A Neural-Network Model of the Input/Output Characteristics of a High-Power Backward Wave Oscillator

Chaouki Abdallah, *Senior Member, IEEE*, Wei Yang, Edl Schamiloglu, *Senior Member, IEEE*, and Larald D. Moreland

***Abstract*— This paper discusses an approach to model the input/output characteristics of the Sinus-6 electron beam accelerator-driven backward wave oscillator. Since the Sinus-6 is extremely fast to warrant the inclusion of dynamical effects, and since the sampling interval in the experiment is not fixed, a static continuous neural network model is used to fit the experimental data. Simulation results show that such a simple nonlinear model is sufficient to accurately describe the input/output behavior of the Sinus-6-driven backward wave oscillator (BWO) and that the fitted output waveforms are basically noiseless. This model will be used to control the BWO in order to maximize the radiated power and the efficiency. This paper is also intended to introduce high-power microwave researchers to control concepts that may enhance the outputs of a wide spectrum of sources.**

I. Introduction

THE UNIVERSITY of New Mexico Pulsed Power and Plasma Science Laboratory, in collaboration with the systems group are currently engaged in an experimental/theoretical study of methods to identify and control the high-power repetitively pulsed electron beam accelerator known as the Sinus-6. Initial experimentation with the Sinus-6-driven backward wave oscillator (BWO) has been reported elsewhere [1], and has yielded input/output data which are used in this research. This paper focuses on the model identification for the Sinus-6 BWO.

It is well known that a feedback control system typically consists of a "plant" and a "controller," where the plant is generally expressed as a mathematical model which describes the behavior of the real-world physical system (Sinus-6 BWO), and the goal of the controller is to use the relevant plant information to obtain an overall behavior which satisfies some performance objectives. Obviously, how good a performance we obtain depends on our knowledge of the controlled plant. The model we search for is also dictated by, and may dictate, the control approach that we may eventually select for the Sinus-6 [2]. Many control approaches are available to us, a sample of which are classical control, learning control, and robust control. These different approaches require different mathematical models. For example, a linear-time-invariant (LTI) nominal model is required for the classical control approach, both an LTI nominal model and uncertainty bound model are required for the robust control approach, while nonlinear models which generally carry more information about the plant can be used for learning control approaches. In any case, all these approaches require the knowledge of a nominal model. The initial stage of our research therefore focuses on obtaining a nominal model which can describe the behavior of the Sinus-6 BWO.

Manuscript received September 18, 1995; revised February 13, 1996. This work was supported through a High Energy Microwave Devices Consortium funded by an AFOSR/DOD MURI grant and administered through Texas Tech University, Lubbock.

The authors are with the Electrical and Computer Engineering Department, University of New Mexico, Albuquerque, NM 87131 USA (e-mail: edl@eece.unm.edu).

Publisher Item Identifier S 0093-3813(96)04640-1.

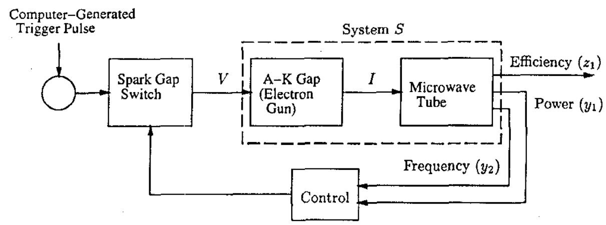


Fig. 1. Block diagram of system.

Due to the complexity of obtaining a physics-based model of high-power BWO's, researchers utilize fully electromagnetic particle-in-cell (PIC) codes like MAGIC [3] in order to simulate certain aspects of the operation of these devices. In this paper, we choose instead to build a model based on the input/output data with the physics providing guidance, but little influence. Our paper is thus in the spirit of [4] where a neural network model was used to control a tokamak plasma.

This paper is organized as follows. Section II describes the identification problem and setup. Section III contains the identification results and their interpretation, while Section IV contains our conclusions.

