Theoretical Justification of Space-Mapping-Based Modeling Utilizing a Database and On-Demand Parameter Extraction

Slawomir Koziel, *Member, IEEE*, John W. Bandler, *Fellow, IEEE*, and Kaj Madsen

Abstract—We present a theoretical justification of a recently introduced surrogate modeling methodology based on space mapping (SM) that relies on an available database and on-demand parameter extraction. Fine-model data, the so-called base set, is assumed available in the region of interest. To evaluate the surrogate, we perform parameter extraction with weighting coefficients dependent on the distance between the point of interest and base points. We provide theoretical results showing that the new methodology can assure any accuracy that is required (provided the base set is dense enough), which is not the case for our benchmark SM modeling methodology. Illustrative examples emphasizing differences between modeling methodologies are provided.

Index Terms—Computer-aided design (CAD), electromagnetic (EM) modeling, microwave circuits, space mapping (SM), surrogate modeling.

I. INTRODUCTION

STATISTICAL analysis and yield optimization are crucial to manufacturability-driven designs in a time-to-market development environment and demand fast accurate device and component models. Full-wave electromagnetic (EM) simulations of microwave structures offer accuracy at the cost of CPU effort. High CPU cost is undesirable from the point-of-view of direct statistical analysis and design. The space mapping (SM) concept introduced by Bandler *et al.* [1], [2] addresses this issue.

SM assumes that a high fidelity CPU-intensive "fine" model is accompanied by a low fidelity or "coarse" model. The "coarse" model can be a simplified representation such as an equivalent circuit with empirical formulas. SM modeling [3]–[8] exploits the speed of the coarse model and the accuracy of the fine model to develop fast accurate enhanced models (surrogates) valid over a wide range of parameter values.

The standard SM modeling methodology [9], [10] is based on setting up the surrogate model using a small amount of fine-model data (usually $2n+1$ points, where n is the number of design variables). Extraction of the model parameters is performed over the whole set of this data. This methodology is simple and provides accuracy that is good enough for some applications. It has, however, a number of limitations, in particular, a limited capability to model nonlinearity of the fine model, limited performance for higher dimensional problems, and difficulty handling a large amount of the fine-model data. The last drawback is particularly important because in order to increase accuracy of the surrogate model over some limit, we have to provide more and more fine-model data. The only way to utilize this data in a standard SM model is to increase the number of model parameters, which makes the parameter-extraction process longer and more difficult, leaving alone the problem of model definition that would allow us to properly introduce linear and nonlinear terms to follow fine-model nonlinearity.

To alleviate the foregoing difficulties, a new SM-based surrogate modeling methodology has been introduced in [11]. It requires an available database and performs on-demand parameter extraction. To evaluate the surrogate, we perform parameter extraction with weighting coefficients dependent on the distance between the point of interest and the base points. In other words, this methodology uses local fine-model information.

In this study, we provide a theoretical justification of the method [11] and show that this methodology can assure any required accuracy provided that the base set is dense enough. We also give a matching error estimate for the surrogate model with respect to the fine model.

II. SM MODELING WITH VARIABLE WEIGHT COEFFICIENTS [11]

Here, in order to set up the notation necessary for our subsequent theoretical considerations, we briefly recall the SM modeling methodology [11].

Let $\boldsymbol{R}_f\colon X_f \to R^m$ and $\boldsymbol{R}_c\colon X_c \to R^m$ denote the fine and coarse model response vectors, where $X_f \subseteq R^n$ and $X_c \subseteq R^n$ are design variable domains of the fine and coarse models, respectively. For example, $\boldsymbol{R}_f(\boldsymbol{x})$ and $\boldsymbol{R}_c(\boldsymbol{x})$ may represent the magnitude of a transfer function at m chosen frequencies. We denote by $X_R \subseteq X_f$ the region of interest in which we want enhanced matching between the surrogate and the fine model. We assume that X_R is an n-dimensional interval in R^n with center at reference point $\boldsymbol{x}^0 = [x_{0.1} \ldots x_{0.n}]^T \in R^n$

$$X_R = [\boldsymbol{x}^0 - \boldsymbol{\delta}, \boldsymbol{x}^0 + \boldsymbol{\delta}] = [x_{0.1} - \delta_1, x_{0.1} + \delta_1] \times \cdots \times [x_{0.n} - \delta_n, x_{0.n} + \delta_n] \quad (1)$$

S. Koziel is with the Simulation Optimization Systems Research Laboratory, Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada L8S 4K1 (e-mail: koziels@mcmaster.ca).

J. W. Bandler is with the Simulation Optimization Systems Research Laboratory, Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada L8S 4K1 and also with Bandler Corporation, Dundas, ON, Canada L9H 5E7 (e-mail: bandler@mcmaster.ca).

K. Madsen is with Informatics and Mathematical Modelling, Technical University of Denmark, DK-2800 Lyngby, Denmark (e-mail: km@imm.dtu.dk).

where $\boldsymbol{\delta} = [\delta_1 \ldots \delta_n]^T$ determines the size of X_R. We use $X_R(\boldsymbol{x}^0, \boldsymbol{\delta})$ to denote the region of interest defined by $\boldsymbol{x}^0$ and $\boldsymbol{\delta}$. Suppose we have the base set $X_B = \{\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^N\} \subset X_R(\boldsymbol{x}^0, \boldsymbol{\delta})$, where N is the number of base points such that the fine-model response is known at all points $\boldsymbol{x}^j$, $j = 1, 2, \ldots, N$. We do not assume any particular location of the base points.

We define a generic surrogate model $\boldsymbol{R}_s: X_f \times M_{m \times m} \times M_{n \times n} \times M_{n \times 1} \times M_{m \times 1} \to R^m$ as [11]

$$\boldsymbol{R}_s(\boldsymbol{x}, A, B, c, d) = A \cdot \boldsymbol{R}_c(B \cdot \boldsymbol{x} + c) + d \qquad (2)$$

with matrices $A = \text{diag}\{a_1, \ldots, a_m\}$, $B \in M_{n \times n}$, $c \in M_{n \times 1}$, and $d \in M_{m \times 1}$ ($M_{k \times l}$ denotes the set of $k \times l$ real matrices) found using the parameter extraction

$$(A, B, c, d)$$
$$= \arg \min_{(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta})} \sum_{k=1}^{N} w_k(\boldsymbol{x}) \| \boldsymbol{R}_f(x^k) - \boldsymbol{R}_s(\boldsymbol{x}^k, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta}) \|. \quad (3)$$

The flexibility of the generic model (2) can be adjusted by imposing constraints on the parameter-extraction procedure. It can also be enhanced by introducing additional parameters (see [11]).

