Behavioral Modeling of a C-Band Ring Hybrid Coupler Using Artificial Neural Networks

Erdem DEMIRCIOGLU [1], *Murat H. SAZLI* [2]

[1] R&D – Satellite Design Department, Turksat AS, Golbasi, 06830, Ankara
[2] Electronics Engineering Department, Ankara University, Tandogan, 06100, Ankara

edemircioglu@turksat.com.tr, sazli@eng.ankara.edu.tr

Abstract. *Artificial Neural Networks (ANNs) gained importance on the RF microwave (MW) design area and behavioral modeling of MW components in the past few decades. This paper presents a cost effective neural network (NN) approach to overcome design, modeling and optimization problems of an 180° ring hybrid coupler operating in C-Band. The proposed NN model is trained by data sets obtained from electromagnetic (EM) simulators and neural test results are compared with simulator findings to determine the network accuracy. Moreover, necessary trade-offs are applied to improve the networks' performance. Finally correlation factors, which are defined as comparison criteria between EM-simulator and proposed neural models, are calculated for each trade-off case.*

Keywords

Artificial Neural Networks (ANNs), Ring Hybrid Coupler (RHC), C-band, optimization, behavioral modeling.

1. Introduction

EM-simulators are the design and optimization tools which include specific algorithms and mathematical methods dedicated to solve RF circuit problems. In recent decades, an ANN, knowledge-aided design (KAD), approach has been developed for the modeling and optimization of RF MW components. A neural model for a device or circuit can be established by using data sets which are acquired by measurement and simulation results, through a process called training. Once the network is trained, this network can be used for circuit design to provide instant answers for tasks it learned. Successful implementations of linear and nonlinear device turns neural applications into a research area for the modeling of various MW components [1, 2], and MW circuit design [3, 4, 5]. Recent works have shown that NN can accurately model components, such as microstrip interconnects [1, 3], vias [3, 5], spiral inductors [4], FET devices [1, 6], power transistors and power amplifiers [7], coplanar waveguide components [3], packaging and interconnects [8], microstrip circuit design [9], MW filter design [10], etc.

This paper discusses the neural modeling of an 180° RHC for efficient and robust behavioral estimations under necessary trade-offs. RHC design parameters are obtained by conformal mapping based approximation formulas. Then Ansoft Designer, which employs Method of Moments (MoM) as the EM problem solver, is used as EM-simulator to analyze and optimize the hybrid design. Subsequently EM-simulator data sets are utilized for establishing ANN. Finally neural modeling and simulator results are compared for each case. Both neural design and EM - ANN comparisons are simulated in MATLAB by a 2.33 GHz Intel Xenon x64 processor with 2.33 GHz, 4.00 GB RAM.

2. 180° Ring Hybrid Coupler Design

A 3 dB, 180° Ring Hybrid Coupler, also called as the "rat-race coupler", is a high-power capable, four-port device, optimized to sum two in-phase combined signals with essentially no loss or to equally split an input signal with no resultant phase difference between outputs and inputs. The fourth port is match terminated. A ring hybrid has many applications in RF microwave world such as mixers, phase shifters, amplifiers, etc. Fig. 1 demonstrates an 180° ring hybrid coupler where detailed equations and operation process can be found in the literature [11].

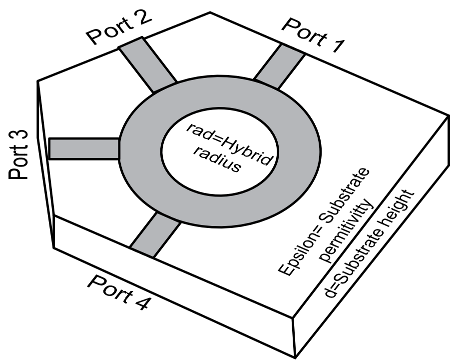

Fig. 1. 180° Ring hybrid coupler.

The simplest and most effective way of representing MW components' characteristic is the scattering parameters commonly known as S-parameters. The effective dielectric constant and characteristic impedance of a ring hybrid coupler, which is designed using the microstrip substrate, can be calculated using well-known conformal mapping based approximation methods such as Wheeler, Schneider or Hammerstad and Jensen [12, 13].

Wheeler Approximation: The analysis and synthesis equations are derived based on conformal mapping approximations of the dielectric boundary with parallel conductor strips separated by a dielectric sheet [14].

The characteristic line impedance for wide microstrip which has width over depth ratio of ($W/d > 3.3$) can be expressed as:

$$Z_L(W,d,\varepsilon_r)=\frac{\eta_O}{2\sqrt{\varepsilon_r}}\frac{1}{\frac{W}{2d}+\frac{1}{d}\ln 4+\frac{\varepsilon_r+1}{2\pi\varepsilon_r}\ln(\frac{\pi e}{2}(\frac{W}{2d}+0.94))+\frac{\varepsilon_r-1}{2\pi\varepsilon_r^2}\cdot\ln\frac{\pi^2 e}{16}} \quad (1)$$