II. Identification

Identification in controls terminology refers to the process of obtaining a mathematical model that can explain the input/output behavior of a physical system. A block diagram of the experimental setup is shown in Fig. 1. (Detailed information on the experiment can be found in [1].) The block labeled System S in the figure is identified as the mathematical model in our experiment. The model of the high-power BWO consists of an A–K gap (electron gun) delivering an intense electron beam current I that is guided through a slow wave structure by a strong axial magnetic field. The only input into this system is the cathode potential V and the two measured outputs are the microwave power y_1 and the microwave frequency y_2. The microwave conversion efficiency z_1 is obtained by dividing the output microwave power by the input beam power $V \times I$.

In the most simple control objective, the output from the system would be fed back into the input to adjust the voltage applied to the system to maximize the power or the efficiency, or adjust the output frequency. From the research described in [1], we have access to a set of input cathode voltage V and output microwave efficiency z_1, power y_1, and frequency y_2. Note that y_1 and y_2 represent data that are physically measurable, and available for feedback. On the other hand, z_1 denotes a signal that is calculated and is required to satisfy some performance objectives, while not being available for feedback. As stated earlier, the task of identifying the system consists of obtaining a mathematical model which describes the behavior of the Sinus-6 so that it may be later controlled. We are then interested in finding a mathematical model which can predict the future behavior of the Sinus-6. In the feedback control configuration of Fig. 1, we cannot predict what the control input signal V to the plant is going to be, since it is the sum of an external command signal and the output signal of the controller. Generally, the control input signal V to the plant is different in the feedback control configuration from the experimental input data collected in an open-loop fashion. We therefore require that our mathematical model "generalize" to explain the unknown control input signal set. This is the so-called *model validation problem.* Ljung [5] gives an overview of model validation in the standard identification framework. He suggests that one can take part of the experimental data for identification purposes, while keeping the remaining part for validation. This is known as the *cross validation approach.* This approach works best when there is an abundance of data. In this study, we only had access to a total of 318 experimental data for four experiments. Therefore, all of our experimental data is used in the identification procedure. The judgment of the "goodness" of our model will be evaluated by future experiments. It is the ability to extrapolate that could justify the use of neural networks.

There is a large body of literature on system identification. Many identification approaches and methods are available to us, depending on the type and format of the available data. For our particular case, the sampling interval of the experimental data is not fixed. More importantly, due to the extremely fast dynamics of our system, a static neural network model is sufficiently rich to explain the experimental data for the Sinus-6. Note that our identification scheme relies on open-loop input signals since we do not currently have a feedback controller on the Sinus-6 BWO. The input signals and the different A–K gap settings were chosen to be representative of the normal operation of the high-power microwave (HPM) source. It is of course conceivable that the closed-loop input generated by a feedback controller may be outside the range of the open-loop signals. We plan to address this issue should it arise by iterating our modeling/control steps.

When discussing neural networks, we are typically referring to a system built by linearly combining a large collection of simple computing devices (i.e., nodes), each of which performs a nonlinear transformation σ (in general, a sigmoid function) on its inputs. These inputs are either external signals supplied to the system, or the outputs of other nodes (see, for example, Fig. 2). Neural network models have two important characteristics. First, since they consist of many nodes, individual nodes carry out only a small amount of the overall computational task. Thus the computational load is *distributed* throughout the network. Second, the large number of *parallel* connections typically found in these systems provide many paths from input to output. These factors combine to make neural networks a very robust model of computing. In theory, damage to a few weights or nodes will not adversely affect the overall performance of the network. In fact, practical implementations indicate that the performance of neural networks tends to degrade gracefully as weights or nodes are destroyed [6], [7]. A neural network with m inputs, p outputs, one hidden layer containing L nodes (similar to the ones used in this paper) may be compactly described by

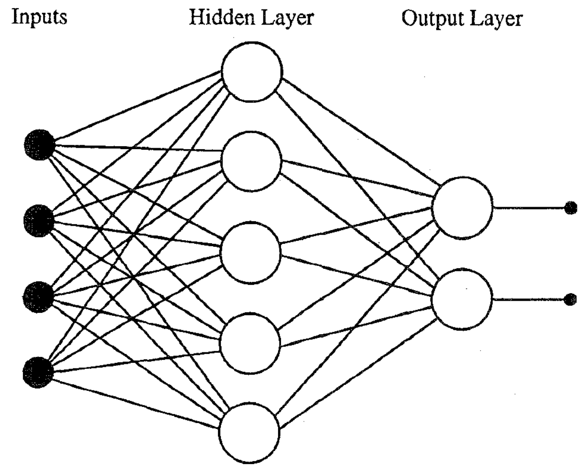


Fig. 2. A neural network architecture.

