

Neural Inverse Space Mapping (NISM) Optimization for EM-Based Microwave Design

John W. Bandler,^{1,†} Mostafa A. Ismail,^{1,‡} José E. Rayas-Sánchez,^{1,§} Qi-Jun Zhang²

¹ Simulation Optimization Systems Research Laboratory and Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, Canada L8S 4K1

² Department of Electronics, Carleton University, 1125 Colonel By Drive, Ottawa, Canada K1S 5B6

Received 3 January 2002; accepted 19 July 2002

ABSTRACT: We present neural inverse space mapping (NISM) optimization for electromagnetics-based design of microwave structures. The inverse of the mapping from the fine to the coarse model parameter spaces is exploited for the first time in a space mapping algorithm. NISM optimization does not require up-front EM simulations, multipoint parameter extraction, or frequency mapping. It employs a simple statistical parameter extraction procedure. The inverse of the mapping is approximated by a neural network whose generalization performance is controlled through a network growing strategy. We contrast our new algorithm with neural space mapping (NSM) optimization. © 2003 Wiley Periodicals, Inc. *Int J RF and Microwave CAE* 13: 136–147, 2003.

Keywords: neural network applications; space mapping; optimization methods; design automation; EM optimization; neural space mapping; microwave circuits; microstrip filters; neural modeling

I. INTRODUCTION

Neural networks have been extensively used for modeling microwave devices and circuits, in many different ways [1, 2]. In contrast, the use of neural networks for design by optimization is at an earlier stage: a few variations in the use of neural networks have been reported. The most widely used strategy [3–7] consists of generating a neuromodel of the microwave circuit within a certain training region of the design parameters and then applying conventional optimization to this model to find a solution that yields the

desired response. Full wave EM simulations are typically employed to generate the training data. The generalization ability of the neuromodel is controlled during the training process by using validation data and testing data, also obtained from EM simulations.

More effective approaches incorporate prior knowledge into the neural network training scheme, following an EM-ANN approach [8, 9] or a neural space mapping (NSM) approach [10, 11]. NSM optimization also avoids the use of validation and testing data.

An elegant new algorithm for EM-based design of microwave circuits is presented here: neural inverse space mapping (NISM) optimization [12]. This is the first space mapping (SM) algorithm that explicitly uses the inverse of the mapping from the fine to the coarse model parameter spaces. NISM follows an aggressive formulation by not requiring up-front fine model evaluations to start building the mapping. A simple procedure for parameter extraction avoids the need for multipoint matching and frequency mappings.

Correspondence to: Dr. John Bandler; e-mail: bandler@mcmaster.ca.

[†]Also affiliated with Bandler Corporation, P.O. Box 8083, Dundas, Ontario, Canada L9H 5E7.

[‡]Now with Com Dev International Ltd., 155 Sheldon Drive, Cambridge, ON, Canada N1R 7H6.

[§]Now with ITESO, Periférico Sur 8585, 45090 Tlaquepaque, Jal., Mexico.

Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/mmce.10067

A neural network whose generalization performance is controlled through a network growing strategy approximates the inverse of the mapping at each iteration. The NISM step consists simply of evaluating the current neural network at the optimal coarse solution. We show that this step is equivalent to a quasi-Newton step, while the inverse mapping remains essentially linear and gradually departs from a quasi-Newton step as the amount of nonlinearity in the inverse mapping increases.

We compare our new algorithm with NSM optimization [11] by solving the same microwave design problems: a bandstop microstrip filter with open stubs and an HTS microstrip filter.

II. NEURAL INVERSE SPACE MAPPING (NISM)

A. Notation

Let the vectors \mathbf{x}_c and \mathbf{x}_f represent the design parameters of the coarse and fine models, respectively ($\mathbf{x}_c, \mathbf{x}_f \in \mathcal{R}^n$). We denote the optimizable fine model responses at point \mathbf{x}_f and frequency ω by $\mathbf{R}_f(\mathbf{x}_f, \omega) \in \mathcal{R}^r$ where r is the number of responses to be optimized. For example, if the responses to be optimized are $|S_{11}|$ and $|S_{21}|$, then $r = 2$. The vector $\mathbf{R}_f(\mathbf{x}_f) \in \mathcal{R}^m$ denotes the fine model responses at the F_p sample frequency points, where $m = rF_p$. Similarly, $\mathbf{R}_c(\mathbf{x}_c) \in \mathcal{R}^m$ denotes the corresponding coarse model responses to be optimized.

Additionally, we denote the characterizing fine model responses at point $\mathbf{x}_f \in \mathcal{R}^n$ and frequency ω by $\mathbf{R}_{fs}(\mathbf{x}_f, \omega) \in \mathcal{R}^R$, which includes the real and imaginary parts of all the available characterizing responses in the model (considering symmetry). For example, for a two-port reciprocal network they include $\text{Re}\{S_{11}\}$, $\text{Im}\{S_{11}\}$, $\text{Re}\{S_{21}\}$, and $\text{Im}\{S_{21}\}$, and $R = 4$. The vector $\mathbf{R}_{fs}(\mathbf{x}_f) \in \mathcal{R}^M$ denotes the characterizing fine model responses at all the F_p frequency points, where $M = RF_p$. Similarly, $\mathbf{R}_{cs}(\mathbf{x}_c) \in \mathcal{R}^M$ denotes the corresponding characterizing coarse model responses.

B. Flow Diagram: An Overview

A flow diagram for NISM optimization is shown in Figure 1. We start by performing regular minimax optimization on the coarse model to find the optimal coarse solution \mathbf{x}_c^* that yields the desired response. The characterizing fine model responses \mathbf{R}_{fs} at the optimal coarse solution \mathbf{x}_c^* are then calculated.

We realize parameter extraction, which consists of finding the coarse model parameters that makes the

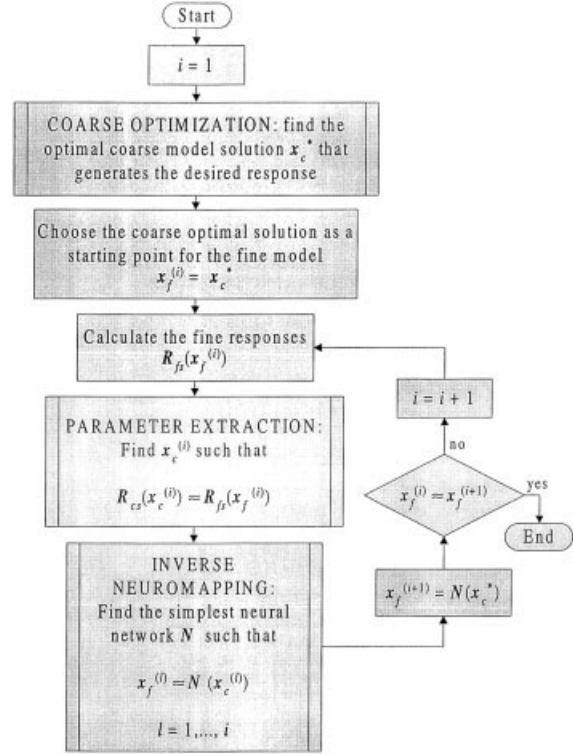


Figure 1. Flow diagram for neural inverse space mapping (NISM) optimization.

characterizing coarse responses \mathbf{R}_{cs} as close as possible to the previously calculated \mathbf{R}_{fs} .

We continue by training the simplest neural network N that implements the inverse of the mapping from the fine to the coarse parameter space at the available points.

