Development of Knowledge Based Artificial Neural Network Models for Microwave Components

*[+]P.M. Watson, *K.C. Gupta, and *R.L. Mahajan

* Center for Advanced Manufacturing and Packaging of Microwave,
Optical, and Digital Electronics (CAMPmode)
Unviersity of Colorado at Boulder
Boulder, CO 80309

[+] AFRL/SNDM
Sensors Directorate
Wright-Patterson Air Force Base, OH 45433

Abstract

Artificial neural networks (ANNs) provide fast and accurate models for microwave modeling, simulation, and optimization. This paper addresses the use of prior knowledge (or existing models) for reducing the complexity of the input/output relationships that an ANN has to learn. This reduction of input/output complexity allows an accurate ANN model to be developed with less training data, which is very advantageous when training data is expensive/ time-consuming to obtain, such as with EM simulation. Two simple methods of incorporating prior knowledge into ANN training are demonstrated and compared: the difference method and the prior knowledge input (PKI) method. As an example, a 2-port microstrip via model has been developed by using a closed-form expression for the via's inductance as prior knowledge.

I. Introduction

Accurate and efficient models for circuit components are essential for cost-effective circuit design. Models are generally developed using analytical, electromagnetic simulation, and/or measurement based methods. In recent years, empirical models for microwave components based on artificial neural networks (ANNs) have received much attention [1-3] as an alternative to standard empirical modeling techniques, such as polynomial fitting and look-up tables. ANN models provide a general framework for modeling complex input/output mappings between multiple inputs and outputs. ANN models can be much faster than original EM models, more accurate than polynomial fitted and other empirical models, allow more input dimensions than look-up table models and are easier to develop when a new component/technology is introduced [4].

A potential drawback of ANN modeling is the amount of training data that needs to be provided in order to obtain an accurate model. Training data must be provided to characterize the component to be modeled over a desired range of operation and for different combinations of geometrical and physical model inputs. The difficulty arises when training data is expensive or difficult to obtain. An approach to reducing this data is through reducing the complexity of the input/output mapping that must be learned by the ANN. To this end we propose using prior knowledge (existing models) about the component to be modeled. Prior knowledge, for example, can be in the form of analytical equations, empirical models, or already trained ANN models. These existing models are models which contain information about the component to be modeled but do not give the required accuracy over the desired range of operation.

III. Use of Prior Knowledge for ANN Modeling

For a chemical vapor deposition in a horizontal reactor, Marwah [5] and Marwah and Mahajan [6] propose using different model modification techniques to convert a previously trained physical neural network model (called the source model) to a model suitable for a modified processing environment (called the target model). Three different techniques, namely the difference method, the

0-7803-4471-5/98/$10.00 (c) 1998 IEEE

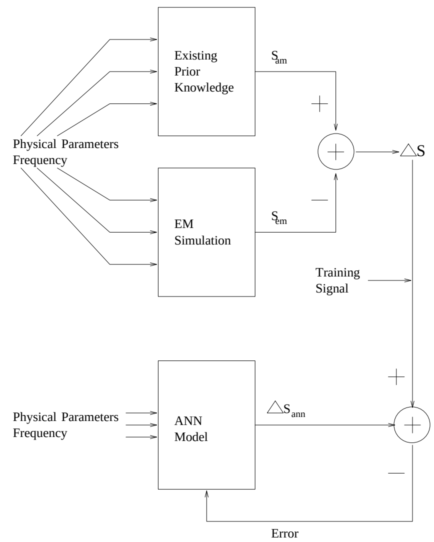
source weights method, and the prior knowledge input (PKI) method are evaluated. The PKI method was shown to out perform the other two methods.

To gain better insight into the dynamics of the above mentioned model modifier techniques, the output/input behavior was monitored during the training process. It was noted that in the difference method, the difference between the source and the target was not a simpler function of the inputs as compared to the mapping directly from the source to the target. As a result, no benefit was expected to result from this modifier approach. This was supported by the training results, which showed that the percent relative error on the target points was the same as that obtained by training the target model without the help of the source model. The source weight technique resulted in a similar performance as the difference method. With the PKI method, on the other hand, the source function converged towards the target function continuously as the training proceeded. Trained on one-fourths of the points used for the source model, the target model achieved the same accuracy as the source model. Similar techniques have not been investigated earlier for their performance in modeling of microwave components.

