2014 International Joint Conference on Neural Networks (IJCNN)
July 6-11, 2014, Beijing, China

Splitted Neural Networks for better performance of antenna optimization

Linh Ho Manh[1], Francesco Grimaccia[1], Marco Mussetta[1], Paola Pirinoli[2], Riccardo E. Zich[1]
[1]Politecnico di Milano, Department of Energy, Milano, Italy
E-mail: linh.ho@polimi.it
[2]Politecnico di Torino, DET, Torino, Italy

***Abstract*—In recent years, evolutionary algorithms have been successfully adopted for the optimization of various electromagnetic problems. One of the most common electromagnetic application is in the framework of microstrip antennas, thanks to the advantage of being low cost and low profile. In order to reduce the computational effort of the electromagnetic optimization, a suitable equivalent model by ANN has been created in order to substitute the commercially available full-wave analysis solvers. With the aim of reducing committed error level, a new solution of multiple neural networks instead of one network is presented. In addition, efficiency of new training scheme is also shown in Numerical results section. The effectiveness of proposed techniques will be illustrated by optimizing a particular type of antenna, namely proximity coupled feed.**

I. Introduction

Thanks to the advantage of being low-cost and low-profile, microstrip antenna has been successfully adopted in a wide range of applications. In the literature, typical proximity coupled-feed microstrip antennas have been carefully studied by full-wave spectral analysis in [1]. The dual rectangular ring configuration, reported in Figure 1, yields more degrees of freedom to the designers but introduces in the same time more complexity. This structure has been first optimized in [2], and later on has been considered as an electromagnetic (EM) benchmark for comparing the effectiveness of different optimizers in [3]. In order to reduce the computational effort of traditional approaches (Figure 2), a fast and accurate model has been first introduced in [4], where this simplified equivalence has been embedded and directly managed by global optimizers, as shown in Figure 3.

In [4], ANN systematic pattern has been trained by a Gradient Descent Method: the well-known Error Backward Propagation. In this context, the aforementioned antenna becomes the benchmark test for the proposed approach to deal with a typical EM problem, in terms of effectively managing both the non-linear complexity and the ANN dimensions and characteristics. A new solution where separated neural networks in order to reduce error commited by the surrogate model have been adopted is here proposed. The implemented training rule is the Second-derivative Levenberg-Marquandt algorithm. The new proposed approach shows significant improvements both in terms of time convergency and accuracy. All the numerical results will be presented in detail in the next sections.

II. Optimization tool and model description

A. Meta-Particle Swarm Optimization

The construction of conventional PSO is based on a model of social interaction between independent individuals (particles) and uses social knowledge (i.e., the experience accumulated during the evolution) [5]. A specific class of PSO that namely Meta-PSO has been used in literature [2] and its effectiveness can be further enhanced by keeping swarms apart from each other and widening in this way the global research [6],[7]. This ability can be fulfilled by introducing an inter-swarm repulsion: it allows the best swarm to keep exploring the surroundings of the current best position, whereas other swarms are repelled and obliged to extend the search in other points of the space, hence improving the possibility of escaping from the local minimum. In the Differentiated Meta-PSO, velocities V and positions X of particles are updated at

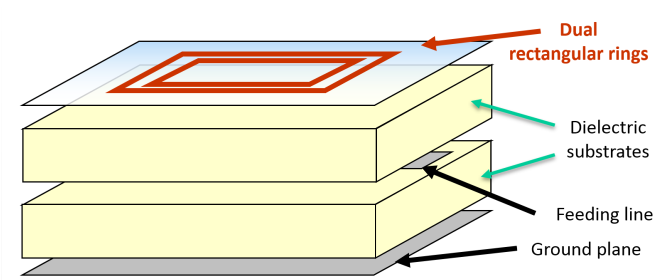

Fig. 1: The considered test antenna

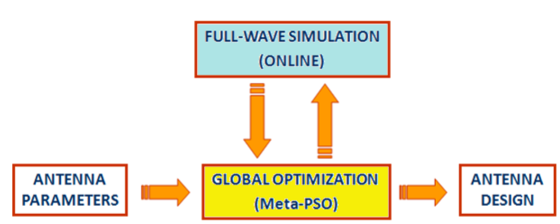

Fig. 2: Block diagram of a traditional antenna optimization scheme.