The new point in the fine model parameter space is then calculated by simply evaluating the neural network at the optimal coarse solution. If the maximum relative change in the fine model parameters is smaller than a previously defined amount we finish, otherwise we calculate the characterizing fine model responses at the new point and continue with the algorithm.

C. Parameter Extraction

The parameter extraction procedure at the i th NISM iteration is formulated as the optimization problem

$$\mathbf{x}_c^{(i)} = \arg \min_{\mathbf{x}_c} U_{PE}(\mathbf{x}_c) \quad (1a)$$

$$U_{PE}(\mathbf{x}_c) = \|\mathbf{e}(\mathbf{x}_c)\|_2^2 \quad (1b)$$

$$\mathbf{e}(\mathbf{x}_c) = \mathbf{R}_{fs}(\mathbf{x}_f^{(i)}) - \mathbf{R}_{cs}(\mathbf{x}_c) \quad (1c)$$

We solve (1) using the Levenberg-Marquardt algorithm for nonlinear curve fitting available in the Matlab™ Optimization Toolbox [13].

We normally use \mathbf{x}_c^* as the starting point for solving (1). This might not be a good starting point when an extremely severe matching problem is being solved, one that has some poor local minimum around \mathbf{x}_c^* . If the algorithm is trapped in a poor local minimum, we change the starting point for (1) by taking a small random perturbation $\Delta\mathbf{x}$ around \mathbf{x}_c^* until we find an acceptable local minimum, i.e., until we obtain good matching between fine and coarse models.

The maximum perturbation Δ_{\max} is obtained from the maximum absolute sensitivity of the parameter extraction objective function at \mathbf{x}_c^* as

$$\Delta_{\max} = \frac{\delta_{PE}}{\|\nabla U_{PE}(\mathbf{x}_c^*)\|_{\infty}} \quad (2)$$

Let $\mathbf{rand} \in \mathcal{R}^n$ be a vector whose elements take random values between 0 and +1 every time it is evaluated. The values of the elements of $\Delta\mathbf{x}$ are calculated as

$$\Delta x_k = \Delta_{\max}(2\mathbf{rand}_k - 1), \quad k = 1, \dots, n \quad (3)$$

A value of $\delta_{PE} = 0.03$ is used in our implementation. Many other values of δ_{PE} could be used in (2), as we use it only to escape from a poor local minimum.

A similar strategy for statistical parameter extraction was proposed in [14], where an exploration region is first created by predefining a fixed number of starting points around \mathbf{x}_c^* .

The algorithm for realizing parameter extraction is stated as follows

<p>Algorithm: Parameter Extraction</p> <pre> begin solve (1) using \mathbf{x}_c^* as starting point while $\ \mathbf{e}(\mathbf{x}_c^{(i)})\ _{\infty} > \epsilon_{PE}$ calculate $\Delta\mathbf{x}$ using (2) and (3) solve (1) using $\mathbf{x}_c + \Delta\mathbf{x}$ as starting point end</pre>

A value of $\epsilon_{PE} = 0.15$ is used in our implementation, assuming that all the response values are normalized.

D. Inverse Neuromapping

When training the neural network N that implements the inverse mapping we solve the optimization problem

$$\mathbf{w}^* = \arg \min_{\mathbf{w}} U_N(\mathbf{w}) \quad (4a)$$

$$U_N(\mathbf{w}) = \|\llbracket \dots \mathbf{e}_l^T \dots \rrbracket^T\|_2^2 \quad (4b)$$

$$\mathbf{e}_l = \mathbf{x}_f^{(l)} - N(\mathbf{x}_c^{(l)}, \mathbf{w}), \quad l = 1, \dots, i \quad (4c)$$

where i is the current NISM iteration and vector \mathbf{w} contains the internal parameters (weights, bias, etc.) of the neural network N . The starting point $\mathbf{w}^{(0)}$ for solving (4) is a unit mapping, i.e., $N(\mathbf{x}_c^{(l)}, \mathbf{w}^{(0)}) = \mathbf{x}_c^{(l)}$, for $l = 1, \dots, i$. We use the Scaled Conjugate Gradient (SCG) algorithm available in the Matlab™ Neural Network Toolbox [15] for solving (4). Notice that the time consumed in solving (4) is negligible because no coarse or fine model simulations are needed.

To control the generalization performance of the neural network N , we follow a network growing strategy [16], in which case we start with a small perceptron to match the initial points and then add more neurons only when we are unable to meet a small error.

We initially assume a two-layer perceptron given by

$$N(\mathbf{x}_c, \mathbf{w}) = \mathbf{x}_f = \mathbf{W}^o \mathbf{x}_c + \mathbf{b}^o \quad (5)$$

where $\mathbf{W}^o \in \mathcal{R}^{n \times n}$ is the matrix of output weighting factors, $\mathbf{b}^o \in \mathcal{R}^n$ is the vector of output bias elements, and vector \mathbf{w} contains \mathbf{b}^o and the columns of \mathbf{W}^o . The starting point is obtained by making $\mathbf{W}^o = \mathbf{I}$ and $\mathbf{b}^o = \mathbf{0}$.

If a two-layer perceptron is not sufficient to make the learning error $U_N(\mathbf{w}^*)$ small enough, then we use a three-layer perceptron with h hidden neurons given by

$$N(\mathbf{x}_c, \mathbf{w}) = \mathbf{W}^o \Phi(\mathbf{x}_c) + \mathbf{b}^o \quad (6a)$$

$$\Phi(\mathbf{x}_c) = [\varphi(s_1) \quad \varphi(s_2) \quad \dots \quad \varphi(s_h)]^T \quad (6b)$$

$$\mathbf{s} = \mathbf{W}^h \mathbf{x}_c + \mathbf{b}^h \quad (6c)$$

where $\mathbf{W}^o \in \mathcal{R}^{n \times h}$, $\mathbf{b}^o \in \mathcal{R}^n$, $\Phi(\mathbf{x}_c) \in \mathcal{R}^h$ is the vector of hidden signals, $\mathbf{s} \in \mathcal{R}^h$ is the vector of activation potentials, $\mathbf{W}^h \in \mathcal{R}^{h \times n}$ is the matrix of hidden weighting factors, $\mathbf{b}^h \in \mathcal{R}^h$ is the vector of hidden bias elements, and h is the number of hidden neurons. In this work we use hyperbolic tangents as nonlinear activation functions, i.e., $\phi(\cdot) = \tanh(\cdot)$. Vector \mathbf{w} contains \mathbf{b}^o , \mathbf{b}^h , the columns of \mathbf{W}^o , and the columns of \mathbf{W}^h .

Our starting point for solving (4) using (6) is also a unit mapping, which is approximately obtained by

making $\mathbf{b}^o = \mathbf{0}$, $\mathbf{b}^h = \mathbf{0}$, $\mathbf{W}^h = 0.1[\mathbf{I} \ \mathbf{0}]^T$ and $\mathbf{W}^o = 10[\mathbf{I} \ \mathbf{0}]$, assuming that the training data has been scaled between -1 and $+1$. Notice that we consider $h \geq n$ in order to achieve the unit mapping.