The weighting coefficients in (3) are functions of $\boldsymbol{x}$. I Coefficients w_k are calculated according to

$$w_k = w_k(\boldsymbol{x}; C, \lambda)$$
$$= \frac{\exp\left(-\dfrac{\|\boldsymbol{x} - \boldsymbol{x}^k\|^2}{C \cdot \lambda^2}\right)}{\sum\limits_{j=1}^{N} \exp\left(-\dfrac{\|\boldsymbol{x} - \boldsymbol{x}^j\|^2}{C \cdot \lambda^2}\right)}, \qquad k = 1, 2, \ldots, N \quad (4)$$

where $\boldsymbol{x}$ is the evaluation point, and $\lambda = \lambda(\boldsymbol{\delta}, N)$ is a *characteristic distance* depending on the size of the region of interest and the number of base points

$$\lambda(\boldsymbol{\delta}, N) = \frac{2}{n N^{1/n}} \sum_{i=1}^{n} \delta_i. \qquad (5)$$

If the base points are uniformly distributed in X_R, $\lambda = \lambda(\boldsymbol{\delta}, N)$ is just an average distance between neighboring points. Constant $C > 0$ determines how fast the weighting coefficients decrease with an increase of base-point distance from $\boldsymbol{x}$. Reference [11] contains a discussion on the implementation details of this method.

III. MODELING ERROR VERSUS CHARACTERISTIC DISTANCE OF THE BASE SET

It is intuitively obvious that modeling accuracy, according to the methodology presented in Section II, depends on the characteristic distance λ, in particular, that accuracy improves with decreasing λ. Here, we provide theoretical results showing that our methodology can assure any accuracy that is required (provided the base set is dense enough, i.e., λ is small enough), which is not the case for the standard methodology [9], [10]. We also give an error estimate for the surrogate model with respect to the fine model. Let us start with the following remarks.

Remark 1: It follows that the standard SM modeling technique [9], [10] is a particular case of the new technique (2)–(4), as it can be obtained from (2)–(4) by choosing a standard (i.e., star-distribution-like) X_B and letting $C \to \infty$ in (4).

Remark 2: It should be emphasized that if the coarse model $\boldsymbol{R}_c$ is continuous, then the surrogate model (2) is continuous with respect to $\boldsymbol{x}$ regardless of the fact that evaluation of the surrogate requires a separate parameter extraction for every argument. This follows from the fact that both w_k and matrices A, B, c, and d are continuous functions of $\boldsymbol{x}$. The latter assumes that parameter extraction (3) has a unique solution for any values of the weighting coefficients w_k, $k = 1, 2, \ldots, N$, and both $\boldsymbol{R}_f$ and $\boldsymbol{R}_c$ are continuous; the uniqueness assumption can be replaced by the assumption of regularity of the solution to (3) with respect to the weighting coefficients, e.g., ordinary or Lipschitz continuity.

As mentioned before, one can expect that the modeling accuracy depends on the characteristic distance λ. In fact, it is possible to prove a rigorous result showing that the modeling error can be arbitrarily small as $\lambda \to 0$.

For the purpose of theoretical considerations, we shall consider a slightly generalized formulation of the SM-based surrogate model. Let $\bar{\boldsymbol{R}}_s : X_R \times X_p \to R^m$ be a generic surrogate model, where X_R is the region of interest and X_p is a parameter domain. For any given base set X_B, the actual surrogate model response at any $\boldsymbol{x} \in X_R$ is defined as

$$\boldsymbol{R}_s(\boldsymbol{x}) = \bar{\boldsymbol{R}}_s(\boldsymbol{x}, \boldsymbol{p}(\boldsymbol{x})) \qquad (6)$$

where

$$\boldsymbol{p}(\boldsymbol{x}) = \arg \min_{\boldsymbol{z} \in X_p} \sum_{k=1}^{N} w_k(\boldsymbol{x}) \| \boldsymbol{R}_f(\boldsymbol{x}^k) - \bar{\boldsymbol{R}}_s(\boldsymbol{x}^k, \boldsymbol{z}) \| \qquad (7)$$

with coefficients w_k defined by (4) for given constant C. For simplicity, we assume that no sensitivity information is used in the model. Note that the model $\bar{\boldsymbol{R}}_s$ is a compact way of writing $\boldsymbol{R}_s$, in which we represented all the parameters (in the case of $\boldsymbol{R}_s$, parameters are matrices A, B, c, and d) by a single parameter vector $\boldsymbol{p}$. We also introduce single-point parameter extraction denoted as

$$\boldsymbol{p}^s(\boldsymbol{x}) = \arg \min_{\boldsymbol{z} \in X_p} \| \boldsymbol{R}_f(\boldsymbol{x}) - \bar{\boldsymbol{R}}_s(\boldsymbol{x}, \boldsymbol{z}) \|. \qquad (8)$$

Definition (8) should be understood in the following sense: if $\boldsymbol{x} \in X_B$, then $\boldsymbol{p}^s(\boldsymbol{x})$ is the limit of $\boldsymbol{p}(\boldsymbol{x})$ obtained for $C \to 0^+$. This means that $\boldsymbol{p}^s(\boldsymbol{x})$ is uniquely defined (if $\boldsymbol{p}(\boldsymbol{x})$ are unique for any nonzero C) even if problem (8) has nonunique solutions. In order to assure the existence of the limit, we need to extend definition (4) for $C = 0$ and $\boldsymbol{x} \in X_B$ as follows: $w_k = 1$ if $\boldsymbol{x}_k = \boldsymbol{x}$ and $w_k = 0$ if $\boldsymbol{x}_k \neq \boldsymbol{x}$. We also need to assume continuity of the functions involved, in particular, the continuity of solution to (7) with respect to parameter C for $C > 0$.

Assumption 1: Suppose that the following conditions hold.
(i) X_R and X_p are compact sets.
(ii) $\boldsymbol{R}_f$ and $\bar{\boldsymbol{R}}_s$ are continuous on X_R and $X_R \times X_p$, respectively.

(iii) For any base set X_B and any $\boldsymbol{x} \in X_B$, the solution to the problem $\boldsymbol{p}^s(\boldsymbol{x}) = \arg \min_{\boldsymbol{z} \in X_p} \|\boldsymbol{R}_f(\boldsymbol{x}) - \bar{\boldsymbol{R}}_s(\boldsymbol{x}, \boldsymbol{z})\|$ [see explanation under (8)] is unique. Moreover, the solution to (8) is uniformly continuous with respect to perturbation in the following sense: if $d: X_R \times X_p \to R$ is the (bounded) perturbation function, then for any $\varepsilon > 0$ there is a $\delta > 0$ (independent of $\boldsymbol{x}$) such that $\|\boldsymbol{p}^s(\boldsymbol{x}) - \arg \min_{\boldsymbol{z} \in X_p}(\|\boldsymbol{R}_f(\boldsymbol{x}) - \bar{\boldsymbol{R}}_s(\boldsymbol{x}, \boldsymbol{z})\| + d(\boldsymbol{x}, \boldsymbol{z}))\| < \varepsilon$ provided that $\sup\{\| d(\boldsymbol{v}, \boldsymbol{y}) \| : \boldsymbol{v} \in X_R, \boldsymbol{y} \in X_p\} < \delta$.

