6 dB total transmissivity between 24-72 GHz. (d) Broadband equal amplitude splitter (-3 dB – -3 dB). (e) Phase difference between 2 output ports is close to 0°. (f) Equal split is achieved with almost perfect amplitude balance and greater than -1.5 dB total transmissivity between 25-79 GHz.

As the training framework has been laid out, architecture optimization for CNN layers could be carried out. As opposed to an exhaustive search, several layer parameters and properties were fixed according to established guidelines [1], [11]. An important rule of thumb is that a deeper network is preferable over a wider network [11]. In parallel to this, we deployed more than 10 convolutional layers and 4-6 FC layers in all candidate architectures. Each convolutional and FC layer was followed by batch normalization and leaky ReLU operation [1]. In addition to these, FC layers are appended with a dropout layer after ReLU operation [1]. Convolutional layers have a width of 64 filters while FC layers contain 1000 neurons, except for the output layer, whose width should match the number of terms in the S-Matrix. Here 9 frequency points between 24-80 GHz were predicted. As passive structures are reciprocal, it is sufficient to consider only a part of the S-Parameters (6 entries per frequency). Since each entry in

Authorized licensed use limited to: Princeton University. Downloaded on April 25,2025 at 18:29:43 UTC from IEEE Xplore. Restrictions apply.

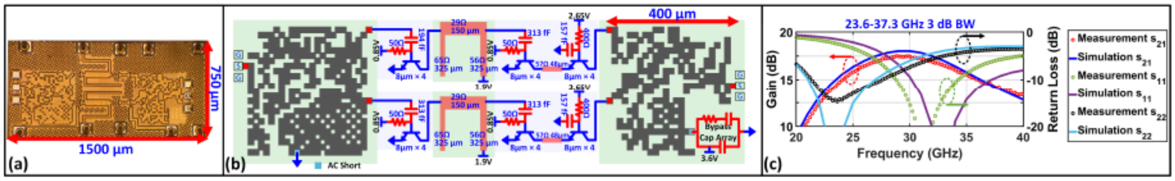

Fig. 4. (a) Die photo. (b) Circuit schematic. (c) Measured and simulated S-Parameters, showing a 3 dB BW of 23.6-37.3 GHz with a peak gain of 17.5 dB.

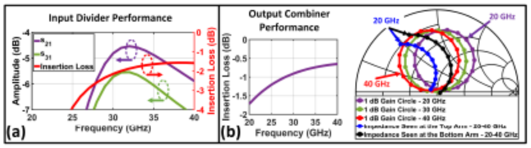

Fig. 5. Input divider performance. Power delivered to top and bottom arms of the input stage was denoted with s21 and s31. (b) Output combiner performance. Low loss output combiner provides close to optimum impedance across frequency to both top and bottom arms (black and blue contours). This was demonstrated with 1 dB gain circles.

the S-Matrix consists of real and imaginary parts, output layer has 108 neurons. Filter sizes for convolutional layers decrease as we go deeper from the input. A coarse architecture optimization was conducted by adjusting the number of FC layers and starting filter size. Finally, we choose 11 convolutional layers with 6 FC layers. Filter sizes vary between 12×12 to 3×3. Error over S-Matrix components for this choice is illustrated in Fig. 1-b. Here error terms are expressed as the average of $|S_{MN_{Sim}} - S_{MN_{Pred}}|$. The average error for different S-Matrix terms, do not exceed ≈0.28.

B. Design Examples

As design examples, un-equal and equal power splitters for 50Ω ports were optimized over a wide frequency range in Fig. 3. Following cost function was minimized for synthesis:

$$C = \sum_{i=1}^{M} a_1(i) \times \left| s_{32}^{\sim}(i) - |s_{32}^{\diamond}(i)| \right| + a_2(i) \times \left| s_{31}^{\sim}(i) - |s_{31}^{\diamond}(i)| \right| + a_3(i) \times \left| \mathrm{Arg}\left(\frac{s_{31}^{\diamond}(i)}{s_{32}^{\diamond}(i)}\right) - \varphi(i) \right| + a_4(i) \times \left| \frac{s_{32}^{\sim}(i)}{s_{31}^{\sim}(i)} \left| \frac{s_{31}^{\diamond}(i)}{s_{32}^{\diamond}(i)} \right| - \frac{s_{31}^{\sim}(i)}{s_{32}^{\sim}(i)} \left| \frac{s_{32}^{\diamond}(i))}{s_{31}^{\diamond}(i)} \right| \right| \quad (1)$$

In this equation, summation is performed over discrete frequency points where S-Parameters are predicted by CNN. $s_{32}^{\sim}(i)$ and $s_{31}^{\sim}(i)$ indicates targeted S-Parameter amplitude across frequency bins. $s_{32}^{\diamond}(i)$ and $s_{31}^{\diamond}(i)$ are CNN predicted complex valued S-Parameters. $\varphi(i)$ denotes the target absolute phase difference between $s_{32}^{\diamond}(i)$ and $s_{31}^{\diamond}(i)$. $a_1(i) \ldots a_4(i)$ are frequency dependent weighting coefficients. While the first 3 terms in this equation are easy to understand, we observed that adding the last term enhances the amplitude balance. As an example, in Fig. 3-a, $s_{32}^{\sim}(i) = 0.86$, $s_{31}^{\sim}(i) = 0.5$ and $\varphi(i) = 0$ for the frequency range of interest.