The algorithm for finding the simplest inverse neuromapping is stated as

Algorithm: Inverse Neuromapping

```

begin
  solve (4) using (5)
   $h = n$ 
  while  $U_N(\mathbf{w}^*) > \epsilon_L$ 
    solve (4) using (6)
     $h = h + 1$ 
end
    
```

In our implementation we use $\epsilon_L = 1 \times 10^{-4}$. Notice that the algorithm for finding the inverse neuromapping uses a two-layer perceptron during at least the first $n + 1$ NISM iterations, since the points $(\mathbf{x}_c^{(i)}, \mathbf{x}_f^{(i)})$ can be mapped with a linear mapping for $i = 1 \dots n + 1$. A three-layer perceptron is needed only when we exceed $n + 1$ NISM iterations and the mapping is significantly nonlinear.

E. Termination Criterion

As illustrated in the flow diagram of Figure 1, we stop NISM optimization when the new iterate is close enough to the current point. We do this by testing the relative change in the fine model parameters. If the expression

$$\|\mathbf{x}_f^{(i+1)} - \mathbf{x}_f^{(i)}\|_2 \leq \epsilon_{end}(\epsilon_{end} + \|\mathbf{x}_f^{(i)}\|_2) \quad (7)$$

is true, we end NISM optimization taking $\mathbf{x}_f^{(i)}$ as the solution, otherwise we continue. We use $\epsilon_{end} = 5 \times 10^{-3}$ in our implementation. Notice that the fine model is not evaluated at the point $\mathbf{x}_f^{(i+1)}$.

III. NATURE OF THE NISM STEP

In this section we show that the NISM step, $\mathbf{x}_f^{(i+1)} = N(\mathbf{x}_c^*)$, is equivalent to a quasi-Newton step while the inverse mapping built during NISM optimization remains linear (i.e., while a two-layer perceptron is enough to approximate the inverse mapping). We also show that the NISM step gradually departs from a quasi-Newton step as the amount of nonlinearity needed in the inverse mapping increases.

A. Jacobian of the Inverse Mapping

From (5), the Jacobian \mathbf{J}_N of the inverse mapping $N(\mathbf{x}_c)$ when a two-layer perceptron is employed is given by

$$\mathbf{J}_N = \mathbf{W}^o \quad (8)$$

When a three-layer perceptron is used, the Jacobian \mathbf{J}_N is obtained from (6) as

$$\mathbf{J}_N = \mathbf{W}^o \mathbf{J}_\Phi \mathbf{W}^h \quad (9)$$

where $\mathbf{J}_\Phi \in \mathcal{R}^{h \times h}$ is a diagonal matrix given by $\mathbf{J}_\Phi = \text{diag}(\phi'(s_j))$, with $j = 1 \dots h$. We use (8) and (9) to demonstrate the nature of the NISM step $\mathbf{x}_f^{(i+1)} = N(\mathbf{x}_c^*)$.

B. NISM Step versus Quasi-Newton Step

A general space mapping optimization problem can be formulated as solving the system of nonlinear equations

$$\mathbf{f}(x_f) = \mathbf{P}(x_f) - \mathbf{x}_c^* = 0 \quad (10)$$

where $\mathbf{x}_c = \mathbf{P}(x_f)$ is the mapping function that makes the coarse model behave as the fine model, i.e., $\mathbf{R}_c(\mathbf{P}(x_f)) \approx \mathbf{R}_f(x_f)$. A Newton step for solving (10) is given by

$$\mathbf{x}_f^{(i+1)} = \mathbf{x}_f^{(i)} - \mathbf{J}_P^{-1} \mathbf{f} \quad (11)$$

where $\mathbf{J}_P \in \mathcal{R}^{n \times n}$ is the Jacobian of the mapping function $\mathbf{P}(x_f)$. This can be stated in an equivalent manner by using the Jacobian $\mathbf{J}_N \in \mathcal{R}^{n \times n}$ of the inverse of the mapping $\mathbf{x}_f = N(\mathbf{x}_c)$

$$\mathbf{x}_f^{(i+1)} = \mathbf{x}_f^{(i)} - \mathbf{J}_N \mathbf{f} \quad (12)$$

Approximating \mathbf{J}_N directly involves the same computational effort as approximating \mathbf{J}_P , but calculating the next step using (12) is more efficient than using (11).

If a two-layer perceptron is being used, we substitute (8) in (12) to obtain

$$\mathbf{x}_f^{(i+1)} = \mathbf{x}_f^{(i)} - \mathbf{W}^o(\mathbf{x}_c^{(i)} - \mathbf{x}_c^*) \quad (13)$$

which can be expressed using (5) as

$$\mathbf{x}_f^{(i+1)} = \mathbf{W}^o \mathbf{x}_c^* - (\mathbf{x}_f^{(i)} - \mathbf{b}^o) + \mathbf{x}_f^{(i)} = N(\mathbf{x}_c^*) \quad (14)$$

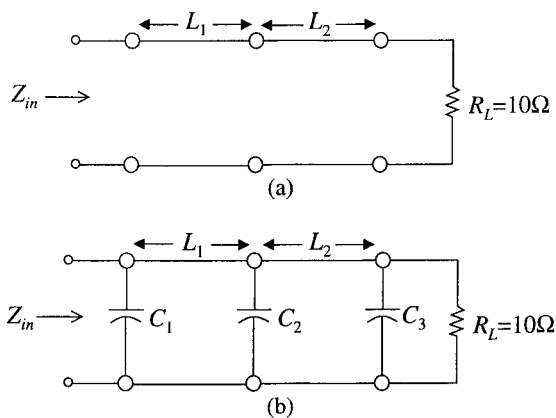


Figure 2. Two-section impedance transformer test problem: (a) “coarse” model, (b) “fine” model.

From (12) and (14) we conclude that while the inverse mapping built during NISM optimization remains linear, the NISM step is equivalent to a quasi-Newton step. Notice that we do not use any of the classical updating formulae to calculate an approximation of the inverse of the Jacobian; this is done by simply evaluating the current neural network at the optimal coarse solution.

If a three-layer perceptron is being used, we substitute (9) in (12) to obtain

$$\mathbf{x}_f^{(i+1)} = \mathbf{x}_f^{(i)} - \mathbf{W}^o \mathbf{J}_\Phi \mathbf{W}^h (\mathbf{x}_c^{(i)} - \mathbf{x}_c^*) \quad (15)$$

Adding and subtracting $\mathbf{W}^o \mathbf{J}_\Phi \mathbf{b}^h$ to (15)

$$\mathbf{x}_f^{(i+1)} = \mathbf{W}^o \mathbf{J}_\Phi (\mathbf{W}^h \mathbf{x}_c^* + \mathbf{b}^h) - \mathbf{W}^o \mathbf{J}_\Phi (\mathbf{W}^h \mathbf{x}_c^{(i)} + \mathbf{b}^h) + \mathbf{x}_f^{(i)} \quad (16)$$

Substituting (6c) in (16)

$$\mathbf{x}_f^{(i+1)} = \mathbf{W}^o \mathbf{J}_\Phi \mathbf{s}(\mathbf{x}_c^*) - \mathbf{W}^o \mathbf{J}_\Phi \mathbf{s}(\mathbf{x}_c^{(i)}) + \mathbf{x}_f^{(i)} \quad (17)$$

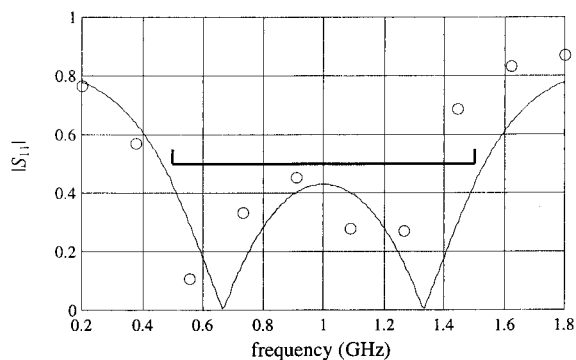


Figure 3. Coarse (—) and fine (○) model responses at \mathbf{x}_c^* for the two-section impedance transformer.

