**Project:** **EST3, CHAMELEON RF**

**Project Number:** FP6/2004/IST/4-027378

**Title**: Neural Network Approach to Modelling Parameters Variation in the Frequency Domain


**Author(s):** Andrzej Rusiecki





**Affiliation(s):** Laboratorul de Metode Numerice – LMN, CIEAC, Universitatea Politehnica Bucuresti
**www.**lmn.pub.ro


**Date:** January 2007

**Acknowledgements:**

I'd like to thank for the scientific support I've received from Prof. **Daniel IOAN,** www.lmn.pub.ro/~daniel, my research supervisor at LMN

The work was financed by **Marie Curie EST3 project** http://est3.lmn.pub.ro/

The work technically supported by the FP6/IST/Chameleon project http://www.chameleon-rf.org

# Neural Network Approach to Modelling Parameters Variation in the Frequency Domain

Andrzej Rusiecki, *Bucharest, January 2007*

*Abstract:* In this report some remarks considering modelling parameters variation with artificial neural networks (NN) trained on Chamy software-based data are presented. The dedicated network structure is proposed, discussed and tested in a few cases.

## 1. Introduction

The current version of the Chamy3d software is able to produce frequency characteristics of the analyzed device for the assumed set of parameters. Potentially, it can compute the device behaviour for the given conditions. In this report we propose to use a sample of the data, generated by the Chamy tool, to build a neural network model that can automatically allow obtaining device responses for the other sets of parameters. The advantages of such approach are: shorter computation time (in the case of the trained network, the response is given immediately after signal is propagated through the structure) and simple use (neural networks work as a "black-box" modelling tool).

The use of artificial neural networks in modelling different properties of electronic devices is not a new idea. NN were used in the various fields from simple prediction of device characteristics [4], automatic building RF/microwave models [5], [7], to constructing models based on time-domain large signals measurements [6]. Also modelling of the statistical behaviour for the parameters variation can be found in literature [8].

In this report we focus on modelling parameters variation, which from the NN point of view, may be considered as a problem of modelling a family of functions. To build such models we can, in fact, disregard the source of our data and focus on the techniques leading to the proper network behaviour.

## 2. Problem definition

The data generated by the Chamy tool are: real and imaginary part of $Z$ for the given frequency and assumed set of parameters. For simplicity (and because the software is not yet able to produce more data automatically) we examine the case of 2 parameters variation.

The problem may be then defined as follows: the set of training samples is given. Every sample contains 3-dimensional argument vector (2 device parameters and frequency) and two-dimensional output vector. We can divide the training set into the subsets corresponding to the each parameters case (2 parameters fixed). Moreover, it is not difficult to notice that such subsets can be also considered as generated by the separate functions. Now, if we take separately also two elements of the output vector, we get subsets of points, each generated by a certain function $F:R \rightarrow R$. These subsets constitute characteristics, examples of which are shown in the figures 5 and 6.

The simplest approach in NN modelling is to choose network architecture, then, fitting size of inputs and outputs to the training data, try to learn the structure. However, it obviously suffers from many difficulties, such as unknown type of correlation the network can build between the data fitting exactly to the training set (e.g. so-called overfitting phenomenon), potentially huge size of such network, and long computation time. This is why a dedicated approach to every case should be proposed.

For the considered task, to make the network size as small and effective as possible, we propose to decompose the main problem into the set of smaller problems. The idea is based on the fact that our task can be described as approximation of the family of functions. The functions from the family have similar shape in the sense that they differ only in parameters (in this case no more than 2). Obviously, the analytical formula of the functions is unknown.

## 3. Dedicated network structure

As it is shown in the Fig.1 we decompose our task into parametric and nonparametric problem. Term 'nonparametric' means here that we gather knowledge of the general shape of functions that can be then parameterized to obtain solutions for the given data (in the regular meaning the NN learning problem is clearly parametric). The nonparametric estimation should be ideally done by the network Net2, which is to approximate the representative function from the family (real or imaginary part of Z) for the assumed and preset parameters. Then its output should be combined with the parameters inputs pre-processed by the Net1, and fed to the output block. The output junction is responsible for solving parametric problem and calculating function values for the given parameters. The question is what should be the outputs of networks Net1 and Net2, and how to construct the last block.