$$\begin{aligned} \alpha &= Bu; \quad \text{where } B \in \mathbb{R}^{L\times m} \\ \Sigma(\alpha) &= \sigma(\alpha) \in \mathbb{R}^{L\times 1} \\ y &= C\Sigma; \quad \text{where} \quad C \in \mathbb{R}^{p\times L} \end{aligned} \tag{1}$$

where u is the input to the neural network (V in our case) and y is the output of the network ($[y_1\ y_2]^T$ in our case). The output is then $y = C\sigma[(Bu)]$. Note that B and C are matrices of weights to be learned or programmed, and that the notation $\sigma(x)$ for $x \in \mathbb{R}^L$ denotes $\sigma(x) = [\sigma(x_1)\cdots\sigma(x_L)]^T$. In this work $\sigma(x) = \tanh(x) = (e^x - e^{-x})/(e^x + e^{-x})$.

It is now known that a one-hidden layer static network is capable of approximating an arbitrary (continuous) function. In the next section we use such neural network models to fit the experimental input/output data for the Sinus-6 driven BWO when the outputs are power, frequency, and efficiency. Considering the expense of controller design for nonlinear systems, and due to some strong linear trends in the power and frequency data, we also investigate a linear fit to the experimental data. This may be considered as a special case of (1) where the nonlinear sigmoid degenerates to a linear term $\sigma(\alpha) = A\alpha$ and the total input/output mapping degenerates into

$$y = CABu = Fu. \tag{2}$$

In our particular case, we are therefore searching for a model of the form

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} F_1(V,W) \\ F_2(V,W) \end{bmatrix}$$
$$z_1 = G(V,W) \tag{3}$$

or more compactly as

$$y = \begin{bmatrix} y_1 \\ y_2 \\ z \end{bmatrix} = \begin{bmatrix} F_1(V,W) \\ F_2(V,W) \\ G(V,W) \end{bmatrix} = F(V,W) \tag{4}$$

where W are weights that will be learned from the experimental data, and F_1, F_2 and G are given structures (linear or nonlinear).

III. Identification Results

A neural network approach has been used to fit the experimental input/output data for the Sinus-6 BWO. (See [1] for a detailed description of the experimental setup and Fig. 1 for a block diagram description.) The experimental data were collected in four separate experiments, where the A–K gap was adjusted to four different values. The A–K gap determines the electron beam diode impedance. We shall denote these four experiments as E_1, E_2, E_3, and E_4. The four intervals were divided into 95 sampling points for the first experiment, 102 sampling points for the second experiment, 78 sampling points for the third experiment, and 43 sampling points for the fourth experiment. The experimental data consist of the cathode voltage input V, the current I, and the two outputs: total peak power y_1, frequency y_2. The RF generation efficiency z_1 was calculated from the formula

$$z_1 = \frac{y_1}{V \times I}. \tag{5}$$

Both nonlinear neural network models with five weights in the hidden layer and a linear neural network model are used to fit the experimental input/output data. The objective of the fit is to minimize the following performance objective:

$$J = \frac{1}{N} \sum_{i=1}^{N} [F(V,W) - y]^2 \tag{6}$$

by a choice of the weights W. In general, this is accomplished by a gradient descent procedure of updating the weights as described, for example, in [8]. In this research, we have used the *backpropagation* training algorithm implemented in the *Neural Network* toolbox of MATLAB® [9]. Several questions related to the choice of models are discussed below.

A. Nonlinear Neural Network Models

Can a single nonlinear network explain the behavior of the Sinus-6 BWO over the four operating conditions E_1, E_2, E_3, and E_4? In other words, should we have an A–K gap-dependent model of the input–output behavior, or is a single neural network appropriate? The neural networks used have one hidden layer with five nonlinear nodes that compute the function $\sigma(x) = \tanh(x)$, and one linear output node. More explicitly, the networks are performing the following

TABLE I
PARAMETERS OF LINEAR MODELS

	E_1	E_2	E_3	E_4	Combined
y_1	$0.7746V - 81.064$	$1.175V - 309.833$	$1.1851V - 300.446$	$0.593V - 70.53$	$1.079V - 258.38$
y_2	$0.0008V + 9.3326$	$0.0009V + 9.225$	$0.0014V + 8.98$	$0.0016V + 8.755$	$0.0013V + 8.991$