TABLE I. Fine Model Parameters for the Two-Section Impedance Transformer at Each NISM Iteration

i	$\mathbf{x}_f^{(i)T}$
1	[90 90]
2	[84.1990 83.0317]
3	[79.3993 73.7446]

Expanding the term $\mathbf{J}_\Phi \mathbf{s}(\mathbf{x}_c)$ we obtain

$$\mathbf{J}_\Phi \mathbf{s}(\mathbf{x}_c) = [\phi'(s_1)s_1 \quad \cdots \quad \phi'(s_h)s_h]^T. \quad (18)$$

Since we are using hyperbolic tangents as nonlinear activation functions, when a small amount of nonlinearity is present (e.g., $s_j < 0.1$), $\phi(s_j) \approx s_j$, and $\phi'(s_j)s_j \approx s_j \approx \phi(s_j)$, for $j = 1, \dots, h$, and using (6b) we express (18) as

$$\mathbf{J}_\Phi \mathbf{s}(\mathbf{x}_c) \approx \Phi(\mathbf{x}_c) \quad (19)$$

Substituting (19) in (17)

$$\mathbf{x}_f^{(i+1)} \approx \mathbf{W}^o \Phi(\mathbf{x}_c^*) - \mathbf{W}^o \Phi(\mathbf{x}_c^{(i)}) + \mathbf{x}_f^{(i)} \quad (20)$$

Adding and subtracting \mathbf{b}^o to (20) and using (6a) we express (20) as

$$\begin{aligned} \mathbf{x}_f^{(i+1)} &\approx \mathbf{W}^o \Phi(\mathbf{x}_c^*) + \mathbf{b}^o - \mathbf{W}^o \Phi(\mathbf{x}_c^{(i)}) - \mathbf{b}^o + \mathbf{x}_f^{(i)} \\ &= \mathbf{N}(\mathbf{x}_c^*) \end{aligned} \quad (21)$$

In conclusion, the NISM step gradually departs from a quasi-Newton step as the amount of nonlinearity needed in the inverse mapping increases.

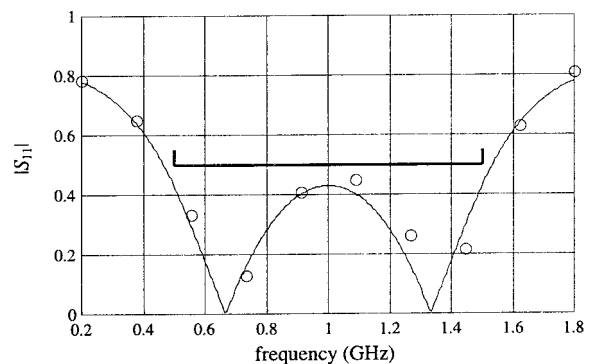


Figure 4. Optimal coarse model response (—) and fine model response (○) at the NISM solution for the two-section impedance transformer.

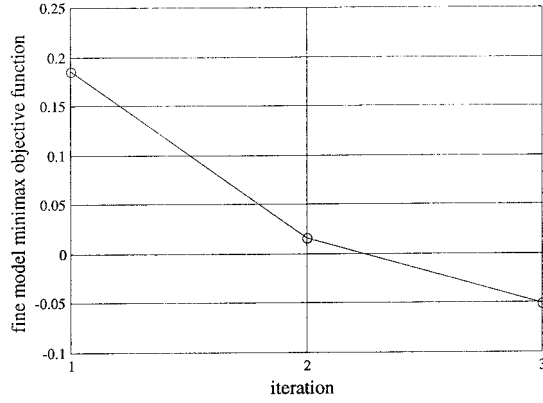


Figure 5. Fine model minimax objective function values for the two-section impedance transformer at each NISM iteration.

IV. EXAMPLES

A. Two-Section Impedance Transformer

As an illustrative case, consider the classical test problem of designing a capacitively loaded 10:1 two-section impedance transformer [17]. The proposed “coarse” and “fine” models are shown in Figure 2. The coarse model consists of ideal transmission lines; the fine model consists of capacitively loaded ideal transmission lines, with $C_1 = C_2 = C_3 = 10$ pF. The design specifications are $|S_{11}| \leq 0.50$ for frequencies between 0.5 and 1.5 GHz.

The electrical lengths of the two transmission lines at 1.0 GHz are selected as design parameters. The characteristic impedances are kept fixed at the following values: $Z_1 = 2.23615 \Omega$, $Z_2 = 4.47230 \Omega$. Both models were implemented in OSA90/hope [18]. The optimal coarse solution is $\mathbf{x}_c^* = [90 \ 90]^T$ (degrees). The coarse and fine model responses at \mathbf{x}_c^* are shown

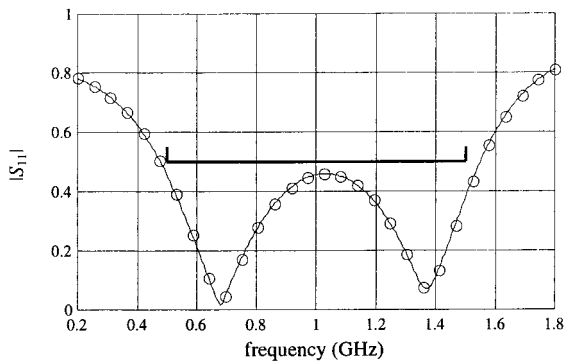


Figure 6. Fine model response at the NISM solution (\circ) and at the direct minimax solution (—) for the two-section impedance transformer.

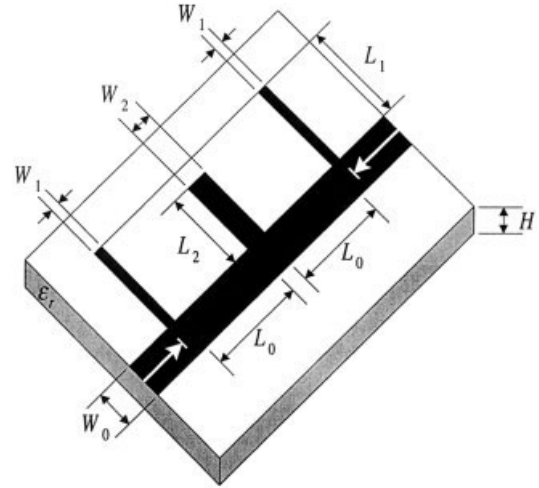


Figure 7. Bandstop microstrip filter with quarter-wave resonant open stubs.

in Figure 3. We used only 10 frequency points from 0.2 to 1.8 GHz for the fine model.

NISM optimization requires only three “fine” model evaluations to solve this problem. The values of the fine model parameters at each iteration are shown in Table I. A two-layer perceptron was enough to approximate the inverse mapping at all NISM iterations. The fine model response at the NISM solution is compared with the optimal coarse model response in Figure 4. The fine model minimax objective function values at each NISM iteration are shown in Figure 5.

Because both the coarse and fine models are fast to evaluate, we applied direct minimax optimization to the fine model, obtaining $\mathbf{x}_f^* = [79.2651 \ 74.2322]^T$ after 64 fine model evaluations. In Figure 6 we compare the fine model response at this solution with the optimal NISM response, where a remarkable match is observed.