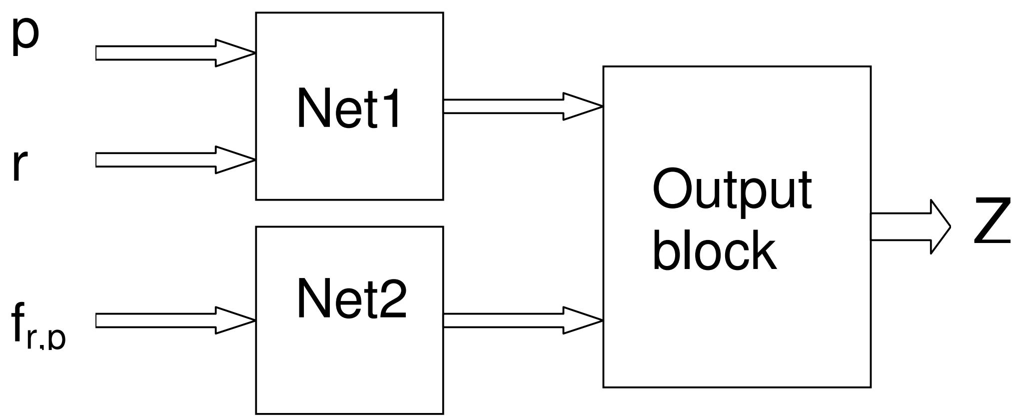



**Figure 1: Decomposed problem – general view**

One approach could be using a product or summation as the output junction. Such structure for the reference parameters *r, p* and imaginary part of Imag($Z$) = $Z_i$ as the goal, is shown in the Fig. 2. Then, automatically, the output of the Net2 is the approximated output for the set parameters $y_{Net2}=Z_{i,rp}$ and the Net1 output is simply scaled $y_{Net1}=Z_i(r,p,f)/ Z_{i,rp}(f)$ (in the case of the product). Disadvantages of this approach are, however, clearly evident. First of all, it is not certain that the dependency $y_{Net1}(r,p,f)$ is simpler than just $Z_i(r,p,f)$ (though, it is so in the case of the tested device, especially if we consider imaginary part of $Z$). The second reason is that the error of both networks is multiplied here, which might result with very high level of inaccuracy.

This is why another, more general and easy to implement, new approach is introduced. Because we have training data with targets, the "learning with a teacher" strategy should be used. This induces the use of supervised feedforward or recurrent neural networks. With the problem of function approximation feedforward sigmoid [1] and radial basis function (RBF) networks deal in the best way because they can be considered as universal approximators [2]. However, the RBF networks [3] interpolate rather than approximate training data (even when achieving higher accuracy than sigmoid networks) which makes the sigmoid transfer function networks the best choice here.