TABLE II
PARAMETERS OF NONLINEAR MODELS FOR POWER OUTPUT TRAINED WITH SORTED DATA

y_1	E_1	E_2	E_3	E_4	Combined
W_1	−3.8208 −0.0136 110.5902 0.4310 0.1132	−0.6665 0.4013 0.1912 0.0102 0.0967	2.5549 2.6283 −0.0077 −0.468 31.6747	−0.2208 1.0666 0.3381 0.1507 −0.0065	0.1948 0.0411 −0.0086 −3.3343 −3.6212
b_1	6.1981 7.0324 16.8845 −53.969 −192.1514	11.1357 −22.8421 13.5039 −5.2089 8.7669	59.3318 −83.3248 3.4993 −0.4675 8.5825	−24.6393 −20.9996 −13.1203 25.5305 2.5336	−5.3918 −87.758 4.2614 −17.2114 15.6948
W_2^T	−4.4018 −91.3155 25.454 −5.7953 24.7047	−84.4896 7.3816 −20.9698 149.9427 −6.2397	−142.9513 249.5083 −190.1057 51.4265 −146.4174	24.8633 −2.9252 −8.0299 −6.4535 −115.9052	−11.4091 −1.0139 −162.4521 36.3118 −14.8563
b_2	323.7953	227.314	334.9528	197.9138	311.2597

TABLE III
PARAMETERS OF NONLINEAR MODELS FOR FREQUENCY OUTPUT TRAINED WITH SORTED DATA

y_2	E_1	E_2	E_3	E_4	Combined
W_1	−0.0675 −0.0947 0.2941 −0.0314 0.0369	−0.0272 −0.0183 0.0317 −0.0368 −0.0361	−0.0749 −0.0429 0.0188 0.0157 0.0265	0.2008 0.0204 0.0618 0.1092 −0.0516	0.0232 −0.0233 −0.0213 0.0226 0.0213
b_1	44.8811 12.0024 17.1557 16.7242 −19.7406	15.15 17.4402 −19.1209 17.6166 17.3137	−10.5713 17.3944 −9.4089 −15.4691 −13.4234	−19.1059 −10.4735 −15.9889 −19.1383 19.6815	−13.3277 9.5801 8.7238 −12.1546 −8.689
W_2^T	−0.0089 0.0402 0.1985 −0.2943 −0.222	0.0386 −0.0309 0.0079 0.1473 −0.1904	0.1773 −0.0388 0.2108 0.0198 −0.0947	0.1542 0.0409 0.0894 0.0591 −0.0331	0.0182 0.1137 −0.5746 0.0549 −0.2783
b_2	9.5841	9.745	9.4697	9.2346	9.5531

operations:

$$y = W_2 \tanh(W_1 u + b_1) + b_2 \tag{7}$$

where

$$\begin{aligned} W_2 &= [w_2^1 \quad w_2^2 \quad w_2^3 \quad w_2^4 \quad w_2^5] \\ W_1 &= [w_1^1 \quad w_1^2 \quad w_1^3 \quad w_1^4 \quad w_1^5]^T \\ b_1 &= [b_1^1 \quad b_1^2 \quad b_1^3 \quad b_1^4 \quad b_1^5]^T \end{aligned} \tag{8}$$

and where b_2 is a scalar. In fact, we can write y more explicitly as

$$y = \sum_{i=5} w_2^i \tanh(w_1^i u + b_1^i) + b_2. \tag{9}$$

The notation $\tanh(x)$ for an n-dimensional vector x denotes the vector

$$[\tanh(x_1) \cdots \tanh(x_n)]^T.$$

Tables I–IV contain the learned parameters under different conditions.