(iv) For any base set X_B and any $\boldsymbol{x} \in X_B$, the solution to the problem $\boldsymbol{p}^s(\boldsymbol{x}) = \arg \min_{\boldsymbol{z} \in X_p} \|\boldsymbol{R}_f(\boldsymbol{x}) - \bar{\boldsymbol{R}}_s(\boldsymbol{x}, \boldsymbol{z})\|$ is such that $\|\boldsymbol{R}_f(\boldsymbol{x}) - \boldsymbol{R}_s(\boldsymbol{x})\| = \|\boldsymbol{R}_f(\boldsymbol{x}) - \bar{\boldsymbol{R}}_s(\boldsymbol{x}, \boldsymbol{p}^s(\boldsymbol{x}))\| = 0$.

Remark 3: Assumption 1 is quite natural. First of all, X_R is usually a multidimensional closed rectangle so it is compact. Similarly, X_p can be made compact by setting suitable bounds for the surrogate model parameters. Continuity of the fine and coarse model [which usually implies continuity of the surrogate, e.g., as in (3)] is also typical. Assumption 1 (iii) is more complex, although it also typically follows, at least for continuous perturbation functions. Eventually, Assumption 1 (iv) is always satisfied if the surrogate model allows output SM, i.e., the transformation of the coarse model image (e.g., in the case of model (2) this can be done by either of the matrices $\boldsymbol{A}$ or $\boldsymbol{d}$).

Theorem 1: Suppose that Assumption 1 is satisfied. The surrogate model (6) and (7) is then arbitrarily accurate in the following sense. For any $\varepsilon > 0$, there is a $\delta > 0$ such that for any base set X_B that satisfies the condition $\forall_{\boldsymbol{x} \in X_R} \exists_{\boldsymbol{x}_b \in X_B} \|\boldsymbol{x} - \boldsymbol{x}_b\| < \delta$ (i.e., X_B is uniform enough), we have $\| \boldsymbol{R}_f(\boldsymbol{x}) - \boldsymbol{R}_s(\boldsymbol{x}) \| < \epsilon$ for any $\boldsymbol{x} \in X_R$, provided that the constant C in (4) is sufficiently small.

Proof: Let us take $\varepsilon > 0$. We shall show that there exist $\delta > 0$ and $C > 0$ such that the assertion of the theorem holds. For any point $\boldsymbol{x}$ in X_R, any base set X_B, and any point $\bar{\boldsymbol{x}} \in X_B$, we have

$$\begin{aligned} \|\boldsymbol{R}_f(\boldsymbol{x}) - \boldsymbol{R}_s(\boldsymbol{x})\| &= \|\boldsymbol{R}_f(\boldsymbol{x}) - \bar{\boldsymbol{R}}_s(\boldsymbol{x}, \boldsymbol{p}(\boldsymbol{x}))\| \\ &\leq \|\boldsymbol{R}_f(\boldsymbol{x}) - \bar{\boldsymbol{R}}_s(\boldsymbol{x}, \boldsymbol{p}^s(\bar{\boldsymbol{x}}))\| \\ &\quad + \|\bar{\boldsymbol{R}}_s(\boldsymbol{x}, \boldsymbol{p}^s(\bar{\boldsymbol{x}})) - \bar{\boldsymbol{R}}_s(\boldsymbol{x}, \boldsymbol{p}(\boldsymbol{x}))\|. \end{aligned} \quad (9)$$

From Assumption 1 (ii), $\boldsymbol{R}_f$ and $\bar{\boldsymbol{R}}_s$ are continuous on X_R and $X_R \times X_p$, respectively, and because both X_R and $X_R \times X_p$ are compact [Assumption 1 (i)], they are, in fact, uniformly continuous. Thus, there is a $\delta_1 > 0$ (independent of $\boldsymbol{x}$) such that if $\|\boldsymbol{x} - \bar{\boldsymbol{x}}\| < \delta_1$, then

$$\|(\boldsymbol{R}_f(\boldsymbol{x}) - \bar{\boldsymbol{R}}_s(\boldsymbol{x}, \boldsymbol{p}^s(\bar{\boldsymbol{x}}))) - (\boldsymbol{R}_f(\bar{\boldsymbol{x}}) - \bar{\boldsymbol{R}}_s(\bar{\boldsymbol{x}}, \boldsymbol{p}^s(\bar{\boldsymbol{x}})))\| < \frac{\varepsilon}{2}. \quad (10)$$

However, according to Assumption 1 (iv), we have

$$\boldsymbol{R}_f(\bar{\boldsymbol{x}}) - \bar{\boldsymbol{R}}_s(\bar{\boldsymbol{x}}, \boldsymbol{p}^s(\bar{\boldsymbol{x}})) = 0. \quad (11)$$

Let X_B be any base set that satisfies $\forall_{\boldsymbol{x} \in X_R} \exists_{\bar{\boldsymbol{x}} \in X_B} \|\boldsymbol{x} - \bar{\boldsymbol{x}}\| < \delta_1$. For any $\boldsymbol{x} \in X_R$, there is then an $\bar{\boldsymbol{x}} \in X_B$ such that

$$\|\boldsymbol{R}_f(\boldsymbol{x}) - \bar{\boldsymbol{R}}_s(\boldsymbol{x}, \boldsymbol{p}^s(\bar{\boldsymbol{x}}))\| < \frac{\varepsilon}{2}. \quad (12)$$

We would now like to find a condition under which the term $\|\bar{\boldsymbol{R}}_s(\boldsymbol{x}, \boldsymbol{p}^s(\bar{\boldsymbol{x}})) - \bar{\boldsymbol{R}}_s(\boldsymbol{x}, \boldsymbol{p}(\boldsymbol{x}))\|$ is smaller than $\varepsilon/2$. Due to the uniform continuity of $\bar{\boldsymbol{R}}_s$ on $X_R \times X_p$, there is a $\delta_2 > 0$ (independent of $\boldsymbol{x}$) such that

$$\|\bar{\boldsymbol{R}}_s(\boldsymbol{x}, \boldsymbol{p}^s(\bar{\boldsymbol{x}})) - \bar{\boldsymbol{R}}_s(\boldsymbol{x}, \boldsymbol{p}(\boldsymbol{x}))\| < \frac{\varepsilon}{2}$$

whenever

$$\|\boldsymbol{p}^s(\bar{\boldsymbol{x}}) - \boldsymbol{p}(\boldsymbol{x})\| < \delta_2. \quad (13)$$