In Fig. 3, genetic algorithm was used as the optimization tool. A population (size of 4096) evolved over 100 generations. While 8 best performing population members were directly transferred from generation N to N+1, the rest of the population was obtained through crossing over between 2 parents and random mutations. Parent selections were performed repeatedly for each member of the generation N+1. To select a pair of parents, a subset of generation N was randomly sampled (256 out of 4096) and 2 best performing samples were picked. After the crossing over, random mutations were performed. It is worth noting that processing each generation takes around ≈3 seconds including S-Parameter prediction. As a result, the optimization algorithm emulates simulation of ≈400K EM structures within 5 minutes. This exploratory strategy and large amounts of samples could help push boundaries for mm-wave designs. As we will show in the next section, it is possible to use this method for the design of active circuits.

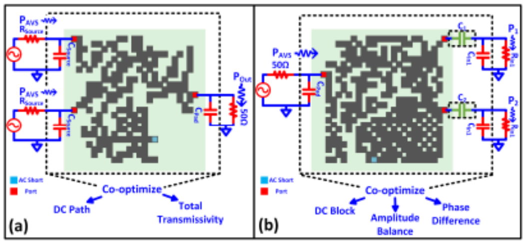

Fig. 6. (a) Evaluation of population members for output combiner synthesis. Once all the parasitic capacitances are lumped into the pixelated structure, transmissivity from real source impedances to 50Ω output load was calculated. (b) Evaluation of population members for input divider synthesis. A pair of series capacitance is optimized along with pixelated structure. To find the optimal capacitance values, predicted S-Parameters are cascaded with a set of capacitance combinations, and cost values are found for each combination. Minimum cost value is then assigned to population member that is being evaluated. The cost calculation for this sweep is performed by lumping the parasitic capacitances into the EM structure.

III. MM-WAVE AMPLIFIER IMPLEMENTATION

To showcase the application of this method for mm-wave systems, we designed a 2-stage mm-wave amplifier that targets the FR-2 spectrum. Chip was designed with Global Foundries 90 nm SiGe BiCMOS process. Fig. 4-a shows the die photo. Detailed schematic of the circuit is available in Fig. 4-b. While high performance BJTs in this process have a BV_{CEO} of ≈1.7 V, presence of a base resistor (BV_{CER}) allows us to use V_{CE} voltages around 1.8-1.9 V. As it can be seen, we rely on an inversely designed splitter (400×400 µm) to divide the input signal into 2 arms while simultaneously performing input matching. Fig. 5-a illustrates input matching network (IMN) performance. Within the operation band, amplitude imbalance is less than 1.3 dB while insertion loss varies between 1.5-4 dB. Interstage MN was designed with π-section

Authorized licensed use limited to: Princeton University. Downloaded on April 25,2025 at 18:29:43 UTC from IEEE Xplore. Restrictions apply.

transmission line stubs. The output stage consists of 2 common emitter stacks that are optimized for the frequency band of interest. Here, output combiner (400×400 μm) provides close to optimum impedance across 24-40 GHz, as shown in Fig. 5-b, while the insertion loss is better than -1.2 dB.

A. *Design of Output Matching Network*

Once output stage stacks have been designed, desired impedance target for high gain could be decomposed into a parallel RC circuit, where R and C could be frequency dependent. For the ease of analysis, all the parasitic capacitances (including GSG pad capacitance) are lumped to the CNN predicted S-Parameters of the pixelated structure during the cost function evaluation for each of the population members. Under this condition, total transmissivity from 2 input ports to output was calculated, assuming optimal phase and equal amplitude excitation [14]. Transmissivity values that are below -1 dB were penalized between 24-40 GHz. As a side objective, connection between AC short (implemented with a bypass capacitor array as V_{CC} feed) location and input ports were checked. If no direct connection was found, extra penalty incurred. Fig. 6-a illustrates these.