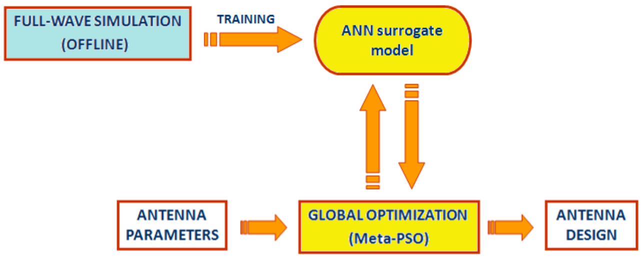

Fig. 3: Block diagram of ANN-assisted optimization.

each time step $k > 1$ according to the following equation (1):

$$V_{j,i}{}^{(k+1)} = \omega^{(k)} V_{j,i}{}^{(k)} + \phi\eta_1(P_{j,i} - X_{j,i}{}^{(k)}) + \phi\eta_3(S_j - X_{j,i}{}^{(k)}) + \delta_{j,i}{}^{(L1)}\phi\eta_2(G - X_{j,i}{}^{(k)}) \quad (1)$$

while the updating rule for the velocity of the position remains the same (2):

$$X_{j,i}{}^{(k+1)} = X_{j,i}{}^{k} + V_{j,i}{}^{(k+1)} \quad (2)$$

B. Extracting the training data

Sampling the target data and using the neural networks are two discrete steps. Regarding this ”Regular meothodology”, the desired outputs have been obtained by a full-wave analysis from formally chosen geometrical inputs in the possible region of interest. For each parameter 5 values have been considered, and more variables to be optimized also means that the needed training set grows exponentially, as reported in Table I

The knowledge extracted from physical models may be used as the target data for training the Artificial Neural Network. As can be seen from Table I, the more inputs to be optimized, the more training data need to be extracted. This possibly leads to the huge amount of target data, which can be tackled by the use of ”Irregular training” presented in the next section.

C. Optimization scheme

Geometrical antenna parameters are directly managed by Meta-PSO optimizer, each set of inputs represents one specific antenna configuration. Once ANNs are trained succesfully, they will be used as substitution model for full-wave analysis. Since ANN architecture only deals with binary and simple activation function, this surrogate model saves a critical amount of execution time with respect to very complex even commercialy available or in house university developed electromagnetic solvers. The best results ever found by ANN will be validated by full-wave analysis in order to check the accuracy of the simplified model.

TABLE I: Computational cost for different problems

Assessments	3 inputs	4 inputs	5 inputs
Number of samples	25	125	625
Time consumption	40 mins	3.5 hour	18 hours

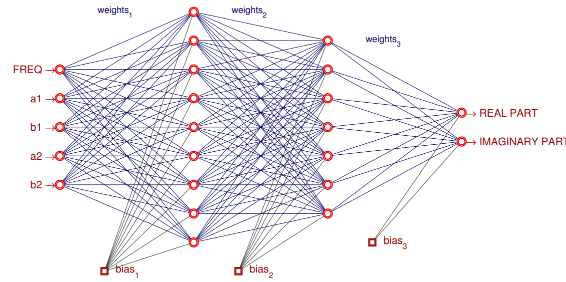

Fig. 4: The considered multilayered perceptron structure, with 5 inputs and 2 outputs neurons and 2 hidden layers of 9 and 7 neurons, respectively.

III. Artificial Neural Network

A. Feed forward multi layer topology

An artificial neural network consists of a pool of simple processing units (neurons or cells) which communicate by sending signals to each other over a large number of weighted connections [8]. It is also known that ANN is a self-adaptive modeling tool that changes its structure on the basis of external or internal information that flows through the network during the learning phase [9],[10]. The resulting network structure is the one reported in Figure 4, where the input nodes correspond to the geometrical parameters of the antenna, as defined in [4]. The input composition in each neuron is made by a nonlinear weighted sum:

$$f(x) = k(x) \sum_i w_i g_i(x) \quad (3)$$

where $k(x)$ is sigmoid, a nonlinear activation function, described as

$$k(x) = \frac{1}{1 + e^{-x}} \quad (4)$$

B. Error Backward Propagation

Error Backward Propagation (EBP) is based on the gradient descent algorithm [11]. EBP propagates error backwards through the network to allow the error derivatives for all network weights to be efficiently computed. The objective is to minimizing the mean-squared error betwween the network outputs, $f(x_i)$ and target data y_i.In general, Backpropagation algorithms update weights between layers based on the gradient of error function:

$$E = \frac{1}{2}\|f(x,i) - y(i)\| \quad (5)$$

C. Levenberg-Marquandt Training

However, when dealing with large-scale problem with huge amount of data set, EBP algorithm is not adequate to handle that kind of sophisticated problem. In order to tackle this issue, a second-order algorithm namely Levenberg-Marquandt (LM) is adopted[12]. LM algorithm uses second-order derivative of total error function for weight updates. In this technique, the Hessian matrix is used in order to save more training time of

