A Neural Network Model for CAD and Optimization of Microwave Filters

P. Burrascano †, M. Dionigi, C. Fancelli, M. Mongiardo

Istituto di Elettronica
†Istituto di Energetica
Via G. Duranti, 93, 06131, Perugia, Italy.

Abstract— **Improvement of the performance/cost ratio for modern microwave filters requires manufacturing-oriented design, hence accommodating full-wave tolerance analyses and yield optimization which are very computer-insensive.**
The use of neural networks for reducing the design effort of microwave filters, although still in its infancy, seems to provide a rather promising option. Once properly selected and trained, neural networks can approximate the filter response at a very modest fraction of the computer resources used by the full-wave rigorous model, hence enabling systematic application of manufacturing-oriented design.
In this contribution we present the solution of the major important choices related to the effective selection of a neural network suitable for approximating the behavior of a typical microwave filter. For illustration we consider the example of a standard four-pole E-plane metal-insert filter operating in X-band.

I. Introduction

THE current trend in the design of microwave filters is to achieve the goal of yield optimization, i.e. to maximize the number of filters which satisfy a defined acceptance criteria with respect to the total number of filters produced. This means that the designer has to optimize not only the "ideal" filter, i.e. the one with the theoretical geometrical dimensions, but the optimization is to be carried out considering an entire statistic of the possible outputs of a given manufacturing process. Needless to say, yield optimization requires a considerable effort and is only feasible at a computer level.

Depending on the selected implementation, filters may be realized on different waveguiding structures: microstrip lines, coplanar waveguides, slotlines, hollow metallic waveguides, dielectric waveguides and several other types of transmission lines [1, Chap. 1]. Moreover, a filter may present a rather simple geometry, or it may require the use of tuning screws or other devices which make the structure a fairly complex one. As a result electromagnetic tools are necessary for the EM rigorous filter analysis.

Unfortunately, to perform yield optimization by using electromagnetic packages such as FEM or FDTD does not appear feasible and even when using very efficient modal analysis it has been found convenient to employ sophisticated techniques, like the adjoint network method [2], [3], in order to expedite computations. Recently, the introduction of *Space Mapping* has alleviated the problem [4].

However, even space mapping still makes use of computer-intensive EM full-wave simulators. But essentially, what we really need is a tool which relates the geometrical dimensions (input) to the filter frequency response (output). A neural network seems to be an ideal candidate for such a task. In fact, it has been shown that after proper training, a neural model can define any nonlinear mapping between an input and an output vector spaces. In our application we can associate the geometrical dimensions (input) to the filter response (output) with a very modest numerical effort. In order to obtain such a result a proper model for the neural network has to be selected, possibly tailored on the particular component that we are modeling (i.e. the filter).

Previous investigations and applications of neural networks for microwave applications have not considered in detail filters examples [5]. The specificity of the filter response calls for an appropriate neural network architecture: the individuation of such a network, and its behavior, is the topic of the present work.

II. Neural networks: a brief description

Neural networks can help to further the understanding of brain functions: engineers are, however, mostly interested in understanding how neural networks compare with different processing techniques for problem solving. Neural networks, because of their massively parallel structure, can perform computations at a very high rate if implemented on a dedicated hardware; because of their adaptive nature, they can learn the characteristics of input signals and adapt to changes in the data; because of their non linear nature they can perform functional approximation and signal filtering operations which are beyond optimal linear techniques. *This latter aspect was the starting point for the present work:*

our objective is to define an approximator of the non linear functional between the geometrical dimensions of the waveguide filter considered and its frequency response in the neighborhood of the point obtained from the direct application of a synthesis procedure. The ability of neural networks to define non linear models is expected to yield an accurate model in a dynamic range of the geometric parameters wider than the one allowed by the usual linear methods.

A. Feedforward layered networks and error back propagation

Typically feedforward layered networks consist of a set of source nodes which constitute the input layer, one or more hidden layers of computation nodes and an output layer of computation nodes. Each processing node (neuron) performs a sum of the signal components at its input; this sum is thus fed into a block performing a differentiable non linear processing (usually a sigmoidal logistic function) of the type

$$f(v) = \frac{1}{1 + \exp(-v)} \tag{1}$$

The input signal propagates through the network in a forward direction, on a layer-by-layer basis. These networks are usually referred to as multilayer perceptrons (MLPs). The processing task to be performed is described by a set of input-output data (the training set): the network is specialized to solve this task by modifying its parameters by means of an iterative optimization procedure (the learning procedure): the error back propagation, described in [6], is a highly popular learning algorithm. It basically consists, at each iteration, of two passes through the different layers of the network: a forward pass and a backward pass. In the forward pass an input pattern is applied to the input nodes and its effect propagates through the layers of the network: a set of outputs is produced (the actual response of the network at that stage). The actual response is subtracted from the desired (target) response to produce an error signal which is propagated backward through the network (the backward pass). In this phase the parameters of the network are adjusted following a gradient descent procedure.