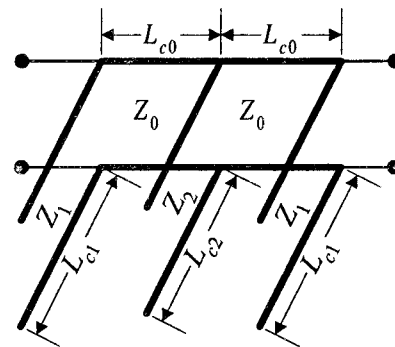


Figure 8. Coarse model for the bandstop microstrip filter with open stubs.

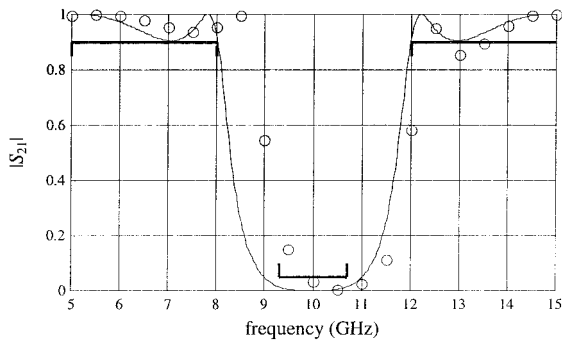


Figure 9. Coarse and fine model responses at the optimal coarse solution for the bandstop filter with open stubs: OSA90/hope™ (—) and *em*™ (○).

B. Bandstop Microstrip Filter with Open Stubs

We apply NISM optimization to a bandstop microstrip filter with quarter-wave resonant open stubs [11], whose physical structure is illustrated in Figure 7. L_1 , L_2 are the open stub lengths and W_1 , W_2 the corresponding widths. An alumina substrate with thickness $H = 25$ mil, width $W_0 = 25$ mil and dielectric constant $\epsilon_r = 9.4$ is used for a 50Ω feeding line.

The specifications are $|S_{21}| \leq 0.01$ in the stopband and $|S_{21}| \geq 0.9$ in the passband, where the stopband lies between 9.3 and 10.7 GHz, and the passband includes frequencies below 8 GHz and above 12 GHz. The design parameters are $\mathbf{x}_f = [W_1 \ W_2 \ L_0 \ L_1 \ L_2]^T$.

Sonnet's *em*™ [19] driven by Empipe™ [18] was employed as the fine model, using a high-resolution grid with a $1 \text{ mil} \times 1 \text{ mil}$ cell size. The coarse model, illustrated in Figure 8, consists of simple ideal transmission lines for modeling each microstrip section and classical formulas [20] to calculate the characteristic impedance and the effective dielectric constant of each transmission line. It is seen that $L_{c2} = L_2 + W_0/2$, $L_{c1} = L_1 + W_0/2$, and $L_{c0} = L_0 + W_1/2 + W_2/2$. We use OSA90/hope™ [18] built-in transmission line elements TRL.

The following optimal coarse model solution is found for L_0 , L_1 , and L_2 of quarter-wave lengths at 10 GHz: $\mathbf{x}_c^* = [6.0 \ 9.0 \ 106.4 \ 110.1 \ 108.8]^T$ (mils), as in

TABLE II. Fine Model Parameters for the Bandstop Filter With Open Stubs at Each NISM Iteration

i	$\mathbf{x}_f^{(i)T}$
1	[6 9 106 110 109]
2	[7 11 103 112 111]
3	[9 20 95 115 115]
4	[9 19 95 115 114]

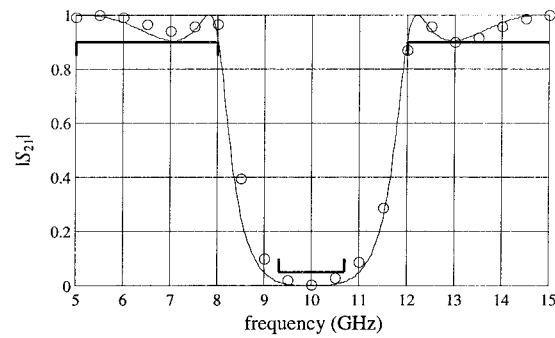


Figure 10. Coarse model response (—) at the optimal coarse solution and fine model response (○) at the NISM solution for the bandstop microstrip filter with open stubs.

[11]. The coarse and fine model responses at the optimal coarse solution are shown in Figure 9.

NISM optimization requires only four fine model evaluations to solve this problem. The sequence of iterates is shown in Table II (all the points are on the grid, to avoid interpolation). A two-layer perceptron was enough to approximate the inverse mapping at all NISM iterations. The fine model response at the NISM solution is compared with the optimal coarse model response in Figure 10. The fine model minimax objective function values at each NISM iteration are shown in Figure 11.

The same problem was solved in [11] using NSM optimization. NSM required 13 fine model evaluations to find the solution whose response is shown in Figure 12. NISM optimization requires fewer fine model evaluations and arrives at a solution closer to the solution of the original optimization problem.

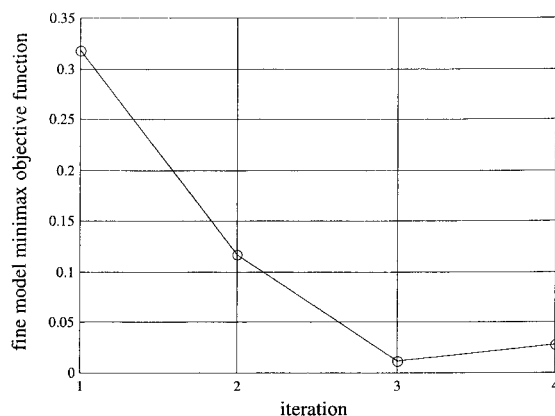


Figure 11. Fine model minimax objective function values for the bandstop microstrip filter at each NISM iteration.

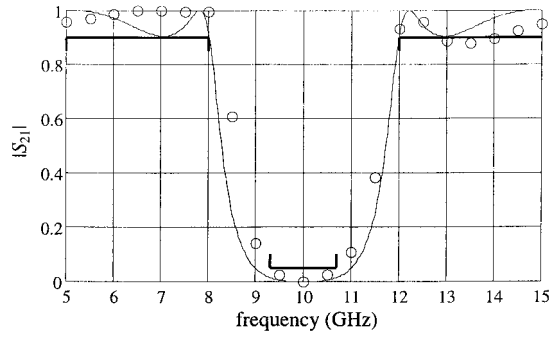


Figure 12. Coarse model response (—) at the optimal coarse solution and fine model response (○) at the NSM solution, obtained in [11], for the bandstop microstrip filter with open stubs.

C. High-Temperature Superconducting Microstrip Filter

We apply NISM optimization to a high-temperature superconducting (HTS) quarter-wave parallel coupled-line microstrip filter and contrast our results with those obtained by using NSM optimization on the same problem [10, 11]. The physical structure of the HTS filter is illustrated in Figure 13.

L_1 , L_2 , and L_3 are the lengths of the parallel coupled-line sections and S_1 , S_2 , and S_3 are the gaps between the sections. The width W is the same for all the sections as well as for the input and output lines of length L_0 . A lanthanum aluminate substrate with thickness H and dielectric constant ϵ_r is used.

