Reliable Microwave Modeling by Means of Variable-Fidelity Response Features

Slawomir Koziel, *Senior Member, IEEE*, and John W. Bandler, *Life Fellow, IEEE*

***Abstract*—In this work, methodologies for low-cost and reliable microwave modeling are presented using variable-fidelity response features. The two key components of our approach are: 1) a realization of the modeling process at the level of suitably selected feature points of the responses (e.g., S-parameters versus frequency) of the structure at hand and 2) the exploitation of variable-fidelity EM simulation data, also for the response feature representation. Due to the less nonlinear dependence between the coordinates of the feature points on the geometrical parameters of the structure of interest, the amount of training data can be greatly reduced. Additional cost reduction is obtained by means of generating the majority of the training data at a coarse-discretization EM simulation level and exploiting the correlations between the EM models of various fidelities. We propose two ways of combining the low- and high-fidelity data sets: 1) an external approach, through space mapping (simpler to implement) and 2) an internal approach, using co-kriging (more flexible and potentially offering better accuracy). The operation and performance of our modeling techniques are demonstrated by three microstrip filter examples and a compact rat-race coupler. A comprehensive verification and comparisons with several benchmark techniques, as well as application examples (filter optimization) are also provided.**

***Index Terms*—Co-kriging, computer-aided design, feature-based modeling, kriging, microwave component modeling, space mapping (SM), surrogates modeling.**

I. Introduction

ACCURATE evaluation of the electrical performance of microwave structures can be obtained through high-fidelity full-wave electromagnetic (EM) analysis. Unfortunately, this comes at considerable computational cost, particularly for complex devices/circuits and when interactions (EM couplings) with the environment are included. Consequently, the utilization of EM simulations in a design process, e.g., for parametric optimization, uncertainty quantification, or tolerance-aware design (design centering), may be prohibitive, as multiple evaluations of the structure at hand are involved. Executing such tasks in a reasonable timeframe requires fast and accurate replacement models (surrogates).

Fast surrogate models can be constructed using two classes of techniques: 1) response surface approximation (RSA) [1] and 2) physics-based surrogate modeling [2], [3]. The first relies on approximating sampled high-fidelity EM simulation data. The most popular methods include neural networks [4], [5], kriging interpolation [6], support vector regression [7], [8], radial-basis function interpolation [9], the Cauchy method [10], as well as Gaussian process regression (GPR) [11], [12]. The advantages of the RSA surrogates include their versatility (as data-driven models they are easily transferable between various problem domains) and speed (once established, the RSA model is computationally cheap to evaluate). On the other hand, approximation models are not suitable for handling multi-dimensional parameter spaces. If the number of parameters exceeds just a few, the amount of training data necessary to ensure sufficient accuracy of the model (typically, below 5% of the relative RMS error [13]) grows very quickly so that the effort for model construction may not be practically justified (unless the model is to be reused under various design scenarios) or even feasible.

The second class of techniques for constructing fast surrogates—physics-based modeling—relies on appropriate correction of an underlying low-fidelity model such as an equivalent circuit (popular method: space mapping (SM) [14], [15]). Physics-based surrogates require less training data and-due to the problem-specific knowledge embedded in the low-fidelity model-offer better generalization. However, they are less generic, more complex to implement, and their applicability is typically limited to cases when fast low-fidelity models are available; their accuracy depends on the reliability of the low-fidelity model; and it might not be straightforward to accommodate additional training data (if available) [16]. To some extent, these issues can be alleviated by combining SM with an approximation-based correction layer (e.g., [17], [18]).

Reduction of the number of training points for approximation-based surrogates can be achieved by realizing the modeling process in an alternative representation of the system responses, where the dependence of the alternative responses on the designable parameters is less nonlinear. This approach has been explored, e.g., in the shape-preserving response prediction (SPRP) technique [19] or in [20] for inverse modeling of filters. A recent modeling technique [21] utilizes the concept of feature points

Manuscript received June 19, 2015; revised September 03, 2015; accepted October 20, 2015. Date of publication November 11, 2015; date of current version December 02, 2015. This work was supported in part by the Icelandic Centre for Research (RANNIS) under Grant 130450051 and Grant 141272051, and in part by the Natural Sciences and Engineering Research Council of Canada under Grant RGPIN7239-11 and Grant STPGP447367-13, and in part by Bandler Corporation. This paper is an expanded version from the IEEE MTT-S International Microwave Symposium, Phoenix, AZ, USA, May 17–22, 2015.

S. Koziel is with the School of Science and Engineering, Reykjavík University, 101 Reykjavík, Iceland, and also with the Faculty of Electronics, Telecommunications, and Informatics, Gdańsk University of Technology, 80-233 Gdańsk, Poland (e-mail: koziel@ru.is).

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

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMTT.2015.2495266

similar to those in SPRP but with considerably simpler implementation (achieved by abandoning the use of so-called reference designs [21]). Feature-based modeling has been demonstrated to ensure good accuracy using a fraction of the training points required by conventional methods [21].

As indicated in [21], further reduction of the setup cost of the surrogate model can be achieved by using variable-fidelity EM simulation within the feature-based modeling framework. In [22], densely sampled low-fidelity EM model data was supplemented by sparsely sampled high-fidelity EM data with the two data sets blended (at the level of response features) into a single surrogate using co-kriging [23].

In this work, we provide extensive numerical validation of this approach. In addition, in Section II, we propose an alternative (external) way of combining low- and high-fidelity feature-based surrogates using SM. The latter is simpler and easier to implement than [22] while ensuring similar accuracy. Our test cases provided in Section III include three microstrip filter examples and a compact rat-race coupler. Application of the variable-fidelity feature-based models for design optimization is also demonstrated.

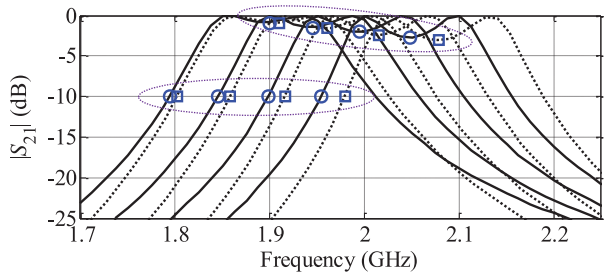

Fig. 1. Family of $|S_{21}|$ responses for a microstrip bandpass filter evaluated along a selected line segment $(1-t)\boldsymbol{x}^a + t\boldsymbol{x}^b$, $0 \le t \le 1$: high-fidelity model $\boldsymbol{R}_f$ (—) and low-fidelity model $\boldsymbol{R}_{cd}(\cdots)$. Selected feature points and groups of corresponding points marked (o) for $\boldsymbol{R}_f$ and □ for $\boldsymbol{R}_{cd}$.