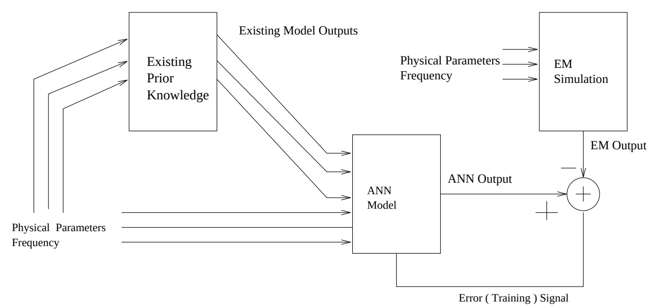
In this paper, we present and compare two of the three techniques mentioned above for incorporating prior knowledge (or existing models) into ANN model development using EM simulation. These are the difference method and the PKI method. In the difference method, the ANN is trained on the difference between the EM simulation output and the existing model (source model) output, shown in Fig. 1a. This method is expected to give good results when the difference has a simpler input/output mapping as a function of the inputs than the target data. A simpler input/output mapping requires less training data to characterize. For the PKI method, the source model outputs are used as inputs for the ANN model in addition to the other inputs, shown in Fig. 1b. In this case, the input/output mapping that must be learned by the ANN is that between the output response of the existing model and that of the target model. For the case when the target outputs are the same as the existing model outputs, the learning problem is reduced to a one-to-one mapping. Note that conventional two-layer neural networks along with backpropagation training [7] are used with both the difference and PKI methods, which is advantageous for a user.

III. Two-Port GaAs Microstrip Ground Via

The geometry of the two-port broadband GaAs microstrip via is shown in Fig. 2. The height of the substrate, the dielectric constant, and all loss parameters are considered constant for this example. Frequency, the



(a)



(b)

Fig. 1 Schemes for using prior knowledge for artificial neural network training, (a) difference method and (b) PKI method.

width of the incoming microstrip lines, W_l, the side of the square shaped via pad, W_p, and the diameter of the via hole to ground, D_{via}, are the variable input parameters for the EM-ANN model. Input variable ranges are given in Table 1. Model outputs are the magnitudes and phases of S_{11} and S_{21}.

An existing model in equation form for the inductance of a microstrip grounding via is given in [8]. The existing model was found to give reasonable results at lower frequencies (<15 GHz), but as frequency increased, errors between the model and EM simulation also increased. Inaccuracies of the model, especially at higher frequencies may be due to pad inductance, pad capacitance, discontinuity effects, and radiation from the via-hole [9,10].

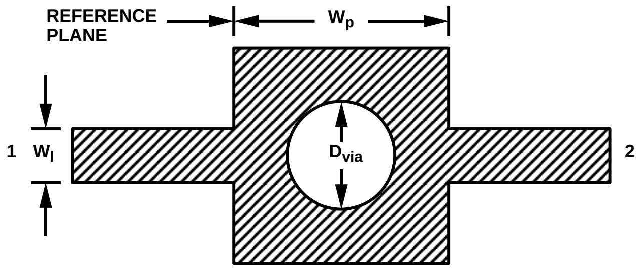


Fig. 2 Two-port GaAs microstrip grounding via. Substrate thickness = 4 mil, ε_r=12.9, tanδ=0.002, σ_{metal}=4.1x10^7, and t_{metal} = 0.1 mil.

Table 1 Variable input parameters for GaAs microstrip ground via modeling.