The specifications are $|S_{21}| \geq 0.95$ in the passband and $|S_{21}| \leq 0.05$ in the stopband, where the stopband includes frequencies below 3.967 GHz and above

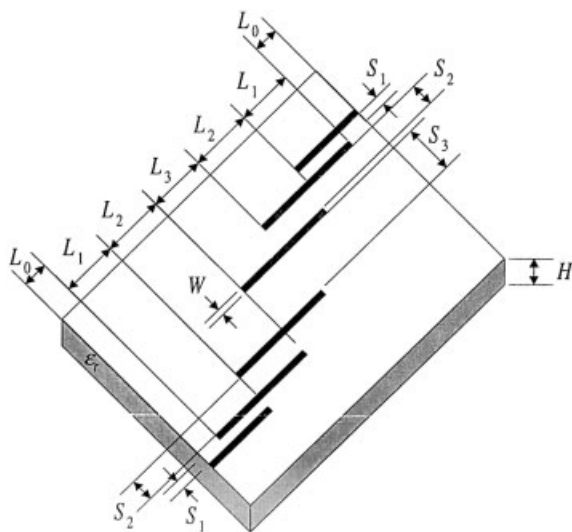


Figure 13. HTS quarter-wave parallel coupled-line microstrip filter.

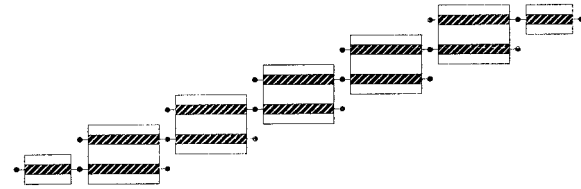


Figure 14. Schematic representation of the coarse model for the HTS filter.

4.099 GHz, and the passband lies in the range [4.008 GHz, 4.058 GHz]. The design parameters are $\mathbf{x}_f = [L_1 L_2 L_3 S_1 S_2 S_3]^T$. We take $L_0 = 50$ mil, $H = 20$ mil, $W = 7$ mil, $\epsilon_r = 23.425$, loss tangent = 3×10^{-5} ; the metalization is considered lossless.

Sonnet's *em*TM [19] driven by EmpipeTM [18] was employed as the fine model, using a high-resolution grid with a 1 mil \times 1 mil cell size. OSA90/hopeTM [18] built-in linear elements MSL (microstrip line), MSCL (two-conductor symmetrical coupled microstrip lines), and OPEN (open circuit) connected by circuit theory over the same MSUB (microstrip substrate definition) are taken as the coarse model, whose schematic representation is illustrated in Figure 14.

The following optimal coarse model solution is used, as in [10,11]: $\mathbf{x}_c^* = [188.33 \ 197.98 \ 188.58 \ 21.97 \ 99.12 \ 111.67]^T$ (in mils). The coarse and fine model responses at the optimal coarse solution are shown in Figure 15. Only 14 frequency points per frequency sweep are used for the fine model.

After only three fine model simulations, the optimal NISM solution was found. The sequence of fine model parameters at each NISM iteration is shown in Table III (all the points are on the grid, to avoid interpolation). A two-layer perceptron was enough to approximate the inverse mapping at all NISM iterations. Figure 16a compares the optimal coarse response with the fine model response at the

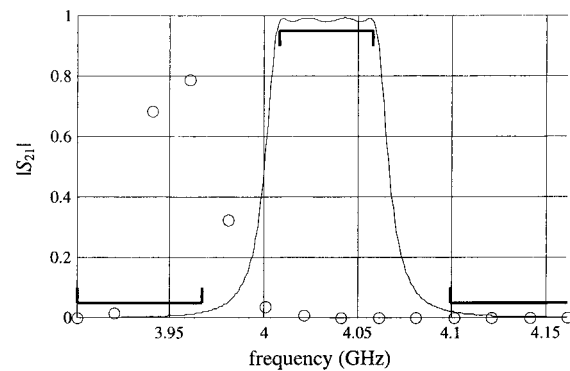


Figure 15. Coarse and fine model responses at the optimal coarse solution for the HTS filter: OSA90/hopeTM (—) and *em*TM (○).

TABLE III. Fine Model Parameters for the HTS Microstrip Filter at Each NISM Iteration

i	$\mathbf{x}_f^{(i)T}$
1	[188 198 189 22 99 112]
2	[187 196 187 21 84 92]
3	[186 194 185 20 80 89]

NISM solution using a fine frequency sweep. A more detailed comparison in the passband is shown in Figure 16b. The fine model minimax objective function values at each NISM iterations for this problem are shown in Figure 17.

Figure 18 shows the results obtained by applying NSM optimization [11] to the same problem, where the optimal NSM solution was found after 14 fine model evaluations.

For all the previous examples, parameter extraction was successfully performed in just one attempt at every NISM iteration. That was not the case for the HTS filter, where the parameter extraction objective function has many poor local minima around \mathbf{x}_c^* . Our proposed algorithm for parameter extraction overcame this problem. We applied NISM optimization to

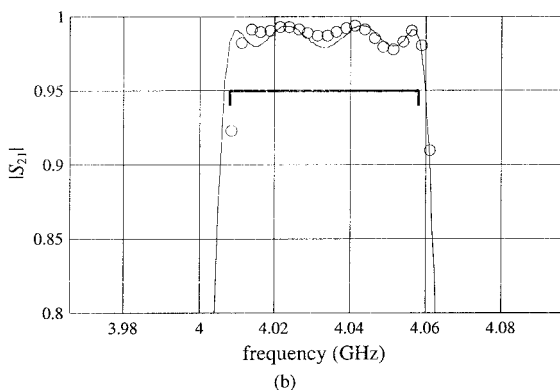
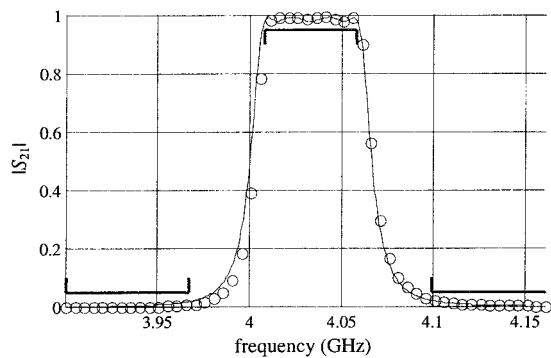


Figure 16. Coarse model response at \mathbf{x}_c^* (—) and fine model response at \mathbf{x}_f^{NISM} (○) for the HTS filter: (a) in the complete range of interest, (b) in the passband.

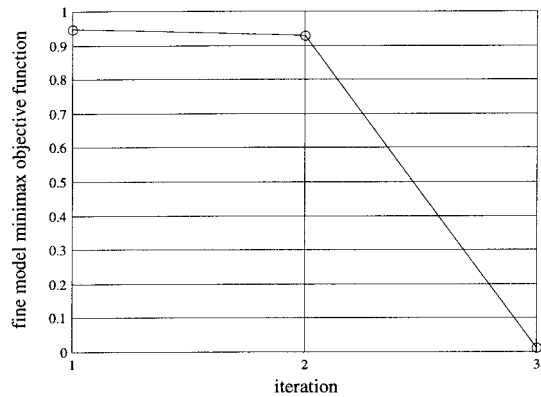


Figure 17. Fine model minimax objective function values for the HTS microstrip filter at each NISM iteration.

the HTS filter five times to test the statistical parameter extraction results. In Table IV we show the number of attempts needed for successful parameter extraction at each NISM iteration for the five optimizations. Exactly the same sequence of points illustrated in Table III was predicted by each of the five optimizations.