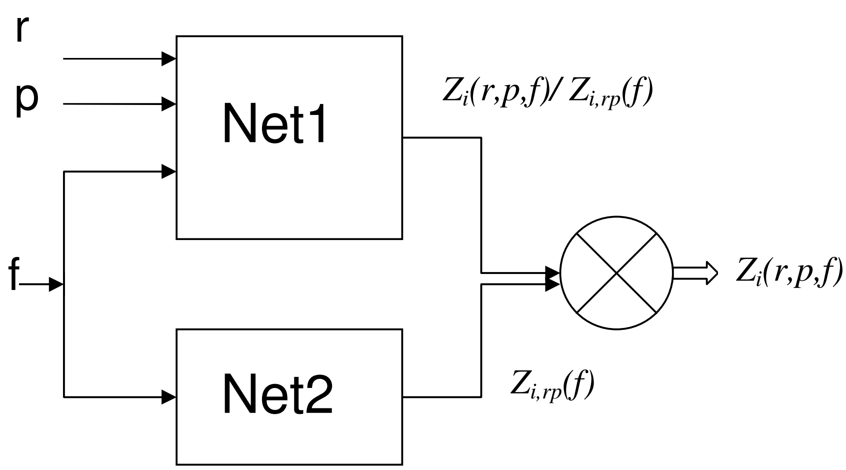



**Figure 2: Example of dedicated approach with 'product' block**

As it was mentioned before, we intend to create the network structure that could be relatively small and effective. To make it so, we follow the idea of decomposition of the problem. The proposed network structure is shown in the Fig. 3. As it can be noticed, the clue is to use:

i) a small network to approximate the $Z_{i,rp}$ in the $f$ domain for the fixed parameters;

ii) a bigger two-layer network with parameters connected to the inputs and $Z_{i,rp}$ fed to the second hidden layer.

In this structure, similarly to what was proposed in the Fig.1, we decompose our problem and the second layer of the main network combines the information of the function shape with the parametric model. Instead of the Net2 also another approximation technique could be applied.

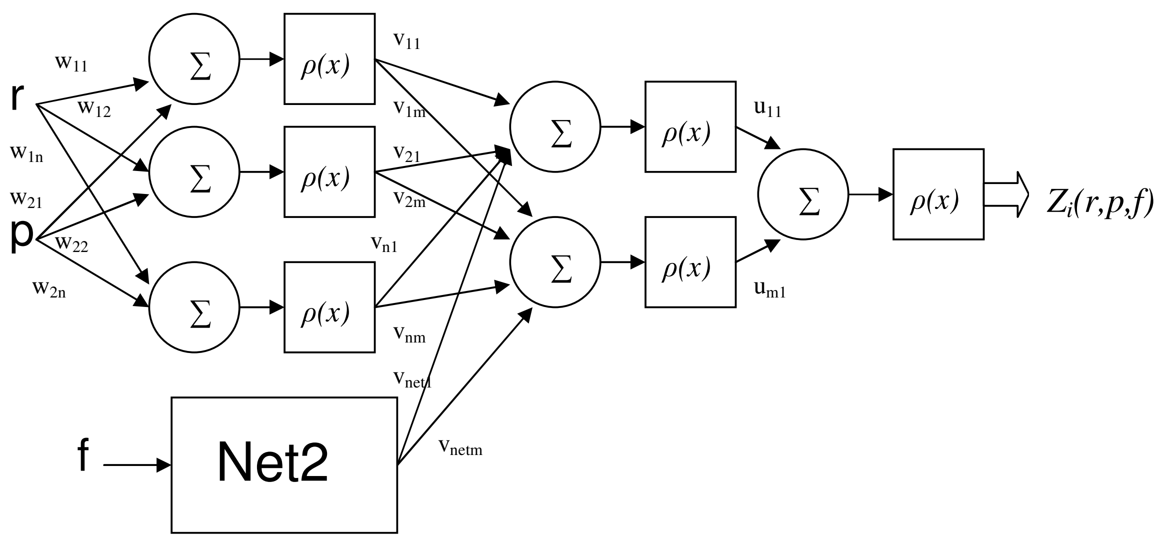



**Figure 3: Dedicated NN structure**

## 4. Case study

Because obtaining data from the current version of the Chamy tool needs many manual operations, we examined here only one example coming from the benchmark set device and one constructed artificially as a group of characteristics.

The first testing example is L-Shaped Interconnect Element (L-dev) [10]. The layout of the L-Shape device is shown in Fig. 4 and the geometrical parameters with materials data in the tables 1-3 (after [10]). In figures 5 and 6 exemplary data for the different values of parameters are presented. The device has two electrical terminals: one current excited (zmax face) and a ground terminal (xmax face). The computational domain was divided into 768 nodes.

In this example to train and test the NN we generated $Z$ values in 20 frequencies for 49 sets of parameters. The varying parameters were $p_2$ and $r_2$.