The results of this study for the case where the input is cathode voltage and the measured output is frequency are shown in Fig. 3, where the top part of the figure shows the performance of four separate nonlinear neural networks trained on the four experiments E_1, E_2, E_3, and E_4, and the bottom part of the figure shows the performance of a single nonlinear

TABLE IV
PARAMETERS OF NONLINEAR MODELS FOR EFFICIENCY OUTPUT TRAINED WITH SORTED DATA

z_1	E_1	E_2	E_3	E_4	Combined
W_1	0.0268	0.0175	0.0725	−0.0341	0.0231
	−0.0287	0.0327	−0.0266	0.0345	0.007
	0.0255	−0.0291	−0.0235	0.0336	0.001
	0.0243	−0.029	−0.0178	−0.0407	−0.00249
	−0.019	0.0324	0.0242	−0.0455	0.0048
b_1	−16.1853	−16.4042	−12.0745	16.4413	−10.9904
	13.7635	−14.8839	10.953	−16.6445	−14.1524
	−12.131	17.7466	13.6433	−21.8055	−2.585
	−11.5699	17.7088	17.371	18.787	11.8571
	12.5369	−14.7289	−14.0382	21.7969	−14.5578
W_2^T	−0.0193	0.0003	−0.0308	−0.305	0.5002
	−0.0091	−0.4369	−0.0118	−0.3064	0.7581
	0.2374	−0.0251	−0.0855	0.0156	−3.078
	−0.2618	0.0369	0.0001	0.0271	0.4596
	0.0066	0.4458	−0.0961	−0.0178	1.4762
b_2	0.1721	0.1385	0.1536	0.1029	−0.6211

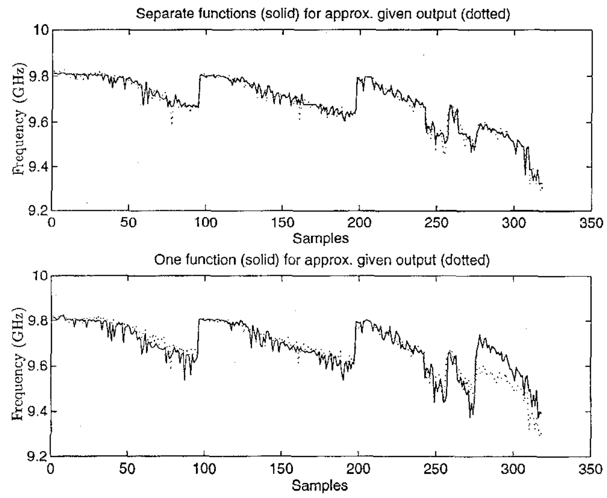


Fig. 3. Experimental frequency output and that learned by the neural network.

neural network trained on the total data. Note that the four experiments correspond to the horizontal axis labeling E_1: 1–95, E_2: 96–197, E_3: 198–276, and E_4: 277–320. As can be seen from these results, a simple neural network efficiently models the experimental input/output relations for E_1, E_2, and E_3, while there are some modeling errors for E_4. The problem is much more severe if one tries to use a single neural network for the efficiency data [10].

B. Noise

Are the collected data noisy? It may be possible that the neural networks in attempting to fit the data are also trying to fit some noise. Therefore, we need to test for the existence and the amount of noise in our data. One way to check this is to sort the input voltage data by their magnitude while keeping the corresponding outputs. Note that this amounts only to changing the order of input–output pairs and does not change the static relations between input and output. In this fashion, we can assume that the input is basically a continuously increasing function. Therefore, if a sorted output is also basically continuous, we can conclude that the output is noiseless. This idea also suggests that we can train the neural network based on the sorted experimental input/output data mentioned above. Fig. 4 shows the results of four neural

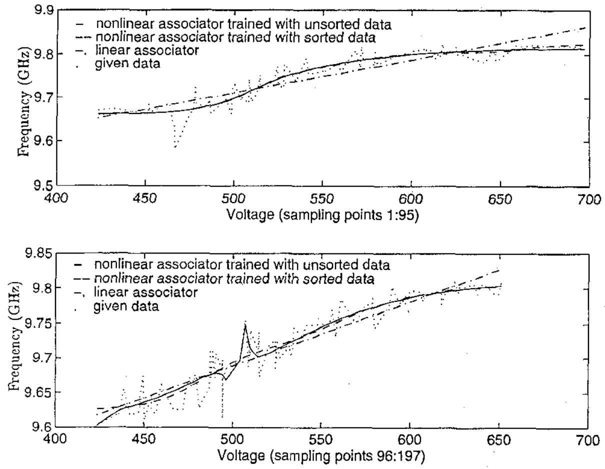


Fig. 4. Experimental frequency output and that learned by the neural network.