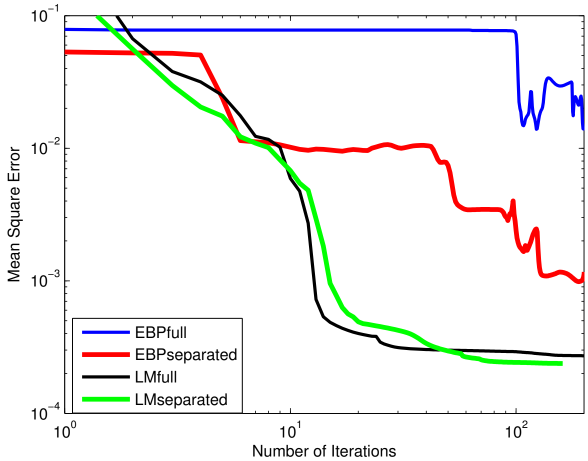

Fig. 5: Training error level as a function of number of iterations versus proposed methods.

ANN. Training is a time and memory consuming process and it is the most critical phase in the ANN setup [13]. The update rule for the weights of LM method can be illustrated as:

$$w_{k+1} = w_k - \left({J_k}^T J + \mu I\right)^{-1} J_k e_k \tag{6}$$

The idea is to create multiple and smaller neural networks and train them at the same time, such that training effort will be minimzed. In order to check out the robustness of proposed method, numerical comparisons are presented in the next section.

IV. Preliminary Results

In this optimization scheme, first reflection coefficients are retrieved by full-wave analysis and then they are used as training set data for ANN training.It is also worth noting that antenna radiation is a lossy process and return loss is always a complex number. Reflection coefficient is separated into two part: Real and Imaginary before being recombined to produce Amplitude which is the main interest in terms of bandwidth optimization problem. In stead of one network of two outputs, we separeate it into two distinguished networks with one output for each: Real and Imaginary. The dimension of splitted neural network is reduced: 5,4 neurons for first and second hidden layer respectively. The original neural network has the dimension of 9 and 7 neurons for first and second hidden layer.

Figure 5 illustrates the robustness of proposed method: LM training for two separated network. For what concerns the EBP algorithm, the division of NN decline significantly error committed to the value of 0.001. However, as reported in Figure 5, LM is proved to be more effective in minimizing the error grade. Both full network approach and separated one demonstrate the great improvement in solution accuracy. The best result is achieved by implementing LM-2 outputs.

The effectiveness of a proper ANN has been observed, considering both its numerical efficiency and the error introduced by the model. Regarding the 3D plot in Figure 6, the axis of Frequency remains unchanged since we need to investigate the structure in fixed frequency range (from 1.5 GHz to 3.5 GHz) with the resolution of 400 steps. The other interval is geometrical parameter b_1 that has been discretized with 17 samples each. What attained from the full-wave approach is considered as target value for the training of Neural Network. The color bar of left subfigure stand for the change of the amplitude from 0 to 1 while that of the right subfigure (from 0 to 0.2) indicates the error introduced by ANN approximation. As can be seen in the plots, at some certain cases, the largest error introduced by ANN model is approximately 0.1; this difference can be neglected since we exploit ANN as an effective tool to minimize the computation effort.

V. New Scheme of Extracting Target Data

As aforementioned, regular sampling of target data for training Neural Network may far exceed the required amount. In order to strictly control the training process, training process is stopped once we know ANN model is robust enough.In new proposed "Irregular scheme", the geometrical parameters are arbitrarily selected by optimizer to generate different antenna configurations. The data from unsatisfied antenna structures will be used as the training set for surrogate model. Artificial Neural Network in this context can be considered as a black box that can auto-correct itself on the basis of prior knowledge. After a certain of time inserting the new training data, an appropriate ANN can be found.

Efficiency comparison is shown in Figure 7, a maximum value of error at some certain values of interval is recorded less than 0.08 in the interval of [0,1]. This fact again confirms that the differences between ANN approximations and physical characterization are negligible . Although the error is slightly higher in this new scheme with respect to the regular one, when data are processed by the optimizer it still can find the antenna configurations satisfying the design constraints. The best results ever found by ANN model again have been validated by full-wave analysis. It also worth noting that new scheme saves a large amount of computational time and, more importantly, ANN surrogate model and global optimizer combines smoothly to form a new and hybrid class of optimization.