III. Function approximation by learning

Approximation theory deals with the problem of approximating or interpolating a continuous, multivariate function $f(x)$ by an approximating function $F(w,x)$ having a fixed number of parameters w belonging to some set P. For a choice of a specific F, the problem is then to find the set of parameters W that provides the best possible approximation for f on the set of the available "examples" (learning step). It is fundamental to choose an approximating function F that can represent f as well as possible. The problem of learning a mapping between an input an output space is equivalent to the problem of estimating the system that transforms inputs into outputs given a set of examples of input-output pairs (supervised learning).

The problem of approximating a function of several variables by MLPs has been studied by many authors: [7], [8], claim that a three layer feedforward layered networks with sigmoid units in the hidden layer can approximate continuous or other kinds of functions defined on compact sets in $\mathcal{R}^n$. Their results can be summarized as follows (universal approximation theorem): *Let $\phi(.)$ be a non constant, bounded and monotone-increasing continuous function. Let ℓ_p denote the p-dimensional unit hyper cube $[0,1]^p$. The space of continuous functions on ℓ_p is denoted by $C(\ell_p)$. Then, given any function $f \in C(\ell_p)$ and $\epsilon > 0$, there exist an integer M and sets of real constants α_i, θ_i, and w_{ij} , where $i = 1, \ldots, M$ and $j = 1, \ldots, p$ such that we may define*

$$F(x_1, x_2, \ldots, x_p) = \sum_{i=1}^{M} \alpha_i \phi \left(\sum_{j=1}^{p} w_{ij} x_j - \theta_i \right) \tag{2}$$

as an approximate realization of the function $f(.)$; that is

$$|F(x_1, x_2, \ldots, x_p) - f(x_1, x_2, \ldots, x_p)| < \epsilon \tag{3}$$

for all $\{x_1, x_2, \ldots, x_p\} \in \ell_p$. This theorem is directly applicable to MLPs: we first note that the logistic function is a non constant, bounded and monotone-increasing continuous function; it therefore satisfies the conditions imposed on $\phi(.)$. Then we note that (2) represents the output of a MLP consisting of p input nodes and a single hidden layer of M neurons with thresholds θ_i, and input-to-hidden weights w_{ij} ; the output neuron is a linear one and the hidden-to-output weights α_i define the coefficients of the combination.

A. Practical application of the universal approximation theorem

Several fundamental questions must be addressed in order to make a practical use of MPLs as approximators. A first question is the following: how many samples are needed to achieve a given degree of accuracy? It is well known that the answer depends on the dimensionality d of the data space and on the degree of smoothness p of the class of functions that has to be approximated [9].

Other fundamental questions are related to both the learning strategy and neural network topological structure. The universal approximation theorem is infact an existence one: it states that a single layer network is

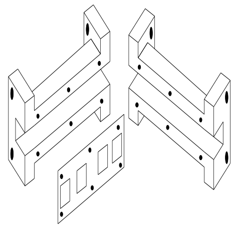

Fig. 1. Geometry of the E-plane metal-insert filter

sufficient for a MLP to compute a uniform e approximation to a given training set represented by the input output couples $\{x_1, x_2, \ldots, x_p\}$ and $f(x_1, x_2, \ldots, x_p)$.

The problem of model complexity selection arises in all parametric modeling techniques: it has been approached in the technical literature by using statistical identification procedures, as for instance the AIC [10] or the MDL [11]; in the specific case of neural models, the use of the above criteria, or their extensions, has been proposed [12], [13], [14].

IV. Application: Learning the E-plane Metal-insert Filter Neural Model

The artificial neural technique has been applied to the simulation of a typical waveguide filter with the goal of obtaining a reliable, accurate, and fast description of the filter electrical response in presence of tolerance errors due to the manufacturing process. For sake of definitess we have selected as an example the E-plane metal-insert waveguide filter shown in Fig. 1. The relative geometrical parameters are the septa lengths and spacing and the waveguide width, resulting in a total number of 11 geometrical parameters; together with the frequency band they represent the input to the neural network; the output being the return loss, s_{11}, computed on the frequency band of interest.