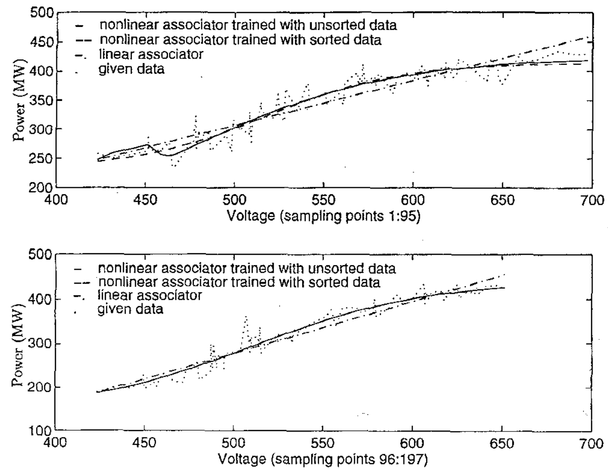


Fig. 5. Experimental total power output and that learned by the neural network.

networks trained on the four experiments phases E_1, E_2, E_3, and E_4 with sorted and unsorted microwave power data. Fig. 5 presents the same information when the output is efficiency. From these results, we can see that:

1) The waveforms obtained from both the unsorted data and sorted data are basically noiseless.
2) The waveforms obtained from sorted data are better than the waveforms trained with unsorted data in the sense that the former are less noisy than the latter.

According to results from neural networks, the order in which the training data is presented to the network should have very little effect on the learned parameters, especially as the number of training samples and training time increase [11]. However, we propose to sort the experimental input/output data before training a neural network for the purpose of removing any measurement noise.

C. Nonlinear Versus Linear

Can a linear neural network be used? This investigation is motivated by the fact that the linear model can simplify the

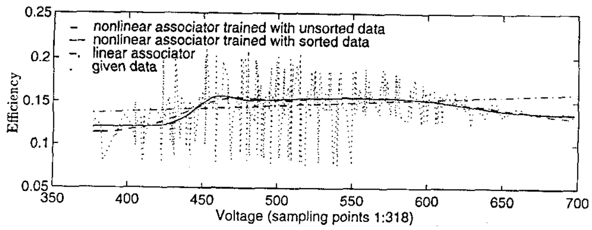


Fig. 6. Experimental efficiency output and that learned by the neural network.

control system design. The learned parameters are given in the tables displayed earlier and the results are shown in Figs. 4 and 5. It turns out that, for the case in which the outputs are frequency and total power, four linear neural networks can be used to approximate four experiment phases E_1, E_2, E_3, and E_4. But, for the case in which the output is the efficiency, it cannot. This is obvious because, from (5), we see that RF efficiency is not a linear function of the voltage. Instead, we obtain a bilinear fit of the efficiency by taking the ratio of y_1 by the product $V \times I$, where I itself is fitted linearly as a function of V, as shown in Fig. 6. Note that the fit in Fig. 6 is much worse than the previous ones since the errors in fitting V and I combine to produce larger efficiency errors. This will be corrected in future experiments by attempting a more complex linear fit of the efficiency.

IV. Conclusions and Future Work

In this paper we have reported on an effort to identify the input/output characteristics of the Sinus-6 electron beam-driven BWO. In addition, we introduce some identification and control systems concepts to the field of HPM tubes. These concepts are well known to the control systems community, but have not yet been fully exploited within the HPM community.

At this stage of our research, we have obtained both linear and nonlinear models to explain the input/output behavior of the Sinus-6 when the input is the cathode voltage. We will next be considering the more realistic model when both the cathode voltage and current (or more accurately, the pressure and A–K gap) are the physical inputs to our system. Using the resulting mathematical model, we can then design a controller to maximize both the efficiency and the power in addition to operating with enhanced frequency agility [12].

References

[1] L. Moreland, E. Schamiloglu, R. Lemke, S. Korovin, V. Rostov, A. Roitman, K. Hendricks, and T. Spencer, "Efficiency enhancement of high power vacuum BWO using nonuniform slow wave structures," *IEEE Trans. Plasma Sci.*, vol. 22, pp. 554–565, 1994.