Input Parameter	Minimum Value	Maximum Value
Frequency	5 GHz	55 GHz
W_l/W_p	0.3	1.0
D_{via}/W_p	0.2	0.8
W_l/H_{sub}	0.1	2.0

EM simulations were performed from 5 GHz to 55 GHz in 10 GHZ steps on 45 via structures within the ranges given in Table 1, generating 270 input/output vectors (termed examples throughout the rest of this paper). The via structures simulated to provide training data were chosen using design of experiments (DOE) central composite techniques [11]. Originally, fifteen vias were simulated for training, 14 for simultaneous testing or additional training, and 16 for verification. However, with the use of prior knowledge, it was found that less training data is sufficient. Therefore, in the following model development, more of the simulated data has been used for simultaneous testing.

Initial model development used only 7 via structures (42 examples) for training, 22 via structures (132 examples) for testing, and 16 via structures (96 examples) for model verification. EM-ANN models were developed using regular training methods (no use of existing knowledge), the difference method, and the PKI method. The training procedure used for EM-ANN model development has previously been published in [1,2].

Model average error and standard deviation are shown in Table 2, Table 3, and Table 4 for regular training, difference training, and PKI training, respectively. Looking at verification dataset errors, both the difference method and the PKI method provide more accurate models than using regular training methods. In addition, the PKI method provides better accuracy for $|S_{11}|$ and comparable accuracy on other parameters.

To further demonstrate the advantages of incorporating existing knowledge into training, EM-ANN models were developed using 15 training vias (90 examples), 14 test structures (84 examples), and again 16 verification structures (96 examples). Model average error and standard deviation are shown in Table 5, Table 6, and Table 7 for regular training, difference training, and PKI training, respectively.

With more training data, error results for regular training improve. However, verification dataset error results using the difference method and PKI method improve also and still provide better accuracy than regular training. What is more important is that when comparing verification dataset errors, the accuracy of the models trained with only 7 via structures using the difference method and PKI method show comparable or better accuracy than the model developed using 15 training via structures and regular training. In other words, when existing knowledge is used for model development, fewer EM simulations are needed for a required model accuracy.

IV. Concluding Remarks

Use of prior knowledge (existing models) has been shown to reduce the amount of training data needed for ANN model development. This is particularly useful when input data is expensive/time-consuming to obtain. Two simple methods have been demonstrated for incorporation of prior knowledge into ANN training: the difference method and the PKI method. Both methods are simple to implement and are applicable to standard 2 layer networks using the error backpropagation training algorithm, which has been studied extensively.

Both the difference method and the PKI method show increased accuracy over regular training methods using no prior knowledge. Also, the PKI method shows slightly better accuracy than the difference method for the examples considered.

References

[1] P.M. Watson and K.C. Gupta, "EM-ANN Models for Microstrip Vias and Interconnects in Multilayer Circuits," *IEEE Trans. on Microwave Theory and Tech.*, Vol. 44, No. 12, Dec. 1996, pp. 2495-2503.

[2] P.M. Watson and K.C. Gupta, "Design and Optimizaton of CPW Circuits Using EM-ANN Models for CPW Components," *IEEE Trans. on Microwave Theory and Tech.*, Vol. 45, No. 12, Dec. 1997, pp. 2515-2523.

[3] G.L. Creech, et al., "Artificial Neural Networks for Fast and Accurate EM-CAD of Microwave Circuits," *IEEE Trans. on Microwave Theory and Tech.*, Vol. 45, No. 5, May 1997, pp. 794-802.

[4] Q.J. Zhang, F. Wang, and M.S. Nakhla, "Optimization of High-Speed VLSI Interconnects: A review," *Int. J. of Microwave and Millimeter-Wave CAD*, Vol. 7, pp. 83-107, 1997.

[5] M. Marwah, "Neural Network Modeling Techniques for Electronics Manufacturing Processes," M.S. Thesis, U. of Colorado, Boulder, CO, 1996.

[6] M. Marwah and R.L. Mahajan, "Building Equipment Models Using Neural Network Models and Model Transfer Techniques," submitted to *J. Semiconductor Manufacturing*, 1997.

[7] C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press Inc., New York, 1996.

[8] M.E. Goldfarb and R.A Pucel, "Modeling Via Hole Grounds in Microstrip,"*IEEE Microwave and Guided Wave Letters*, Vol. 1, No. 6, pp. 135-137, June 1991.