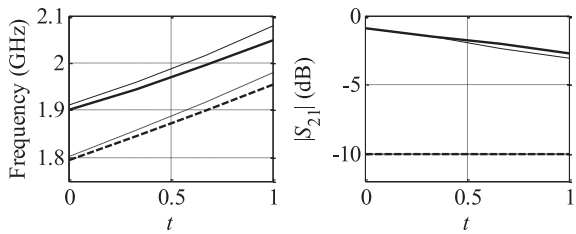

Fig. 2. Selected feature point plots between designs $\boldsymbol{x}^a(t=0)$ and $\boldsymbol{x}^b(t=1)$: (a) frequency and (b) levels. They correspond to two feature points: center frequency of the filter (– – –) and −10 dB level on the left-hand side of the passband (—); thick and thin lines are used for high- and low-fidelity model feature points, respectively.

II. Variable-Fidelity Feature-Based Modeling

In this section, we formulate the surrogate modeling problem, recall the concept of feature-based surrogates, and redefine the concept within a variable-fidelity setting.

A. Surrogate Modeling

We denote by $\boldsymbol{R}_f : X \to R^m$, $X \subseteq R^n$ the response vector of the microwave device of interest. In particular, $\boldsymbol{R}_f(\boldsymbol{x})$ may represent $|S_{21}|$ at m chosen frequencies, ω_1 to ω_m, i.e., $\boldsymbol{R}_f(\boldsymbol{x}) = [R_f(\boldsymbol{x},\omega_1)\ldots R_f(\boldsymbol{x},\omega_m)]^T$. $\boldsymbol{R}_f$ is assumed to be evaluated using high-fidelity EM analysis. Consequently, it is computationally expensive. The task is to build a fast surrogate (replacement) model $\boldsymbol{R}_s$ that represents $\boldsymbol{R}_f$ in X. Given sufficient accuracy of $\boldsymbol{R}_s$, it can be used in place of $\boldsymbol{R}_f$ for solving design tasks that require multiple, high-cost evaluations of the latter.

Let $X_T = \{\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^N\} \subset X$ be the training set so that the responses of the high-fidelity model $\boldsymbol{R}_f(\boldsymbol{x}^k)$ at $\boldsymbol{x}^k$, are known. Conventional response surface modeling attempts to directly model $\boldsymbol{R}_f(\boldsymbol{x},\omega_j)$, $j = 1,\ldots,m$. In many cases, the surrogate is created as an ensemble of RSA models constructed for individual frequencies, i.e., obtained by approximating the data sets $\{(\boldsymbol{x}^k, \boldsymbol{R}_f(\boldsymbol{x}^k,\omega_j)\}_{k=1,\ldots,N}$, for $j = 1,\ldots,m$. Sometimes [23], frequency is treated as an additional designable parameter, so that the RSA surrogate is constructed by approximating the data pairs $\{([(\boldsymbol{x}^k)^T\omega_l]^T, \boldsymbol{R}_f(\boldsymbol{x}^k,\omega_l)\}_{k=1,\ldots,N;l\in\{1,\ldots,m\}}$.

B. Variable-Fidelity Response Features

Given the high nonlinearity of typical responses of microwave devices with respect to their designable parameters, particularly for filters, the direct modeling of the high-fidelity model responses $\boldsymbol{R}_f(\boldsymbol{x},\omega_j)$ is a challenging task that requires large data sets using $\{(\boldsymbol{x}^k, \boldsymbol{R}_f(\boldsymbol{x}^k,\omega_j)\}_{k=1,\ldots,N}$ (cf. Section II-A) and which is virtually impossible in high-dimensional design spaces.

The key concept behind the modeling techniques considered here that allow reduction of the number of training points in the modeling process are certain response features [21]. The feature points (cf. Fig. 1) may include points corresponding to specific response levels (e.g., −10 dB, −3 dB), as well as those allocated in between fixed-level points (e.g., uniformly spaced in frequency). As indicated in Fig. 2 the dependence of the feature points on the design parameters is much less nonlinear than those of the original responses (here, S-parameters), and thus easier to model. Feature-based modeling was originally introduced in [21]. It relies on extracting the feature points from the sampled EM-simulated responses at the training locations, constructing the RSA models of the individual feature points, and synthesizing the surrogate response at a design of interest from these RSA models.

In this work, we utilize training data acquired from variable-fidelity simulations: from sparsely-sampled $\boldsymbol{R}_f$ points $X_{Tf} = \{\boldsymbol{x}_f^1, \boldsymbol{x}_f^2, \ldots, \boldsymbol{x}_f^{Nf}\}$ and from densely-sampled data obtained from coarse-discretization EM simulations (low-fidelity model $\boldsymbol{R}_{cd}$), $X_{Tc} = \{\boldsymbol{x}_c^1, \boldsymbol{x}_c^2, \ldots, \boldsymbol{x}_c^{Nc}\}$. Although $\boldsymbol{R}_{cd}$ and $\boldsymbol{R}_f$ are misaligned (cf. Fig. 1), they are also well correlated so that the initial surrogate model obtained from the $\boldsymbol{R}_{cd}$ data can be enhanced by using a few $\boldsymbol{R}_f$ points to construct the accurate, final surrogate model. We use the notation $\boldsymbol{f}_{k.f}{}^j = [\omega_{k.f}{}^j\ l_{k.f}{}^j]^T$, $j = 1,\ldots,K$, and $k = 1,\ldots,N_f$ to denote the jth feature point of $\boldsymbol{R}_f(\boldsymbol{x}_f^k)$, and $\boldsymbol{f}_{k.c}{}^j = [\omega_{k.c}{}^j l_{k.c}{}^j]^T$ to denote the jth feature point of $\boldsymbol{R}_{cd}(\boldsymbol{x}_c^k)$; $\omega_{k.f}{}^j$ and $l_{k.f}{}^j$ denote the frequency and magnitude (level) components of $\boldsymbol{f}_{k.f}{}^j$ (similarly for $\boldsymbol{f}_{k.c}{}^j$).

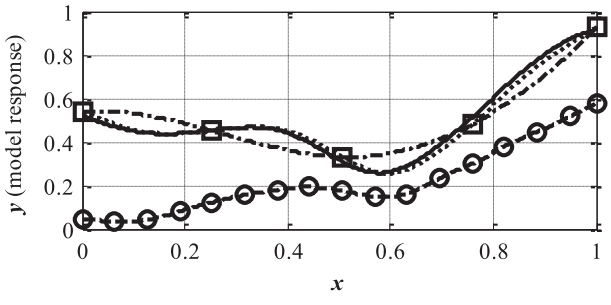

Fig. 3. Co-kriging concept: $\boldsymbol{R}_f$ model (—), $\boldsymbol{R}_c$ model (– – –), $\boldsymbol{R}_f$ model samples (□), $\boldsymbol{R}_c$ model samples (○). Kriging interpolation of $\boldsymbol{R}_f$ model samples ($-\cdot-$) is not an adequate representation of the $\boldsymbol{R}_f$ model (limited data set). Co-kriging interpolation ($\cdots$) of blended $\boldsymbol{R}_c$ and $\boldsymbol{R}_f$ data provides better accuracy at a lower computational cost.