However,

$$\boldsymbol{p}^s(\bar{\boldsymbol{x}}) = \arg \min_{\boldsymbol{z} \in X_p} \|\boldsymbol{R}_f(\bar{\boldsymbol{x}}) - \bar{\boldsymbol{R}}_s(\bar{\boldsymbol{x}}, \boldsymbol{z})\| \quad (14)$$

and (15), shown at the bottom of this page, in which we have taken into consideration that the weighting coefficients are normalized, i.e., $\sum_{k=1}^N w_k(\boldsymbol{x}) = 1$ for any $\boldsymbol{x}$. Let us define a function $\bar{d}: X_R \times X_p \to R$ as $\bar{d}(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{R}_f(\boldsymbol{x}) - \bar{\boldsymbol{R}}_s(\boldsymbol{x}, \boldsymbol{y})\|$. Using this function, we can write (16), shown at the bottom of this page. Let us now define function $d: X_R \times X_p \to R$ as $d(\boldsymbol{x}, \boldsymbol{z}) = \sum_{k=1}^N w_k(\boldsymbol{x}) [\bar{d}(\boldsymbol{x}^k, \boldsymbol{z}) - \bar{d}(\bar{\boldsymbol{x}}, \boldsymbol{z})]$. Using this function, we can rewrite (16) as

$$\boldsymbol{p}(\boldsymbol{x}) = \arg \min_{\boldsymbol{z} \in X_p} \left\{ \|\boldsymbol{R}_f(\bar{\boldsymbol{x}}) - \bar{\boldsymbol{R}}_s(\bar{\boldsymbol{x}}, \boldsymbol{z})\| + d(\boldsymbol{x}, \boldsymbol{z}) \right\}. \quad (17)$$

$$\begin{aligned} \boldsymbol{p}(\boldsymbol{x}) &= \arg \min_{\boldsymbol{z} \in X_p} \sum_{k=1}^N w_k(\boldsymbol{x}) \|\boldsymbol{R}_f(\boldsymbol{x}^k) - \bar{\boldsymbol{R}}_s(\boldsymbol{x}^k, \boldsymbol{z})\| \\ &= \arg \min_{\boldsymbol{z} \in X_p} \left\{ \left(\sum_{k=1}^N w_k(\boldsymbol{x}) \right) \|\boldsymbol{R}_f(\bar{\boldsymbol{x}}) - \bar{\boldsymbol{R}}_s(\bar{\boldsymbol{x}}, \boldsymbol{z})\| + \sum_{k=1}^N w_k(\boldsymbol{x}) \left(\|\boldsymbol{R}_f(\boldsymbol{x}^k) - \bar{\boldsymbol{R}}_s(\boldsymbol{x}^k, \boldsymbol{z})\| - \|\boldsymbol{R}_f(\bar{\boldsymbol{x}}) - \bar{\boldsymbol{R}}_s(\bar{\boldsymbol{x}}, \boldsymbol{z})\| \right) \right\} \\ &= \arg \min_{\boldsymbol{z} \in X_p} \left\{ \|\boldsymbol{R}_f(\bar{\boldsymbol{x}}) - \bar{\boldsymbol{R}}_s(\bar{\boldsymbol{x}}, \boldsymbol{z})\| + \sum_{k=1}^N w_k(\boldsymbol{x}) \left(\|\boldsymbol{R}_f(\boldsymbol{x}^k) - \bar{\boldsymbol{R}}_s(\boldsymbol{x}^k, \boldsymbol{z})\| - \|\boldsymbol{R}_f(\bar{\boldsymbol{x}}) - \bar{\boldsymbol{R}}_s(\bar{\boldsymbol{x}}, \boldsymbol{z})\| \right) \right\} \end{aligned} \quad (15)$$

$$\boldsymbol{p}(\boldsymbol{x}) = \arg \min_{\boldsymbol{z} \in X_p} \left\{ \|\boldsymbol{R}_f(\bar{\boldsymbol{x}}) - \bar{\boldsymbol{R}}_s(\bar{\boldsymbol{x}}, \boldsymbol{z})\| + \sum_{k=1}^N w_k(\boldsymbol{x}) \left[\bar{d}(\boldsymbol{x}^k, \boldsymbol{z}) - \bar{d}(\bar{\boldsymbol{x}}, \boldsymbol{z}) \right] \right\} \quad (16)$$

We would like to take advantage of Assumption 1 (iii). In particular, there is a $\delta_3 > 0$ such that if

$$\sup\{|d(\boldsymbol{v},\boldsymbol{y})| : \ \boldsymbol{v} \in X_R, \ \boldsymbol{y} \in X_p\} < \delta_3 \qquad (18)$$

then

$$\|\boldsymbol{p}^s(\bar{\boldsymbol{x}}) - \arg \min_{\boldsymbol{z} \in X_p} (\|\boldsymbol{R}_f(\bar{\boldsymbol{x}}) - \bar{\boldsymbol{R}}_s(\bar{\boldsymbol{x}},\boldsymbol{z})\| + d(\bar{\boldsymbol{x}},\boldsymbol{z}))\| < \delta_2 \quad (19)$$

for any $\bar{\boldsymbol{x}}$ that belongs to some base set. What we need to do now is to show that there is a base set X_B and constant C for which (18) holds. Let $M = \sup\{\bar{d}(\boldsymbol{x},\boldsymbol{y}) : (\boldsymbol{x},\boldsymbol{y}) \in X_R \times X_p\}$ (M is well-defined because $\bar{d}$ is continuous on the compact set). Let $\delta_4 > 0$ be such that $|\bar{d}(\boldsymbol{x},\boldsymbol{z}) - \bar{d}(\boldsymbol{y},\boldsymbol{z})| < \delta_3/2$ for all $\boldsymbol{z} \in X_p$ whenever $\|\boldsymbol{x} - \boldsymbol{y}\| < \delta_4$.

Let X_B be any base set such that $\forall_{\tilde{\boldsymbol{x}} \in X_R} \exists_{\bar{\boldsymbol{x}} \in X_B} \|\tilde{\boldsymbol{x}} - \bar{\boldsymbol{x}}\| < \delta_4$. For such X_B and any $\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}} \in X_B$ satisfying $\|\bar{\boldsymbol{x}} - \boldsymbol{x}\| < \delta_4$, $\|\bar{\boldsymbol{y}} - \boldsymbol{x}\| < \delta_4$, we have $|\bar{d}(\bar{\boldsymbol{x}},\boldsymbol{z}) - \bar{d}(\bar{\boldsymbol{y}},\boldsymbol{z})| < \delta_3/2$ for all $\boldsymbol{z} \in X_p$. We divide X_B into two subsets: $X_{B.\text{inner}} = \{\tilde{\boldsymbol{x}} \in X_B : \|\tilde{\boldsymbol{x}} - \boldsymbol{x}\| < \delta_4\}$ and $X_{B.\text{outer}} = \{\tilde{\boldsymbol{x}} \in X_B : \|\tilde{\boldsymbol{x}} - \boldsymbol{x}\| \geq \delta_4\}$. Obviously, we have $X_{B.\text{inner}} \cup X_{B.\text{outer}} = X_B$, and $X_{B.\text{inner}} \cap X_{B.\text{outer}} = \emptyset$. Let us denote by $\bar{\boldsymbol{x}}$ the point from X_B that is closest to $\boldsymbol{x}$. We define