[9] V.K. Sadhir, I.J. Bahl, and D.A. Willems, "CAD Compatible Accurate Model of Microwave Passive Lumped Elements for MMIC Applications," *Int. J. MIMICAE*, Vol. 4, NO. 2, pp. 148-162, 1994.

[10] G. Cerri, M. Mongiardo, and T. Rozzi, "Radiation from Via-Hole Grounds in Microstrip Lines," *MTT-S Int. Microwave Symp. Dig.*, 1994, pp. 341-344.

[11] D.C. Montgomery, *Design and Analysis of Experiments*, John Wiley and Sons, Inc., 1991.

Table 2 Error results for the 2-port microstrip via, **regular training**. (7 train structures, 4 inputs, 4 outputs, 5 hidden neurons, 49 weights)

	$\|S_{11}\|$	$\angle S_{11}$ (°)	$\|S_{21}\|$	$\angle S_{21}$ (°)
Train/test				
Average error	0.0076	2.000	0.0314	2.575
Standard dev.	0.0114	2.275	0.0389	3.000
Verification				
Average error	0.0066	1.677	0.0226	2.104
Standard dev.	0.0085	1.831	0.0244	2.197

Table 3 Error results for the 2-port microstrip via, **difference method**. (7 train structures, 4 inputs, 4 outputs, 8 hidden neurons, 76 weights)

	$\|S_{11}\|$	$\angle S_{11}$ (°)	$\|S_{21}\|$	$\angle S_{21}$ (°)
Train/test				
Average error	0.0042	1.313	0.0094	2.084
Standard dev.	0.0058	1.803	0.0089	3.174
Verification				
Average error	0.0041	0.941	0.0083	1.477
Standard dev.	0.0043	0.908	0.0061	1.387

Table 4 Error results for the 2-port microstrip via, **PKI method**. (7 train structures, 8 inputs, 4 outputs, 5 hidden neurons, 69 weights)

	$\|S_{11}\|$	$\angle S_{11}$ (°)	$\|S_{21}\|$	$\angle S_{21}$ (°)
Train/test				
Average error	0.0035	1.209	0.0092	1.587
Standard dev.	0.0087	1.780	0.0122	1.890
Verification				
Average error	0.0023	0.949	0.0066	1.526
Standard dev.	0.0023	1.011	0.0077	1.426

Table 5 Error rsults for the 2-port microstrip via, **regular training**. (15 train structures, 4 inputs, 4 outputs, 13 hidden neurons, 121 weights)

	$\|S_{11}\|$	$\angle S_{11}$ (°)	$\|S_{21}\|$	$\angle S_{21}$ (°)
Train/test				
Average error	0.0020	0.528	0.0065	0.620
Standard dev.	0.0023	0.448	0.0059	0.544
Verification				
Average error	0.0041	0.714	0.0101	1.061
Standard dev.	0.0049	0.504	0.0089	0.929

Table 6 Error results for the 2-port microstrip via, **difference method**. (15 train structures, 4 inputs, 4 outputs, 12 hidden neurons, 112 weights)

	$\|S_{11}\|$	$\angle S_{11}$ (°)	$\|S_{21}\|$	$\angle S_{21}$ (°)
Train/test				
Average error	0.0013	0.628	0.0036	0.731
Standard dev.	0.0014	0.502	0.0035	0.526
Verification				
Average error	0.0026	0.709	0.0047	0.983
Standard dev.	0.0032	0.524	0.0038	0.839

Table 7 Error results for the 2-port microstrip via, **PKI method**. (15 train structures, 8 inputs, 4 outputs, 11 hidden neurons, 147 weights)

	$\|S_{11}\|$	$\angle S_{11}$ (°)	$\|S_{21}\|$	$\angle S_{21}$ (°)
Train/test				
Average error	0.0017	0.538	0.0032	0.662
Standard dev.	0.0014	0.563	0.0026	0.742
Verification				
Average error	0.0021	0.782	0.0038	1.087
Standard dev.	0.0024	0.604	0.0026	0.947