C. Variable-Fidelity Feature-Based Modeling: The Internal Approach

Multi-fidelity feature-based modeling is a two-step process. First, we construct approximation surrogates $s_{\omega.j}(\boldsymbol{x})$ and $s_{l.j}(\boldsymbol{x})$, $j = 1, \ldots, K$, corresponding to the feature points. In the internal approach presented in this section, the construction of $s_{\omega.j}(\boldsymbol{x})$ and $s_{l.j}(\boldsymbol{x})$ is based on both N_f high-fidelity and N_c low-fidelity training points and their corresponding feature points $\{\boldsymbol{f}_{1.f}{}^j, \boldsymbol{f}_{2.f}{}^j, \ldots, \boldsymbol{f}_{N_f.f}{}^j\}$, and $\{\boldsymbol{f}_{1.c}{}^j, \boldsymbol{f}_{2.c}{}^j, \ldots, \boldsymbol{f}_{N_c.c}{}^j\}$, $j = 1, \ldots, K$. Blending these two types of data sets is realized through co-kriging [20] as described in the next paragraph.

Let $\boldsymbol{f}(X_{Tc})$ be the set of responses associated with the training set X_{Tc} (i.e., low-fidelity feature points $\boldsymbol{f}_{1.c}$ through $\boldsymbol{f}_{Nf.c}$). The kriging interpolant is given as [13]

$$\boldsymbol{s}_{KR}(\boldsymbol{x}) = M\alpha + r(\boldsymbol{x}) \cdot \Psi^{-1} \cdot (\boldsymbol{f}(X_{Tc}) - F\alpha) \tag{1}$$

where M and F are Vandermonde matrices of the test point $\boldsymbol{x}$ and the base set X_{Tc}, respectively; α is determined by generalized least squares (GLS), $r(\boldsymbol{x})$ is an $1 \times N_f$ vector of correlations between the point $\boldsymbol{x}$ and the base set X_{Tf}, where the entries are $r_i(\boldsymbol{x}) = \psi(\boldsymbol{x}, \boldsymbol{x}_f{}^i)$, and Ψ is a $N_f \times N_f$ correlation matrix, whose entries are given by $\Psi_{i,j} = \psi(\boldsymbol{x}_{KR}{}^i, \boldsymbol{x}_{KR}{}^j)$. We use the exponential correlation function $\psi(\boldsymbol{x}, \boldsymbol{x'}) = \exp(\Sigma_{k=1,\ldots,n} - \theta_k |x_k - x'_k|)$. The regression function is constant, $F = [1 \ldots 1]^T$ and $M = (1)$. Co-kriging is a type of kriging where the $\boldsymbol{R}_f$ and $\boldsymbol{R}_{cd}$ model data are combined to enhance the prediction accuracy (cf. Fig. 3). Co-kriging is a two-step process: first a kriging model $\boldsymbol{s}_{KRc}$ of the coarse data $(X_{Tc}, \boldsymbol{c}(X_{Tc}))$ is constructed and, on the residuals of the fine data $(X_{Tf}, \boldsymbol{R}_d)$, a second kriging model $\boldsymbol{s}_{KRd}$ is applied, where $\boldsymbol{R}_d = \boldsymbol{f}(X_{Tf}) - \rho \cdot \boldsymbol{c}(X_{Tf})$; $\boldsymbol{c}(X_{Tf})$ can be approximated as $\boldsymbol{c}(X_{Tf}) \approx \boldsymbol{s}_{KRc}(X_{Tf})$. The co-kriging interpolant is defined as [25]

$$\boldsymbol{s}_{CO}(\boldsymbol{x}) = M\alpha + r(\boldsymbol{x}) \cdot \Psi^{-1} \cdot (\boldsymbol{R}_d - F\alpha). \tag{2}$$

Definitions of M, F, $r(\boldsymbol{x})$, and Ψ can be found in [25].

The multi-fidelity feature-based surrogate (the internal approach) is defined as

$$\boldsymbol{R}_s(\boldsymbol{x}) = [R_s(\boldsymbol{x}, \omega_1) \ldots R_s(\boldsymbol{x}, \omega_K)]^T \tag{3}$$

with

$$R_s(\boldsymbol{x}, \omega_j) = I(\Omega(\boldsymbol{x}), L(\boldsymbol{x}), \omega_j) \tag{4}$$

where $L(\boldsymbol{x}) = [s_{l.1}(\boldsymbol{x}) \ldots s_{l.K}(\boldsymbol{x})]$ and $\Omega(\boldsymbol{x}) = [s_{\omega.1}(\boldsymbol{x}) \ldots s_{\omega.k}(\boldsymbol{x})]$ are predicted feature point locations corresponding to $\boldsymbol{x}$, the design being considered; ω_j, $j = 1, \ldots, m$, is a discrete set of frequencies at which the response of the structure is being evaluated (cf. Section II-A). According to the internal approach, both $s_{\omega.j}(\boldsymbol{x})$ and $s_{l.j}(\boldsymbol{x})$ are implemented through co-kriging as in (2). $I(\Omega, L, \omega)$ denotes a function that interpolates the level vector L and frequency vector Ω into the response at a given frequency ω_j.

Because $\omega_k^j(\boldsymbol{x})$ and $l_k^j(\boldsymbol{x})$ (i.e., frequencies and levels of the feature points) are less nonlinear than the original responses $\boldsymbol{R}_f(\boldsymbol{x}, \omega_j)$, a substantially smaller number of training points is necessary to ensure faithful modeling. Also, excellent correlation between $\boldsymbol{R}_{cd}$ and $\boldsymbol{R}_f$ (cf. Fig. 2) allows for further reduction of the surrogate model setup cost because a very limited number of $\boldsymbol{R}_f$ samples is sufficient to elevate the $\boldsymbol{R}_{cd}$-based kriging model to high-fidelity accuracy through co-kriging.

D. Variable-Fidelity Feature-Based Modeling: The External Approach

The internal approach presented in Section II-C combines the low- and high-fidelity training data at the level of the feature points. In the external approach outlined below, the high-fidelity data is included through a SM correction of the initial feature-based surrogate model $\boldsymbol{R}_{s.c}(\boldsymbol{x})$ obtained as in (1), (2), however, using the low-fidelity training data only. The SM correction is realized at the level of the original responses as follows [15]:

$$\boldsymbol{R}_s(\boldsymbol{x}) = \boldsymbol{A} \cdot \boldsymbol{R}_{s.c.F}(\boldsymbol{B} \cdot \boldsymbol{x} + \boldsymbol{c}; \boldsymbol{F}) \tag{5}$$

where $\boldsymbol{R}_{s.c.F}$ is a frequency scaled model $\boldsymbol{R}_{s.c}(\boldsymbol{x})$ such that $\boldsymbol{R}_{s.c.F}(\boldsymbol{x}; \boldsymbol{F}) = [R_{s.c}(\boldsymbol{x}, f_0 + f_1\omega_1) \ldots R_{s.c}(\boldsymbol{x}, f_0 + f_1\omega_m)]^T$, with $R_{s.c}(\boldsymbol{x}, \omega_j)$, $j = 1, \ldots, m$, being components of $\boldsymbol{R}_{s.c}(\boldsymbol{x})$. Here, the model parameters $\boldsymbol{A}$ (diagonal matrix), $\boldsymbol{B}$ (square matrix), $\boldsymbol{c}$ (column vector), and $\boldsymbol{F}$ are obtained from the standard parameter extraction procedure needed by SM, namely,

$$[\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{c}, \boldsymbol{F}] = \arg \min_{[\overline{\boldsymbol{A}}, \overline{\boldsymbol{B}}, \overline{\boldsymbol{c}}, \overline{\boldsymbol{F}}]} \sum_{k=1}^{N_f} \| \boldsymbol{R}_f(\boldsymbol{x}_f^k) - \overline{\boldsymbol{A}} \cdot \boldsymbol{R}_{s.c.F}(\overline{\boldsymbol{B}} \cdot x_f^k + \overline{\boldsymbol{c}}; \overline{\boldsymbol{F}}) \|. \tag{6}$$

If the correlation between the low- and high-fidelity models is sufficient (which is normally the case when both models are based on EM analysis), the correction given by (5) and (6) should significantly improve the accuracy of the surrogate.