The different input components of each pattern are associated to the geometrical dimensions of the filter under consideration. Their values have been scaled to the same dynamical range (-10,10 in our experiments) to ensure that the neural model assigns the same relevance to the different geometrical dimensions. After training phase, the neural model approximates the return loss with respect to its theoretical behaviour. We denote these deviations as "approximation error". It is interesting to represent this quantity in the frequency band of interest by considering its mean and standard deviation as obtained by considering a fairly large set of 65 filter responses. A typical response for the filter considered is illustrated in Fig. 2; it is apparent that the error becomes fairly large at certain frequencies. However, it turns out that the approximation error is particularly severe just at the nominal filter resonances. This is quite natural because even a modest change in the geometrical dimensions causes a slight shift in the resonance which, in turn, generates a severe error in the actual value. However, for practical purposes this type of error is not particularly relevant. Based on this observation, after our early experiments, we modified the learning procedure in order to improve the approximation accuracy in the range of interest. Fig. 2 refers to the results obtained in this way: namely, the MLP implements a non linear compression/decompression during the training phase. The choice of neural model's topology has been performed by using the AIC procedure, as proposed in [13], [14]. The AIC procedure, and exaustive experiments, have shown that a topology consisting of 50 neurons in the hidden layer (15,50,1) is capable of providing fairly good results with a relatively modest numerical effort. The training set has been defined by assuming as nominal manufacturing tolerance 10 microns, and we have trained the network with a training set of 65 filter responses sampled in 300 frequency points. The training set has been produced as a random gaussian distribution of filter geometrical parameters around the nominal design with a standard deviation of 40 microns, considering in such a way the case of very large manufacturing errors. The number of filter curves considered for training purposes has been defined by taking into account the number of free parameters in the neural model. A number of 65 filter responses sampled in 300 frequency points, proved to give a sufficient description of the mapping while allowing a reasonable computational effort (2 hour on a PC pentium 200 MHz). After the learning phase the trained network gives its response in a few milliseconds, and allows to approximate any filter response in the range of geometrical values we considered (40 microns for each dimension, four times the nominal tolerances). To give an idea of the generalization ability allowed by the trained network, fig. 3 allows a comparison between the responses obtained by using a fullwave analysis and the neural model; it is worth noting that these filters were not considered in the training set of the neural model. It is moreover apparent that the neural network response is reliable and accurate even for the frequency points corresponding to the zeroes of the reflection coefficient. The computational cost associated to the modal analysis procedure is over 100 times more expensive that the one associated to the neural model. This clarifies why the use of a neural model can be particularly well suited for an intensively repetitive procedure.

V. Conclusions

A neural model of an E-plane metal insert filter has been defined which takes into account manufacturing errors. A suitable procedure has been defined to select neural model's topological complexity and the number

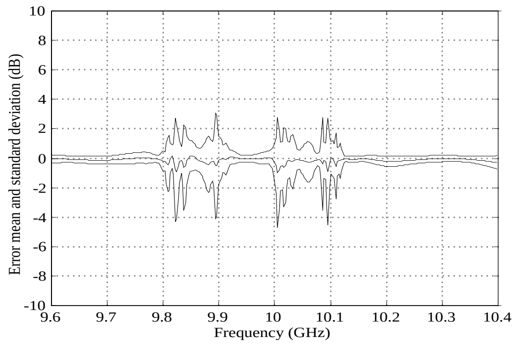

Fig. 2. Manufacturing errors change the return loss with respect to the theoretical one; these deviations are synthetically represented by the mean and relative standard deviation.

of samples required to ensure an accurate description of filter response, even in the presence of very large manufacturing errors. The results we reported in the paper show that our procedure meets the original requirements posed by means of a model whose recursive computational cost, after the learning phase, is about two orders of magnitude lower than the one associated to a full wave model analysis: this aspect makes the neural modelling approximation procedure an interesting candidate for repetitive on line controls. Other relevant perspectives are opened by the availability of a flexible and accurate model of the filter: we are currently exploring a number of these aspects, and namely (1) a further improvement of neural model accuracy by means of a preliminary neural clusterization of filter curves typology; (2) the automatic estimate of filter response sensitivity to geometrical dimensions inaccuracies, in order to obtain information about the manufacturing accuracy needed for each geometrical dimension in order to maintain the overall filter response within a desired tolerance; (3) inverting the neural model in order to obtain a new set of "nominal" geometrical dimensions which take into account manufacturing tolerances.