The effective dielectric constant for wide strips which have width over depth ratio of ($W/d > 1.3$) is given as:

$$\varepsilon_{r_{eff}}=\varepsilon_r\left(\frac{E-D}{E}\right)^2 \quad (2)$$

where $$D=\frac{\varepsilon_r-1}{2\pi\varepsilon_r}\cdot\left(\ln\left(\frac{\pi e}{2}\left(\frac{W}{2d}+0.94\right)\right)-\frac{1}{\varepsilon_r}\ln\frac{e\pi^2}{16}\right) \quad (3)$$

and $$E=\frac{1}{2}\cdot\frac{W}{d}+\frac{1}{\pi}\cdot\ln(\pi e\frac{W}{d}+16.0547)\ . \quad (4)$$

The characteristic impedance for narrow microstrip with ($W/d \le 3.3$) can be stated as:

$$Z_L(W,d,\varepsilon_r)=\frac{\eta_O}{\pi\sqrt{2(\varepsilon_r+1)}}\cdot\left(\ln\left(\frac{4d}{W}+\sqrt{\left(\frac{4d}{W}\right)^2+2}\right)-\frac{1}{2}\cdot\frac{\varepsilon_r-1}{\varepsilon_r+1}\left(\ln\frac{\pi}{2}+\frac{1}{\varepsilon_r}\ln\frac{4}{\pi}\right)\right) \quad (5)$$

The effective dielectric constant for narrow microstrip with ($W/d \le 3$) depending on the characteristic impedance is given as following equation.

$$\varepsilon_{r_{eff}}=\frac{\varepsilon_r+1}{2}+\frac{\eta_O}{2\pi Z_L}\cdot\frac{\varepsilon_r-1}{2}\left(\ln\frac{\pi}{2}+\frac{1}{\varepsilon_r}\ln\frac{4}{\pi}\right). \quad (6)$$

The effective dielectric constant which is independent of the characteristic impedance, for narrow microstrip ($W/d \le 1.3$) is given as;

$$\varepsilon_{r_{eff}}=\frac{1+\varepsilon_r}{2}\left(\frac{A}{A-B}\right)^2 \quad (7)$$

where $$A=\ln\left(8\frac{d}{W}\right)+\frac{1}{32}(\frac{W}{d})^2 \quad (8)$$

and $$B=\frac{1}{2}\cdot\frac{\varepsilon_r-1}{\varepsilon_r+1}\left(\ln\frac{\pi}{2}+\frac{1}{\varepsilon_r}\ln\frac{4}{\pi}\right). \quad (9)$$

Schneider Approximation: In this approach, effective dielectric constant and the characteristic impedance formulas are obtained by rational function approximation with accuracy of ±2.5% for $0 \le W/d \le 10$ which is the range of importance for most engineering applications [15].

$$Z_L=\frac{\eta_O}{\sqrt{\varepsilon_{r_{eff}}}}\cdot\left\{\begin{array}{ll}\frac{1}{2}\cdot\ln\left(8\frac{d}{W}+\frac{W}{4d}\right) & ,for\ \frac{W}{d}\le 1\\ \frac{1}{\frac{W}{d}+2.42-0.44\cdot\frac{d}{W}+(1-\frac{d}{W})^6} & ,for\ \frac{W}{d}\ge 1\end{array}\right\}, \quad (10)$$

$$\varepsilon_{r_{eff}}=\frac{\varepsilon_r+1}{2}+\frac{\varepsilon_r-1}{2}\cdot\frac{1}{\sqrt{1+10\frac{d}{W}}}. \quad (11)$$

Hammerstad Approximation: The characteristic impedance and effective dielectric constant equations provides errors at least less than those caused by physical tolerances and is better than 0.01% for $W/d \le 1$ and 0.03% for $W/d \le 1000$ [16].

$$Z_L(W,d,\varepsilon_r)=\frac{Z_{L1}(W,d)}{\sqrt{\varepsilon_r}} \quad (12)$$

where $$Z_{L1}(W,d)=\frac{\eta_O}{2\pi}\cdot\ln\left(f_u\frac{d}{W}+\sqrt{1+\left(\frac{2d}{W}\right)^2}\right), \quad (13)$$

$$f_u=6+(2\pi-6)\cdot\exp(-(30.666\cdot\frac{d}{W})^{0.7528}), \quad (14)$$

$$\varepsilon_{r_{eff}}(W,d,\varepsilon_r)=\frac{\varepsilon_r+1}{2}+\frac{\varepsilon_r-1}{2}\left(1+10\frac{d}{W}\right)^{-ab}, \quad (15)$$

$$a(u)=1+\frac{1}{49}\ln(\frac{u^4+(u/52)^2}{u^4+0.432})+\frac{1}{18.7}\ln\left(1+(\frac{u}{18.1})^2\right), \quad (16)$$

$$b(\varepsilon_r)=0.564\cdot\left(\frac{\varepsilon_r-0.9}{\varepsilon_r+3}\right)^{0.053} \quad (17)$$

where $u = W/d$ and $\eta_0 = 120\pi$.

In this work, the mathematical approximations are employed to analyze the various transmission line characteristics and these analysis results are inserted in RHC synthesis process. Mathematically obtained design parameters are utilized and some tuning optimizations are applied through EM-simulator for accurate design. The simulations are repeated for a variety of substrate permittivity and height, ring radius and operating frequency values.