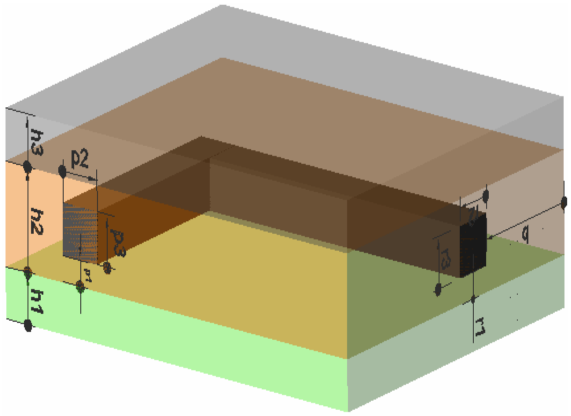
In the second example artificially data were generated according to the function:

$$y(x) = e^{\frac{-(rx)^2}{p}},$$

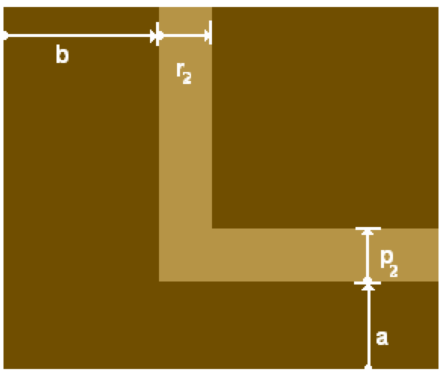
where $r$ and $p$ are parameters, $x$ is independent and $y$ dependent variable. Just as in the previous case, $y$ values were calculated for the different sets of parameters. This example represent situation when the dependence between parameters and function shape is not as trivial as in the case of the L-Shaped Device. Exemplary characteristics are shown in Fig. 7.

## 5. Simulation results

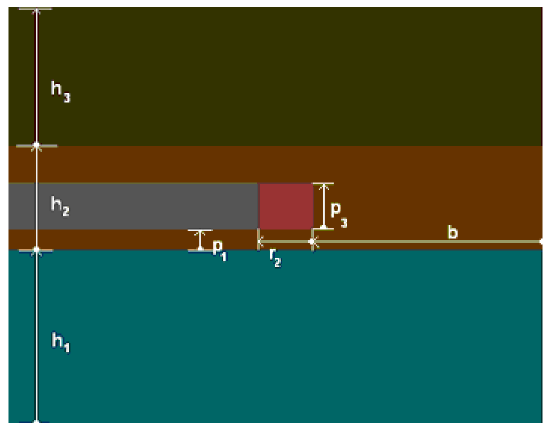
The comparison of the performance of regular networks and our dedicated networks tested on the aforementioned examples is gathered in the tables 4 and 5. The first table contains the mean MSE (mean squared error) for one hundred networks trained on different permutations of data sets. In the second table the median MSE values are shown. The data were divided into the training and testing set in the relation 1:3. The NN were trained on the data belonging to the training sets and tested on the rest of the data. The networks responses for the testing data not used in the training process are presented in figures 8-10. Both tested structures were based on the two-layer sigmoid transfer function feedforward network architecture. The numbers of the neurons in both hidden layers were set to $n=m=5$. To train the NN a conjugated gradient algorithm in the Polak-Ribierre variant [9], with 550 epochs was applied.




- L Shape – 3D view




- L Shape – 2D view (xOz)




- L Shape – 2D view (xOy)

**Figure 4: L – Shape Device**

Looking at the tables 4 and 5 one can notice that the lowest error level was achieved in each case for the new dedicated NN structure. The error for the regular network was up to 300% higher. Closer look in the figures reveals however, that even for the dedicated network, its output differs slightly from the desired values. Obviously, for the training data the network achieves almost exact fitting. It clearly demonstrates that the dedicated network simulates test examples successively.