VI. Numerical Results and Conclusion

After the optimization run, the resulting geometrical configurations are validated by full-wave analysis. Figure 8 shows the comparisons between the different uses of ANN (by LM training). It demonstrates that all proposed methods have a good match with target data. However, the 2 output solution appears to be the best since the output data is closer to outputs since the absolute difference between the target data and ANN outcome is just 0.0005.

The aim of implementing multiple neural networks is to reduce mean-square error therefore ANN convergence will be speed up. This approach shows a great prospect to be employed in "Irregular sampling"

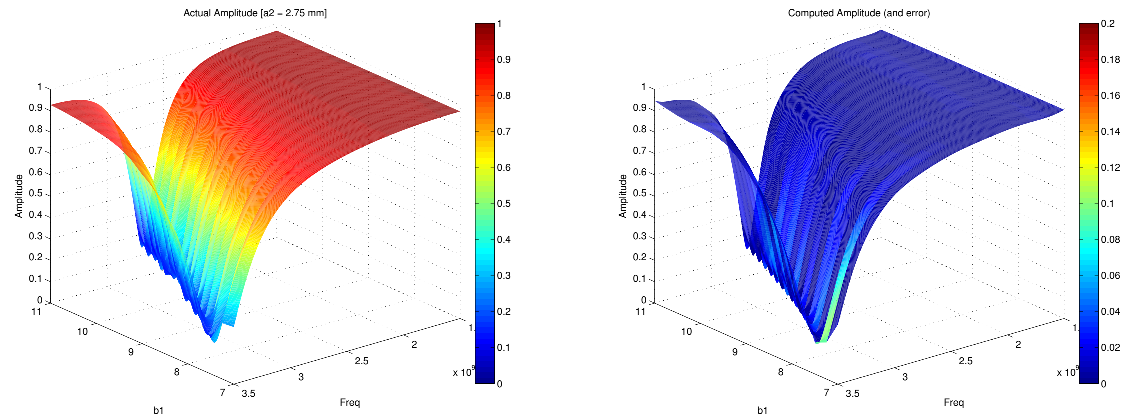

Fig. 6: Reflection coefficient amplitude versus b_1 of the proposed antenna configuration, computed with the full wave (left) and reconstructed with the ANN (right) in the certain frequency range.

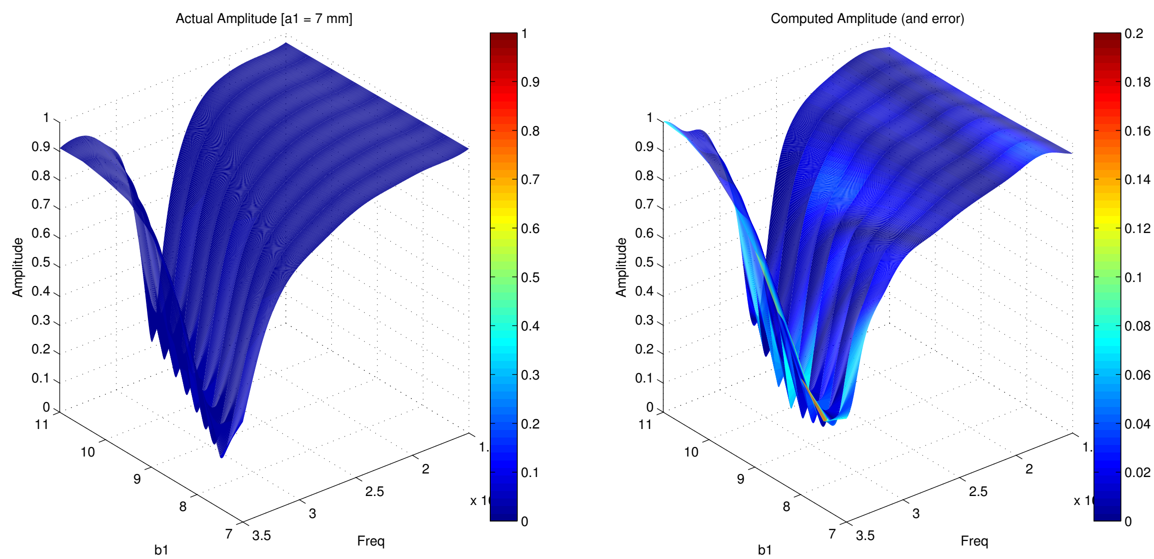

Fig. 7: Numerical comparison between the two ANN training schemes

TABLE II: Computational cost for different approaches

Methods	Conventional	Regular sampling	Irregular sampling
Number of assessments	300	125	90
Time consumption	10 hours	4.7 hours	3 hours

REFERENCES

[1] Pirinoli, P.; Vecchi, G.; Orefice, M., "Full-wave spectral analysis and design of annular patch antenna with electromagnetically coupled microstrip feed line," Antennas and Propagation, IEEE Transactions on , vol.52, no.9, pp.2415,2423, Sept. 2004.