3. Artificial Neural Networks (ANNs)

ANNs are information processing systems, which are utilized to learn the input-output relationship characteristics of the device under consideration. The ANNs design is inspired from the human brain's ability to learn from observations. NNs must be first trained to model electrical behavior of linear and nonlinear complex components or systems. These trained neural models can be used to design, model and optimize the focused device by providing fast simulation answers compared with computationally loaded numerical solutions, or toughly obtained analytical and limited experimental results [17]. The ANN architectures and learning algorithms are the most important factors for developing neural models. The selected architecture and algorithm vary depending on the focused problem. The multilayer perceptron neural networks (MLPNNs) offer limited complexity and common approximation capabilities. Thus they are the most widely used NN architectures [18, 19]. In this study, MLPNN models with feed-forward network architectures and various training algorithms are utilized to solve the modeling problem of passive microwave devices, which are stripline and microstrip line type ring hybrid couplers.

3.1 Multilayer Perceptron Neural Networks (MLPNNs) Models

MLPNNs consist of an input layer, one or more hidden layers of computation nodes, and an output layer. The input signal propagates through the network in a forward direction, on a layer-by-layer basis. MLPNNs have been applied successfully to solve some difficult and diverse problems by training them in a supervised manner with a highly popular algorithm known as the error back-propagation algorithm [20]. Fig. 2 represents the MLPNN architecture.

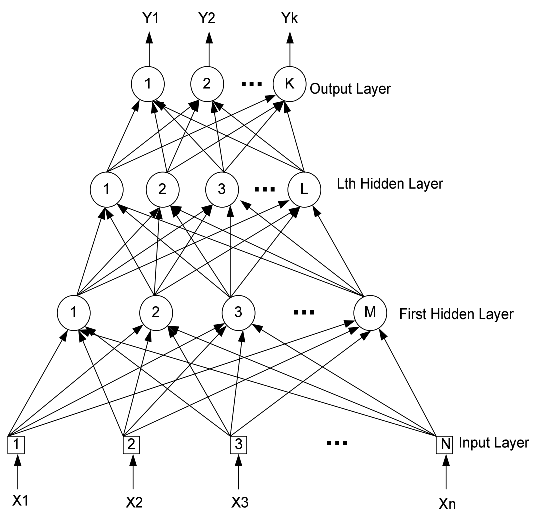

Fig. 2. MLPNN architecture.

Neural models can be constructed to estimate the linear and nonlinear devices' behavior. In this study, the generated ANN is applied to behavioral modeling of a linear ring hybrid coupler. Input parameters of the ring hybrid are defined as simulation frequency, ring hybrid radius, substrate height, and permittivity. Then scattering parameters such as S_{11}, S_{12}, S_{13}, and S_{14} present the device response. Sample data generated using EM simulations will be used in training process.

Once NN is trained, then the neural model can be used for predicting the output values corresponding to input variables. In NN testing stage, an independent set of input-output samples, called testing data which covers the whole definition space and equally distributed over the regression surface between training data, is used to test the neural model accuracy. When the network outputs are continuous functions of the inputs, modeling problem is known as regression or function approximation, which is the most common case in microwave design area.

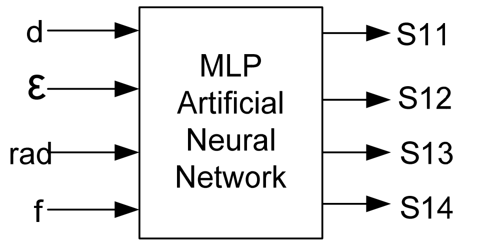

Fig. 3. ANN illustration of the discussed modeling problem.

Fig. 3 represents the proposed neural model for the microstrip line type ring hybrid coupler modeling problem, which is relatively complex compared with single input modeling problems. The network outputs are the ring coupler S-parameters S_{11}, S_{12}, S_{13}, S_{14}, which clearly demonstrate the fundamental characteristics of a symmetrical MW device. The model inputs are taken as dielectric substrate height d, dielectric substrate permittivity ε, simulation frequency range f, ring hybrid coupler radius *rad*.

3.2 Backpropagation Training Algorithms

Standard backpropagation is a gradient descent algorithm, as is the Widrow-Hoff learning rule, in which the network weights are moved along the negative of the gradient of the performance function. Properly trained backpropagation networks tend to give reasonable answers when presented with inputs that they have never seen. Typically, a new input leads to an output similar to the correct output for input vectors used in training that are similar to the new input being presented. This generalization property makes it possible to train a network on a representative set of input/target pairs and get good results without training the network on all possible input/output pairs. Once the network weights and biases are initialized, it is ready for training.

