A Mesh-less Method Based on Artificial Neural Network for Solving Poisson Equation

Hai-Li Zhang
School of Electronic Engineering
University of Electronic Science and Technology of China
Chengdu 611731, China
1540350232@qq.com

Xiao-Yang He and Ming-Yao Xia
School of Electronics Engineering and Computer Science
Peking University
Beijing 100871, China
myxia@pku.edu.cn

***Abstract*—A new mesh-less method based on artificial neural network is presented for solving Poisson equation. Artificial neural network functions are used as trial functions instead of traditional superposition of basis functions. The least square method is adopted to construct the error function according to the electromagnetic problems. Minimization of the error function is implemented by adjusting the neural network parameters using gradient descent method. The new algorithm avoids the storage and solution of massive matrix equations.**

Keywords—Poisson equation; artificial neural network; mesh-less method.

I. INTRODUCTION

Typical methods in computational electromagnetics to solve the Poisson equation include finite difference method (FDM) [1], finite element method (FEM) [2], and the method of moments (MoM) [3].

The meshless method is a kind of numerical algorithm based on nodes rather than on meshes. In 2001, Ho [4] used minimized least square (MLS) method to solve the Poisson equation. The weak form of the Poisson equation was employed to find the potential distribution of metal slots. In 2003, Yang [5] proposed the wavelet meshless Galerkin hybrid method to analyze the electric field problems; Lu [6] used the local Petrov-Galerkin meshless method to calculate the static potential distribution. In 2004, Zhao [7] proposed the MLPG-FEM hybrid algorithm to calculate the static field potential distribution; Lai [8] proposed RBF to solve the two-dimensional time domain electromagnetic problem. The application of various meshless methods in the field of computational electromagnetics has made great advances in recent years and has become an attractive research topic. In this work, we propose a new meshless method by using the artificial neural network to solve electrostatic problems.

II. FORMULAITON

A. The Principle of Artifical Neural Network

Artificial neural network (ANN), based on the basic principle of biologicaol neural networks, simulates a mathmatical model of the brain's nervous system's handling of complex information [9]. Neural networks are composed of a large number of simple interconnected units. The complex ANN is a non-linear system that can express complex logic and nonlinear relationships.

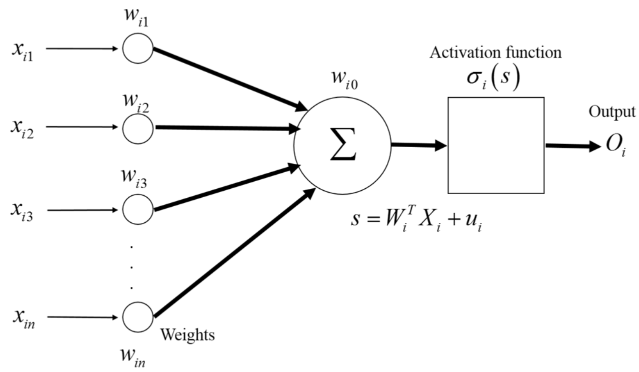

Fig. 1. Illustration of working manner of a neuron.

As shown in Fig.1, the input of the first neural network unit is a list of vectors $X_i = [x_{i1}, x_{i2}, \cdots, x_{in}]^T$. The connection between every two nodes represents a weighted value for the signal passing through the connection. The neural network simulates the signal processing in this way. Each input quantity gets a weight before the neural unit is input, and the weight vector is $W_i = [w_{i1}, w_{i2}, \cdots, w_{in}]^T$. Each node represents a specific output function σ , called an Activation Function. The output of the network depends on the structure of the network, the way the network is connected, weights and activation functions. The input and output of a single neuron can be described as:

$$O_i = \sigma_i \left[w_{i0} + \sum_{j=1}^{n} w_{ij} x_j \right] \tag{1}$$

Activation function plays an important role in the structure of neural networks. The nonlinear activation function can introduce the nonlinear feature to the neural network system. In this paper, we adopt the sigmoid function as activation function:

$$\sigma(x) = \frac{1}{1+e^{-x}} \tag{2}$$

Single-layer neural networks can express very few problems. For complex multidimensional problems, the use of multi-layer neural networks is necessary.