$$d_1(\boldsymbol{x},\boldsymbol{z}) = \sum_{\{k:\boldsymbol{x}^k \in X_{B.\text{inner}}\}} w_k(\boldsymbol{x})[\bar{d}(\boldsymbol{x}^k,\boldsymbol{z}) - \bar{d}(\bar{\boldsymbol{x}},\boldsymbol{z})] \qquad (20)$$

$$d_2(\boldsymbol{x},\boldsymbol{z}) = \sum_{\{k:\boldsymbol{x}^k \in X_{B.\text{outer}}\}} w_k(\boldsymbol{x})[\bar{d}(\boldsymbol{x}^k,\boldsymbol{z}) - \bar{d}(\bar{\boldsymbol{x}},\boldsymbol{z})]. \qquad (21)$$

Clearly, $d(\boldsymbol{x},\boldsymbol{z}) = d_1(\boldsymbol{x},\boldsymbol{z}) + d_2(\boldsymbol{x},\boldsymbol{z})$. Due to our definition of $X_{B.\text{inner}}$, we have

$$|d_1(\boldsymbol{x},\boldsymbol{z})| < \sum_{\{k:\boldsymbol{x}^k \in X_{B.\text{inner}}\}} \frac{w_k(\boldsymbol{x})\delta_3}{2} \leq \frac{\delta_3}{2}. \qquad (22)$$

Let us now set the constant C such that for any k such that $\boldsymbol{x}_k \in X_{B.\text{outer}}$, we have $w_k < \delta_3/4NM$. For this C, we have

$$|d_2(\boldsymbol{x},\boldsymbol{z})| < \sum_{\{k:\boldsymbol{x}^k \in X_{B.\text{outer}}\}} \frac{\delta_3}{4NM}|\bar{d}(\boldsymbol{x}^k,\boldsymbol{z}) - \bar{d}(\bar{\boldsymbol{x}},\boldsymbol{z})|$$

$$< \sum_{\{k:\boldsymbol{x}^k \in X_{B.\text{outer}}\}} \frac{\delta_3}{4NM}2M \leq \frac{\delta_3}{2}. \qquad (23)$$

Let $\delta = \min\{\delta_1,\delta_4\}$ and let X_B be any base set for which $\forall_{\tilde{\boldsymbol{x}} \in X_R} \exists_{\bar{\boldsymbol{x}} \in X_B} \|\tilde{\boldsymbol{x}} - \bar{\boldsymbol{x}}\| < \delta$. For this set, it follows from (22) and (23) that $|d(\boldsymbol{x},\boldsymbol{z})| \leq |d_1(\boldsymbol{x},\boldsymbol{z})| + |d_2(\boldsymbol{x},\boldsymbol{z})| < \delta_3$ on X_p. At the same time, it follows from (9), (12), (13), and (19) that, for our choice of X_B, we have $\|\boldsymbol{R}_f(\boldsymbol{x}) - \boldsymbol{R}_s(\boldsymbol{x})\| < \varepsilon$. Since the above reasoning is valid for any $\boldsymbol{x} \in X_R$, this ends the proof of the theorem.

Remark 4: Adjustment of the decrease rate of coefficients w_k (by changing C) plays a crucial role in the proof of Theorem 1. In particular, it is not possible in general for the standard model to be arbitrarily accurate, regardless of the base set used (i.e., even if the characteristic distance of the base set is very small).

Remark 5: Results equivalent to Theorem 1 can be proven for a different choice of the formula that determines weighting
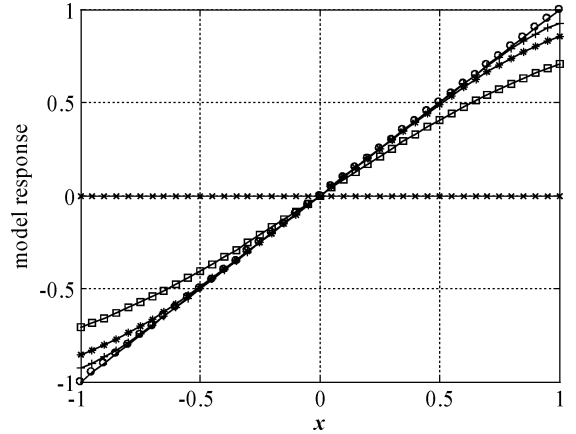

Fig. 1. Example 1: fine-model response ($\circ$), standard SM surrogate model responses for $k = 1, 2, 4$ ($\times$), and the new SM surrogate model responses for $k = 1$ ($\square$), $k = 2$ ($*$), and $k = 4$ ($+$).

coefficients w_k [cf. (4)]. The only requirement is that the formula allows changing the decrease rate of w_k while moving away from the evaluation point (e.g., by proper adjustment of the control parameters).

Remark 6: Good coarse models allow us to obtain very accurate surrogate models even for sparse base sets. However, it follows from Theorem 1 that even for poor coarse models, we still have the property of making the surrogate model error arbitrarily small provided that the base set is sufficiently "dense" and one can obtain perfect matching for single-point parameter extraction (this can be guaranteed by any kind of output SM). This follows from the fact that there is no assumption about "quality" of the coarse model in the Theorem: the basic analytical condition is continuity.

Example 1: Let $X_R = [-1, 1]$, $X_p = R$, and $\boldsymbol{R}_f: X_R \to R$, $\bar{\boldsymbol{R}}_s : X_R \times X_p \to R$ are defined as

$$\boldsymbol{R}_f(\boldsymbol{x}) = \boldsymbol{x} \qquad (24)$$

$$\bar{\boldsymbol{R}}_s(\boldsymbol{x},\boldsymbol{p}) = \boldsymbol{p}. \qquad (25)$$

Let

$$X_{B.k} = \left\{-\frac{k}{k}, -\frac{k-1}{k}, \ldots, -\frac{1}{k}, \frac{0}{k}, \frac{1}{k}, \ldots, \frac{k-1}{k}, \frac{k}{k}\right\}. \qquad (26)$$

For the standard model (i.e., the one with all weighting coefficients the same and equal to 1), for any integer $k > 0$, we have $R_s(\boldsymbol{x}) = \bar{R}_s(\boldsymbol{x},p(\boldsymbol{x})) = \bar{R}_s(\boldsymbol{x},0) \equiv 0$, which implies that the modeling error is the same, and does not depend on the choice of k in the base set (26). On the other hand, the modeling error for the new model can be reduced to zero in the limit $k \to \infty$ according to Theorem 1. As an illustration, Fig. 1 shows the fine-model response, as well as the standard and the new SM surrogate model responses for $k = 1, 2$, and 4.