Levenberg-Marquardt (LM) training algorithm is a least-squares estimation method based on the maximum neighborhood idea. LM presents adequate characteristics in convergence time and ability to handle small networks. The preeminent aspects of Gauss-Newton technique and steepest-descent method are combined in this algorithm without many of their limitations. The error function can be expressed as below [21].

$$E(w) = \sum_{i=1}^{m} e_i^2(w) = \| g(w) \|^2 \tag{18}$$

and
$$e_i^2(w) = (y_{di} - y_i)^2 . \tag{19}$$

where $g(w)$ is the function containing individual error terms, y_{di} is the desired value of the output neuron i, and y_i is the actual output of that neuron. It is presumed that, $g(w)$ and its Jacobian $\boldsymbol{J}_g$ matrix are known at point w. Jacobian matrix contains the first derivatives of the network errors with respect to biases and weights. The weight vector $\mathbf{w}$ is calculated while the error function is minimized. The subsequent weight vector $\mathbf{w}_{k+1}$ can be derived from the preceding weight vector $\mathbf{w}_k$ as given below.

$$w_{k+1} = w_k + \delta w_k \tag{20}$$

and
$$\delta w_k = -(J_g^T g(w_k))(J_g^T J_g + \lambda I)^{-1} \tag{21}$$

where k is the number of iterations, $\boldsymbol{J}_g$ is the Jacobian matrix of $g(\mathbf{w}_k)$ which is computed by taking derivative of $g(\mathbf{w}_k)$ with respect to $\mathbf{w}_k$, λ is the Marquardt parameter and $\boldsymbol{I}$ is the identity matrix.

Conjugate Gradient of Polak-Ribière (CGP) training algorithm updates the weight and bias values according to conjugate gradient backpropagation proposed by Polak-Ribière, on condition that, network weights, inputs and transfer functions have derivative functions. In this algorithm, the line search is used to locate the minimum point and the search direction is computed from the new gradient in subsequent iterations. The search direction in each iteration can be determined by updating the weight vector [22].

$$w_{k+1} = w_k + \alpha p_k \tag{22}$$

where
$$p_k = -g_k + \beta_k p_{k-1} , \tag{23}$$

$$\beta_k = \frac{\Delta g_{k-1}^T g_k}{g_{k-1}^T g_{k-1}} \tag{24}$$

and
$$\Delta g_{k-1}^T = g_k^T - g_{k-1}^T . \tag{25}$$

Gradient Descent (GD) training algorithm is a first-order optimization tool which finds the local minimum of the function using gradient descent. This line search minimization procedure smoothen the descent direction in the steepest descent method. The weights and biases are updated in the direction of the negative gradient of the performance function [23].

Conjugate Gradient of Fletcher-Reeves (CGF) training algorithm updates weight and bias values by the Fletcher-Reeves conjugate gradient formulas [24]. This algorithm employs the norm squared of the previous gradient and the norm squared of the current gradient to perform the update procedure and evaluate the weight and bias values [24].

Scaled Conjugate Gradient (SCG) training algorithm combines the model thrust region approach used in LM. SCG offers avoiding time consuming line search process, in contrast to conventional conjugate scaled algorithms which require line search in all iterations. This line search is computationally expensive, because it requires that the network response to all training inputs to be computed several times for each search [25].

Resilient Propagation (RP) training algorithm provides faster convergence than other algorithms and eliminates harmful effects of the magnitudes of the partial derivatives. Then the RP algorithm determines the direction of the weight update by using the sign of the derivative and determines the size of the weight change by a separate update value. The magnitude of the derivative has no effect on the weight update [26].

4. Simulation Results

MLPNN architectures are trained with various back-propagation algorithms for stripline and microstrip line types of hybrid couplers' data. Then, neural model improvement is accomplished by altering training algorithms and network input parameters of substrate dielectric and height, coupler radius and frequency. Moreover target performance and epoch numbers of the training process are varied to find the optimum architecture. The accuracy and reliability of the generated networks are measured using the Pearson Product-Moment correlation coefficient γ which is selected as success criteria between NN and simulator results.

$$\gamma = \frac{\sum (x_i - x)(y_i - y)}{\sqrt{\sum (x_i - x)^2 (y_i - y)^2}} \tag{26}$$

where x_i is the simulator scattering parameter value, y_i is the MLPNN computed value, x is the simulated sample mean and y is the MLPNN computed sample mean.

The design parameters of the microstrip ring hybrid coupler is defined in the range of $2.33 \le \varepsilon_r \le 10.2$, $0.875\ \text{mm} \le d \le 1.578\ \text{mm}$, $5.7029\ \text{mm} \le rad \le 10.243\ \text{mm}$, $3\ \text{GHz} \le f_{\text{sim}} \le 8\ \text{GHz}$, $4.425\ \text{GHz} \le f_{\text{cen}} \le 6.775\ \text{GHz}$. Additionally, stripline type ring hybrid coupler has design parameters ranges of $2.33 \le \varepsilon_r \le 10.2$, $0.875\ \text{mm} \le d \le 1.578\ \text{mm}$, $3.797\ \text{mm} \le rad \le 11.295\ \text{mm}$, $3\ \text{GHz} \le f_{\text{sim}} \le 7\ \text{GHz}$, $4.16\ \text{GHz} \le f_{\text{cen}} \le 6.3\ \text{GHz}$. The simulation frequency sweep is defined by f_{sim} and center frequencies of hybrid couplers are given by f_{cen}.