It should also be noted that the computational cost of surrogate model identification [i.e., solving the parameter extraction process (6)] can be neglected compared to the high-fidelity data acquisition because it is realized at the level of a fast feature-based model $\boldsymbol{R}_{s.c}$.

E. Internal Versus External Approach

Apart from considerable conceptual differences between the internal and external approaches, i.e., blending in the high-fidelity model data at the level of the response features rather than

at the level of the original responses, the external approach is simpler to implement. On the other hand, the internal approach can be more flexible because the inclusion of a larger amount of high-fidelity training data should lead to an improvement of model accuracy. This may not be the case for the external approach (beyond a certain number of points) due to the fixed number of degrees of freedom of the surrogate (5) (the number of model parameters depends on the problem's dimensionality but not on the cardinality of the data set).

III. Verification Examples

In this section, we provide a comprehensive benchmarking of the multi-fidelity feature-based modeling techniques of Section II and demonstrate applications of the feature-based surrogate models in design optimization.

A. Test Cases and Experimental Setup

In order to illustrate the operation and performance of the proposed modeling techniques we consider three microstrip filter examples and a compact coupler. The first example (Filter 1) is the stacked slotted resonators bandpass filter [26] shown in Fig. 4(a). The high-fidelity filter model is simulated in Sonnet ***em*** using a grid of 0.05 mm × 0.05 mm. The low-fidelity model is also simulated in Sonnet on a 2 mm × 2 mm grid. The substrate parameters are thickness $h = 0.635$ mm, and permittivity $\epsilon_r = 9.7$. The designable parameters are $\boldsymbol{x} = [L_1 L_2 W_1 S_1 S_2 D]^T$. The region of interest is defined as the interval $[\boldsymbol{x}^0 - \boldsymbol{d}, \boldsymbol{x}^0 + \boldsymbol{d}]$ with $\boldsymbol{x}^0 = [6.0\ 9.6\ 1.0\ 1.0\ 2.0\ 2.0]^T$ and $\boldsymbol{d} = [0.8\ 0.8\ 0.2\ 0.2\ 0.4\ 0.4]^T$.

The second structure (Filter 2) is the fourth-order ring resonator bandpass filter [27] shown in Fig. 4(b). The high-fidelity filter model is simulated in FEKO using 952 triangular meshes. The low-fidelity FEKO model utilizes 174 meshes. The substrate parameters are thickness $h = 1.52$ mm, and permittivity $\epsilon_r = 4.32$. The designable parameters are $\boldsymbol{x} = [L_1 L_2 L_3 S_1 S_2 W_1 W_2]^T$. The region of interest is defined as the interval $[\boldsymbol{x}^0 - \boldsymbol{d}, \boldsymbol{x}^0 + \boldsymbol{d}]$ with $\boldsymbol{x}^0 = [24.0\ 20.0\ 25.0\ 0.2\ 0.2\ 1.0\ 0.5]^T$ and $\boldsymbol{d} = [1.0\ 1.0\ 1.0\ 0.1\ 0.1\ 0.2\ 0.2]^T$.

The last filter structure considered here (Filter 3) is the microstrip bandpass filter with open stub inverter [28]shown in Fig. 4(c). The high-fidelity filter model is simulated in FEKO using 432 triangular meshes. The low-fidelity FEKO model utilizes 112 meshes. The substrate parameters are thickness $h = 0.508$ mm, and permittivity $\epsilon_r = 2.6$. The designable parameters are $\boldsymbol{x} = [L_1 L_2 L_3 S_1 S_2 W_1]^T$. The region of interest is defined as the interval $[\boldsymbol{x}^0 - \boldsymbol{d}, \boldsymbol{x}^0 + \boldsymbol{d}]$ with $\boldsymbol{x}^0 = [24.0\ 10.0\ 2.0\ 0.6\ 0.2\ 0.5]^T$ and $\boldsymbol{d} = [2.0\ 2.0\ 1.0\ 0.4\ 0.1\ 0.4]^T$.

The final test structure is a folded rat-race coupler (RRC) [29] shown in Fig. 4(d). The structure is implemented on RF-35 substrate ($\epsilon_r = 3.5$, $h = 0.762$ mm, $\tan\delta = 0.018$). The designable parameters are given by: $\boldsymbol{x} = [l_1 l_2 l_3 dw]^T$, with $w_0 = 1.7$, $l_0 = 15$ fixed (all dimensions in millimeters). The high- and low-fidelity models of the structure are both implemented in CST Microwave Studio ($\sim 220,000$ mesh cells, simulation time 15 min for $\boldsymbol{R}_f$ and 8000 mesh cells, simulation time 20 s for $\boldsymbol{R}_{cd}$). The region of interest is defined as the interval $[\boldsymbol{x}^0 - \boldsymbol{d}, \boldsymbol{x}^0 + \boldsymbol{d}]$ with $\boldsymbol{x}^0 = [5.0\ 15.0\ 22.0\ 1.0\ 0.8]^T$ and $\boldsymbol{d} = [1.0\ 2.0\ 2.0\ 0.2\ 0.2]^T$.

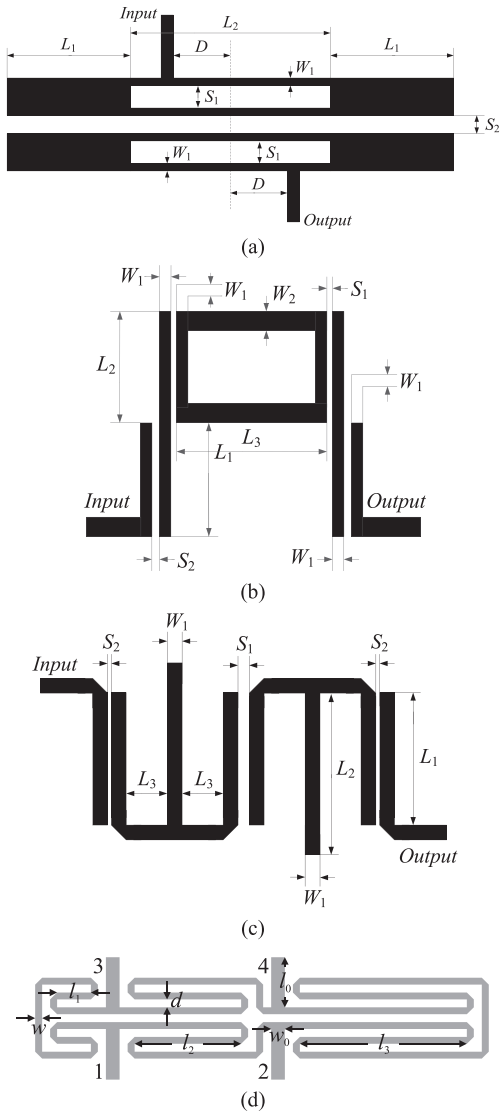

Fig. 4. Filter structures used for feature-based modeling verification: (a) stacked slotted resonators filter [26]; (b) fourth-order ring resonator bandpass filter [27]; (c) bandpass filter with open stub inverter [28]; (d) rat-race coupler [29].