BP (Back-Propagation) neural network is a multi-layer feed-forward neural network, which was pioneered by a group of scientists headed by Rumelhart and Mccelland in 1986 [10]. The BP neural network consists of two processes, namely the forward propagation of information and the back propagation of errors. When the information is transmitted in the forward direction, the input sample data is transmitted from the input layer, processed layer by layer through the hidden layer, and finally passed to the output layer. If the output value obtained by the output layer does not match the expected actual value and the initial error set by the neural network, the back propagation of the error is transferred. When the error propagates backward, the output error is inverted in layers from the hidden layer to the input layer, and the error is evenly distributed to all neuronal units in each layer, so as to obtain the error signal of each layer of the neuron unit. This error signal can be used as a basis for correcting the weight of each unit. The two processes are continuously and cyclically carried out in the neural network so that the weights are constantly adjusted. This is the learning and training process of the network. When the BP neural network operation continues until the error of the network output is reduced to an acceptable level for initial setting, or it terminates when it reaches a predetermined number of learning times.

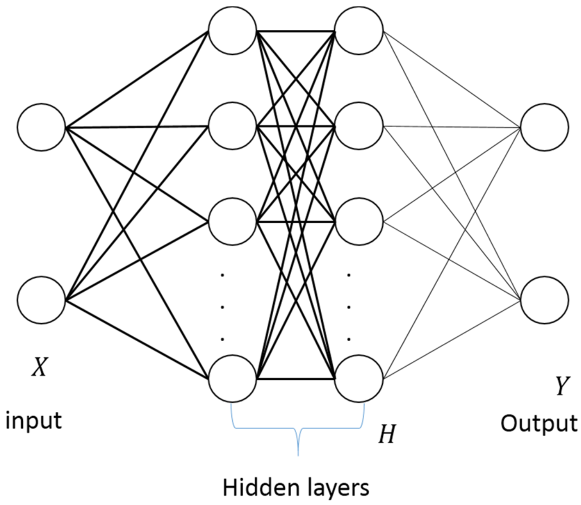

Fig. 2. The structure of a neural network.

For an output of the hidden layers being $H=(h_1,h_2,\ \dots\ ,h_n)^T$, the output of the network can be given as:

$$Y=\sum v_i\sigma(s_i) \tag{3}$$

where

$$s_i=\sum_{j=1}^{n}w_{ij}h_j+u_i \tag{4}$$

and w_{ij} denotes the weight from the input unit j to the hidden unit i, v_i represents weight from the hidden unit i to the output, u_i is the bias of hidden unit i, and $\sigma(s)$ is the sigmoid activation function. Now the derivative of networks output Y with respect to input vector x_k is

$$\begin{aligned}\frac{\partial Y}{\partial x_k}&=\frac{\partial}{\partial x_k}\left[\sum_{i=1}^{N_H}v_i\sigma\left(\sum_{j=1}^{n}w_{ij}h_j+u_i\right)\right]\\&=\sum_{i=1}^{N_H}v_i\sum_{j=1}^{n}\left[w_{ij}\frac{\partial h_j}{\partial x_k}\sigma^{(1)}\left(\sum_{j=1}^{n}w_{ij}h_j+u_i\right)\right]\end{aligned} \tag{5}$$

The derivative of network output Y with respect to the parameters of the original network can be easily obtained as

$$\frac{\partial Y}{\partial v_i}=\sigma_i\left(\sum_{j=1}^{n}w_{ij}h_j+u_i\right) \tag{6}$$

$$\frac{\partial Y}{\partial u_i}=v_i\sigma_i^{(1)}\left(\sum_{j=1}^{n}w_{ij}h_j+u_i\right) \tag{7}$$

$$\frac{\partial Y}{\partial w_{ij}}=x_jv_i\sigma_i^{(1)}\left(\sum_{j=1}^{n}w_{ij}h_j+u_i\right) \tag{8}$$

B. A New Method for Solving Poisson Equation

Differential equations for electromagnetic field problems can be expressed as:

$$F\Phi=f \tag{9}$$

where F is a differential operator, f is excitation function, and Φ is the unknown function.