## 6. Summary

In this report a simple NN-based approach to model sets of similar functions was presented. The approach can be used also for modeling parameters variations of nanoelectronic devices. As it was demonstrated on the testing data generated by the Chamy tool, the new dedicated network presents better performance when compared with the typical

NN structure. Future research could be focused on searching for the NN structures for modeling devices characteristics in the multidimensional parameters space.

**Tabel 1: Layers descrition**

| Layer | Parameters | Nominal [μm] | Min [μm] | Max [μm] | Material |
|---|---|---|---|---|---|
| Domain | Xmax | 15 | 9 | 725 | |
| | Zmax | 15 | 9 | 725 | |
| Si | h1 | 125 | 0 | 725 | Si |
| SiO2 | h2 | 21.131 | 20.074 | 22.187 | SiO2 |
| Air | h3 | 125 | 0 | 725 | AIR |

**Tabel 2: Bricks description**

| Nr. Trick | Layer | Parameters | Symbolic | Nominal [μm] | Min [μm] | Max [μm] | Material |
|---|---|---|---|---|---|---|---|
| 1 | SiO2 | a | a | 6 | 3 | 361 | ALUM |
| | | p1 | p1 | 1.572 | 1.404 | 1.74 | |
| | | p2 | p2 | 3 | 2.85 | 3.15 | |
| | | p3 | p3 | 0.665 | 0.565 | 0.765 | |
| 2 | SiO2 | b | b | 6 | 3 | 361 | ALUM |
| | | r2 | r2 | 3 | 2.85 | 3.15 | |

**Tabel 3: Material properties**

| Material | Name | Type | Properties | | |
|---|---|---|---|---|---|
| | | | μ | ε | σ |
| 1 | Si | Semiconductor | 1 | 11.9 | 10000 |
| 2 | SiO2 | Insulator | 1 | 3.9 | 1e-07 |
| 3 | ALUM | Conductor | 1 | 1 | 6.6e+07 |
| 4 | AIR | Insulator | 1 | 1 | 1e-07 |