Model accuracy is verified using the relative error measure $\|\boldsymbol{R}_f(\boldsymbol{x}) - \boldsymbol{R}_s(\boldsymbol{x})\| / \|\boldsymbol{R}_f(\boldsymbol{x})\|$ expressed in percent and averaged over 100 random test designs. The multi-fidelity feature-based models (both the internal and external versions) are compared with the following modeling methods:

- regular (single-fidelity) feature-based modeling [21];
- generalized shape-preserving response prediction (GSPRP) [19];
- direct kriging interpolation of the high-fidelity data [6];
- response surface modeling using radial-basis functions [9].

The kriging model utilizes a Gaussian correlation function [6], whereas the radial-basis function model uses Gaussian basis functions [6]. The length-scale parameter of the latter is optimized using cross-validation [6]. For all the above modeling

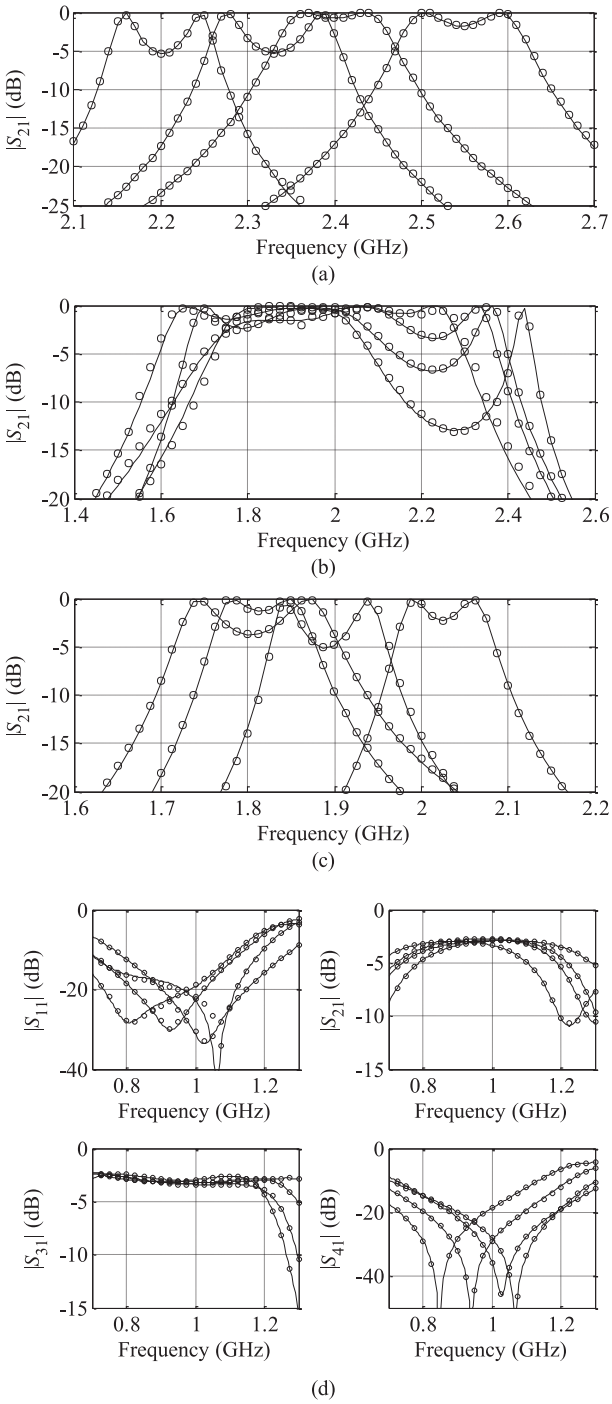

Fig. 5. Model responses at the selected (random) test designs: high-fidelity model (—) and multi-fidelity feature-based model (set up with 20 $\boldsymbol{R}_f$ and 200 $\boldsymbol{R}_{cd}$ points) (o): (a) Filter 1; (b) Filter 2; and (c) Filter 3; and (d) RRC.

methods, five different cases are considered with the number of training points varying between $N = 20$ and $N = 400$.

B. Numerical Results and Comparisons With Benchmark Methods

The results are gathered in Tables I–IV. Fig. 5 shows the high-fidelity and feature-based model responses at selected test points for Filters 1 to 3 and for the RRC. The following observations can be made.

TABLE I
MODELING RESULTS FOR FILTER 1

Modeling Method	Average Error				
	N^* = 20	N = 50	N = 100	N = 200	N = 400
Multi-Fidelity Feature-Based (Internal Approach)[$]	0.5% [35.0][&]	0.4% [65.0][&]	0.35% [115.0][&]	0.3% [215.0][&]	0.25% [415.0][&]
Multi-Fidelity Feature-Based (External Approach)[$$]	2.2%	1.1%	1.0%	0.8%	0.8%
Feature-Based [%%]	3.9%	1.8%	0.9%	0.6%	0.4%
GSPRP [%]	4.0%	1.7%	0.7%	0.5%	0.3%
Direct Kriging[#]	15.6%	11.9%	11.0%	9.3%	6.7%
Direct RBF[##]	17.2%	13.5%	11.8%	9.9%	8.7%

[*] N stands for the number of training points.
[$] (Internal) multi-fidelity feature-based modeling of Section II.*C* (200 $\boldsymbol{R}_{cd}$ samples).
[&] Numbers in brackets refer to the total model setup cost (including $\boldsymbol{R}_{cd}$ points).
[$$] (External) multi-fidelity feature-based modeling of Section II.*D*.
[%] Generalized SPRP modeling [19].
[%%] Feature-based modeling using the procedure of [21].
[#] Direct kriging interpolation of the high-fidelity model $|S_{21}|$ responses.
[##] Direct radial-basis function interpolation of the high-fidelity model data.