Example 2: Consider a two-section impedance transformer example [12]. Both fine and coarse models in this problem are circuit models with design variables being the lengths of the transmission lines. The region of interest is defined by $\boldsymbol{x}^0 = [90\ 90]^T$ (coarse model optimal solution), and size $\boldsymbol{\delta} = [10\ 10]^T$. Fig. 2 shows a comparison of the average modeling error versus characteristic distance λ of the base set for
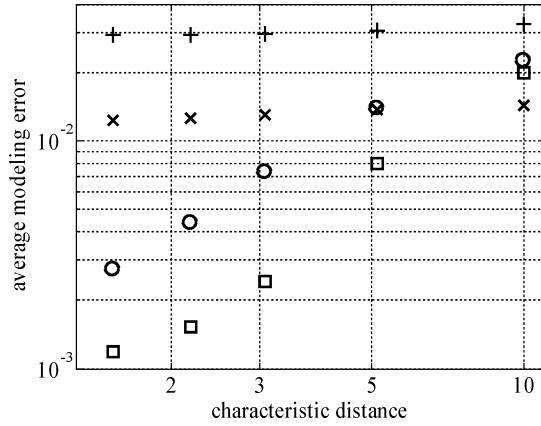
Fig. 2. Example 2: average modeling error versus characteristic distance λ of the base set for: standard SM model with input and output SM ($+$), standard SM model with input, output, frequency, and implicit SM ($\times$), SM model with variable weight coefficients using input and output SM surrogate ($\circ$), and SM model with variable weight coefficients using input, output, frequency, and implicit SM ($\square$).

Fig. 3. Example 3: average modeling error versus scaling factor C for SM model with variable weight coefficients for uniform base sets with nine points ($+$), 16 points ($\times$), 25 points ($\circ$), and 49 points ($\square$).

the standard SM model with the surrogate model using input and output SM ($+$), the standard SM model with the surrogate using input, output, frequency, and implicit SM ($\times$), and the SM model with variable weight coefficients using an input and output SM surrogate ($\circ$), and input, output, frequency, and implicit SM ($\square$). The results were obtained for 100 random test points. Characteristic distances range from approximately 1.5 (which corresponds to a uniform mesh base set with 100 points) to approximately 10 (uniform mesh with four points)

The results clearly show the difference between the standard and new SM modeling technique. The modeling error for the standard SM model is almost independent of the base set size. The accuracy improvement can only be observed for large values of λ; after the number of base points is increased, the standard SM model is not able to exploit all available fine-model information due to the limited flexibility of the surrogate. The only way of increasing the model accuracy in a significant way is to increase the number of model parameters (in our case, by introducing additional degrees of freedom with frequency and implicit SM). In case of SM modeling with variable weight coefficients, modeling error is decreasing while characteristic distance is going down, as predicted in Theorem 1. More results can be found in [11].

Example 3: In this example, we again use the two-section transformer example in order to investigate the dependence of the modeling error on the value of the scaling factor C. As before, the region of interest is defined by $\boldsymbol{x}^0 = [90\ 90]^T$ (coarse model optimal solution) and size $\boldsymbol{\delta} = [10\ 10]^T$. Fig. 3 shows comparison of the average modeling error versus scaling factor C for four different base sets being uniform meshes with nine points ($+$), 16 points ($\times$), 25 points ($\circ$), and 49 points ($\square$). The results were obtained for 100 random test points.

The results show that $C = 1$ is, for the considered example, the optimal value of the scaling factor, which is independent of the density of the base set. This independence of the base set can be explained by the construction of (4), which is used for calculating the weight coefficients necessary to evaluate the SM surrogate model. In particular, (4) contains the square of
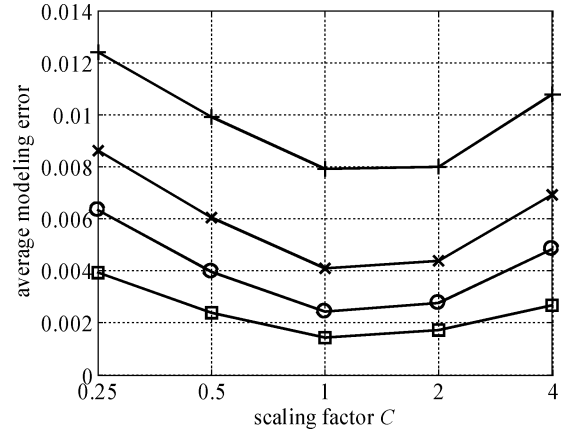
the characteristic distance λ. Since λ is nothing else but the average distance between base points, the density of the base set is already taken into account in (4) and the distribution of the weight factors in the neighborhood of any evaluation point $\boldsymbol{x}$ is invariant with respect to the base set (or, more specifically, to the relative distance between $\boldsymbol{x}$ and base points $\|\boldsymbol{x} - \boldsymbol{x}^k\|/\lambda$). Similar experiments performed for other examples (not shown here) indicate that the optimal value of C is equal or close to 1 for most cases.

The following result gives an error estimate for the surrogate model (2) and (3).

Assumption 2: Suppose that functions $\boldsymbol{R}_f$ and $\bar{\boldsymbol{R}}_s$ satisfy the following conditions:

(i) $\Delta \boldsymbol{R} : X_R \times X_p \to R^m, \Delta \boldsymbol{R}(\boldsymbol{x}, \boldsymbol{p}) = \boldsymbol{R}_f(\boldsymbol{x}) - \bar{\boldsymbol{R}}_s(\boldsymbol{x}, \boldsymbol{p})$ is Lipschitz continuous with respect to the first variable, i.e.,

$$\|\Delta \boldsymbol{R}(\boldsymbol{x}, \boldsymbol{p}) - \Delta \boldsymbol{R}(\boldsymbol{y}, \boldsymbol{p})\| \leq L_1 \|\boldsymbol{x} - \boldsymbol{y}\| \qquad (27)$$

where $L_1 \in R_+$,

(ii) $\bar{\boldsymbol{R}}_s$ is Lipschitz continuous with respect to the second variable, i.e.,

$$\|\bar{\boldsymbol{R}}_s(\boldsymbol{x}, \boldsymbol{p}_1) - \bar{\boldsymbol{R}}_s(\boldsymbol{x}, \boldsymbol{p}_2)\| \leq L_2 \|\boldsymbol{p}_1 - \boldsymbol{p}_2\|. \qquad (28)$$

where $L_2 \in R_+$.