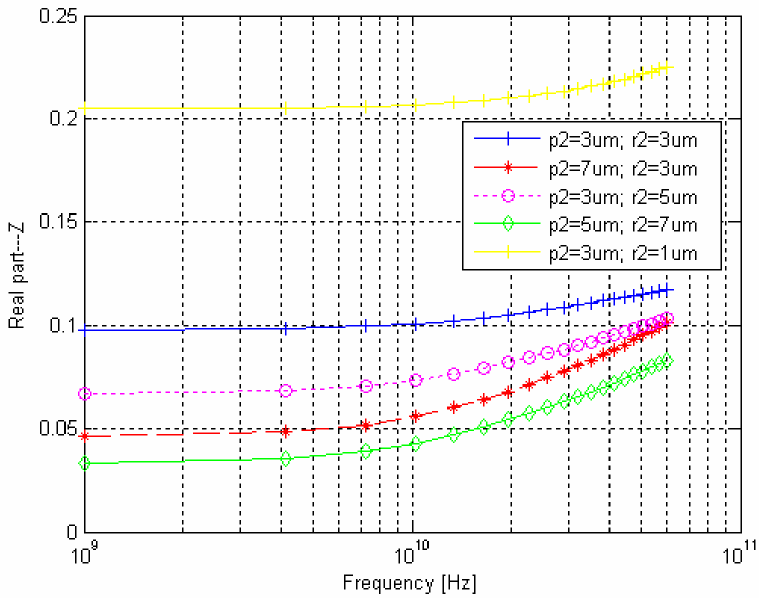



**Figure 5: Testing example 1 – real part of *Z***

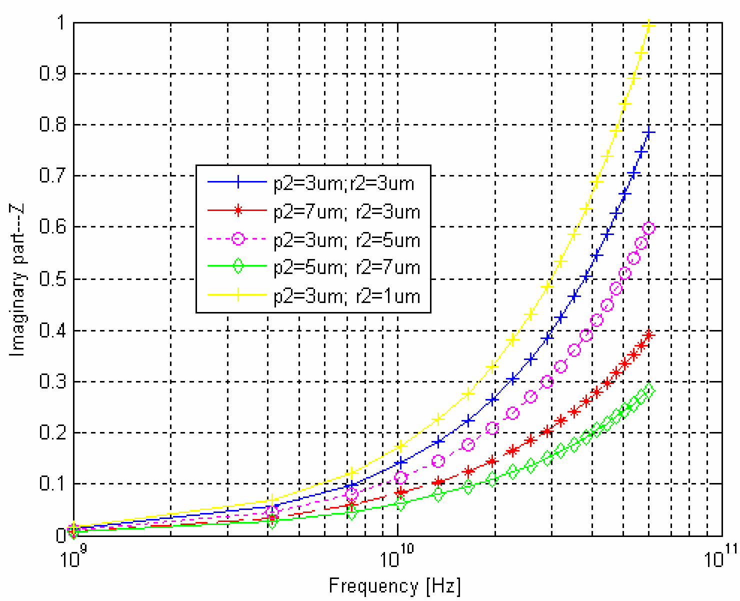



**Figure 6: Testing example 1 – imaginary part of *Z***

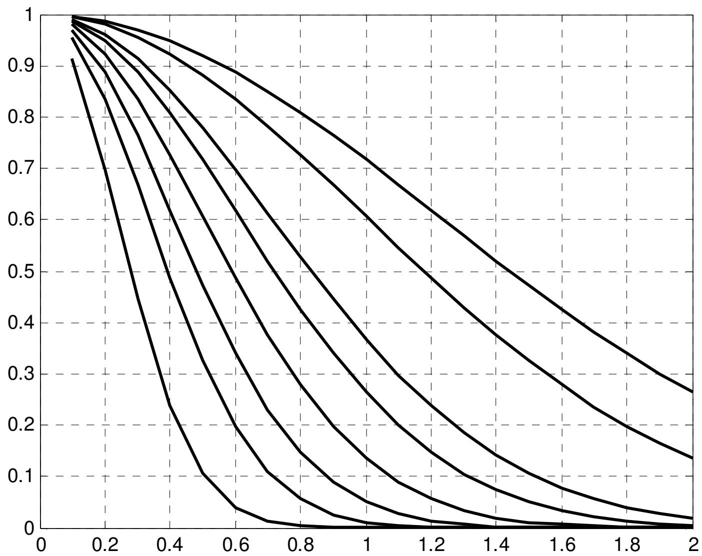



**Figure 7: Testing example 2 – function shape for different parameters set**

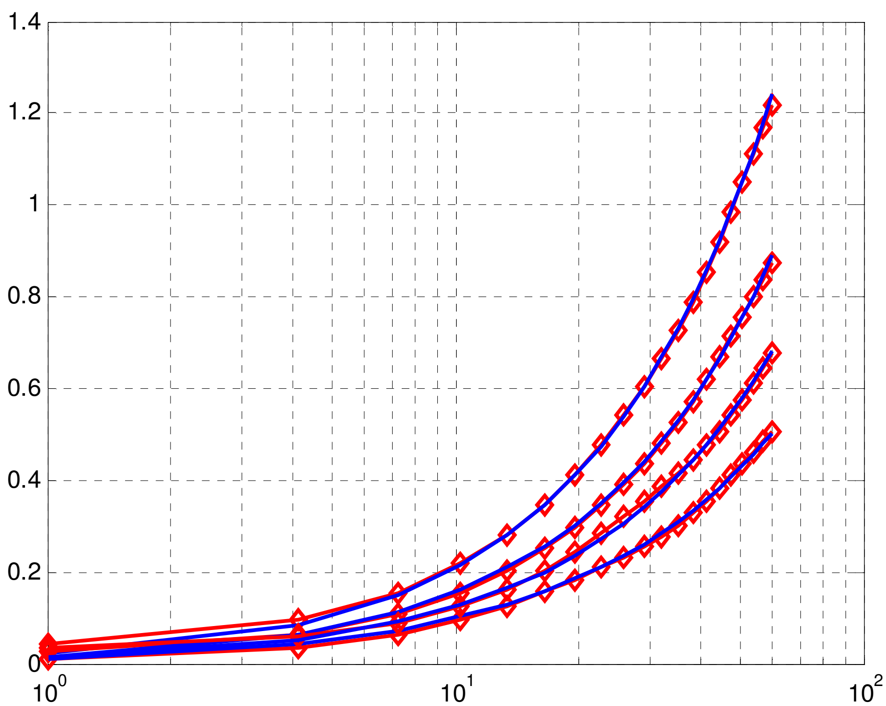



**Figure 8: Test example 1, Imag{*Z*}, real values (blue) and dedicated NN responses (red)**

**Tabel 4: Mean MSE for the testing data**

| **Network type** | **Real{$Z$}** | **Imag{$Z$}** | **Example 2** |
|---|---|---|---|
| Dedicated | 0.00648 | 0.01887 | 0.00063 |
| Regular | 0.01600 | 0.06393 | 0.00105 |