In electromagnetics, Poisson equation is a very important equation. In actual engineering, there are very few cases where analytical solutions can be obtained. So approximate methods are necessary. The most extensive methods are the Ritz method and the Galerkin method.

Both the Ritz method and the Galerkin method expand the trial function $\tilde{\Phi}$ approximately as:

$$\tilde{\Phi}=\sum_{j=1}^{N}c_jv_j \tag{10}$$

where v_j is the basis function which be defined in the solution area, and c_j is the unknown coefficient.

The Ritz method is a variational method. By constructing a functional and finding the minimum value of the functional to determine the minimum value of a variable, an approximate solution can be obtained. The Galerkin method belongs to the residual weighting method and establishes an equation to solve the expansion coefficient.

In this paper, BP neural network functions are adopted as trial functions instead of the basis function. The electrostatic potential in the region of computation with boundary condition can be described as:

$$\nabla\cdot(\varepsilon(r)\nabla\phi(r))=-\rho(r) \tag{11}$$

$$\phi|_{\Gamma 1}=f_1 \tag{12}$$

$$\left.\frac{\partial\phi}{\partial x}\right|_{\Gamma 2}=f_2 \tag{13}$$

We use the least square method to build the error function:

$$\begin{aligned}I&=\int_{\Omega}\left|\nabla\cdot(\varepsilon(r)\nabla\phi(r))+\rho(r)\right|^2\mathrm{d}\Omega\\&+\int_{\Gamma 1}\left|\phi(r)-f_1\right|^2\mathrm{d}\Gamma_1+\int_{\Gamma 2}\left|\frac{\partial\phi(r)}{\partial \mathrm{x}}-f_2\right|^2\mathrm{d}\Gamma_2\end{aligned} \tag{14}$$

It is easy to demonstrate that if $I=0$ the $\phi(r)$ will satisfy the differential function and boundary conditions. According the uniqueness theorem of the electromagnetic field, if the error function $I=0$, the $\phi(r)$ is the only solution.

We discretize the error function, and replace the trial function by the output artificial neural function $Y(r_i)$:

$$I_D = \frac{1}{N_1+N_2+N_3}\sum_{r_i\in\Omega}^{N_1}\left|\nabla\cdot\left(\varepsilon(r_i)\nabla Y(r_i)\right)+\rho(r_i)\right|^2 + \sum_{r_i\in\Gamma 1}^{N_2}\left|Y(r_i)-f_1(r_i)\right|^2 + \sum_{r_i\in\Gamma 2}^{N_3}\left|\frac{\partial Y}{\partial x}(r_i)-f_2(r_i)\right|^2 \tag{15}$$

where N_1 , N_2 , N_3 are numbers of point of solving zone, boundary Γ_1 and boundary Γ_2 .

The next step is to adjust the parameters of ANN function $Y(r_i)$. The method of gradient descent is adopted to minimize the error function I_D . The process of ANN parameter updating is described as:

$$u_i = u_i + \alpha\frac{\partial I_D}{\partial u_i} \tag{16}$$

$$v_i = v_i + \beta\frac{\partial I_D}{\partial v_i} \tag{17}$$

$$w_{ij} = w_{ij} + \gamma\frac{\partial I_D}{\partial w_{ij}} \tag{18}$$

where α , β and γ are the learning rates.

III. Numerical Results

In this paper, we analyze the electric potential of a 2-dimensional coaxial infinitely long circular waveguide. The cross-section of the waveguide is shown in Fig. 3. The radius of the central cylinder is 0.02 meters and the potential is 1. The radius of outer cylindrical shuck is 0.04 meters and the potential is zero.

We have constructed a neural network with 3 hidden layers and every layer has 20 neurons. The potential distribution in the cross-section of the waveguide is obtained by the present method as shown in Fig. 3. Fig. 4 shows the average error as a function of the number of gradient drops. The numerical results show that the complex electromagnetic problems can be expressed through the training parameters ANN function.