[2] Ho Manh Linh, M. Mussetta, F. Grimaccia, R.E. Zich, "Differentiated Meta-PSO for rectangular ring Antenna with Proximity coupled feed", APS 2013, Orlando.

[3] Ho Manh Linh; Mussetta, M.; Grimaccia, F.; Zich, R.E.; Pirinoli, P.; Chien Ngoc Dao, "Multi-layer antenna design with evolutionary optimization," Electromagnetics in Advanced Applications (ICEAA), 2013 International Conference on , vol., no., pp.1533,1536, 9-13 Sept. 2013

[4] Ho Manh Linh; Mussetta, M.; Grimaccia, F.; Zich, R.E.; Pirinoli, P., "Antenna design and optimization based on ANN techniques," Antennas and Propagation (EuCAP), 2014 8th European Conference

[5] Robinson, J.; Rahmat-Samii, Y., "Particle swarm optimization in electromagnetics," Antennas and Propagation, IEEE Transactions on , vol.52, no.2, pp.397,407, Feb. 2004

[6] Selleri, S.; Mussetta, M.; Pirinoli, P.; Zich, R.E.; Matekovits, L., "Some Insight Over New Variations of the Particle Swarm Optimization Method," Antennas and Wireless Propagation Letters, IEEE , vol.5, no.1, pp.235,238, Dec. 2006

[7] Selleri, S.; Mussetta, M.; Pirinoli, P.; Zich, R.E.; Matekovits, L., "Differentiated Meta-PSO Methods for Array Optimization," Antennas and Propagation, IEEE Transactions on , vol.56, no.1, pp.67,75, Jan. 2008.

[8] Dalton, J.; Deshmane, A., "Artificial neural networks," Potentials, IEEE,

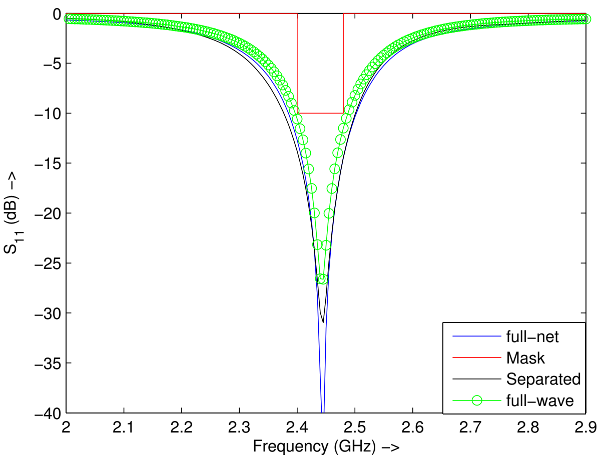

Fig. 8: ANN optimization and full-wave analysis validation.

vol.10, no.2, pp.33,36, April 1991

[9] A. Freni, M. Mussetta, P. Pirinoli, Neural Network Characterization of Reflectarray Antennas, International Journal of Antennas and Propagation, vol. 2012, Article ID 541354, 2012.

[10] Nesil, S.; Gunes, F.; Ozkaya, U., "Phase characterization of a reflectarray unit cell with Minkowski shape radiating element using Multilayer Perceptron Neural Network," Electrical and Electronics Engineering (ELECO), 2011 7th International Conference on , vol., no., pp.II-219,II-222, 1-4 Dec. 2011.

[11] Robustillo, P.; Encinar, J.A.; Zapata, J., "ANN element characterization for reflectarray antenna optimization," Antennas and Propagation (EUCAP), Proceedings of the 5th European Conference on , vol., no., pp.957,960, 11-15 April 2011.

[12] Hao Yu and B. M. Wilamowski, "LevenbergMarquardt Training Industrial Electronics Handbook, vol. 5 Intelligent Systems, 2nd Edition, chapter 12, pp. 12-1 to 12-15, CRC Press 2011.

[13] L. H. Manh, F. Grimaccia, M. Mussetta, and R. E. Zich, "Optimization of a dual ring antenna by means of artificial neural network," Progress In Electromagnetics Research B, Vol. 58, 59-69, 2014.