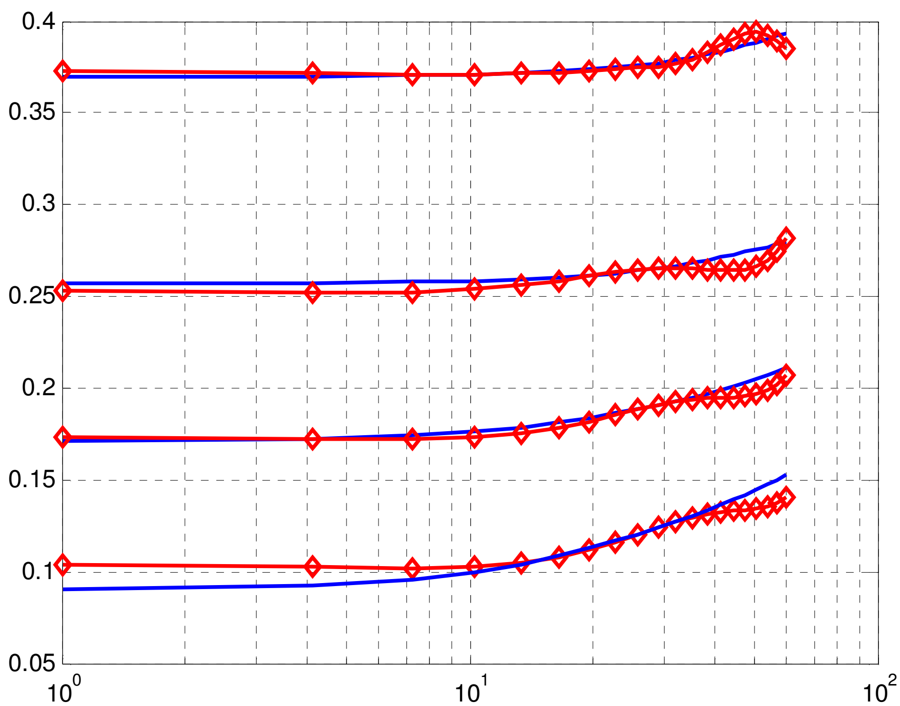



**Figure 9: Test example 1, Re{$Z$}, real values (blue) and dedicated NN responses (red)**

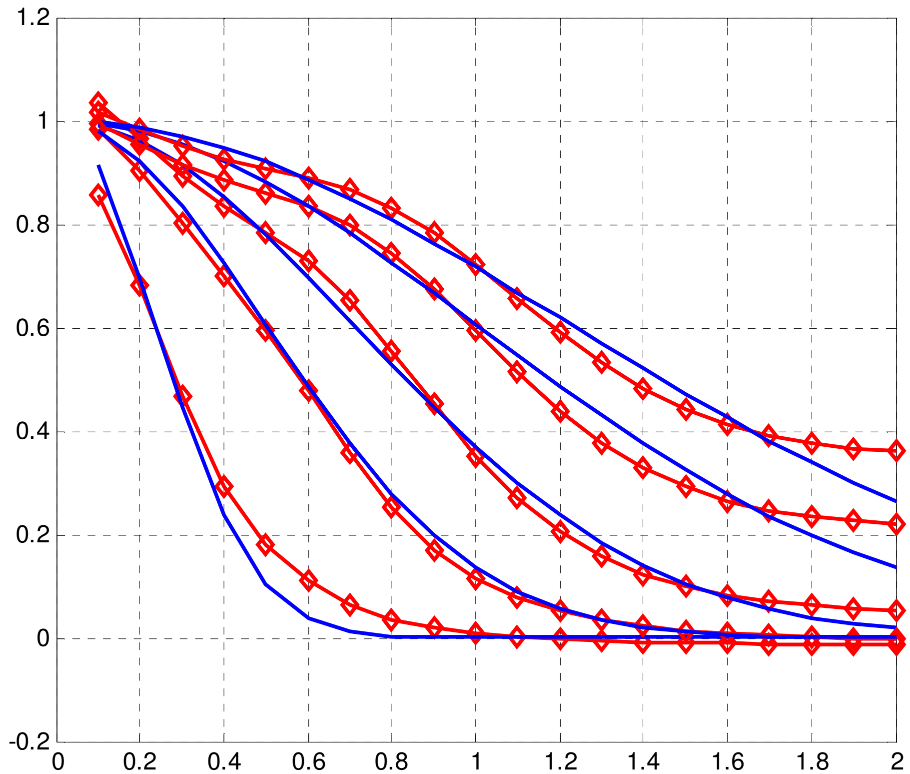



**Figure 10: Test example 2, real values (blue) and dedicated NN responses (red)**

**Tabel 5: Median MSE for the testing data**

| **Network type** | **Real{*Z*}** | **Imag{*Z*}** | **Example 2** |
|---|---|---|---|
| Dedicated | 0.00379 | 0.00800 | 0.00062 |
| Regular | 0.01073 | 0.02016 | 0.00104 |

## References:


[1] G. Cybenko, Approximation by superpositions of sigmoidal functions, Mathematics of Control, Signals, and Systems, vol. 2, pp. 303-314, 1989

[2] K. Hornik, M. Stinchcombe, H. White, Multilayer Feedforward Networks are Universal Approximators, Neural Networks, vol. 2, pp. 359-366, 1989

[3] J. Park, I. W. Sandberg, Universal approximation using radial basis function networks, Neural Computation, vol. 3, pp. 246-257, 1991

[4] B. Rajendran, G. S. Kar, S. Sen & S. K. Ray, Modeling of SiGe/Si heterostructure pMOSFET devices using artificial neural networks *International Conference on Communications, Computers and Devices, December 2000.*

[5] V. K. Devabhaktuni, et al., A Robust Algorithm for Automatic Development of Neural-Network Models for Microwave Applications, IEEE Trans. on Microwave Theory and Techniques, vol. 49, no. 12, pp. 2282-2290 December 2001

[6] D. M. M.-P. Schreurs, et al., Artificial Neural Network Model for HEMTs Constructed from Large-Signal Time-Domain Measurements, 59th ARFTG Conf. Dig., Seattle, WA, June 7, 2002, pp. 31-36

[7] Q.-J. Zhang, K. C. Gupta,V. K. Devabhaktuni, Artificial Neural Networks for RF and Microwave Design - From Theory to Practice, IEEE Trans. on Microwave Theory and Techniques, vol. 51, no. 4, April 2003

[8] H. Taher, et al., Artificial Neural Network to Statistically Model the Variation in Small Signal Equivalent Circuit Model Parameters for a Si/SiGe HBT Process, ARFTG Conference Digest Spring, vol. 63 , pp. 103-106, June 2004.

[9] Hagan, M. T., Demuth, H. B., Beale, M. H.: Neural Network Design, Boston, MA: PWS Publishing, 1996

[10] D. Ioan, et al., Chameleon RF- Report D3.1, *confidential*