V. CONCLUSIONS

We propose neural inverse space mapping (NISM) optimization for EM-based design of microwave structures. The inverse of the mapping is exploited for the first time in a space mapping algorithm. NISM optimization does not require up-front EM simulations, multipoint parameter extraction, or frequency mapping. A simple statistical procedure overcomes the existence of poor local minima during parameter extraction. A neural network whose generalization performance is controlled through a network growing strategy approximates the inverse of the mapping at

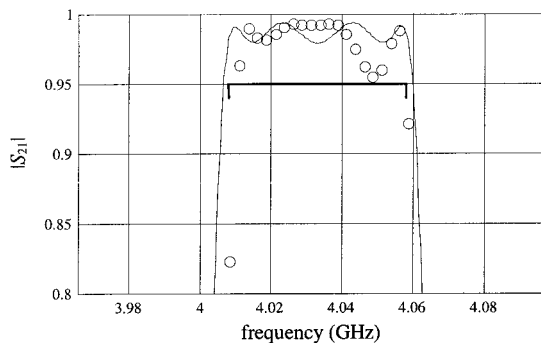


Figure 18. Coarse model response at \mathbf{x}_c^* (—) and fine model response at \mathbf{x}_f^{NISM} (○), obtained in [11], in the passband.

TABLE IV. Parameter Extraction Results for Five NISM Optimizations for the HTS Filter

i	Number of Attempts Needed for Successful PE				
1	12	9	3	10	8
2	3	3	6	7	3
3	1	1	1	1	1

each iteration. The NISM step simply evaluates the current neural network at the optimal coarse solution. We show that this step is equivalent to a quasi-Newton step while the inverse mapping remains essentially linear, and gradually departs from a Newton step as the amount of nonlinearity in the inverse mapping increases. In the examples considered, our new algorithm exhibits superior performance to NSM optimization and the trust region aggressive space mapping algorithm exploiting surrogates [21, 22].

ACKNOWLEDGMENTS

The authors thank Dr. J.C. Rautio, President, Sonnet Software, Inc. (Liverpool, NY) for making *em*TM available. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (grants OGP0007239 and STR234854-00), through the Micronet Network of Centres of Excellence and Bandler Corporation. J.E. Rayas-Sánchez was supported by an Ontario Graduate Scholarship, by ITESO (Instituto Tecnológico y de Estudios Superiores de Occidente, Mexico) and by CONACYT (Consejo Nacional de Ciencia y Tecnología, Mexico). M.A. Ismail was supported by a Nortel Networks Ontario Graduate Scholarship in Science and Technology.

REFERENCES

1. P. Burrascano and M. Mongiardo, A review of artificial neural networks applications in microwave CAD, *Int J RF Microwave CAE [Special Issue]* 9 (1999), 158–174.
2. J.W. Bandler, M.A. Ismail, J.E. Rayas-Sánchez, and Q. J. Zhang, Neuromodeling of microwave circuits exploiting space mapping technology, *IEEE Trans Microwave Theory Tech* 47 (1999), 2417–2427.
3. T.S. Horng, C.C. Wang, and N.G. Alexopoulos, Microstrip circuit design using neural networks, *IEEE MTT-S Int Microwave Symp Digest*, Atlanta, GA, 1993, pp. 413–416.
4. A.H. Zaabab, Q.J. Zhang, and M.S. Nakhla, A neural network modeling approach to circuit optimization and statistical design, *IEEE Trans Microwave Theory Tech* 43 (1995), 1349–1358.
5. A. Veluswami, M.S. Nakhla, and Q.J. Zhang, The application of neural networks to EM-based simulation and optimization of interconnects in high-speed VLSI circuits, *IEEE Trans Microwave Theory Tech* 45 (1997), 712–723.
6. P.M. Watson and K.C. Gupta, Design and optimization of CPW circuits using EM-ANN models for CPW components, *IEEE Trans Microwave Theory Tech* 45 (1997), pp. 2515–2523.
7. P. Burrascano, M. Dionigi, C. Fancelli, and M. Mongiardo, A neural network model for CAD and optimization of microwave filters, *IEEE MTT-S Int Microwave Symp Digest*, Baltimore, MD, 1998, pp. 13–16.
8. P.M. Watson, G.L. Creech, and K.C. Gupta, Knowledge based EM-ANN models for the design of wide bandwidth CPW patch/slot antennas, *IEEE AP-S Int Symp Digest*, Orlando, FL, 1999, pp. 2588–2591.
9. C. Cho and K.C. Gupta, EM-ANN modeling of overlapping open-ends in multilayer lines for design of bandpass filters, *IEEE AP-S Int Symp Digest*, Orlando, FL, 1999, pp. 2592–2595.
10. M.H. Bakr, J.W. Bandler, M. A. Ismail, J.E. Rayas-Sánchez, and Q.J. Zhang, Neural space mapping optimization of EM microwave structures, *IEEE MTT-S Int Microwave Symp Digest*, Boston, MA, 2000, pp. 879–882.
11. M.H. Bakr, J.W. Bandler, M.A. Ismail, J.E. Rayas-Sánchez, and Q.J. Zhang, Neural space mapping optimization for EM-based design, *IEEE Trans Microwave Theory Tech* 48 (2000), 2307–2315.
12. J.W. Bandler, M.A. Ismail, J.E. Rayas-Sánchez, and Q.J. Zhang, Neural inverse space mapping EM-optimization, *IEEE MTT-S Int Microwave Symp Digest*, Phoenix, AZ, 2001, pp. 1007–1010.
13. *Matlab*TM Optimization Toolbox, Version 2, The MathWorks, Inc., Natick, MA, 1999.
14. J.W. Bandler, R.M. Biernacki, S.H. Chen, and D. Omeragić, Space mapping optimization of waveguide filters using finite element and mode-matching electromagnetic simulators, *Int J RF Microwave CAE* 9 (1999), 54–70.
15. *Matlab*TM Neural Network Toolbox, Version 3, The MathWorks, Inc., Natick, MA, 1998.
16. S. Haykin, *Neural networks: a comprehensive foundation*. Prentice-Hall, Upper Saddle River, New Jersey, 1999.
17. J.W. Bandler, *Optimization methods for computer-aided design*, *IEEE Trans Microwave Theory Tech* 17 (1969), 533–552.
18. OSA90/hopeTM and EmpipeTM, Version 4.0, formerly Optimization Systems Associates Inc. (now Agilent Technologies, Santa Rosa, CA), Dundas, ON, Canada, 1997.
19. *em*TM Version 4.0b, Sonnet Software, Inc., Liverpool, NY, 1997.
20. M. Pozar, *Microwave engineering*, John Wiley, Amherst, MA, 1998.

21. M.H. Bakr, J.W. Bandler, K. Madsen, J.E. Rayas-Sánchez, and J. Søndergaard, Space mapping optimization of microwave circuits exploiting surrogate models, *IEEE Trans Microwave Theory Tech* 48 (2000), 2297–2306.
22. J.E. Rayas-Sánchez, Neural space mapping methods for modeling and design of microwave circuits, Ph.D. thesis, McMaster University, Hamilton, Canada, 2001 (www.sos.mcmaster.ca/theses.htm).

BIOGRAPHIES



John W. Bandler was born in Jerusalem on November 9, 1941. He studied at Imperial College of Science and Technology, London, England, from 1960 to 1966. He received B.Sc.(Eng.), Ph.D., and D.Sc.(Eng.) degrees from the University of London in 1963, 1967, and 1976, respectively.

Dr. Bandler joined Mullard Research Laboratories, Redhill, Surrey, England, in 1966.