Fig. 4 illustrates the instantaneous scattering results of the stripline type coupler modeled by a LM trained network which utilizes 1000 sample pairs in training and test processes where equal number of pairs is taken for both processes. Each sample pair includes S-parameters vs. equally distributed frequency points between 2-7 GHz under fixed stripline substrate properties such as *rad, d* and ε. Thus the NN model includes single input with four output neurons. Resonance frequency and 3dB coupling points can be monitored to determine the ring hybrid operating frequency band. The applied NN to solve the stripline example is a single input, multiple outputs network with one hidden layer. Adequate neurons are employed in the hidden layer to find optimum results. As shown in Tab. 1, the results obtained by NN are high enough to accurately model the stripline hybrid coupler.

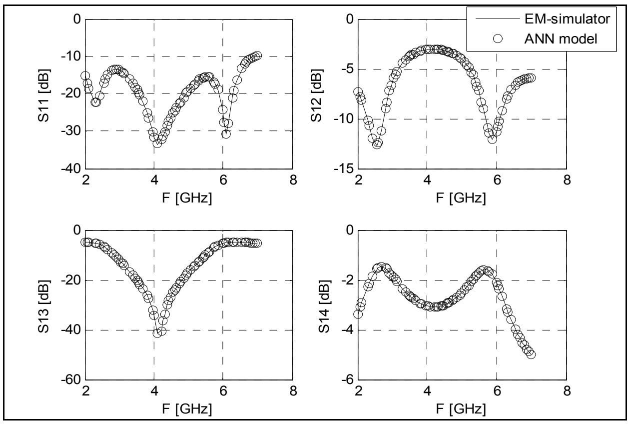

Fig. 4. Stripline type RHC (f_{cen} = 4.16 GHz) scattering parameters: EM simulator vs. NN model.

Scattering Parameters	Correlation factor for single hidden layer
S_{11}	0.9998
S_{12}	0.9999
S_{13}	0.9999
S_{14}	0.9999

Tab. 1. Correlation factors for stripline hybrid design.

This concludes that a single hidden layered NN architecture with Levenberg-Marquardt training algorithm is adequate to solve the single input and multiple outputs modeling problem as a consequence of input layer simplicity. Subsequent step in modeling problem will be more complicated for NN utilization. A microstrip line type RHC modeling is aimed to be solved by the neural model. This neural implementation can be used as generalized design, modeling and optimization tool, when broadened network input parameters such as hybrid radius, operating frequency and substrate properties are utilized. Multiple input variations adapt the proposed network to a modeling and optimization tool which can be used to find scattering values of any RHC design that falls into training range. Thus the necessary coarse and fine tunings for design optimization can be achieved by parametric neural model without having to redo the fullwave EM-simulator.

In microstrip case, four input parameters are altered in order to reach the optimum output results. Number of hidden layers and neurons in these layers are adjusted in accordance with applied learning algorithms to achieve the best test results. The overall data set includes 2500 data samples which cover the whole input parameters range. Test set contains half the data samples and these testing samples are equally distributed over the regression surface between the training samples. Neural models are tested with varying number of training and test sets. Initial trials have shown that as the data sets are increased proposed networks can handle modeling problem with higher accuracy. Then the number of hidden layers is enhanced and multi-hidden layered network is trained by the same training set until aimed results are acquired. All networks are tested with the same sets. According to the computed results, trade-offs such as correlation coefficients and computation time between single hidden layered and multi hidden layered networks are done. Variations on number of neurons in hidden layers, epoch number and target performance are also applied to obtain detailed trade-offs for NN applications. Training algorithms are compared by the criteria's of converging speed and accuracy, memory needs and correlation results.

Fig. 5 – 7 represent the neural modeling results of a microstrip RHC operating at center frequency of 4526 MHz. Networks comprising single to four hidden layers are analyzed to reach the finest scattering estimations. Two and three layered networks offer more adequate results though single and four hidden layered networks present underfitting and overfitting respectively. Constructed NNs are trained with the Levenberg–Marquardt algorithm which converges fast and provides accurate results. However it requires a lot of memory to run, thus additional NN simulations with other training algorithms are employed to find optimal training algorithm and improve NN estimations for the proposed modeling problem. Fig. 5 illustrates the return loss parameter of the modeled coupler.

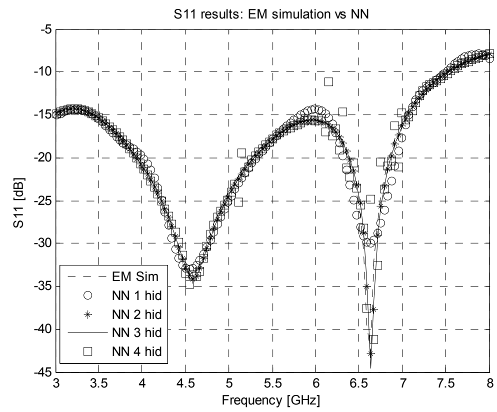

Fig. 5. S_{11} results: EM simulator vs. NN results for LM (f_{cen} = 4526 MHz).