B. *Design of Input Matching Network*

For the IMN, it is desirable to ensure amplitude symmetry and optimal phase difference between 2 outputs of the networks. Moreover, a series capacitor must be present to provide DC isolation. In parallel to these, a sweep over series capacitor values were performed for every member of the population to calculate cost function. For a given series capacitor pair, parasitic capacitances could be lumped into the IMN and transmissivity from input to output ports could be calculated using generalized S-Parameters. Once these transmissivity values are found, the cost function introduced in Section-II was used to perform optimization. Weighting coefficients over the frequency were adjusted to ensure flat frequency response for the full chain. Fig. 6-b illustrates these.

C. *Measurement Results*

In parallel to design specification, a 3 dB bandwidth of 23.6-37.3 GHz was achieved with a peak gain of 17.5 dB (Fig. 4-c). Good agreement is observed when small signal measurement is compared to full die EM simulation (Fig. 4-c). 67 GHz Z-Probes were used for on wafer measurement with an Anritsu VNA (MS4647B). Measured stability metrics μ and μ' are above 1 within the sweep range.

IV. CONCLUSION

We demonstrate a deep learning assisted design method for multiport EM structure synthesis. Once the initial training is completed, EM emulator could be reused for the synthesis of different design targets. Enabled by the fast EM emulator that can accurately predict multiport S-Parameters, we show different design examples that are product of rapid synthesis. As proof of concept, a mm-wave amplifier was designed in 90 nm SiGe BiCMOS. This approach could be used in power combining or backoff efficient power amplifier designs.

ACKNOWLEDGMENT

We would like to acknowledge the Defense Advanced Research Program Agency (DARPA), Army Research Office (W911NF2110314), Air Force Office of Scientific Research (FA9550-22-1-0455, FA9550-23-1-0176), and Office of Naval Research (N00014-22-1-2302, N00014-23-1-2332, N00014-23-1-2592) for funding. The authors also acknowledge GlobalFoundries for fabrication support, and all members of IMRL for technical discussions.

REFERENCES

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

[2] S. Er, E. Liu, M. Chen, Y. Li, Y. Liu, T. Zhao, and H. Wang, "Deep learning assisted end-to-end synthesis of mm-wave passive networks with 3d em structures: A study on a transformer-based matching network," in *2021 IEEE MTT-S International Microwave Symposium (IMS)*, 2021, pp. 66–69.

[3] H. T. Nguyen and A. F. Peterson, "Machine learning for automating the design of millimeter-wave baluns," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 68, no. 6, pp. 2329–2340, 2021.

[4] F. Passos, N. Lourenço, E. Roca, R. Martins, R. Castro-López, N. Horta, and F. V. Fernández, "Pacosyt: A passive component synthesis tool based on machine learning and tailored modeling strategies towards optimal rf and mm-wave circuit designs," *IEEE Journal of Microwaves*, vol. 3, no. 2, pp. 599–613, 2023.

[5] J. Lee, W. Jia, B. S. Rodriguez, and J. S. Walling, "Pixelated rf: Random metasurface based electromagnetic filters," in *2023 21st IEEE Interregional NEWCAS Conference (NEWCAS)*, 2023, pp. 1–5.

[6] E. A. Karahan, A. Gupta, U. K. Khankhoje, and K. Sengupta, "Deep learning based modeling and inverse design for arbitrary planar antenna structures at rf and millimeter-wave," in *2022 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting (AP-S/URSI)*, 2022, pp. 499–500.

[7] A. Gupta, E. A. Karahan, C. Bhat, K. Sengupta, and U. K. Khankhoje, "Tandem neural network based design of multiband antennas," *IEEE Transactions on Antennas and Propagation*, vol. 71, no. 8, pp. 6308–6317, 2023.

[8] P. Terway and N. K. Jha, "Inform: Inverse design methodology for constrained multiobjective optimization," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 42, no. 7, pp. 2200–2213, 2023.

[9] S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, and A. W. Rodriguez, "Inverse design in nanophotonics," *Nature Photonics*, vol. 12, no. 11, p. 659–670, 2018.

[10] S. So, T. Badloe, J. Noh, J. Bravo-Abad, and J. Rho, "Deep learning enabled inverse design in nanophotonics," *Nanophotonics*, vol. 9, no. 5, pp. 1041–1057, 2020.

[11] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, http://www.deeplearningbook.org.

[12] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, "A survey on deep transfer learning," 2018.

[13] C. Shorten and T. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, no. 60, 2019.

[14] C. R. Chappidi and K. Sengupta, "Frequency reconfigurable mm-wave power amplifier with active impedance synthesis in an asymmetrical non-isolated combiner: Analysis and design," *IEEE Journal of Solid-State Circuits*, vol. 52, no. 8, pp. 1990–2008, 2017.

Authorized licensed use limited to: Princeton University. Downloaded on April 25,2025 at 18:29:43 UTC from IEEE Xplore. Restrictions apply.