TABLE II
MODELING RESULTS FOR FILTER 2

Modeling Method	Average Error				
	N^* = 20	N = 50	N = 100	N = 200	N = 400
Multi-Fidelity Feature-Based (Internal Approach)[$]	3.7% [28.3][&]	1.5% [58.3][&]	1.3% [108.3][&]	1.2% [208.3][&]	1.2% [408.3][&]
Multi-Fidelity Feature-Based (External Approach)[$$]	4.7%	3.6%	2.7%	2.6%	2.5%
Feature-Based [%%]	11.6%	5.1%	3.8%	3.6%	2.5%
GSPRP [%]	8.8%	5.4%	4.2%	4.0%	3.0%
Direct Kriging[#]	13.5%	7.8%	6.3%	4.6%	3.6%
Direct RBF[##]	21.2%	19.3%	15.5%	13.3%	11.5%

[*] N stands for the number of training points.
[$] (Internal) multi-fidelity feature-based modeling of Section II.*C* (200 $\boldsymbol{R}_{cd}$ samples).
[&] Numbers in brackets refer to the total model setup cost (including $\boldsymbol{R}_{cd}$ points).
[$$] (External) multi-fidelity feature-based modeling of Section II.*D*.
[%] Generalized SPRP modeling [19].
[%%] Feature-based modeling using the procedure of [21].
[#] Direct kriging interpolation of the high-fidelity model $|S_{21}|$ responses.
[##] Direct radial-basis function interpolation of the high-fidelity model data.

- Both the internal and external multi-fidelity feature-based models ensure excellent accuracy even with a very small number of high-fidelity training samples (specifically, 20 and 50).
- The internal feature-based modeling approach is generally better than the external approach, however, the latter is still considerably better than a single-level feature-based surrogate for a small number of training samples and comparable or better overall.
- Asymptotically (i.e., for the number of high-fidelity training points increasing to 400), both multi-fidelity feature-based modeling methods are comparable or better than the single-fidelity feature-based models and ones based on GSPRP.
- All modeling approaches exploiting the concept of response features are considerably more accurate than

TABLE III
MODELING RESULTS FOR FILTER 3

Modeling Method	Average Error				
	$N^{*}=20$	$N=50$	$N=100$	$N=200$	$N=400$
Multi-Fidelity Feature-Based (Internal Approach)[$]	0.7% [28.3][&]	0.6% [58.3][&]	0.5% [108.3][&]	0.45% [208.3][&]	0.4% [408.3][&]
Multi-Fidelity Feature-Based (External Approach)[$]	1.1%	0.9%	0.7%	0.65%	0.6%
Feature-Based[%%]	7.5%	2.4%	1.2%	0.7%	0.6%
GSPRP[%]	7.7%	2.7%	1.8%	1.4%	1.2%
Direct Kriging[#]	11.6%	8.8%	6.1%	4.8%	3.1%
Direct RBF[##]	16.7%	15.2%	12.7%	10.4%	9.3%

[*] N stands for the number of training points.
[$] (Internal) multi-fidelity feature-based modeling of Section II.*C* (200 $\boldsymbol{R}_{cd}$ samples).
[&] Numbers in brackets refer to the total model setup cost (including $\boldsymbol{R}_{cd}$ points).
[$$] (External) multi-fidelity feature-based modeling of Section II.*D*.
[%] Generalized SPRP modeling [19].
[%%] Feature-based modeling using the procedure of [21].
[#] Direct kriging interpolation of the high-fidelity model $|S_{21}|$ responses.
[##] Direct radial-basis function interpolation of the high-fidelity model data.

TABLE IV
MODELING RESULTS FOR THE RRC

Modeling Method	Average Error[**]				
	$N^{*}=20$	$N=50$	$N=100$	$N=200$	$N=400$
Multi-Fidelity Feature-Based (Internal Approach)[$]	6.9% [24.4][&]	4.4% [74.4][&]	2.9% [124.4][&]	2.1% [224.4][&]	2.0% [424.4][&]
Multi-Fidelity Feature-Based (External Approach)[$$]	10.8%	6.3%	4.5%	4.1%	3.9%
Feature-Based[%%]	8.3%	5.2%	3.9%	3.3%	3.2%
GSPRP[%]	9.8%	5.9%	4.2%	3.4%	3.1%
Direct Kriging[#]	13.3%	8.6%	7.4%	4.8%	4.1%
Direct RBF[##]	19.2%	16.6%	13.3%	9.8%	8.1%

[**] Relative RMS error averaged over the four responses $|S_{11}|$, $|S_{21}|$, $|S_{31}|$, $|S_{41}|$.
[*] N stands for the number of training points.
[$] (Internal) multi-fidelity feature-based modeling of Section II.*C* (200 $\boldsymbol{R}_{cd}$ samples).
[&] Numbers in brackets refer to the total model setup cost (including $\boldsymbol{R}_{cd}$ points).
[$$] (External) multi-fidelity feature-based modeling of Section II.*D*.
[%] Generalized SPRP modeling [19].
[%%] Feature-based modeling using the procedure of [21].
[#] Direct kriging interpolation of the high-fidelity model $|S_{21}|$ responses.
[##] Direct radial-basis function interpolation of the high-fidelity model data.

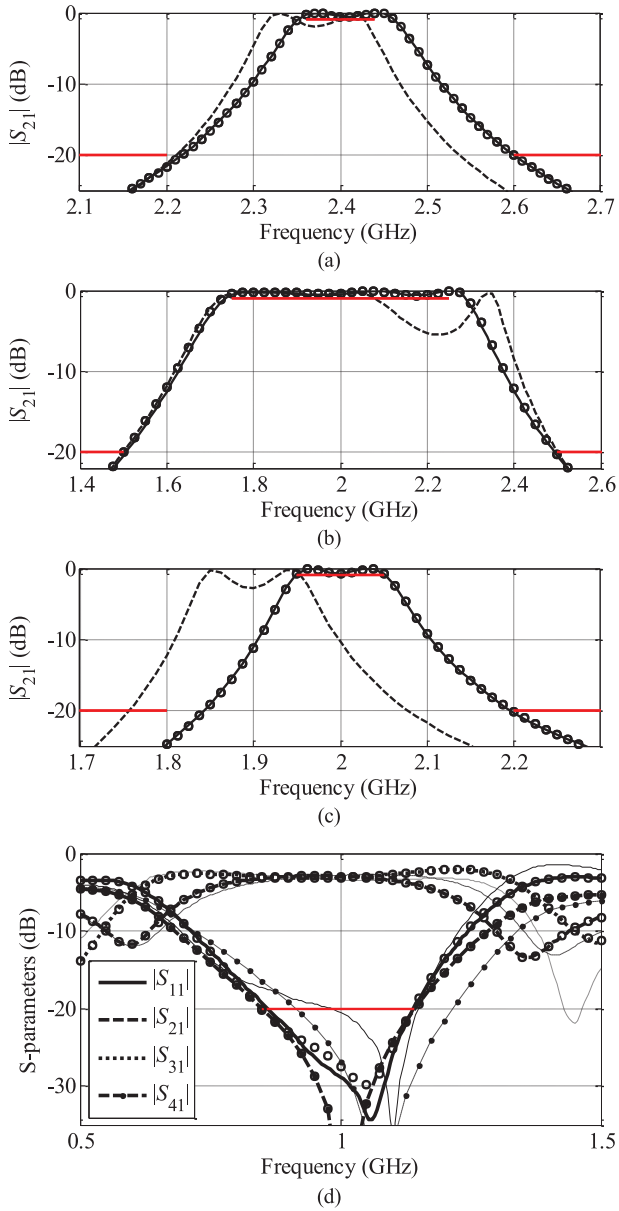

Fig. 6. Optimization results: (a) Filter 1; (b) Filter 2; (c) Filter 3; (d) RRC. For (a)–(c), responses shown as (– – –) and (—) for the multi-fidelity feature-based surrogate model responses (the internal approach) at the initial design and at the optimized design; corresponding verification by high-fidelity model responses shown as (o). For (d), thin lines show surrogate responses at the initial design, thick lines show the surrogate responses as the optimized design; corresponding verification by high-fidelity model responses shown as (o); 280-MHz bandwidth of the optimized coupler marked with horizontal line.

conventional modeling techniques working directly with the original system responses, here, S-parameters versus frequency.