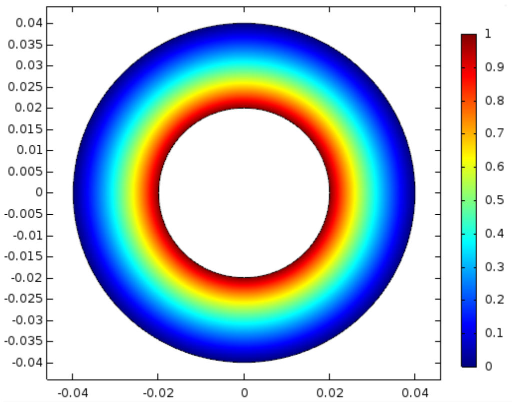

Fig. 3. The potential distribution in cross-section of co-axial cable.

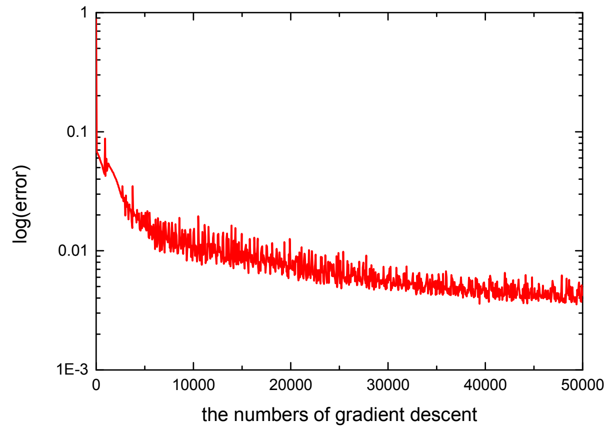

Fig. 4. The convergent property of the ANN training.

IV. Conclusion

In this paper, a new mesh-less method based on artificial neural network (ANN) to solve Poisson equation is presented. It gives the results as a weighted superposition of the ANN outputs. There is no mesh dependence, which reduces the difficulties caused by mesh distortion. The meshless method lies on node distribution sampling, and the adaptation is very strong. It can fit arbitrary regions and multi-scale structural problems. The pre-processing of the meshless method is easy for analysis of complex three-dimensional structures as long as node position information is provided. Neural activation functions rather than traditional basis functions are used. A special error function is constructed, and the gradient descent method is used to minimize the error function. The method can avoid the storage and solution massive matrix equations. It has certain prospects for solution of general electromagnetic problems.

Acknowledgment: This work is supported by NSFC under Project 61531001 and 61271032.

References

[1] A. Taflove and S. C. Hagness, Computational Electrodynamics: The Finite-Difference Time-Domain Method, 3rd ed. Artech House, 2005.
[2] J.-M. Jin, The Finite Element Method in Electromagnetics, 3rd ed. Wiley-IEEE Press, 2014.
[3] R. F. Harrington, Field Computation by Moment Methods. Wiley-IEEE Press, 1993.
[4] S. L. Ho, S. Yang, J. M. Machado, et al. Application of a meshless method in electromagnetic. IEEE Trans. Mag., 2001, 37(5):3198-3202.
[5] S. Y. Yang, G. Z. Ni, J. R. Cardoso, et al. A combined wavelet-element free Galerkin method for numerical calculations of electromagnetic fields. IEEE Trans. Mag., 2003, 39(3): 1413-1416.
[6] H. F. Lu, H. Jiang. Application of the meshless local Petrov-Galerkin method in electromagnetics. IEEE Trans. Mag., 2003, 39(3):799-801.
[7] M. L. Zhao, Y. F. Nie, C. W. Zuo. A new coupled MLPG-FE method for electromagnetic field computations. The 3rd International Conference on Computational Electromagnetics and its Applications, Beijing, China, 2004:29-32.
[8] Artificial Neural Networks. Encyclopedia of Microfluidics & Nanofluidics: 23-33.
[9] J. L. Mcclelland and D. E. Rumelhart. A Distributed Model of Human Learning and Memory Parallel Distributed Processing. MIT Press, 1986:170-215.