From 1967 to 1969 he was a Postdoctorate Fellow and Sessional Lecturer at the University of Manitoba, Winnipeg, Canada. He joined McMaster University, Hamilton, Canada, in 1969. He has served as Chairman of the Department of Electrical Engineering and Dean of the Faculty of Engineering. He is currently Professor Emeritus in Electrical and Computer Engineering, directing research in the Simulation Optimization Systems Research Laboratory. He is a member of the Micronet Network of Centres of Excellence.

Dr. Bandler was President of Optimization Systems Associates Inc. (OSA), which he founded in 1983, until November 20, 1997, the date of acquisition of OSA by Hewlett-Packard Company (HP). OSA implemented a first-generation yield-driven microwave CAD capability for Raytheon in 1985, followed by further innovations in linear and nonlinear microwave CAD technology for the Raytheon/Texas Instruments Joint Venture MIMIC Program. OSA introduced the CAE systems RoMPE™ in 1988, HarPE™ in 1989, OSA90™ and OSA90/hope™ in 1991, Empipe™ in 1992, Empipe3D™, and EmpipeExpress™ in 1996. OSA created the product *empath*™ in 1996, which was marketed by Sonnet Software, Inc. (Liverpool, NY). Dr. Bandler is President of Bandler Corporation, which he founded in 1997.

Dr. Bandler served as an associate editor of the *IEEE Transactions on Microwave Theory and Techniques* from 1969 to 1974, and has continued serving as a member of the editorial board. He was guest editor of the special issue of the *IEEE Transactions on Microwave Theory and Techniques on Computer-Oriented Microwave Practices* (1974) and guest coeditor of the special issue of the *IEEE Transactions on Microwave Theory and Techniques on Process-Oriented Microwave CAD and Modeling* (1992). He joined the editorial boards of the *International Journal of Numerical Modeling* in 1987, the *International Journal of Microwave and Millimeter-Wave Computer-Aided Engineering* in 1989, and *Optimization and Engineering* in 1998. He was a guest editor of the *International Journal of Microwave and Millimeter-Wave Computer-Aided Engineering* special issue "Optimization-Oriented Microwave CAD" (1997), and a guest editor of the *IEEE Transactions on Microwave Theory and Techniques* special issue "Automated Circuit Design Using Electromagnetic Simulators" (1997). He is guest coeditor of the upcoming *IEEE Transactions on Microwave Theory and Techniques* special issue on electromagnetics-based optimization of microwave components and circuits (2004). He is currently cochair of the MTT-1 Technical Committee on Computer-Aided Design.

Dr. Bandler published more than 350 papers from 1965 to 2002.

He contributed to *Modern Filter Theory and Design* (Wiley-Interscience, 1973) and to *Analog Methods for Computer-Aided Analysis and Diagnosis* (Marcel Dekker, Inc., 1988). Four of his papers have been reprinted in *Computer-Aided Filter Design* (IEEE Press, 1973), one in each of *Microwave Integrated Circuits* (Artech House, 1975), *Low-Noise Microwave Transistors and Amplifiers* (IEEE Press, 1981), *Microwave Integrated Circuits*, 2nd ed. (Artech House, 1985), *Statistical Design of Integrated Circuits* (IEEE Press, 1987), and *Analog Fault Diagnosis* (IEEE Press, 1987).

Dr. Bandler is a Fellow of the Royal Society of Canada, the Institute of Electrical and Electronics Engineers, the Institution of Electrical Engineers (Great Britain), and the Engineering Institute of Canada; he is a member of the Association of Professional Engineers of the Province of Ontario (Canada) and the MIT Electromagnetics Academy. He received the Automatic Radio Frequency Techniques Group (ARFTG) Automated Measurements Career Award in 1994.



Mostafa A. Ismail was born in Cairo, Egypt, on May 21, 1968. He received a B.Sc. degree in electronics and communications engineering and a master's degree in engineering mathematics from Cairo University, Egypt, in 1991 and 1995, respectively. He received a Ph.D. degree from McMaster University, Hamilton, Ontario, Canada, in 2001.

From October 1991 to August 1997 he was a teaching assistant in the Department of Engineering Mathematics and Physics, Faculty of Engineering, Cairo University. He joined the Simulation Optimization Systems Research Laboratory and the Department of Electrical and Computer Engineering at McMaster University as graduate student in 1997. He was granted a one-year Nortel Networks' Ontario Graduate Scholarship in Science and Technology (OGSST) for the academic year 2000–2001. In 2001 he joined Com Dev Ltd., Cambridge, Ontario, Canada where he is currently an advanced member of technical staff in corporate research and development. His work includes computer-aided design and modeling of microwave circuits, electromagnetic optimization, efficient optimization of waveguide circuits, computer-aided tuning, device modeling, and parameter extraction.



José Ernesto Rayas-Sánchez was born in Guadalajara, Jalisco, Mexico, on December 27, 1961. He received a B.Sc. degree in electronics engineering from the Instituto Tecnológico y de Estudios Superiores de Occidente (ITESO), Guadalajara, Mexico, in 1984, and a master's degree in electrical engineering from the Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM), Monterrey, Mexico, in 1989. He obtained a Ph.D. degree in electrical engineering at McMaster University, Hamilton, Ontario, Canada, in 2001.

He has been Professor in the Electrical and Computer Engineering Department, ITESO, since 1989. He was granted in 1997 a sabbatical leave to pursue doctoral studies in the Simulation Optimization Systems Research Laboratory, McMaster University. His research focuses on the development of novel methods and techniques for computer-aided modeling, design, and optimization of analog wireless electronic circuits and devices exploiting space mapping (SM) and artificial neural networks (ANN).

Dr. Rayas-Sánchez was the recipient of a 1997–2000 Consejo Nacional de Ciencia y Tecnología (CONACYT) scholarship presented by the Mexican government, as well as a 2000–2001 Ontario Graduate Scholarship (OGS) presented by the Ministry of Training for Colleges and Universities in Ontario. He was also the recipient of a 2001–2002 CONACYT Repatriation Grant presented by the Mexican government.

Dr. Rayas-Sánchez is a senior member of the IEEE.



Qi-jun Zhang received a B.Eng. degree from East China Engineering Institute, Nanjing, China, in 1982, and a Ph.D. degree in electrical engineering from McMaster University, Hamilton, Canada, in 1987.

He was with the System Engineering Institute, Tianjin University, Tianjin, China, in 1982 and 1983. From 1988 to 1990, he was with Optimization Systems Associates Inc. (OSA), Dundas, Ontario, Canada, developing advanced microwave optimization software. He joined the Department of Electronics, Carleton University, Ottawa, Canada, in 1990 where he is presently a professor. His research interests are neural network and optimization methods for high-speed/high-frequency circuit design, and he has published over 150 papers in the area. He is a coauthor of *Neural Networks for RF and Microwave Design* (Artech House, Boston, 2000), a coeditor of *Modeling and Simulation of High-Speed VLSI Interconnects* (Kluwer, Boston, 1994), a contributor to *Analog Methods for Computer-Aided Analysis and Diagnosis* (Marcel Dekker, New York, 1988), a guest coeditor for a special issue on high-speed VLSI interconnects for the *International Journal of Analog Integrated Circuits and Signal Processing* (Kluwer, Boston, 1994), and twice a guest editor for the special issues on Applications of ANN to RF and Microwave Design for the *International Journal of RF and Microwave CAE* (Wiley, 1999, 2002).

Dr. Zhang is a member of the Association of Professional Engineers of the Province of Ontario, Canada.