The above conjectures are consistent throughout all the test cases considered in this paper. It should also be emphasized that the benchmarking is quite comprehensive, both with respect to the competitive methods (four techniques) and with respect to the training set size (from 20 to 400 samples).

Perhaps the most important point is that—according to the authors' knowledge—multi-fidelity feature-based modeling (both the internal and external approaches) are the only methods that result in excellent (and practically usable) accuracy for an extremely small number of training points (20 and 50 samples).

At the same time, one should bear in mind the limitations of the method, namely, the necessity of maintaining consistency of the feature points across the entire training set. For certain structures, such as the ones utilized in this work, as well as other structures with well-defined response "shapes" (e.g., coupling structures, narrow-band antennas, certain integrated photonic components such as microring resonators) it is easy to achieve. For other structures, such as high-order filters, feature-based modeling may be the method of choice for local modeling for, e.g., statistical design purposes [30].

C. Application Examples: Design Optimization

As an additional verification, the multi-fidelity feature-based surrogate models have been utilized for parametric optimization of the Filters 1 through 3 as well as the RRC.

The following design specifications are considered.

- Filter 1: $|S_{21}| \geq -1$ dB for 2.35 GHz $\leq \omega \leq$ 2.45 GHz, $|S_{21}| \leq -20$ dB for $\omega \leq$ 2.2 GHz and $\omega \geq$ 2.6 GHz.
- Filter 2: $|S_{21}| \geq -1$ dB for 1.75 GHz $\leq \omega \leq$ 2.25 GHz, $|S_{21}| \leq -20$ dB for $\omega \leq$ 1.5 GHz and $\omega \geq$ 2.5 GHz.
- Filter 3: $|S_{21}| \geq -1$ dB for 1.95 GHz $\leq \omega \leq$ 2.05 GHz, $|S_{21}| \leq -20$ dB for $\omega \leq$ 1.8 GHz and $\omega \geq$ 2.2 GHz.
- RRC: Obtain an equal power split, i.e., $|S_{21}| = |S_{31}|$ at the operating frequency of $f_0 = 1$ GHz; simultaneously maximize the -20 dB bandwidth (symmetrically) around f_0 for $|S_{11}|$ and $|S_{41}|$.

In all cases, the multi-fidelity feature-based surrogate constructed using the internal approach and 50 high-fidelity training samples (200 low-fidelity samples) has been used in the process. The initial and final responses obtained by optimizing the feature-based surrogate model (the final design is verified by the high-fidelity model) are shown in Fig. 6. The design specifications for the filter structures are marked using thick horizontal lines. Because of the very good accuracy of the surrogates, no further design tuning was found necessary. In the case of RRC, the bandwidth of the optimized coupler is 280 MHz with the power split error < 0.2 dB at 1 GHz.

IV. Conclusion

Variable-fidelity feature-based techniques for low-cost surrogate modeling of microwave structures have been proposed. Reduction of the computational cost associated with setting up surrogate models has been achieved by combining two basic components: 1) the exploitation of certain feature points, which allows us to move the modeling process to an alternative representation of the system response, where the functional landscape is much less nonlinear than for the original responses (in particular, the frequency-dependent S-parameters) and 2) the utilization of variable-fidelity EM simulations, where an initial surrogate created with densely sampled coarse-discretization EM simulation data is enhanced by sparsely sampled high-fidelity EM data. Two approaches to blending the variable-fidelity EM data into the final surrogate have been proposed, i.e., an internal one (based on co-kriging at the level of the feature points), and an external one (based on SM).

As demonstrated by three microstrip filters and a rat race coupler example and comparisons with several benchmark techniques, both of our multi-fidelity feature-based approaches outperform not only conventional approximation modeling methods but also feature-based approaches that exploit single-fidelity EM simulations. Significant improvement of the predictive power of the surrogate is especially observed for small high-fidelity training sets. This opens new opportunities for construction of quasi-global surrogates for applications such as parametric design optimization (also demonstrated in this work).

According to our knowledge, no surrogate modeling technique reported in the literature so far exhibits comparable performance. At the same time, one needs to bear in mind the limitations of the method, namely, the necessity of maintaining consistency of the feature point sets across the surrogate model domain. Consequently, our method is less versatile than general-purpose approximation techniques. On the other hand, with careful definition of the response features, as well as for numerous cases where the system response is well-defined in terms of its shape (microwave couplers, low-order filters, narrow-band antennas, phased array antennas, various classes of integrated photonic devices, wireless power transfer systems, etc.) but also higher-order filters in terms of local modeling for statistical/robust design application and uncertainty quantification, multi-fidelity feature-based modeling may be the method of choice for rapid construction of fast, accurate and reusable surrogates.

References

[1] T. W. Simpson, J. Peplinski, P. N. Koch, and J. K. Allen, "Metamodels for computer-based engineering design: Survey and recommendations," *Eng. With Comput.*, vol. 17, no. 2, pp. 129–150, Jul. 2001.

[2] J. W. Bandler, Q. S. Cheng, S. A. Dakroury, A. S. Mohamed, M. H. Bakr, K. Madsen, and J. Søndergaard, "Space mapping: The state of the art," *IEEE Trans. Microw. Theory Techn.*, vol. 52, no. 1, pp. 337–361, Jan. 2004.

[3] S. Koziel and J. W. Bandler, "Recent advances in space-mapping-based modeling of microwave devices," *Int. J. Numer. Modelling*, vol. 23, no. 6, pp. 425–446, Nov. 2010.

[4] H. Kabir, Y. Wang, M. Yu, and Q. J. Zhang, "Neural network inverse modeling and applications to microwave filter design," *IEEE Trans. Microw. Theory Techn.*, vol. 56, no. 4, pp. 867–879, Apr. 2008.

[5] Y. Cao, X. Chen, and G. Wang, "Dynamic behavioral modeling of nonlinear microwave devices using real-time recurrent neural network," *IEEE Trans. Electron Devices*, vol. 56, no. 5, pp. 1020–1026, May 2009.

[6] N. V. Queipo, R. T. Haftka, W. Shyy, T. Goel, R. Vaidynathan, and P. K. Tucker, "Surrogate-based analysis and optimization," *Progr. Aerosp. Sci.*, vol. 41, no. 1, pp. 1–28, Jan. 2005.