Theorem 2: Suppose that Assumption 2 is satisfied and X_B is a base set. Let δ be such that $\forall_{\boldsymbol{x} \in X_R} \exists_{\bar{\boldsymbol{x}} \in X_B} \|\boldsymbol{x} - \bar{\boldsymbol{x}}\| < \delta$. Let $M = \max \{\|\boldsymbol{R}_f(\boldsymbol{x}) - \bar{\boldsymbol{R}}_s(\boldsymbol{x}, \boldsymbol{p}(\boldsymbol{x}))\| : \boldsymbol{x} \in X_B\}$. Suppose further that the function $\boldsymbol{p}(\boldsymbol{x}) = \arg \min_{\boldsymbol{z} \in X_p} \sum_{k=1}^{N} w_k(\boldsymbol{x}) \| \boldsymbol{R}_f(\boldsymbol{x}^k) - \bar{\boldsymbol{R}}_s(\boldsymbol{x}^k, \boldsymbol{z})\|$ [cf. (7)] is Lipschitz continuous, i.e., $\|\boldsymbol{p}(\boldsymbol{x}) - \boldsymbol{p}(\boldsymbol{y})\| \leq L_3 \|\boldsymbol{x} - \boldsymbol{y}\|$, where $L_3 \in R_+$. We then have the following estimate for the modeling error on X_R

$$\|\boldsymbol{R}_f(\boldsymbol{x}) - \boldsymbol{R}_s(\boldsymbol{x})\| \leq M + K\delta, \qquad \boldsymbol{x} \in X_R \qquad (29)$$

where $K = L_1 + L_2 L_3$.

Proof of Theorem 2: Let $\boldsymbol{x}$ be any point in X_R, and $\bar{\boldsymbol{x}} \in X_B$ be such that $\|\boldsymbol{x} - \bar{\boldsymbol{x}}\| < \delta$. We have

$$
\begin{aligned}
&\|R_f(\boldsymbol{x}) - R_s(\boldsymbol{x})\| \\
&= \|R_f(\boldsymbol{x}) - \bar{R}_s(\boldsymbol{x}, p(\boldsymbol{x}))\| \\
&\leq \|R_f(\boldsymbol{x}) - \bar{R}_s(\boldsymbol{x}, p(\bar{\boldsymbol{x}}))\| + \|\bar{R}_s(\boldsymbol{x}, p(\bar{\boldsymbol{x}})) - \bar{R}_s(\boldsymbol{x}, p(\boldsymbol{x}))\| \\
&\leq \|(R_f(\boldsymbol{x}) - \bar{R}_s(\boldsymbol{x}, p(\bar{\boldsymbol{x}}))) - (R_f(\bar{\boldsymbol{x}}) - \bar{R}_s(\bar{\boldsymbol{x}}, p(\bar{\boldsymbol{x}}))) \\
&\quad + (R_f(\bar{\boldsymbol{x}}) - \bar{R}_s(\bar{\boldsymbol{x}}, p(\bar{\boldsymbol{x}})))\| + \|\bar{R}_s(\boldsymbol{x}, p(\bar{\boldsymbol{x}})) - \bar{R}_s(\boldsymbol{x}, p(\boldsymbol{x}))\| \\
&\leq \|\Delta R(\boldsymbol{x}, p(\bar{\boldsymbol{x}})) - \Delta R(\bar{\boldsymbol{x}}, p(\bar{\boldsymbol{x}}))\| \\
&\quad + \|R_f(\bar{\boldsymbol{x}}) - \bar{R}_s(\bar{\boldsymbol{x}}, p(\bar{\boldsymbol{x}}))\| + \|\bar{R}_s(\boldsymbol{x}, p(\bar{\boldsymbol{x}})) - \bar{R}_s(\boldsymbol{x}, p(\boldsymbol{x}))\|.
\end{aligned}
\tag{30}
$$

According to Assumption 2 (i), we have

$$
\|\Delta R(\boldsymbol{x}, p(\bar{\boldsymbol{x}})) - \Delta R(\bar{\boldsymbol{x}}, p(\bar{\boldsymbol{x}}))\| \leq L_1 \|\boldsymbol{x} - \bar{\boldsymbol{x}}\|.
\tag{31}
$$

From the assumption of the theorem, we get

$$
\|R_f(\bar{\boldsymbol{x}}) - \bar{R}_s(\bar{\boldsymbol{x}}, p(\bar{\boldsymbol{x}}))\| \leq M.
\tag{32}
$$

Finally, it follows from Assumption 2 (ii) and the assumptions of the theorem that we have

$$
\begin{aligned}
\|\bar{R}_s(\boldsymbol{x}, p(\bar{\boldsymbol{x}})) - \bar{R}_s(\boldsymbol{x}, p(\boldsymbol{x}))\| & \\
\leq L_2 \|p(\bar{\boldsymbol{x}}) - p(\boldsymbol{x})\| & \leq L_2 L_3 \|\boldsymbol{x} - \bar{\boldsymbol{x}}\|.
\end{aligned}
\tag{33}
$$

Now, we have from (30)–(33) that

$$
\|R_f(\boldsymbol{x}) - R_s(\boldsymbol{x})\| \leq M + (L_1 + L_2 L_3) \cdot \delta.
\tag{34}
$$

This ends the proof of the theorem.

Remark 7: Theorem 2 says, in fact, that the modeling error is proportional to the characteristic distance of the base set and to the error at base points. On the other hand, for any fixed δ, one can reduce M to as small a value as desirable by proper choice of C (provided that Assumption 1 (iv) is satisfied, i.e., $\|R_f(\boldsymbol{x}) - \bar{R}_s(\boldsymbol{x}, p^s(\boldsymbol{x}))\| = 0$ at the base points), although this would affect the constant L_3 (usually in an undesirable way if C is too small).

IV. CONCLUSION

A theoretical justification of the recently published SM-based modeling methodology with variable weight coefficients has been presented. We have provided theoretical results showing that the new methodology can assure any accuracy that is required (provided that the base set is dense enough), which is not the case for the standard methodology. We have also given an error estimate for the surrogate model with respect to the fine model. Examples have demonstrated the fundamental differences between the standard and novel modeling method.

REFERENCES

[1] J. W. Bandler, R. M. Biernacki, S. H. Chen, P. A. Grobelny, and R. H. Hemmers, "Space mapping technique for electromagnetic optimization," *IEEE Trans. Microw. Theory Tech.*, vol. 42, no. 12, pp. 2536–2544, Dec. 1994.

[2] J. W. Bandler, R. M. Biernacki, S. H. Chen, R. H. Hemmers, and K. Madsen, "Electromagnetic optimization exploiting aggressive space mapping," *IEEE Trans. Microw. Theory Tech.*, vol. 43, no. 12, pp. 2874–2882, Dec. 1995.

[3] M. H. Bakr, J. W. Bandler, and N. Georgieva, "Modeling of microwave circuits exploiting space derivative mapping," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Anaheim, CA, Jun. 1999, pp. 715–718.