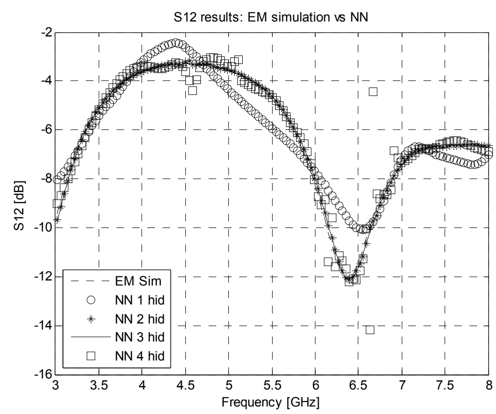

Fig. 6. S_{12} results: EM simulator vs. NN results for LM (f_{cen} = 4526 MHz).

Additionally, Fig. 6 and Fig. 7 represent the through port scattering parameters.

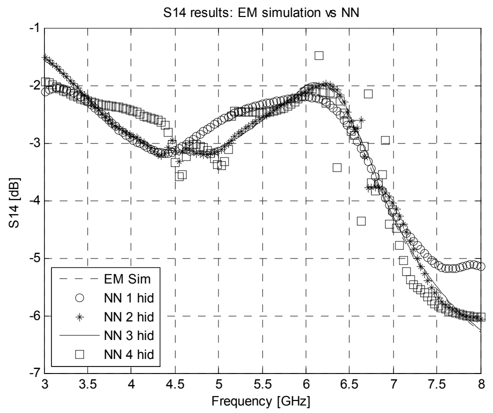

Fig. 7. S_{14} results: EM simulator vs. NN results for LM (f_{cen} = 4526 MHz).

Number of neurons in each hidden layer is varied for the improvement of LM training algorithm results. Finally tuned networks with highest accuracies are presented in Tab.2 compliant with the correlation factors (CF).

Scattering Parameters	CF for single hidden layered NN	CF for two hidden layered NN	CF for three hidden layered NN	CF for four hidden layered NN
S_{11}	0.9787	0.9963	0.9964	0.9717
S_{12}	0.9711	0.9992	0.9991	0.9895
S_{13}	0.9920	0.9988	0.9988	0.9925
S_{14}	0.9567	0.9996	0.9973	0.9355

Tab. 2. Correlation factors for MLPNN trained by LM algorithm.

Furthermore, correlation factors are computed among successful learning algorithms which applied to solve the modeling problem and grouped in Tab. 3 according to highest correlation factor values.

Scattering Parameters	LM	CGP	GD	CGF	SCG	RP
S_{11}	0.9963	0.9963	0.9662	0.9954	0.9962	0.9962
S_{12}	0.9992	0.9988	0.9824	0.9954	0.9983	0.9983
S_{13}	0.9988	0.9988	0.9879	0.9986	0.9988	0.9988
S_{14}	0.9996	0.9974	0.9695	0.9771	0.9937	0.9946

Tab. 3. Comparison of correlation factors for successful algorithms.

Fig. 8 – 10 illustrate the comparison of NN training algorithms for the RHC operating at the center frequency of 5269 MHz. The plots are zoomed into the frequency range of 4 GHz – 6.5 GHz in order to demonstrate the small neural solution variations. The modeled RHC provides adequate return loss results as shown in Fig. 8.

Fig. 8 and Fig. 10 clearly show the estimation capabilities of NNs trained by algorithms achieving highly successful approximations.

As shown Fig. 10 in Levenberg – Marquardt, resilient backpropagation, and conjugate gradient of Polak – Ribière algorithms offer the preeminent estimation values among other successful algorithms for the RHC modeling problem. However LM algorithm converges much slower than others for relatively large networks.

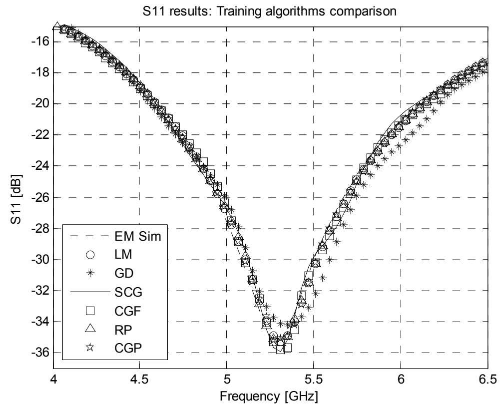

Fig. 8. Comparison of training algorithms for S_{11} (f_{cen} = 5269 MHz).

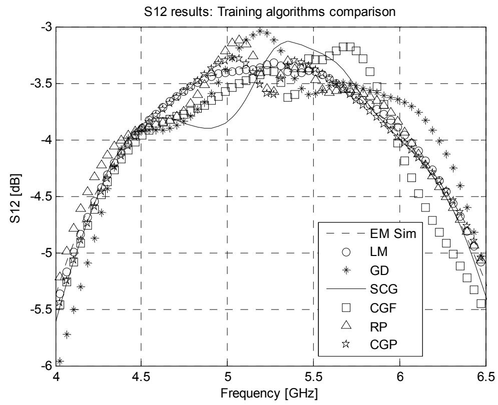

Fig. 9. Comparison of training algorithms for S_{12} (f_{cen} = 5269 MHz).