[7] L. Xia, J. Meng, R. Xu, B. Yan, and Y. Guo, "Modeling of 3-D vertical interconnect using support vector machine regression," *IEEE Microw. Wirel. Comp. Lett.*, vol. 16, no. 12, pp. 639–641, Dec. 2006.

[8] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statist. Comput.*, vol. 14, no. 3, pp. 199–222, Aug. 2004.

[9] M. D. Buhmann and M. J. Ablowitz, *Radial Basis Functions: Theory and Implementations*. Cambridge, U.K.: Cambridge University, 2003.

[10] A. G. Lamperéz, P. K. Sarkar, and M. S. Palma, "Generation of accurate rational models of lossy systems using the Cauchy method," *IEEE Microw. Wirel. Comp. Lett.*, vol. 14, no. 10, pp. 490–493, Oct. 2014.

[11] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, 2006.

[12] J. P. Jacobs and J. P. De Villiers, "Gaussian-process-regression-based design of ultrawide-band and dual-band CPW-fed slot antennas," *J. Electromagn. Waves Appl.*, vol. 24, pp. 1763–1772, 2010.

[13] I. Couckuyt, F. Declercq, T. Dhaene, H. Rogier, and L. Knockaert, "Surrogate-based infill optimization applied to electromagnetic problems," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 20, no. 5, pp. 492–501, Sep. 2010.

[14] 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 Techn.*, vol. 49, no. 1, pp. 67–79, Jan. 2001.

[15] J. W. Bandler, Q. S. Cheng, and S. Koziel, "Simplified space mapping approach to enhancement of microwave device models," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 16, no. 5, pp. 518–535, Sep. 2006.

[16] S. Koziel, J. W. Bandler, and K. Madsen, "Theoretical justification of space-mapping-based modeling utilizing a data base and on-demand parameter extraction," *IEEE Trans. Microw. Theory Techn.*, vol. 54, no. 12, pp. 4316–4322, Dec. 2006.

[17] S. Koziel and J. W. Bandler, "A space-mapping approach to microwave device modeling exploiting fuzzy systems," *IEEE Trans. Microw. Theory Techn.*, vol. 55, no. 12, pp. 2539–2547, Dec. 2007.

[18] S. Koziel and J. W. Bandler, "Modeling of microwave devices with space mapping and radial basis functions," *Int. J. Numer. Modell.*, vol. 21, no. 3, pp. 187–203, May 2008.
[19] S. Koziel and L. Leifsson, "Generalized shape-preserving response prediction for accurate modeling of microwave structures," *IET Microw., Ant. Prop.*, vol. 6, no. 12, pp. 1332–1339, Sep. 2012.
[20] E. Menargues, S. Cogollos, V. E. Boria, B. Gimeno, and M. Guglielmi, "An efficient computer-aided design procedure for interpolating filter dimensions using least squares methods," in *Eur. Microw. Integr. Circuits Conf.*, 2012, pp. 250–253.
[21] S. Koziel, J. W. Bandler, and Q. S. Cheng, "Low-cost feature-based modeling of microwave structures," presented at the IEEE MTT-S Int. Microw. Symp., Tampa, FL, USA, 2014.
[22] S. Koziel and J. W. Bandler, "Accurate modeling of microwave structures using variable-fidelity response features," presented at the IEEE MTT-S Int. Microw. Symp., Phoenix, AZ, USA, 2015.
[23] M. C. Kennedy and A. O'Hagan, "Predicting the output from complex computer code when fast approximations are available," *Biometrika*, vol. 87, pp. 1–13, 2000.
[24] J. P. Jacobs and S. Koziel, "Two-stage framework for efficient Gaussian process modeling of antenna input characteristics," *IEEE Trans. Antennas Propag.*, vol. 62, no. 2, pp. 706–713, Feb. 2014.
[25] S. Koziel, S. Ogurtsov, I. Couckuyt, and T. Dhaene, "Variable-fidelity electromagnetic simulations and co-kriging for accurate modeling of antennas," *IEEE Trans. Antennas Propag.*, vol. 61, no. 3, pp. 1301–1308, Mar. 2013.
[26] C. L. Huang, Y. B. Chen, and C. F. Tasi, "New compact microstrip stacked slotted resonators bandpass filter with transmission zeros using high-permittivity ceramics substrate," *Microw. Opt. Tech. Lett.*, vol. 50, no. 5, pp. 1377–1379, May 2008.
[27] M. K. M. Salleh, G. Pringent, O. Pigaglio, and R. Crampagne, "Quarter-wavelength side-coupled ring resonator for bandpass filters," *IEEE Trans. Microw. Theory Techn.*, vol. 56, no. 1, pp. 156–162, Jan. 2008.
[28] J. R. Lee, J. H. Cho, and S. W. Yun, "New compact bandpass filter using microstrip $\lambda/4$ resonators with open stub inverter," *IEEE Microw. Guided Wave Lett.*, vol. 10, no. 12, pp. 526–527, Dec. 2000.
[29] S. Koziel, A. Bekasiewicz, P. Kurgan, and J. W. Bandler, "Expedited multi-objective design optimization of miniaturized microwave structures using physics-based surrogates," presented at the IEEE MTT-S Int. Microw. Symp., Phoenix, AZ, USA, 2015.
[30] S. Koziel and J. W. Bandler, "Rapid yield estimation and optimization of microwave structures exploiting feature-based statistical analysis," *IEEE Trans. Microw. Theory Techn.*, vol. 63, no. 1, pp. 107–114, Jan. 2015.

Slawomir Koziel (M'03–SM'07) received the M.Sc. and Ph.D. degrees in electronic engineering, the M.Sc. degree in theoretical physics, and the M.S. degree in mathematics from Gdańsk University of Technology, Gdańsk, Poland, in 1995, 2000, 2000, and 2002, respectively, and the Ph.D. degree in mathematics from the University of Gdańsk, Gdańsk, Poland, in 2003.

He is currently a Professor with the School of Science and Engineering, Reykjavík University, Reykjavík, Iceland. He is also a Visiting Professor with Gdańsk University of Technology. His research interests include CAD and modeling of microwave circuits, simulation-driven design, surrogate-based optimization, space mapping, circuit theory, analog signal processing, evolutionary computation, and numerical analysis.

John W. Bandler (S'66–M'66–SM'74–F'78–LF'06) studied at Imperial College, London, U.K., and 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 McMaster University, Hamilton, ON, Canada, in 1969, where he is now a Professor Emeritus. He founded Optimization Systems Associates Inc. in 1983 and sold it to Hewlett-Packard in 1997. He is President of Bandler Corporation, Dundas, ON, Canada.

Dr. Bandler is a Fellow of several societies, including the Royal Society of Canada and the Canadian Academy of Engineering. In 2004, he was a recipient of the IEEE MTT-S Microwave Application Award. He was a recipient of the IEEE Canada McNaughton Gold Medal in 2012, the Queen Elizabeth II Diamond Jubilee Medal in 2012, the IEEE MTT-S Microwave Career Award in 2013, and the McMaster University's Faculty of Engineering Research Achievement Award in 2014.