[4] J. W. Bandler, N. Georgieva, M. A. Ismail, J. E. Rayas-Sánchez, and Q. J. Zhang, "A generalized space mapping tableau approach to device modeling," *IEEE Trans. Microw. Theory Tech.*, vol. 49, no. 1, pp. 67–79, Jan. 2001.

[5] J. W. Bandler, M. A. Ismail, J. E. Rayas-Sánchez, and Q. J. Zhang, "Neuromodeling of microwave circuits exploiting space mapping technology," *IEEE Trans. Microw. Theory Tech.*, vol. 47, no. 12, pp. 2417–2427, Dec. 1999.

[6] L. Zhang, J. J. Xu, M. Yagoub, R. T. Ding, and Q. J. Zhang, "Neuro-space mapping technique for nonlinear device modeling and large signal simulation," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Philadelphia, PA, Jun. 2003, pp. 173–176.

[7] V. K. Devabhaktuni, B. Chattaraj, M. C. E. Yagoub, and Q.-J. Zhang, "Advanced microwave modeling framework exploiting automatic model generation, knowledge neural networks, and space mapping," *IEEE Trans. Microw. Theory Tech.*, vol. 51, no. 7, pp. 1822–1833, Jul. 2003.

[8] L. Zhang, J. Xu, M. C. E. Yagoub, R. Ding, and Q.-J. Zhang, "Efficient analytical formulation and sensitivity analysis of neuro-space mapping for nonlinear microwave device modeling," *IEEE Trans. Microw. Theory Tech.*, vol. 53, no. 9, pp. 2752–2767, Sep. 2005.

[9] S. Koziel, J. W. Bandler, A. S. Mohamed, and K. Madsen, "Enhanced surrogate models for statistical design exploiting space mapping technology," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Long Beach, CA, Jun. 2005, pp. 1609–1612.

[10] J. W. Bandler, Q. S. Cheng, and S. Koziel, "Implementable space mapping approach to enhancement of microwave device models," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Long Beach, CA, Jun. 2005, pp. 1139–1142.

[11] S. Koziel and J. W. Bandler, "Space-mapping-based modeling utilizing parameter extraction with variable weight coefficients and a data base," presented at the IEEE MTT-S Int. Microw. Symp., San Francisco, CA, Jun. 2006.

[12] M. H. Bakr, J. W. Bandler, K. Madsen, and J. Søndergaard, "An introduction to the space mapping technique," *Optim. Eng.*, vol. 2, no. 4, pp. 369–384, Dec. 2001.

Slawomir Koziel (M'03) was born in Poland, in 1970. He received the M.Sc. and Ph.D. (with honors) degrees in electronic engineering from Gdansk University of Technology, Gdansk, Poland, in 1995 and 2000, respectively, and the M.Sc. degree in theoretical physics and mathematics and Ph.D. degree in mathematics (with honors) from the University of Gdansk, Gdansk, Poland, in 2000, 2002, and 2003, respectively.

He is currently a Research Associate with the Simulation Optimization Systems Research Laboratory, Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada. He has authored or coauthored over 90 papers. His research interests include SM-based modeling and optimization, circuit theory, analog signal processing, active filter design, evolutionary computation, and numerical analysis.

John W. Bandler (S'66–M'66–SM'74–F'78) was born in Jerusalem on November 9, 1941. He received the B.Sc. (Eng.), Ph.D., and D.Sc. (Eng.) degrees from the University of London, London, U.K., in 1963, 1967, and 1976, respectively.

He joined Mullard Research Laboratories, Redhill, Surrey, U.K., 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, ON, Canada, in 1969. He was 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 has authored or coauthored over 385 papers. He was a member of the Micronet Network of Centres of Excellence. He was President of Optimization Systems Associates Inc. (OSA), which he founded in 1983, until November 20, 1997, the date of acquisition of OSA by the Hewlett-Packard Company. OSA implemented a first-generation yield-driven microwave computer-aided design (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, and Empipe3D and EmpipeExpress in 1996. OSA created the product *empath* in 1996, which was marketed by Sonnet Software Inc. He is President of Bandler Corporation, Dundas, ON, Canada, which he founded in 1997. He joined the Editorial Boards of the *International Journal of Numerical Modelling* in 1987, the *International Journal of Microwave and Millimeterwave 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 on "Optimization-Oriented Microwave CAD" (1997). He was Guest Coeditor of the *Optimization and Engineering* Special Issue on "Surrogate Modelling and Space Mapping for Engineering Optimization" (2001).

Dr. Bandler is a Fellow of the Canadian Academy of Engineering, the Royal Society of Canada, the Institution of Electrical Engineers (IEE), U.K., and the Engineering Institute of Canada. He is a member of the Association of Professional Engineers of the Province of Ontario, Canada, and the Massachusetts Institute of Technology (MIT) Electromagnetics Academy. He received the Automatic Radio Frequency Techniques Group Automated Measurements Career Award in 1994 and the IEEE Microwave Theory and Techniques Society (IEEE MTT-S) Microwave Application Award in 2004. He was an associate editor of the IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES (1969–1974) and has continued serving as a member of the Editorial Board. He was guest editor of the IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES Special Issue on "Computer-Oriented Microwave Practices" (1974) and "Automated Circuit Design Using Electromagnetic Simulators" (1997) and guest coeditor of the Special Issue on "Process-Oriented Microwave CAD and Modeling" (1992) and "Electromagnetics-Based Optimization of Microwave Components and Circuits" (2004). He was chair of the MTT-1 Technical Committee on Computer-Aided Design.

Kaj Madsen was born in Denmark, in 1943. He received the cand.scient. degree in mathematics from the University of Aarhus, Aarhus, Denmark, in 1968, and the Dr.Techn. degree from the Technical University of Denmark (DTU), Lyngby, Denmark, in 1986.

From 1968 to 1988, he was a Lecturer in numerical analysis, apart from the 1973–1974 academic year, when he was with AERE Harwell, Didcot, U.K. Most of his career has been spent with the Department for Numerical Analysis, DTU. From 1981 to 1983, he was with the Computer Science Department, Copenhagen University, Copenhagen, Denmark. During Summer 1978, he visited McMaster University, Hamilton, ON, Canada. In 1988, be became a Full Professor. Since the 1990s, he has arranged several international workshops on linear programming, parallel algorithms, surrogate modeling, and SM. In 1993, he joined the Department of Mathematical Modelling, DTU, and during 1995–2000 was Head of that department. In 2000, he took an active part in forming the new department Informatics and Mathematical Modelling, DTU, which includes computer science and applied mathematics. Since January 2001, he has been Head of that department. His primary fields of interest in teaching and research are nonlinear optimization, including SM techniques and global optimization, and validated computing using interval analysis.