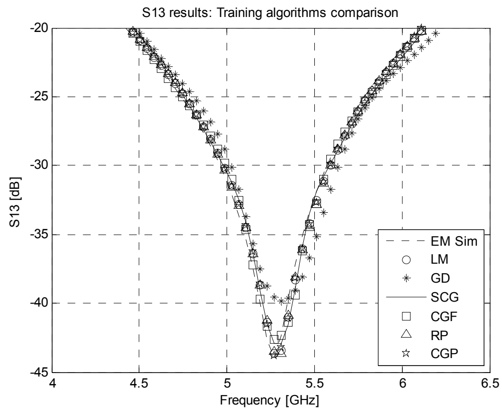

Fig. 10. Comparison of training algorithms for S_{13} (f_{cen} = 5269 MHz)

In Tab. 4, the training algorithms with less successive approximation values are compared. The quasi-Newton method, gradient descent with momentum approach, one step secant and variable learning rate backpropagation training algorithms provide relatively appalling correlation factors. Tab. 5 and 6 illustrate the computation time measurements for training algorithms under the same activation functions, number of hidden layers and neurons are given.

In each model measurement, 3 hidden layers with 20 neurons are trained for 2000 epochs. The four inputs and four outputs network with many hidden layers causes the LM algorithm training time increment.

Scattering Parameters	BFG	GDM	GDX	OSS
S_{11}	0.7574	0.9787	0.8191	0.8557
S_{12}	0.9254	0.9759	0.9072	0.9113
S_{13}	0.9474	0.9934	0.9857	0.9807
S_{14}	0.9257	0.9460	0.7574	0.5897

Tab. 4. Correlation factors for relatively low successful algorithms.

LM	CGP	GD	CGF	SCG	RP
4231.25	109.937	65.343	43.296	122.328	65.187

Tab. 5. Computation time in seconds for successful algorithms.

BFG	GDM	GDX	OSS
2739.062	62.796	92.046	157.968

Tab. 6. Computation time in seconds for less successful algorithms.

5. Conclusions

In this work, the MLPNNs are employed as a design, modeling and optimization tool for C-Band ring hybrid couplers. Well-known conformal mapping based mathematical approximations mentioned in Section 2 are exploited for the theoretical component design phase. Then the theoretical foundations are simulated and some tunings are applied using the EM-simulator to generate necessary training and test data sets. A parametric neural model is stated using ring hybrid radius, operating frequency points, substrate properties such as dielectric permittivity and thickness. Optimum networks are generated by varying number of hidden layers, neurons in these layers and training algorithms. Finally, obtained networks are used for modeling and optimization such that once the ANN is trained for given input data ranges then the parametric model can be used to find scattering values of any design that falls into given range. Thus the fine and coarse tunings for design optimization, which is initially done by EM-simulator, can be succeeded by ANN evolvement within simulation range.

As the number of input parameters is increased, sufficient number of new hidden layers must be added to the network architecture to overcome the problem complexity. However complex architectures with many hidden layers and neurons have a vital drawback of high computation time and memory need. Moreover the transfer functions of the hidden layer neurons are varied in order to achieve desired output values using accurate neural applications. These trade-offs must be done when the neural models are realized to solve microwave modeling problems. Time and memory requirements for both the fullwave EM-simulator and neural solutions must be compared to determine precise scattering parameters. The neural training algorithms are also evaluated to optimize the network outputs. Convergence speed, memory needs and accuracy are the main assessment constraints of the training algorithms. Thus the most accurate and fast converging algorithms are selected to train the networks under adequate neuron and hidden layer numbers condition.

Consequently, neural solutions of RF microwave design and optimization problems offer many advantages comparing to EM-simulators which require high level processors to overcome the background calculations' complexity. Improved time consumption in numeric computations and cost are the main drawbacks in EM-simulators. Thus parametric neural models can be applied for design optimization inside training range without having to redo the time and processor consuming fullwave EM-simulator analyses. This may force neural solutions as an alternative design, modeling and optimization way of microwave devices under the adequate selection of training algorithms and network architectures.

References

[1] WANG, F., ZHANG, Q. J. Knowledge based neural models for microwave design. *IEEE Transactions on Microwave Theory and Techniques*, 1997, vol. 45, no. 12, p. 2333 – 2343.

[2] CREECH, G. L., PAUL, B. J., LESNIAK, C. D., JENKINS, T. J., CALCATERA, M. C. Artificial neural networks for accurate microwave CAD applications. *IEEE International Microwave Symposium Digest*, San Francisco, CA, 1996, p. 733 - 736.

[3] WATSON, P. M., GUPTA, K. C. Design and optimization of CPW circuits using EM-ANN models for CPW components. *IEEE Transactions on Microwave Theory and Techniques*, 1997, vol. 45, no. 12, p. 2515 - 2523.

[4] ZAABAB, A. H., ZHANG, Q. J., NAKHLA, M. S. Neural network modeling approach to circuit optimization and statistical design. *IEEE Transactions on Microwave Theory and Techniques*, 1995, vol. 43, no. 6, p. 1349 – 1358.

[5] WATSON, P. M., GUPTA, K. C., MAHAJAN, R. L. Development of knowledge based artificial neural network models for microwave components. *IEEE International Microwave Symposium Digest*, Baltimore, MD, 1998, p. 9 - 12.

[6] KOTHAPALLI, G. Artificial neural networks as aids in circuit design. *Microelectronics Journal*, 1995, vol. 26, p. 569 - 578.