[2] R. Smith and M. Dahleh, Eds., *The Modeling of Uncertainty in Control Systems*, vol. 192. London: Springer-Verlag, 1994.

[3] B. Goplen, L. Ludeking, D. Smithe, and G. Warren, "User-configurable magic for electromagnetic pic calculations," *Comp. Phys. Commun.*, vol. 87, pp. 54–86, 1995.

[4] C. Bishop, P. Haynes, M. Smith, T. Todd, D. Trotman, and C. Windsor, "Real-time control of tokomak plasma using neural networks," in *Advances in Neural Information Processing Systems 7*, G. Tesauro, D. Touretzky, and T. Leen, Eds. Cambridge, MA: MIT Press, 1995, pp. 1007–1014.

[5] L. Ljung, *System Identification: Theory for the User*, 1st ed. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[6] D. Rumelhart and J. McClelland, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations*, vol. 1. Cambridge, MA: MIT Press, 1986.

[7] D. Rumelhart and J. McClelland, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Psychological and Biological Models*, vol. 2. Cambridge, MA: MIT Press, 1986.

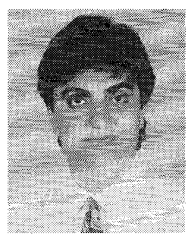
[8] P. Baldi, "Gradient descent learning algorithm overview: A general dynamical systems perspective," *IEEE Trans. Neural Networks*, vol. 6, no. 1, pp. 182–195, 1995.

[9] H. Demuth and M. Beale, *Neural Network Toolbox*. Natick, MA: The Math Works, 1993.

[10] C. Abdallah, W. Yang, E. Schamiloglu, and L. Moreland, "A neural network model of a high power backward-wave oscillator," Elect. Comput. Eng. Dep., Univ. of New Mexico, Tech. Rep. EECE95-005, 1995.

[11] S. Haykin, *Neural Networks: A Comprehensive Foundation*. New York: Macmillan, 1992.

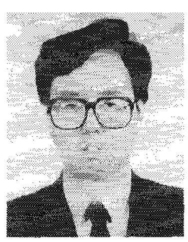
[12] L. D. Moreland, E. Schamiloglu, R. W. Lemke, A. M. Roitman, S. D. Korovin, and V. V. Rostov, "Enhanced frequency agility of high-power relativistic backward wave oscillators," this issue, pp. 852–858.



Chaouki Abdallah (S'81–M'83–SM'95) received the B.E. degree in electrical engineering in 1981 from Youngstown State University, Youngstown, OH, and the M.S. degree in 1982 and the Ph.D. degree in electrical engineering in 1988 from Georgia Tech, Atlanta, GA.

Between 1983 and 1985 he was with SAWTEK, Inc., Orlando, FL. He joined the Department of Electrical and Computer Engineering at the University of New Mexico, Albuquerque, as an Assistant Professor in 1988 and was promoted to Associate Professor in August 1994. His research and teaching interests are in the areas of robust control, adaptive and nonlinar systems, and neural networks. He is co-editor of *Robot Control: Dynamics, Motion Planning, and Analysis* (New York: IEEE, 1993), and co-author of *Control of Robot Manipulators* (New York: Macmillan, 1993) and *Linear Quadratic Control: An Introduction* (Englewood Cliffs, NJ: Prentice-Hall, 1995).

Dr. Abdallah is a member of Tau Beta Pi. He was Exhibit Chairman of the 1990 International Conference on Acoustics, Speech, and Signal Processing (ICASSP) and will serve as the Local Arrangement Chairman for the 1997 American Control Conference (ACC). He was also one of the founders of the Ibero-American Science and Technology Consortium (ISTEC).



Wei Yang was born in JiangXi Province of China. He received the B.S. degree in 1983 and the M.S. degree in 1986, both from Nanjing University of Aeronautics and Astronautics (NUAA), China, in electrical engineering. He is presently pursuing the Ph.D. degree in electrical engineering at the University of New Mexico, Albuquerque.

From 1986 to 1994, he was a Lecturer in the Department of Automatical Control Engineering of NUAA. His research interests include robust multi-objective control, system identification, and digital control.

Edl Schamiloglu (M'90–SM'95), for a biography, see this issue, p. 857.

Larald D. Moreland, for a biography, see this issue, p. 857.