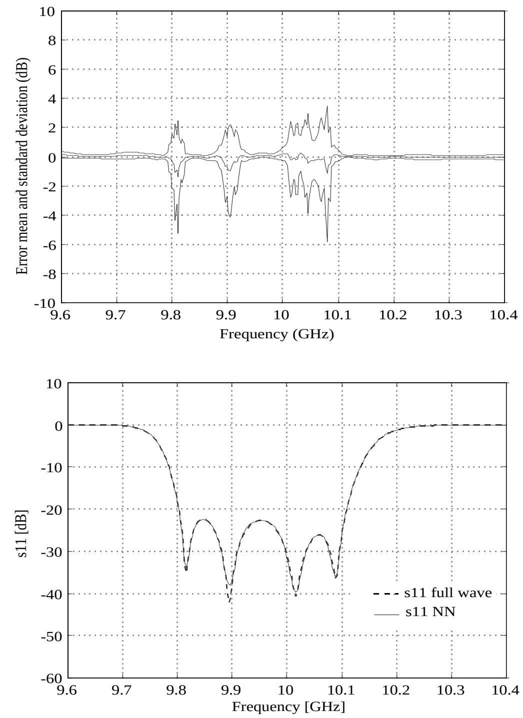

Fig. 3. Statistic obtained for the testing set (upper figure) and relative results obtained for one filter of the testing set (lower figure). The example refers to a manufacturing tolerance of 10 microns and a random gaussian distribution of the geometrical dimensions around the nominal design with a standard deviation of 40 microns.

References

[1] T. Rozzi and M. Mongiardo, *Open Electromagnetic Waveguides.* London: IEE, 1997.

[2] M. Mongiardo and R. Ravanelli, "Automated design of corrugated feeds by the adjoint network method," *IEEE Trans. Microwave Theory Tech.*, May 1997.

[3] F. Alessandri, M. Mongiardo, and R. Sorrentino, "New efficient full wave optimization of microwave circuits by the adjoint network method," *IEEE Microwave Guided Wave Lett.*, vol. 3, pp. 414–416, Nov. 1993.

[4] J. Bandler, R. Biernacki, S. Chen, R. Hemmers, and K. Madsen, "Electromagnetic optimization exploiting aggressive space mapping," *IEEE Trans. Microwave Theory Tech.*, vol. 43, pp. 2874–2882, Nov. 1995.

[5] P. Watson and K. Gupta, "EM-ANN modeling and optimal chamfering of 90 bends with air-bridges," *Int. Microwave Symp. Dig.*, June 1997.

[6] D. Rumelhart, G. Hinton, and R.J.Williams, "Learning internal representations by error propagation," in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (C. . vol. 1: Foundations, ed.), pp. 318–362, Cambridge, MA: MIT Press, 1986.

[7] K. Hornik, M. Stichombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, pp. 359–366, 1989.

[8] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Math Contr., Signals, Syst.*, vol. 2, pp. 303–314, 1989.

[9] C. Stone, "Optimal global rates of convergence for nonparametric regression," *Ann. Stat.*, vol. 10, pp. 1040–1053, 1982.

[10] Akaike, "A new look at the statistical model identification," *IEEE Trans. on Aut. Control*, vol. 19, no. 6, pp. 716–723, 1974.

[11] J. Rissanen, "Stochastic complexity and modelling," *Ann. Statist.*, vol. 14, pp. 1080–1100, 1986.

[12] N. Murata, S. Yoshizawa, and S. Amari, "Network information criterion - determining the number of hidden units for an artificial neural network," *METR 92-05, Technical Report, University of Tokyo*, 1992.

[13] P. Burrascano, "Practical applicability of statistical identification techniques foe network topology selection," *Post-NIPS Workshop on Model Complexity Dig.,Snowmass (Aspen), Colorado*, 1996.

[14] P. Burrascano et al., "Statistical identification techniques for network topology selection," *ISIS-97, (Invited Paper), International Sympoisum on Intelligent Systems Dig., 11-13 Sept. 1997, Reggio Calabria, Italy*, 1996.