[7] HARKOUSS, Y., ROUSSET, J., CHEHADE, H., BARATAUD, D., TEYSSIER, J. P. Modeling microwave devices and circuits for telecommunications system design. In *Proceedings of IEEE International Conference on Neural Networks*. Anchorage (USA, AK), 1998, p. 128 - 133.

[8] VELUSWAMI, A., NAKHLA, M. S., ZHANG, Q. J. The application of neural networks to EM based simulation and optimization of interconnects in high speed VLSI circuits. *IEEE Transactions on Microwave Theory and Techniques,* 1997, vol. 45, no. 5, p. 712 - 723.

[9] HORNG, T., WANG, C., ALEXOPOULOS, N. G. Microstrip circuit design using neural networks. *IEEE International Microwave Symposium Digest*, Atlanta, GA, 1993, p. 413 - 416.

[10] BURRASCANO, P., DIONIGI, M., FANCELLI, C., MONGIARDO, M. A neural network model for CAD and

optimization of microwave filters. *IEEE International Microwave Symposium Digest*, Baltimore, MD, 1998, p. 13 - 16.

[11] RIZZI, P. A. *Microwave Engineering Passive Circuits*. Prentice Hall, 1988.

[12] BAHL, B., GUPTA G. *Microstrip Lines and Slotlines.* Artech House, 1996

[13] SVACINA, J. New method for analysis of microstrip with finite-width ground plane. *Microwave and Optical Technology Letters,* vol. 48, no. 2, 2006, p. 396-399.

[14] WHEELER, H. A. Transmission-Line Properties of Parallel Strips Separated by a Dielectric Sheet. *IEEE Transactions on Microwave Theory and Techniques,* 1965, vol. 13, no. 2, p. 172-185.

[15] SCHNEIDER, M. V. Microstrip lines for microwave integrated circuits, *The Bell System Technical Journal,* 1969*,* vol. 48, p. 1421-1444.

[16] HAMMERSTAD, E., JENSEN, Ø. Accurate models for microstrip computer-aided design. *Symposium on Microwave Theory and Techniques,* 1980*,* p. 407-409.

[17] SELVAN, P. T., RAGHAVAN, S. Multilayer perceptron neural analysis of edge coupled and conductor-backed edge coupled coplanar waveguides. *Progress in Electromagnetics Research B,* 2009, vol. 17, p. 169-185.

[18] KAYA, S., TURKMEN, M., GUNEY, K., YILDIZ C. Neural models for the elliptic- and circular-shaped microshield lines. *Progress in Electromagnetics Research B,* 2008, vol. 6, p. 169-181.

[19] TURKMEN, M., YILDIZ C. Quasi-static models based on artificial neural networks for calculating the characteristic parameters of multilayer cylindrical coplanar waveguide and strip line. *Progress in Electromagnetics Research B,* 2008, vol. 3, p. 1-22.

[20] WANG, F., DEVABHAKTUNI, V. K., XI, C., ZHANG, Q. J. Neural network structures and training algorithms for RF and microwave applications. *International Journal of Microwave and Millimeter-Wave Computer-Aided Engineering,* 1999, vol. 9, no. 3, p. 216 – 240.

[21] LEVENBERG, K. A method for the solution of certain nonlinear problems in least squares. *Quarterly of Applied Mathematics*, 1963, vol. 11, p. 431- 441.

[22] POLAK, E., RIBIERE, G. Note sur la convergence de methods de directions conjuguees. *Reveu Francaise d'Informatique et de Recherche Operationnelle*, 1969, p. 16-35.

[23] HAYKIN, S. *Neural Networks – A Comprehensive Foundation*. Prentice Hall, 1999.

[24] FLETCHER, R., REEVES, C. M., Function minimization by conjugate gradients. *The Computer Journal,* 1964, vol. 7, p. 149-154.

[25] MOLLER, M. F. A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks,* 1993, vol. 6, p. 525-533.

[26] REIDMILLER, M., BRAUN, H., A direct adaptive method for faster backpropagation learning: The Rprop algorithm. In *Proceedings of the IEEE International Conference on Neural Networks.* San Francisco (USA), 1993, p. 586-591.

About Authors ...

Erdem DEMİRCİOGLU was born in Mugla, Turkey. He received his B.Sc. degree from Yildiz Technical University in 2004 and M.Sc. degree from Syracuse University in 2006. He is working as a senior specialist in R&D – Satellite Design Department of Turksat AS and pursuing his doctoral studies in Electronics Department of Ankara University. His research interests include satellite payload systems, artificial intelligence applications in EM and biomedical topics.

Murat H. SAZLI was born in 1973 in Elazig, Turkey. He received the B.Sc. (with highest honors) and M.Sc. degrees from Electronics Engineering Department, Ankara University, in 1994 and 1997, respectively. He received the Ph.D. degree in electrical engineering from Syracuse University in 2003. He was the recipient of "Outstanding Teaching Assistant Award" from Syracuse University in 2002. He is currently an assistant professor and vice chairman of Electronics Engineering Department of Ankara University. His areas of interest include turbo coding and decoding, neural networks and their applications, wireless communications.